



CBS

COPENHAGEN
BUSINESS SCHOOL

HANDELSHØJSKOLEN

MSC IN BUSINESS ADMINISTRATION AND INFORMATION SYSTEMS

MASTER'S THESIS

Big Data – A Source for Organizational Knowledge Creation and Support in Decision Making

An exploratory case study of the concept and practical use of Big Data in Danish local government

AUTHORS

LASSE VINTER

ANDERS MÅRTENSSON

HAND-IN DATE

15th of May 2017

SUPERVISOR

Helle Zinner Henriksen

NUMBER OF PAGES

105 / 237.676 characters

Preface

This thesis is submitted as a partial fulfillment of the requirements for obtaining the Master in Business Administration and Information Systems at Copenhagen Business School (CBS). The project started November 15th 2016 and was completed May 15th 2017, having a work load of 30 ECTS points per author.

The thesis rises from interest to utilize big data as a source for new organizational knowledge creation and support in decision making in Danish local government. We chose to pursue this research project because very little is known about the actual outcome of the practical use of big data in the public sector. The identified gaps in literature justified us with a need to study the actual effects of big data applications to enable a better understanding of the complex concept and the possibilities and limitations it encompasses. Researching the phenomenon proved to be very challenging in spite of the fact that it is rather new, yet we managed to gather the needed empirical material to succeed with the project.

We would like to thank Helle Zinner Henriksen, who has provided us with support and guidance throughout the project.

Copenhagen, May 15th, 2017

Lasse Vinter
lavi12ab

Anders Mårtensson
anma12ao

Abstract

This master's thesis is concerned with the practical use of big data in Danish local government. The primary aim of this thesis is to investigate the actual effects of big data as a source for new organizational knowledge creation and support in decision making, which most literature generally consider it has potential to provide and enable. The secondary aims are to assess and add understanding of the complex concept of big data and to present an example of the practical use of it.

The study attempts to establish insights on how big data creates new organizational knowledge and supports decision making in local government context, seeing that little is known about the actual outcome of the practical use of it in the public sector. Hence, the study operates under the pragmatic paradigm, whose focus is concerned with action, intervention and constructive knowledge. Empirical data have been collected through the qualitative research methods of partially-structured interviews and participant observations to enable a rich understanding of the concept of big data and the practical use of it in Danish local government.

As such, this master's thesis contributes with an in-depth account of how big data is used in Danish local government with the case of Residential Social Monitoring. The findings both support and extend current big data literature. Even though big data has received a great deal of attention in recent years, this study shows a variety of meanings and reflects a wide range of what the concept encompasses. The practical usage at Svendborg Municipality shows that big data is able to create new organizational knowledge as 'large' datasets offer a higher form of intelligence that can generate new insights. With new knowledge, the ability and capacity to make better decisions are now possible, seeing decisions now can be made on a greater informed basis. Apart from using big data for informational purposes and thus to decrease uncertainty, big data also reduces equivocality by creating shared understandings of choice situations. Despite the great opportunities that follows big data, a number of organizational challenges and IT-related requirements exist. To harvest the full potential, these challenges and requirements must be addressed. The study ends by concluding that big data shows great utility for Svendborg Municipality and the rest of the Danish public sector.

Table of Contents

1.0 Introduction	9
1.1 Research Background	10
1.2 Motivation of the Research	11
1.3 Research Aims and Research Questions	12
1.4 Theoretical and Methodological Considerations	13
1.5 Structure of the Thesis	14
2.0 Theoretical Grounding.....	16
2.1 Introduction	16
2.2 The Dimensions of Typology for Literature Reviews	16
2.2.1 Review Ambition.....	17
2.2.2 Breadth of Review	17
2.2.3 Article Identification.....	18
2.2.4 Analytical Approach.....	19
2.3 The Concept of Big Data	21
2.4 Knowledge Creation Theory	24
2.4.1 What is Knowledge?.....	25
2.4.2 Explicit and Tacit Knowledge	25
2.4.3 Big Data – A Source for Organizational Knowledge Creation	28
2.5 Decision Making Theory	32
2.5.1 What is a Decision?	33
2.5.2 Perfect Rationality	35
2.5.3 Bounded Rationality	36
2.5.4 Big Data towards Perfect Rationality	38
3.0 Research Methodology	41
3.1 Introduction	41
3.2 Paradigmatic Underpinnings	41
3.3 Research Design	43
3.4 Geospatial Big Data Analytics	44
3.5 Case: Residential Social Monitoring	45
3.5.1 Technical Components	46

3.5.2 Access to the Case Study.....	47
3.5.3 Selection of Participants	48
3.6 Data Collection.....	49
3.6.1 Partially-Structured Interviews.....	51
3.6.2 Participant Observation	52
3.6.3 Reports and Articles	53
3.7 Data Analysis.....	54
4.0 Empirical Findings	57
4.1 Introduction	57
4.2 The Concept of Big Data	57
4.3 Big Data in the Danish Public Sector	62
4.4 Residential Social Monitoring	69
4.5 Knowledge Creation and Predictive Insights	74
4.6 Decision Making and Strategic Planning	78
4.7 Targeted Social Services and Evidence-Based Practice.....	83
4.8 Data Challenges and Privacy Issues	86
5.0 Discussion	93
5.1 Introduction	93
5.2 Implications for Research and Contributions	93
5.2.1 Findings that Support and Extend the Conceptualization of Big Data	93
5.2.2 Findings that Support and Extend Big Data Literature and Knowledge Creation.....	94
5.2.3 Findings that Support and Extend Big Data Literature and Decision Making	95
5.3 Reflections on the Research Process	96
5.3.1 Conducting Qualitative Research	96
5.3.2 Evaluating Design-Science Research	98
6.0 Concluding Remarks.....	101
6.1 Conclusion.....	101
6.2 Suggestions for Future Research	103
7.0 References	105

List of Figures

Figure 1 – The research methodology	13
Figure 2 – How big data is described in literature	23
Figure 3 – The SECI process	27
Figure 4 – DIKW pyramid in the context of big data management	30
Figure 5 – Environmental change effects on organisational processes	32
Figure 6 – Three separate phases of a decision process	33
Figure 7 – The human as a rational decision maker	36
Figure 8 – The human as a bounded rational decision maker	37
Figure 9 – The thesis's methodological basis.....	41
Figure 10 – Technical components in Residential Social Monitoring	46
Figure 11 – Data fields in FLIS	47

List of Tables

Table 1 – Dimensions of typology for literature review.....	16
Table 2 – Implications of applying different review ambitions	17
Table 3 – Results from database inquiry	18
Table 4 – Abstract of concept matrix	20
Table 5 – Summary of the components and applied approaches.....	20
Table 6 – The study's participants.....	49
Table 7 – An abstract of the interview guide.....	52
Table 8 – An abstract of the secondary material used.....	54
Table 9 – An abstract of the data matrix.	55
Table 10 – Data challenges and privacy issues	92
Table 11 – Reflections on the qualitative evaluation criterias.....	98
Table 12 – Design-science research guidelines	99

List of Appendices

8.A Concept matrix	116
8.B Interview guide	118
8.C Reports and articles.....	120
8.D Data matrix	121
8.E Transcribed interview – Hedvig Hvarregaard Andersen (HHA)	132
8.F Transcribed interview – Sara Thestrup Hansen (STH)	132
8.G Transcribed interview – Jan Carlsson (JC).....	132
8.H Transcribed interview – Mikkel Næsvang (MN).....	132
8.I Transcribed interview – Astrid Krarup Andersen (AKA).....	132
8.J Transcribed interview – Troels Rasmussen (TR).....	132

1.0 Introduction

Big data has received a great deal of attention in recent years. As a result of digitization and datafication (Schönberger & Cukier, 2014), big data is projected to be one of the most important and profound technology in the digital age (Boyd & Crawford, 2012). In the digital age, vast amount of data from and about people, things and interactions between them are being generated. Everything we do is being data traced. It is estimated that the amount of data produced in just two days is as much as the whole human history up to the year of 2003 (Schmidt & Cohen, 2014). Not only is the data revolution fascinating, but new technologies, thanks to increased processing power, are now able to analyze the large amounts of data to extract new insights and new forms of value, in ways that change markets, organizations, the relationships between citizens and governments, and more (Schönberger & Cukier, 2014).

Big data has already swept private businesses for several years (Thapa, 2016). In the Danish industry, big data has been acknowledged as a growth factor and have provided significant effects such as production- and warehouse optimization, better pricing, additional sales, new services, faster product development as well as more concentrated marketing campaigns (Irisgroup, 2013). Yet, big data is not characterized as an emerging technology anymore. Big data was taken down of Gartner's Hype Cycle in 2013, seeing it has become prevalent in our lives (Woodie, 2015). Despite the fact that big data is not exemplified as an emerging technology anymore, it has only just in the recent years reached the public sector, driven first and foremost by management consultancies and technology companies (Thapa, 2016). Branded as 'data-driven government', implying a more encompassing administrative paradigm beyond business-as-usual plus running data analysis (Ibid.), big data is considered to be one of the most important technologies for the future of public administrations (Pollitt, 2014).

The potentials of using big data in the public sector are expected to be great, given the fact that public administrations produce and collect a vast amount of data in order to perform their daily tasks, making the public sector one of the economy's most data-intensive sectors (OECD, 2015). In spite of the vast amount of data available and the analytical opportunities that follow with it, the public sector is becoming increasingly aware of the potential value to be gained from it (Munné, 2016; Bassi et al., 2014) and realizing that data is a strategic asset that needs to be protected and leveraged (O'Brien, 2012). According to a report from McKinsey Global Institute, big data has the potential to increase internal performance and forecasts a total of \$250 billion annual value for European public sector

administrations (Manyika et al., 2011, p. 2). The report further highlights the fact that big data is already widely used throughout the private sector and others need to follow, such as governments, whom also may employ it. Accordingly, mining data¹ in new ways can help facilitate and generate products and services like identification of emerging governmental and societal needs, thus contributing in a variety of ways to improve performance within the public sector and across the economy (OECD, 2015). In terms of this, there seems to be no doubt that the use of big data holds great potential for the public sector and with an expected 15-20% reduction on the operating budgets (Ibid.), the adoption of it may nearly seem like an obligation in times of recession and budgetary constraints. The great potentials of both economic and societal opportunities may explain the massive investments that governments and countries around the world currently initiate to capture the full potential of big data. United States, United Kingdom, France and Japan are only a few of the global economies whom are spending up to \$200 million in infrastructure, tools and resources to manage big data (Gamage, 2016).

1.1 Research Background

Following the global trend of introducing and applying big data in the public sector, the Danish government have recently launched its new digitization strategy, 'Lokal og digital – et sammenhængende Danmark' (KL, 2015a). The strategy is formulated as a political vision for digitalization and for the municipalities' common work towards the year of 2020. A part of the strategy involves that Danish municipalities must pursue to create efficiency and added value by the use of data. It is recognized that data is the most valuable material as a result of the digitization and contains great potential to target service for citizens and businesses, as well as to improve the process of municipal budgeting. Local authorities and the rest of the public sector hold much data about citizens and businesses, which are being registered as part of carrying out their duties. It is a global trend that the amount of data increases in all sectors (Ibid.).

It is expected that the opportunities from clever use of data can lead to better service in terms of more personal and targeted efforts that suits citizens' current life situation and needs. That being said, by sharing data across service areas, it is possible to analyze more complex relationships than before, offering greater degree of evidence. As a result, citizens can reasonably expect the effort that works

¹ The process of analyzing data into useful information and knowledge

the best. For the local council and executive boards, data can help create and support decision making as well as the ability to identify and react to trends in advance (Ibid.).

Despite the opportunities, several challenges and barriers are recognized in the strategy. Firstly, it is recognized that data is not standardized, making it difficult to use in analytical settings. Data standardization is an essential condition to fulfill the potential of digitization so data can be shared across relevant authorities or combined with other data to produce new knowledge. Data standardization is a major task that often requires many partners' participation and support. Secondly, some IT-suppliers takes well paid for the municipalities to access their own data. Therefore, municipalities must constantly be aware of that data can be accessible through open interfaces when making new agreements with suppliers. Thirdly, it is important to ensure that the regulatory environment is lasting, so the possibilities of sharing data across service areas is not unnecessarily constrained by outdated laws that troubles to keep up with the technological developments. Lastly, not only is sharing data a technical, financial and legal challenge, but the ability to exploit new technology is inextricably linked with the ability to create change in the organization and culture. Only when that happens is it possible to create real added value for citizens, businesses and municipalities (Ibid.).

1.2 Motivation of the Research

The idea of becoming a 'data-driven government' by introducing big data makes instinctive sense and the development relates great to both the global trend of applying data analytics and the newly formed Danish digitization strategy. However, despite the increasing interest in big data, a clear understanding of the concept seem yet to be absent in current literature (Chen et al., 2014; Schönberger & Cukier, 2014; Demchenko et al., 2013). Furthermore, while there exist a good deal of literature on the potential contributions and value in terms of improved internal performance and economic savings, as well as the challenges and barriers to adopt big data, little is known about the actual outcome of the practical use of big data in the public sector, "[...] *Big Data is still in its infancy and many important questions regarding the true value of Big Data remain unanswered*" (Desouza & Jacob, 2014). Desouza & Jacob (2014) conclude by requesting scholars to look at how data currently is being used and assess the degree to which it is being underutilized. That the perception of big data is valuable but the true value remains unanswered, is in relation to the findings by Fosso Wamba et al. (2015). From a conducted literature review on the subject, Fosso Wamba et al. (2015) present that only three articles, or seven percent of all publication, were identified focusing on

government, “[...] *very few empirical studies have been conducted to assess the real potential of big data*” (ibid.).

Thus, the gaps in literature justify a need to study the actual effects of big data applications to enable a better understanding of the complex concept and the possibilities and limitations it encompasses. In this thesis, big data is studied in Danish local government, as one level of the public sector. Here, big data is considered a source for new organizational knowledge creation and support in decision making, operationalized through the empirical variables of effectiveness and efficiency, both of which are seen important aspects of measuring the performance of the public sector's output (Andrews & Entwistle, 2010; Marieta et al., 2010).

1.3 Research Aims and Research Questions

Based on the preliminary text, the primary aim of this research is to study the actual effects of how big data creates new organizational knowledge and supports decision making at the common initiative of becoming a data-driven government. The secondary aims of this research are to (1) assess and add understanding to the concept of big data, and (2) provide an example of the practical use.

In order to fulfill these aims of the research, a primary research question has been formulated as follows:

- How does big data create new organizational knowledge and support decision making in Danish local government?

To facilitate answering the primary research question as above, the following secondary research questions have been formulated, which systematically will be addressed throughout the thesis:

- What is big data, what opportunities does it have to offer, and what are its limitations?
- How does big data create new knowledge and support decision making in the case of Residential Social Monitoring?
- Which organizational challenges and IT-related requirements need to be addressed to harvest the full potential of big data in the future?

1.4 Theoretical and Methodological Considerations

The thesis's theoretical foundation takes stand within the organizational areas of knowledge creation and decision making. From the theory of how organizational knowledge is created by Nonaka (1994), big data is considered to create new knowledge as 'large' datasets offers a higher form of intelligence that can generate new insights. Further, with increased information and knowledge, an improved base for decision making is available. From the theory of bounded rationality in decision making by Simon (1945), big data is considered to break the established limits leading to bounded rationality and thus pave the way towards the ideal state of perfect rationality.

The collection of empirical material has been achieved by the completion of six partially-structured interviews and participant observations. The six interviews, all of which had a specific focus, had the intention to obtain information on how big data is understood, how it creates new organizational knowledge and how it supports decision making. Further, secondary material in terms of reports and articles has been used to provide a profound knowledge around the hyped technological phenomenon of big data.



Figure 1 – The research methodology

1.5 Structure of the Thesis

The thesis is structured in six chapters as described below.

Chapter 1 – Introduction

This chapter introduced the subject matter of the research study – the use of big data as a source for new organizational knowledge creation and support in decision making. Additionally, it outlined the justification and motivation of the study, being that big data is a new phenomenon in the public sector and little is known about the actual outcome of the practical use of it. Further, it introduced primary and secondary research questions that are to fulfill the aims of the investigation. Finally, it presents the outline of the thesis to provide a map for an easy navigation of the paper.

Chapter 2 – Theoretical Grounding

This chapter presents the theoretical grounding. It starts by outlining how the literature review is conducted in a systematic approach, based on the typology of Rowe (2014). Secondly, it illustrates how the phenomenon of big data is conceptualized in literature. Thirdly, it introduces the theoretical framework, followed by an in-depth discussion on how big data is linked to this.

Chapter 3 – Research Methodology

This chapter presents the research methodology. It starts by outlining the paradigmatic underpinnings that this study operates under, including the perspectives of ontology and epistemology of such. It then guides the reader through the research design and introduces the geospatial big data project, 'Residential Social Monitoring', being a single case study. This is followed by a description of the various data collection methods used to gather the primary and secondary empirical material. Lastly, a detailed description of the process of analyzing the primary data is presented.

Chapter 4 – Empirical Findings

This chapter presents the analysis of the obtained empirical data. It provides rich insights on how big data is understood, creates new organizational knowledge and supports decision making in Danish local government.

Chapter 5 – Discussion

This chapter discusses the empirical findings of the study. It starts by outlining the implications for research and contributions in terms of whether the empirical findings support or extend current literature on big data. The chapter then moves on and presents our reflections on the research process in regards to conducting qualitative research and design-science inquiries.

Chapter 6 – Concluding Remarks

The final chapter presents the concluding remarks of the study. It presents how the empirical findings accommodate the research aims and answers the research questions that steered this study. The chapter ends by suggesting future research opportunities that we have become aware of along the way.

2.0 Theoretical Grounding

2.1 Introduction

This chapter presents a map of the big data literature to provide an overview of the underlying assumptions that shape the positioning of this research study. The literature review has played a key role in assessing the relevancy of the study and helped us forming the theoretical framework for the research. Before presenting the main results of the literature review, that being to capture the essence of what the literature says about big data as a source for new organizational knowledge creation and support in decision making, we first present and discuss our considerations for a carrying out the literature review in a systematic approach.

2.2 The Dimensions of Typology for Literature Reviews

The following part describes our considerations of the establishment of the theoretical grounding for researching the phenomenon in a systematic approach. From the classification of a literature review goals by Schwarz et al. (2006), the purpose is to (1) to summarize prior research, (2) to examine critically contributions of past studies, (3) to explain the findings of prior research, and (4) to clarify alternative views of extant literature. Inspired by the work of Rowe (2014), the literature review is based on what he refers to as the ‘dimensions of typology for literature reviews’ (Ibid.).

Component	Description
Review ambition (Objective)	Describing (a-theoretically), understanding or explaining
Breadth of review (Focal Point)	Problem, stream or theme, discipline
Article Identification (Systematicity)	Inclusion criteria (search process, type of source, period, discipline), coverage, quality assessment, sources description
Analytical Approach (Argumentative strategy)	‘Logical structures in the argumentation enacted in the paper’ ... ‘the order of the components of the author’s argument’ (de Vaujany et al., 2011, p. 401)

Table 1 – Dimensions of typology for literature review by Rowe (2014)

2.2.1 Review Ambition

Rowe (2014) states that a literature review, “[...] *synthesizes past knowledge on a topic or domain of interest, identifies important biases and knowledge gaps in the literature and processes corresponding future research directions*” (p. 243). He distinguishes between three main types of review ambitions, i.e. for describing, for understanding, and for explaining, which can be grouped under the categories of ‘summative’ and ‘developmental’ reviews (Ibid.). Applying a summative approach, we as researchers may seek to map or classify what has been written in regards to our phenomenon with no contributions to existing theory (Ibid., p. 243). Carrying such an approach would provide novel insights from the process of data collection and analysis and offer a mapping of the existing literature on the distinct topics of big data, organizational knowledge creation and decision making. However, strictly speaking, the literature review may only provide a description of the territories and an indication on how there might a relationship between the distinct topics. Accordingly, a developmental approach can comprehend the missing link, seeing that such review may want to explain why, how and when things in a phenomenon, and thus focus on causal relationships with certain outcomes (Ibid.). As a result, this thesis employs a balance of both review ambitions and strives to explore, challenge, and confirm the relationship between big data, organizational knowledge creation and support in decision making processes.

Review ambition	Outcome
Summative	Describing the territories of big data, organizational knowledge creation and decision making (no contribution to theory)
Developmental	Understand and explain the potential effects of big data on organizational knowledge creation and support in decision making (initial contribution to theory that can be explored, challenged, and confirmed in qualitative or quantitative studies)

Table 2 – Implications of applying different review ambitions

2.2.2 Breadth of Review

For the breadth of review, Rowe (2014) states that, “[...] *all the above types of literature reviews aim at a better understanding of a knowledge domain for which breadth can vary considerably from a specific problem to a stream of research and ultimately to what a discipline has produced*” (p. 245). It is acknowledged that the breadth of the review has to be seen in relation to the ambition of it, considering that working under a summative review it may be characterized as being phenomenon-

centric, in spite of the fact that this ambition seeks to map the existing body of knowledge within the distinct topics of big data, organizational knowledge creation and decision making. Yet, anchored in the predicament of understanding the effects of big data on organizational knowledge creation and support in decision making, the breadth of research may be classified as problem-centric (Ibid., p. 244).

It is acknowledged that the process may not represent a clear distinction between either summative/developmental, neither phenomenon-/problem-centric, but rather a combination to establish a thorough understanding on the subject. Yet, there seem not to be any right or wrong choices when it comes to the ambition and breadth of the review, but rather the approaches are chosen by its common ability to support answering the stated research question. Accordingly, the distinction between summative and developmental review ambition is considered upon how it contributes to the different parts of the study and therefore viable solutions for the review.

2.2.3 Article Identification

The third step in the typology comprehends the process of finding relevant articles within our topic (Rowe, 2014). The adopted process is initially driven by defining keywords that to our understanding seem to grasp the topics of big data, organizational knowledge creation and decision making. Rowe (2014) characterizes such search process as keyword-based search (p. 247). From the table below, an initial identification of potential keywords that seemed relevant were recognized.

Phenomena	Initial keywords	
Big Data – A Source for Organizational Knowledge Creation and Support in Decision Making	Big data Business intelligence Data analytics Decision making Data knowledge Data mining Machine learning Data science	Decision-support Evidence-based management Data-driven decision making Knowledge management Knowledge creation Knowledge discovery Informatics

Table 3 – Results from database inquiry

The identified keywords provided the basis for the initial search process. However, in spite of the fact that the same phenomenon can be captured through different keywords, an open mindset and awareness of such was adopted. Rowe (2014) mentions that such awareness is especially important in the area of management topics, such as decision making, seeing that there is a strong incentive to differentiate oneself by inventing own concepts that in a given well-defined theme there might be dozens of completely different words or expressions for designating the same phenomenon, which makes the keyword approach very useful and necessarily iterative after one realizes one's omissions (p. 247). An example of such identified expression is the keyword 'data-driven decision making', which basically is defined as the practice of which management makes decisions based on the analysis of various data and not just on intuition (Provost & Fawcett, 2013a). Thus, a mindset of continuously building a pool of keywords is the means forward. From the work of Webster & Watson (2002), we started by exploring major journals such as Administrative Science Quarterly and Government Information Quarterly for publications containing our defined keywords. The reason for focusing on the major journals was that it is generally expected that the highest quality articles appear from these. Next, having found the initial pool of research and having reviewed it carefully, we then proceeded to go 'backwards', by scanning the reference lists. This was due to the belief that the most recent research has its basic knowledge from prior research. At last, we went 'forwards' with the ambition to identify articles that reference another articles. The process of going 'forwards' was primarily done by utilizing scholarly search engines such as Google Scholar and CBS Library, seeing that they both provide a function for viewing citations on specific articles. The process of snowballing, i.e. going 'forwards' and 'backwards', have been dominated by the ambition of quality instead of quantity, in spite of the fact that snowballing may generate a wide range of existing studies. It is believed that the applied approach have provided a reasonable picture of the existing body of knowledge within our phenomenon (Webster & Watson, 2002).

2.2.4 Analytical Approach

Having defined the extant literature within the phenomenon of big data as a source for organizational knowledge creation and support in decision making, the final step of a literature review is, from the thoughts of Schwarz et al. (2006), to synthesize and summarize the findings. From the work of Webster & Watson (2002), a concept matrix was developed after the reading was complete to identify patterns between the concepts within the found literature. The rationale of synthesizing and summarizing is ultimately to help others for future research purposes, as Webster & Watson (2002)

state, “[...] *make sense of the accumulated knowledge on a topic*”. An abstract of the concept matrix is demonstrated below. Please see appendix A for the complete table.

	Conceptualization				Opportunities				Challenges			
	3v's (Volume, variety and velocity)	Extremely 'large' datasets	Tools, processes and procedures	Cultural, technological and scholarly phenomenon	New knowledge discovery	Improve decision making	Economic value	Predictive analytics	Demand for highly skilled workers	Technical challenges	Data quality	Privacy issues
Articles												
Schönberger & Cukier (2014)		x			x			x				
McAfee & Brynjolfsson (2012)	x				x	x				x	x	
Provost & Fawcett (2013b)		x				x	x		x			
Fredriksson (2016)		x			x	x	x		x	x		x
Knapp (2013)			x			x						
Boyd & Crawford (2012)				x	x					x		x

Table 4 – Abstract of concept matrix

The table has provided effective means of communicating the major findings and insights that was achieved in a visually and informative way. The crosses represent a situation in which an article contains or relates to one or more of the identified concepts. The selected abstract of the table reflects an understanding of the potential effects of big data on organizational knowledge creation that can support decision making. The components of a literature review by Rowe (2014) and the applied approaches are summarized in the table below.

Component	Applied approach
Review ambition (Objective)	Summative / developmental
Breadth of review (Focal Point)	Phenomenon- / problem-centric
Article Identification (Systematicity)	(1) Keyword-based search (2) Exploring major journals (3) Snowballing ('backwards' and 'forwards')
Analytical Approach (Argumentative strategy)	Concept matrix

Table 5 – Summary of the components and applied approaches

2.3 The Concept of Big Data

The concept big data has received a great deal of attention in recent years. Today, the amount of data is on a completely different level than before, due to the massive distribution of information, the increasing amount of electronic devices, and the overall development of the IT society (Schmidt & Cohen, 2014). However, despite the increasing interest in big data, the perceptions of it are rather different and the concept has no overall definition that is agreed upon by the industry and academic scholars (Fredriksson, 2016). In terms of this, it is necessary to attempt to understand the concept to shed light on the yet unclear concept of big data to further build upon a common definition in relation to further research purposes.

In 2001, Doug Laney presented his thoughts on big data and thereby placed himself as the frontier, trying to make a general definition of at that time, new phenomenon. The definition was based on what he called the 3v's, where big data is viewed in the context of volume, variety and velocity (Laney, 2001). Big data also fits into the category of digital information (Michael & Lupton, 2016), where social media, web pages, store purchases, downloads, consumer feedback, health records, maps, protocols and geographic locations, are just a few of the sources that big data descends from (Fredriksson, 2016). Even though big data is commonly seen in the context of the 3v's, Power (2014) argues that one should also look beside the raw numbers and deal with the change in data management processes, within the organizations.

A different view of big data was in 2012 presented by Boyd and Crawford. Here, big data is presented with three basic assumptions: (1) that computer power and algorithmic accuracy is maximized to gather, analyze, link and compare large datasets; (2) that economic, social, technical and legal claims are based on identification of patterns in large datasets; (3) that large datasets work as the kick starter of using higher intelligence and fostering of knowledge creation that will create insights that was formerly unimaginable. Hence, big data is seen as a cultural, technological and scholarly phenomenon (Boyd & Crawford, 2012). That big data is a technological phenomenon implies that tools, processes and procedures are the focal point in understanding what big data is (Knapp, 2013). These three areas are central in understanding big data, as explained by Carter (2011), who argues that big data is new 'technologies' and 'architectures', which are, "[...] *designed to economically extract value from very large volumes of a variety of data, by enabling high velocity capture, discovery and or analysis*".

The organization can hereby use big data to efficiently create, manipulate and manage very large datasets and also store it properly (Knapp, 2013), by utilizing big data as an analytical tool (Jordan, 2014). These newly developed tools, “[...] *let one manage far larger quantities of data than before, and the data – importantly – need not be placed in tidy rows or classic database tables*” (Schönberger & Cukier, 2014).

Where the variety of data is one dimension, volume is another rather abstract dimension with subjective meanings. However, Ohlhorst (2012) reasons that, “[...] *big data is extremely large datasets*”, which makes it impossible to manage and analyze with standard computing processing power and tools. As Fredriksson (2016) elaborates, “*The bigger the data set, the more difficult it is to gain any value from it*”. Hence, new technologies are required to be developed and new techniques to be used, in the pursue of creating value for the organization (Provost & Fawcett, 2013b). The cohesive agreement among researchers surrounds the dimension of volume presented by Laney (2001), where big data is commonly seen as very ‘large’ datasets. Following the report from McKinsey Global Institute, a manifest of big data being in the category of high volume is presented, “[...] *datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze*” (Manyika et al., 2011). Hence, organizations can be limited to effectively draw decisions based on big data in raw form, since it cannot easily be interpreted by workers and managers (Philips-Wren and Hoskisson, 2015). This is due to the fact that big data, “[...] *exceed or are beyond the capabilities of the organization to store or analyze for the purpose of decision making*” (Fredriksson, 2016). An area which Troester (2012) also agrees, “[...] *the volume, velocity and variety of data exceed an organization’s storage or compute capacity for accurate and timely decision making*”.

To sum up, an extensive variety of definitions exists within the concept of big data. However, it is evident that what Laney in 2001 presented as ‘Volume’, ‘Variety’ and ‘Velocity’, complies well with what is written in the literature. A summarization of the concept of big data in regards to the literature follows in the figure below.

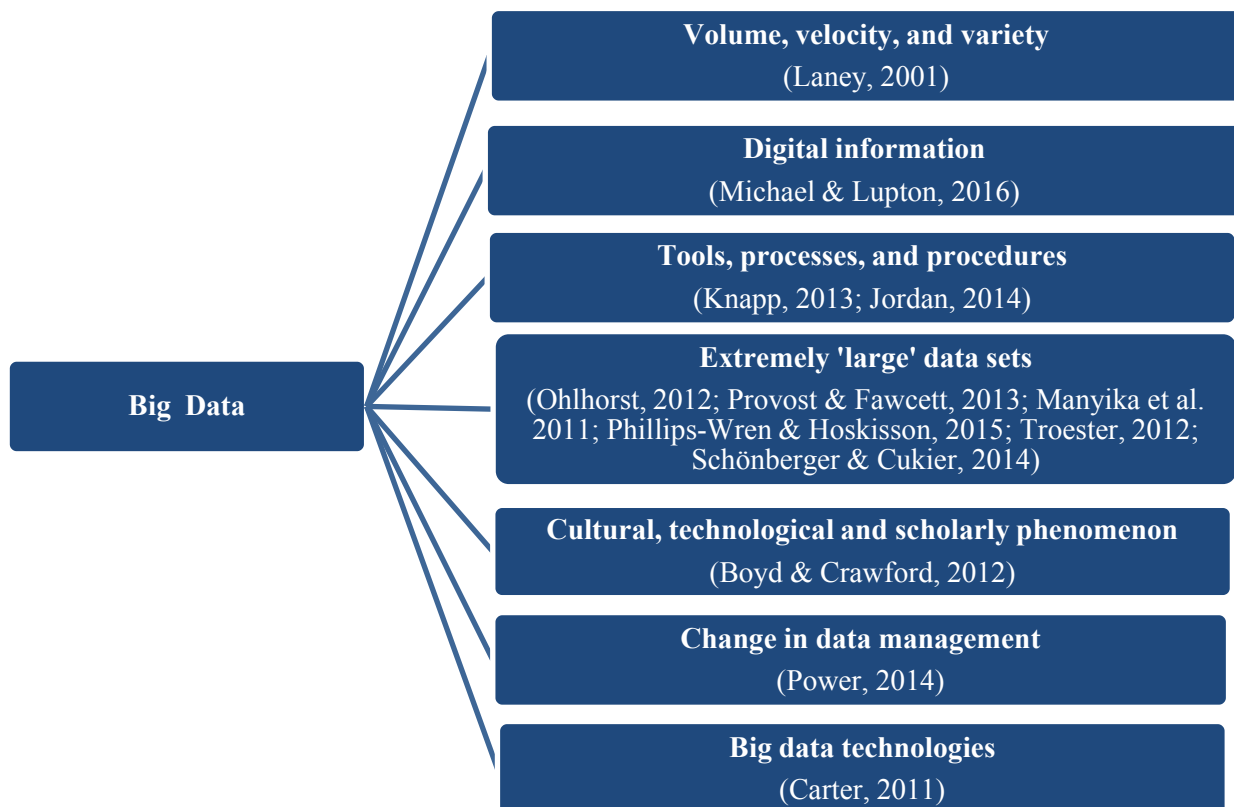


Figure 2 – How big data is described in literature

Despite the variety of definitions above, four resemblances can be derived. Firstly, the amount of data has to be 'large'. Naturally, it raises the question what amount that is. To comprehend such matter, the phrase 'it's all relative'² is used, in spite of the fact that the answer is not expressed in terabyte or petabyte, or whatever follows next. Thus, the amount is subjective and depends on the standard size of datasets within the given industry (Manyika et al., 2011). Secondly, no ordinary software can manage big data. New advanced technologies are required to process the 'large' amount of data. Thirdly, a depth of complexity is present, seeing that the data has to be somewhat unstructured or 'messy', as Schönberger & Cukier (2014) label it. Fourthly, an agreement that some form of machine learning is going to be applied to the data set, such as regression or classification (& Fawcett, 2013b). Lastly, recognition that it is not the quantity of data that is revolutionary, but the fact that we can do something with the data by generating new insights that was not possible before.

² Meaning that one size does not fit all when it comes to human beings

2.4 Knowledge Creation Theory

Any organization that dynamically deals with a changing environment ought not only to process information efficiently but also create information and knowledge. Analyzing the organization in terms of its design and capability to process information imposed by the environment no doubt constitutes an important approach to interpreting certain aspects of organizational activities. (Nonaka, 1994, p. 14)

Organizational theory has for centuries been focusing on a paradigm that conceptualizes any organization as a system that ‘processes’ information or ‘solves’ problems. The latter, a central task of the organization is how it efficiently handles information and deals with decisions in a changing environment. As Toffler (1991) characterizes it, “*We are now living in a knowledge-based society, where knowledge is the source of the highest quality power*”. Furthermore, this paradigm promotes that the solution to gain efficiency lies within the ‘input-process-output’ system of information processing (Nonaka, 1994, p. 14). However, Nonaka (1994) states that a major problem exists within the paradigm, that, “*Information processing is viewed as a problem-solving activity which centers on what is given to the organization – without due consideration of what is created by it*” (p. 14). Thus, one should rather focus on what information and knowledge that is created in the means of the organization’s engagement with the surrounding environment and not the processing of already existing information. Innovation is an example of a key organizational output or knowledge creation, which cannot be explained in the terms of information processing or problem solving. It should rather be viewed in the context of a process, where the organization’s job is to create and define problems and thus to solve them by developing new knowledge (Ibid.). Therefore, the organizations’ ‘raison d’être’³ are to continuously create knowledge (Nonaka et al., 2000, p. 6).

Big data is seen to play an important role in creating knowledge for organizations, which will be described later in this chapter. Yet, first a closer look into the underlying theory.

³ Reason or justification for existence

2.4.1 What is Knowledge?

Knowledge is extensive and contains many different viewpoints, definitions and expressions. Some organizations view knowledge as 'evidence' or 'fact-based'. Others understand it in relation to systems of 'best practices' and some simply describe it as 'past experiences' (Christensen, 2010, p. 160). Nonaka et al. (2000) define knowledge as a 'justified true belief' (p. 7). However, many different criteria exist before something can be categorized as knowledge or if it is rather simply principles and stereotypical opinions, which does not align with the reality (Christensen, 2010, p. 160).

Further, knowledge shall be seen as dynamic, because it regards to the social interaction between organizations and individuals (Nonaka et al., 2000, p. 7) and as a context specific asset that depends on time and place (Hayek, 1945, p. 521). If knowledge is not put in a context, "[...] *it is just information, not knowledge [...] 1234 ABC Street is just information*". However, when used in a context, "*My friend David lives at 1234 ABC Street, which is next to the library*", information becomes knowledge (Nonaka et al., 2000, p. 7). Thus, information is defined as 'a flow of messages' or meanings that is transformed into knowledge when, "[...] *interpreted by individuals and given a context and anchored in the beliefs and commitments of individuals*" (Ibid; Fredriksson, 2016, p. 6). However, when transforming information into knowledge, the interpretation by individuals creates a main problem called 'absorptive capacity' (Cohen & Levinthal, 1990), where knowledge is understood subjectively (Christensen, 2010, p. 162). That knowledge is subjective is furthermore highlighted by Szulanski (1996, 2003), who explains that knowledge is not just something that you wrap up and send. Hence, one should rather accept the limitations that exist and thus the knowledge, which are to be shared, is not complete (Christensen, 2010, p. 162). That information and knowledge is interpreted by individuals who possess subjective meanings, leads to the next aspect of knowledge creation and sharing, namely tacit and explicit knowledge.

2.4.2 Explicit and Tacit Knowledge

One important dimension of the knowledge creation and sharing theory is the view of knowledge being either tacit, i.e. 'non-transferable' or explicit, i.e. 'transferable' (Nonaka, 1994). A distinction of these main concepts is presented below.

Explicit, or codified knowledge, refers to something that can be expressed in formal and structured language and thereby easily shared between organizations, individuals or groups in the form of data, scientific formulae, specifications, manuals and so forth (Nonaka et al., 2000, p. 7).

Another criterion for knowledge to be explicit is that it can be, “[...] *processed, transmitted and stored relatively easily*” (Ibid.). Tacit knowledge on the other hand is the opposite. Here, knowledge is something that regards to our personal quality of action and involvement (Fredriksson, 2016, p. 6), yet also something that is intensely rooted into the actions, procedures, routines, commitment, ideals, values and emotions that might be present (Nonaka et al., 2000; Schon, 1987). Tacit knowledge cannot be easily shared because it, “[...] *indwells in a comprehensive cognizance of the human mind and body*” (Polanyi, 1958; Nonaka, 1994, p. 16) and it is knowledge that a person holds on an individual level from own practical experiences, such as former jobs, courses, etc. (Christensen, 2010, p. 165). Tacit knowledge has been explained as, “*We can know more than we can tell*” and in the context of learning how to ride a bike (Polanyi, 1967). You do not learn it by reading manuals and books, but rather to practice and develop an experience with it. However, any guidance from others with the experience, might help you reach the goal earlier and to be more effective (Christensen, 2010, p. 166).

Knowledge is also a cultural term, where western countries view it as being explicit, and for others, like Japan, knowledge is primarily considered as tacit (Nonaka et al., 1996, p. 205). Therefore, one has to be aware of the cultural differences that might exist when working with e.g. big data towards new knowledge creation. However, Nonaka et al. (2000) emphasize that a combination of both tacit and explicit knowledge has to be recognized, because they are complementary to each other and together play an important role in knowledge creation (p. 8). Hence, new valuable knowledge is created by an interaction and combination of both tacit and explicit knowledge as Nonaka et al. (2000) argue, “*Written speech is possible only after internal speech is well developed*” (p. 8). Regarding to the interaction of tacit and explicit knowledge, Nonaka (1994) has developed a framework for four different modes of knowledge creation, also known as the SECI process (Nonaka et al., 2000). The interplay between the two types of knowledge is called ‘knowledge conversion’, where the modes contain of: (1) ‘socialization’, i.e. from tacit knowledge to tacit knowledge; (2) ‘externalization’, i.e. from explicit knowledge to explicit knowledge; (3) ‘combination’, i.e. from tacit knowledge to explicit knowledge; and (4) ‘internalization’, i.e. from explicit knowledge to tacit knowledge (Nonaka, 1994, p. 19).

Firstly, socialization refers to the shared experiences and interactions between individuals, such as observation, imitation and practice, demonstrations, face-to-face dialogues and informal meetings.

Secondly, externalization includes processes of transforming tacit knowledge to explicit, such as abductive thinking, concept design, notes, diagrams and metaphors. Thirdly, combination occurs when individuals collect explicit knowledge from inside or outside the organization and then combine, edit and process it to create new knowledge inside the organization, such as specifications, databases and synthesized reports. Fourthly, internalization is viewed as ‘learning by doing’, where individuals perform tasks specified explicitly in documents and manuals. This is done in combination with training programs and simulations that are typically held in close cooperation with the rest of the organization (Nonaka et al., 2000).

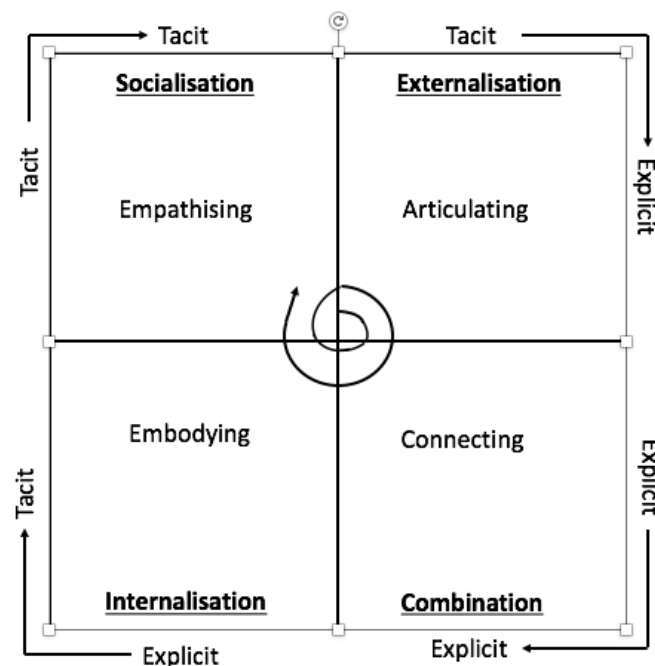


Figure 3 – The SECI process
Source (Nonaka et al., 2000)

The different modes of knowledge conversion shall be seen as a dynamic process, where knowledge enters and evolves in a spiral movement that comprises all four modes (Nonaka, 1994). Hence, the spiral becomes larger and can trigger a new knowledge creation to develop, which then starts a new spiral that undergoes the same process as the former. The SECI process encompasses that knowledge creation is made throughout the organization, starting at the individual level, moving beyond boundaries to departments, divisions and sections, or even to other organizations (Nonaka et al., 1996). Thus, the knowledge creation process “[...] is a never ending process that upgrades itself continuously” (Nonaka et al., 2000).

2.4.3 Big Data – A Source for Organizational Knowledge Creation

Big data is all about seeing and understanding the relations within and among pieces of information that, until very recently, we have struggled to fully grasp (Schönberger & Cukier, 2014, p. 19).

In this section we link the phenomenon of big data to the organizational theory of knowledge creation, seeing that big data in this thesis is considered a source with great potential to create new organizational knowledge. This potential is established because big data implies large datasets and it allows the organization to be more intelligent and engender new insights. Therefore, big data is considered new knowledge (Knapp, 2013). This is also supported by Schönberger & Cukier (2014), who acknowledge big data as being able to, “[...] *extract new insights or create new forms of value*” (p. 6) that can help us understand and make sense of the world that we live in (p. 7). Yet, others argue that big data creates more valuable knowledge than before, due to the possibility of combining traditional data analysis with the new big data analytical tools, “[...] *in order to get richer and improved insights and make smarter decisions*” (Marr, 2015, p. 10). For any organization, the new insights and knowledge gathered from big data on customers, products and operations can result in a new or more efficient value chain process. By utilizing the collected data, it provides the organization with insights and knowledge about customer's interests, passions, affiliations and associations. Thus, these insights can be used in optimizing the ‘customer engagement process’ (Schmarzo, 2013). Thus, put in another context, public institutions or municipalities can also use these new opportunities in their effort of servicing citizens. The possibility of seeing things in real time is also referred to as ‘big data knowledge creation’. Since it helps the organization to use predictive maintenance, product performance recommendations and network optimization and thus be a ‘predictive enterprise’ (Ibid.).

The combination of datasets from both internal and external sources is in relation to big data, seen as a driving force for new information and knowledge creation (Fredriksson, 2016). This fits well into the perspective by Nonaka et al. (2000), where knowledge shall be combined even if it is both tacit and explicit. Thus, it is not just the new advanced technologies developed in the era of big data that fosters new knowledge creation, but rather the data itself and the way we use it (Schönberger & Cukier, 2014, p. 7). Particularly, the explicit or codified view of knowledge is important to highlight, due to the fact that big data is transferable (Fredriksson, 2016). Being transferable, big data follows what Nonaka et al. (2000) explains, “*Explicit knowledge can be expressed in formal and systematic language and shared in the form of data [...] it can be processed, transmitted, and stored*”.

Additionally, big data shares many of the same features of what Nonaka (1994) explains as explicit knowledge, i.e. digital, captured in databases and archives. Seeing big data as digital data, one can argue that it creates explicit knowledge (Fredriksson, 2016) and thus amplifies the relevancy of using knowledge creation theory in relation to big data. Nonaka et al. (1996) hold the argument that explicit knowledge is a set of meaningful information, which is something that we believe can be regarded as the opportunity to create visualization with big data (LaValle et al., 2011).

Information refined from big data can be transformed into new valuable knowledge, patterns and insights that can be used to solve problems, increase profits and productivity (Fredriksson, 2016). Furthermore, the new knowledge derived from big data fosters opportunities and cultivate best practices of working, for both private and public institutions (Elgendy & Elragal, 2014; Ohlhorst, 2012). That new knowledge is created by utilizing big data from outside of the organization and not just the processing of already existing information, aligns well with the view of Nonaka (1994), who states that any organization should not solely gather information and process it to solve problems, but rather create new knowledge with 'what' is given to the organization. From our literature review, big data is not just acknowledged as 'large datasets', but also as a process of understanding, managing and connecting it to a relevant context. Connecting data to relevant context is also acknowledged by Hayek (1945), as a paramount to gain knowledge. This connection is done by the individuals with expertise of using big data analytical tools and different methods of understanding data; therefore we argue that tacit knowledge is still an important aspect of big data management. Thus, information is, *"[...] interpreted by individuals and given a context and anchored in the beliefs and commitments of individuals"* (Nonaka et al. 2000), which then is related to the concept of 'absorptive capacity', i.e. knowledge is understood subjectively (Cohen & Levinthal, 1990).

To visualize how big data management is able to create new organizational knowledge, the framework of Ackoff (1989) & Rowley (2007), namely the Data-Information-Knowledge-Wisdom (DIKW) hierarchy, can be applied. Firstly, data represents 'symbols and properties' of objects and events. Secondly, information comprises processed data. Thirdly, knowledge is seen in relation to 'instructions of how to answer' questions. Fourthly, wisdom deals with the values that involves the 'exercise of judgment' made by the individual (Ackoff, 1989). To this, *"Each of the higher types in the hierarchy includes the categories that fall below it"* (Rowley, 2007), i.e. knowledge is derived from information and information from data. Thus, when big data enters the hierarchy, it undergoes an equivalent complex process to the one found in the SECI model (Nonaka et al., 2000).

In both processes, knowledge is extracted and internalized into the individuals of the organization. Due to the fact that knowledge may answer questions (Ackoff, 1989) is why LaValle et al. (2011) argue that, any organization should ask questions in regard to its strategy, vision or *raison d'être* before analyzing, compromising and combining big data to new knowledge.

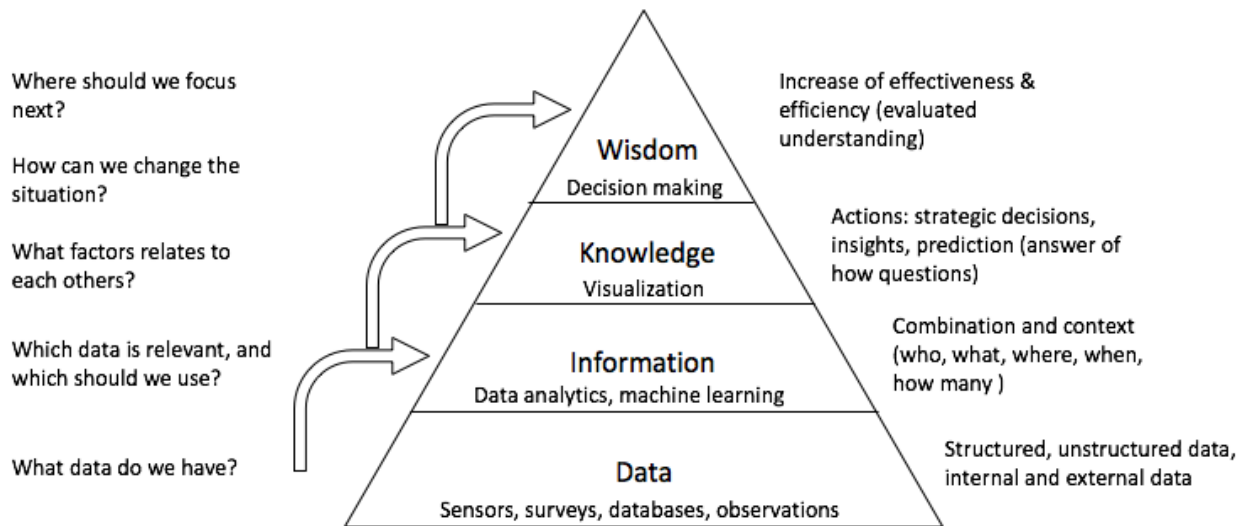


Figure 4 – DIKW pyramid in the context of big data management
Source: Own contributions with inspiration of (Rowley, 2007; Ackoff, 1989)

Firstly, the questions on the left, relates to the principle that any organization should ask questions to its strategy, vision or *raison d'être* (LaValle et al., 2011) before addressing the knowledge that big data may create. Secondly, general examples of how big data may enter the different layers in the DIKW hierarchy are presented in the middle of each layer and on the right. However, the presented interpretation of the framework shall only be seen in the context of our research perspective and not as such a general model for application. Thirdly, it is important to emphasize that the information- and knowledge layer are of great interest, since they refer to the technologies that big data encompass. Such technologies may include analytic software solutions (e.g. Hadoop, MapReduce and big Table) that enable the possibility to break-up, analyze and combine large datasets in new ways. Hence, big data allows new knowledge creation by using text-, speech-, video/image- and combination analytics, including behavior analysis (Marr, 2015). Fourthly, knowledge is gathered by presenting the information found through the data analysis in easy and understandable matter, which is something that visualization encompasses.

With visualization, data becomes “[...] *more accessible and meaningful but one can also better illustrate the relationship between the data*”, which is furthermore important in decision making (Marr, 2015, p. 157). Where visualization previously focused on presenting tables and spreadsheets in forms of charts and graphs, big data allows organizations and decision makers to present data, information and knowledge in brand new ways. The reason for this is the new technological tools as explained by Marr (2015), “*People don’t want to search for the insights locked within the data, they want their insights provided to them, nicely packaged in a way that helps them understand the messages and these new tools can help in that challenge*” (p. 161). Data insights can be provided to decision makers in several ways such as displaying maps, text, data, behavior & emotions and connections (Marr, 2015, p. 162). Fifthly, with new knowledge, decisions makers can gather a shared understanding of situations and thus increase the effectiveness and efficiency of their organization.

To summarize, the above section examined how big data can leverage new knowledge creation in an organizational context. Based on this examination, we propose that big data plays an important role in generating new knowledge for further usage in strategic decisions, including forming of visions and accomplishment of goals. Furthermore, big data helps generate new explicit knowledge, which is highly relevant to examine in any case, where one wants to move from tacit into explicit knowledge that can be used easier throughout the organization. Thus, the collected qualitative data surrounds the presented theory of knowledge creation, regarding to tacit and explicit knowledge.

2.5 Decision Making Theory

Organizational decision making in the organizations of the post-industrial world shows every sign of becoming a great deal more complex than the decision making of the past. As a consequence of this fact, the decision making process, rather than the processes contributing immediately and directly to the production of the organization's final output, will built larger and larger as the central activity in which the organization is engaged. In the post-industrial society, the central problem is...how to organize to make decisions. (Simon, 1973; Huber, 2004)

The environment of organizations has never been changing as much as now, compared to what we have seen in the past (Huber & Glick, 1993). Since, the environment are changing rapidly decision making is a topic that is highly relevant to study for organizations to survive (Ibid.).

The reasons for the environmental change originate from the evolvement of information technology and its increasing effectiveness (Ibid.). There is a great attention to the constant change, where the demand for effective decision making has never been higher. Thus, this environmental transformation puts pressure on the organizations to make more frequent and rapid decisions. Furthermore, the increasing complexity of environments also increases the complexity of the decision to be made by managers (Ibid.). The effects of changing environments on organizational processes are presented in the figure below.

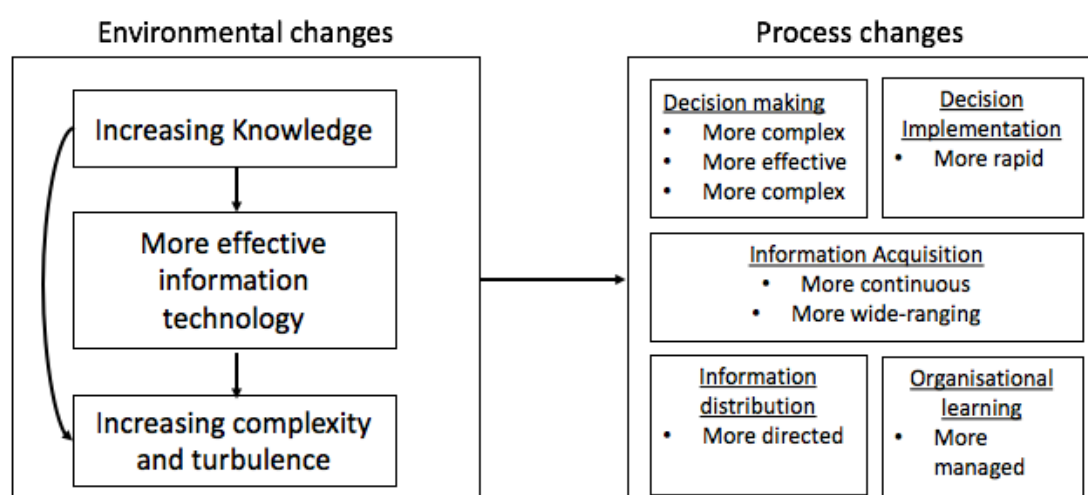


Figure 5 – Environmental change effects on organisational processes.
Source: (Huber & Glick, 1993, p. 8) (Own contribution).

Due to the increasing turbulence of the external environment, organizations face two major challenges surrounding processing the internal information, namely greater volume of data and ambiguity of information (Daft et al., 1993). This argument fits well into the context of big data. However, before moving into how big data supports decision making, we have to look at the main elements of decision making and decisions as such. An elaboration of these concepts will follow next.

2.5.1 What is a Decision?

According to his in-depth literature examination, Eilon (1969) points out that even though much has been written about decision making, a clear definition of a 'decision' is absent. Regardless, being almost 50 years ago since this claim was made, we as researchers believe that it is important to clarify the underlying concept of a decision.

Most literature defines a decision to be a choice between several alternatives (Thorsvik & Jacobsen, 2008). Furthermore, Langley et al. (1995) argue that any decision is surrounded by the relationship between commitment and action. Commitment to action implies that decisions are expected to be followed up on with delivery of tangible solutions and actions that lead the intention of the decision to realization (Thorsvik & Jacobsen, 2008). In other words, decisions are the outcome of a process, which includes several stages of narrowing different alternatives down into one single, which is acted out (Simon, 1997). A short version of the three separate and different decision process phases is shown below:

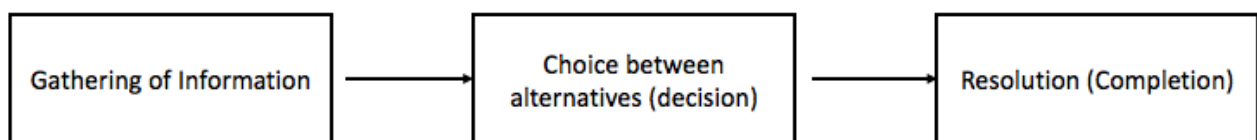


Figure 6 – Three separate phases of a decision process.
Source: (Thorsvik & Jacobsen, 2008, s. 269) (Own contribution).

The decision process can be viewed as a chain; with one decision comes another. Decisions are made by following a specific goal to which a certain behavior is devoted. However, this goal is often just one link of the chain towards further secondary goals. In the end, the links of goals connect the whole chain in reaching the final aim (e.g. optimize processes, products or services) (Simon, 1997). Since decisions cultivate decisions, organizations make them very often related to the different levels of activities that they perform.

All new decisions are typically based on previous experience and choices made. Thus, the activities made within the organization are pushed by a chain of decisions (Thorsvik & Jacobsen, 2008). Simon (1997) explains that the decision making process can be seen as a vertical stream throughout the organization, where different actors, such as operators and supervisors, use their expertise to guide the decision into action. Hence, an important decision is based on several factors such as 'the raw material inputs' or 'decision premises' (Simon, 1997), which contains 'facts', 'values', 'side conditions' and 'constraints' made throughout the process and assembled into a final 'product', namely the decision itself. Thus, *"A decision can be divided into components, each fabricated by specialists and specialized groups, and finally brought together into a coordinated picture"* (Ibid., p. 23). These inputs can be seen as information, which is considered in the specific setting of the decision that is proceeding. Before information can be used in a given perspective, organizations or researchers have to gather, systematize, analyze, interpret and communicate it clearly, so that it can be seen in context of other alternatives, before taking the final decision (Thorsvik & Jacobsen, 2008). Additionally, one shall be aware of the fact that decisions in most circumstances are made with the purpose of affecting a conduct (Ibid.). Information can also be viewed in the context of how different media or information systems deliver information in high- or low richness. The idea behind 'information media richness' follows the statement of Daft et al. (1993), where 'rich media' is characterized by 'high touch' and qualitative data, appropriate for resolving ambiguity. 'Lean media' on the other hand is *"Technology-based, high volume data exchanges"*, suitable for, *"[...] conveying quantitative data with precision and accuracy to large audiences"* (Daft et al., 1993). In terms of strategy formulation, media of low richness can induce managers to think quantitatively and base decisions on rationality. That managers can think quantitatively typically happens when implementing a strategy. Here managers seek to convey facts, hard data measurements and rational expectations to organizational members. This is especially done when using lean i.e. low media richness communication (Ibid.).

The opportunity to base decisions on facts and objectivity leads us to the next important aspect of decision making theory, evolving rationality. However, the expression 'rational' does not only imply that individuals act reasonable, it also suggests how to make a decision and what to do in a given situation, facing a problem (Thorsvik & Jacobsen, 2008).

Rationality is by March (1994), explained with four basic questions:

- (1) The question of alternatives: What actions are possible?
- (2) The question of expectations: What future consequences might follow each alternative? How likely is each possible consequence, assuming that alternative is chosen?
- (3) The question of preferences: How valuable (to the decision maker) are the consequences associated with each of the alternatives?
- (4) The question of the decision rule: How is a choice to be made among the alternatives in terms of the values of their consequences?

Following these questions, the situation you face is evaluated in terms of: Firstly, the characteristics and challenges it holds. Secondly, different alternatives to a solution are outlined. Thirdly, the consequences of different alternatives are evaluated. Fourthly, a comparison is made. Lastly, the alternative with the best solution is chosen (Thorsvik & Jacobsen, 2008). Rationality surrounds us with the belief that any human being is to be seen as either a perfect- or bounded rational actor in decision making. Thus, decisions are considered in the context of 'the being's' governance of reason, which affects our function, work and actions (Ibid.).

2.5.2 Perfect Rationality

The so called 'economic man', which is an ideal model/state of the 'man' in decision making theory, assume that the human being can act with 'perfect rationality' in any given decision situation (Simon, 1997). The model is generally characterized with four different assumptions of how decisions are made (Thorsvik & Jacobsen, 2008):

- (1) Decisions are made based on clear goals.
- (2) The decision maker has complete/perfect information about any alternative solutions and the consequences surrounding them.
- (3) The decision maker can prioritize between alternatives, stating which one is the most favorable and the least.
- (4) The decision maker will always choose the best fitted alternative to a given situation/problem.

The four assumptions can be summarized in the following figure.

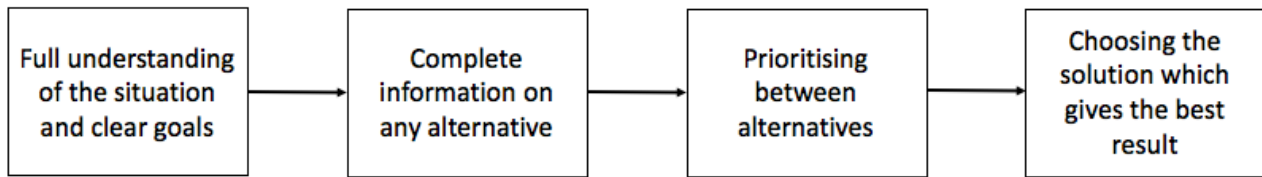


Figure 7 – The human as a rational decision maker
Source: (Thorsvik & Jacobsen, 2008, p. 271) (Own Contribution)

The figure outlines that clear goals, preferences and complete information needs to be in place for the decision maker to be perfect rational. Clear preferences refer to the choice between alternatives, where the decision maker knows what one favors and therefore is able to make a choice of the perfect solution. However, Thorsvik and Jacobsen (2008) also argue that the model must not be seen as ‘the truth’, but rather as guidance to how we preferably should handle decisions. Simon (1957, 1997) elaborates on how this ideal state can be difficult to reach by any human being:

The capacity of human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world – or even for a reasonable approximation to such objective rationality. (p. 198)

2.5.3 Bounded Rationality

Thorsvik and Jacobsen (2008) argue that several challenges exists within perfect rationality and that these are the reason why ‘bounded rationality’, is more likely to occur in any decision making process. Furthermore, humans are said to be only rational for a short amount of time or to certain extent (Shollo, 2013), due to three main areas of limitations, which the ideal model of perfect rationality encompasses: (1) the limitations attached to goals, (2) the limitations of the ability to handle complete information, (3) the limitations of choosing between alternatives (Thorsvik & Jacobsen, 2008). Further, Shollo (2013) also explains that bounded rationality is more likely to occur due to the costs associated with collecting, analyzing and interpreting information, but also because of, “*The complexity of reality exceeds the human capacity to process information*” (p. 49). The limitations stated by Thorsvik and Jacobsen (2008) are worth elaborating on to completely understand what bounded rationality is. The first limitation addresses that perfect rationality involves having clear and well defined goals, and that these always come before action (March, 1994).

However, bounded rationality implies that this only happen in very few cases (Thorsvik & Jacobsen, 2008). The reasons for this can be found in the different personalities, preferences, identities and understanding of values and goals in an organizational matter (Ashforth & Mael, 1989; March, 1994). Another aspect is that bounded rationality refers to the thought of one defining goals after the action has occurred. Hence, it indicates that goals and preferences are outcomes of actions and not the foundation of actions (Thorsvik & Jacobsen, 2008; March & Olsen, 1976; Brunsson, 1989; Weick, 1979). Aronson (1995) argues that 'the man' is more rationalizing than rational. Thus, a postulation is made that 'the man' wants to be viewed as rational and therefore claim a meaning to every action after one has accomplished them (Thorsvik & Jacobsen, 2008).

The second limitation comprehends to how 'the man' cannot handle complete information. This refer to the limited resources of the human brain, which cannot simply gather, process and store data in large amounts and thus does not fulfill the 'requirement' of complete information about every alternative (Thorsvik & Jacobsen, 2008; Simon, 1960; Schott, 1991). Further, research shows that people often base decision on 'rules of thumps' (Kahneman et al., 1982) and experience, rather than actually analyze all alternatives (Thorsvik & Jacobsen, 2008). Information that does not fit into our picture of reality can easily be thrown away and therefore we cannot be perfect rational (Janis, 1983; Espedal, 1997).

The third limitation is regarded to psychological factors when choosing among alternative solutions (Thorsvik & Jacobsen, 2008). Here, choosing between alternatives can be seen as a 'defeat' or 'loss' if one has strong intentions, feelings and/or commitments with a specific alternative and it is not chosen (Staw, 1976; Thorsvik & Jacobsen, 2008). The better the alternative or problem is described and compared to others, the less is the chance of people being physiological affected. However, this seems to be almost impossible (Kahneman, 1997; Kahneman & Tversky, 1984).

A figure of how decision makers usually handle a decision, following the bounded rationality approach, is presented below.

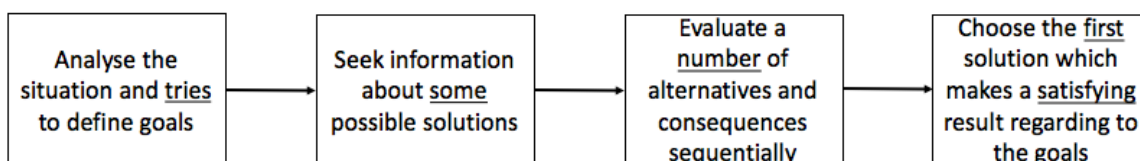


Figure 8 – The human as a bounded rational decision maker
Source: (Thorsvik & Jacobsen, 2008, p. 274) (Own Contribution)

The model assumes that no human being can ever act with perfect rationality. Rather, the person shall be seen as 'the administrative man' (Simon, 1997), who tries to act rational with limited resources in terms of information about alternative solutions that one has. The model of bounded rationality can be characterized with four statements (Thorsvik & Jacobsen, 2008):

- (1) Goals are present, but rather unclear and changing.
- (2) The decision maker does only evaluate a few possible solutions and its consequences.
- (3) The decision maker evaluates solutions sequentially as he gets time and capacity for it.
- (4) The decision maker chooses the first satisfying solution which presents.

2.5.4 Big Data towards Perfect Rationality

What's the point of all that data, anyway? It's to make decisions.
(Regalado, 2014)

In this section we link the phenomenon of big data to the organizational theory of decision making with the main argument that big data fits well into the representation of leveraging better support for decision makers in organizations.

As Power (2014) explains it, "*Big data will lead to better health, better teachers and improved education and better decision making*". These improvements in sectors and areas origins from the new data sources, technologies and analyses that keep developing (Fredriksson, 2016). Further, it is made clear that big data only creates real value for the organization if it is used in support for decision making and if the data provided actually helps the decision makers to make better decisions (Power, 2014; Kudyba, 2014). When big data is used in decision making it enhances governmental programs, actions, solutions and public policy on a more general level and should be used in support of making better decisions, develop strategies and enhance productivity and efficiency (Fredriksson, 2016). The latter is likewise something that Feldman & March (1981) focus on in their work of describing the importance of information in an organizational matter. They argue that the desire to improve decisions is followed by massive investments in information management and information retrieval and that the value of information must be seen in context of how it influences decision making. Vaidhyanathan and Bullock (2014) argue that any decision maker wants to collect as much data as possible, which allows them to make the best decisions (p. 56).

Thus, we argue that big data enables the decision maker to come closer the idea of having complete information and all alternatives presented before making a decision and consequently accomplish two requirements within the perfect rationality perspective (Simon, 1997).

Furthermore, big data is said to be the fuel for decisions, insights and actions, following the massive data processing that now happens in bigger and faster scales (Sorofman, 2013). It is further agreed that big data analytics have changed the way organizations perform their decision making process, due the ability to easily access vast amount of data (Fredriksson, 2016). The organization that incorporates big data experiences a transformation from being a 'retrospective' and 'rear-view mirror' business, which only uses some parts of aggregated or sampled data. To becoming a 'forward looking', 'predictive view' business that uses all available data within or outside of the organization (Schmarzo, 2013). Big data-driven decision making thus holds 'forward looking' recommendations, based on an exploitation of complete data from diverse sources in real-time, correlated and governed with business optimization (Ibid.). At the same time, one still have to be aware of the cognitive biases that might exist within any organization and decision making processes, also known as the area of bounded rationality. However, EY (2016) argues that successful and skilled big data analytics can help overcome the cognitive bias, information 'clouding' and the bounded rationality that exists. Big data enables the organization managers to objectively identify key values drivers, goals and objectives. Combined with the use of advanced scenario-planning and what-if analysis, one can choose between all possible alternatives and exam how the potential decisions may affect them. Thus, it is possible for the organization to 'stress-test' the validity of their decisions and hereby optimizes their strategy (EY, 2016). Based on this, it seems reasonable to say that big data paves the way towards perfect rationality in organizational decision making by addressing the four assumptions regarded to situation/goals, gathering complete information, prioritizing between alternatives and choosing the optimal solution (Thorsvik & Jacobsen, 2008).

Big data can also be used as fuel for decisions, because it creates new knowledge that is considered an important strategic asset to the organization (Grant, 1996). Further, big data analytics implies that data is being collected, stored and analyzed. Thus, creating information and knowledge, which works as the foundation for the decision making process (Manyika et al., 2011; Raghupathi, 2014; Feldman & March, 1981). Hence, it strengthens the connection of big data and new knowledge creation as described previous in this chapter. However, it is not just a matter of using big data in support for decision making; the organization needs to be fit for it. As Gartner (2013) and Power (2014) highlight,

big data needs innovative forms of information processing such as organizational change and resource allocation for enhanced insight and decision making. Here, information technology plays an important role in generating value from big data (Fredriksson, 2016). The way that information technology affects strategic decision making is something that Huber (1990) has made several proposals to. One important aspect is that advanced information technologies must be viewed in regards, “[...] *to more rapid and accurate identification of problems and opportunities, and to more rapid and higher quality decisions*” (Fredriksson, 2016). Thus, the concept of big data management is paramount for the further development of advanced information technologies (Ibid.).

To summarize, the above section examined big data in relation to decision making including the theories of bounded and perfect rationality. Based on this examination, we propose that big data plays an important role in assisting the decision maker. Further, big data paves the way towards the ideal state of perfect rationality. However, several obstacles need to be addressed on the big data journey for organizations. When addressed, decisions can be based exclusively on the premises of data. Thus, this is interesting to have in mind when we commence our qualitative analysis of the Danish municipalities later in this thesis.

3.0 Research Methodology

3.1 Introduction

In the collection of empirical material in this case study of big data in the Danish public sector, a mixture primary and secondary material have been used. This is carried out to achieve the greatest possible insight on the subject and to understand the practical usefulness of big data. Below is the applied approach of the research methodology and subsequently a number of sections for the description of the same. The figure illustrates how the thesis, which constitutes a case study, makes use of interviews and observations, and reports and articles – hereby qualitative data – for illumination of how big data creates new organizational knowledge and supports decision making.



Figure 9 – The thesis's methodological basis

3.2 Paradigmatic Underpinnings

As stated in the motivation for this study, a gap in research exists on the actual outcome of the practical use of big data in the public sector (Fredriksson, 2016; Fosso Wamba et al., 2015). Thus, the need to study the actual effects of big data applications in the public sector to enable a better understanding of the complex concept of big data and the possibilities and limitations it encompasses, is justified. The request for scholars to look at how data currently is being used and assess the degree in which it is being underutilized is further stated by Desouza & Jacob (2014), whom argue that the perception of big data is valuable and actors in the society have started to consider ways on how big data can be used to improve public sector outcomes.

In terms of this, the study operates under the pragmatic paradigm that being the basic belief system or worldview that guides us as researchers (Lincoln & Cuba, 1994), seeing that the paradigm is

concerned with action, intervention and constructive knowledge (Goldkuhl, 2012; Braa & Vidgen, 1999). Despite the fact that interpretivism is considered the most used paradigm within qualitative methods, pragmatism is increasingly gaining position in the research field, especially within IS research (Goldkuhl, 2012). Aiming for intervention and change (Braa & Vidgen, 1999), pragmatism is needed, claimed by Blumer (1969, p. 71),

The essence of society lies in an ongoing process of action – not in a posited structure of relations. Without action, any structure of relations between people is meaningless. To be understood, a society must be seen and grasped in terms of the action that comprises it. (Blumer, 1969, p. 71)

Accordingly, the ontology builds on the elements of action and change, being either subjective or objective in nature. People live and act in a world of constant development and to understand such world, concrete actions are thus to be taken into consideration. Hence, as stated by Goldkuhl (2012), one of the fundamental ideas within pragmatism is that the meaning of an idea or a concept is focused on the practical consequences of such.

With the concept of 'inquiry', which is central to the thoughts of pragmatic research (Dewey, 1938, p. 108), focus is aimed at creating knowledge and insight in the interest of change and improvement (Cronen, 2001, p. 20). Thus, inquiry is the idea of investigation into some part of reality with the purpose of creating knowledge for a controlled change of this part of reality. Further, not only has pragmatism an interest in what 'is', but also for what 'might be', that is an orientation of a future prospective in a not yet realized world (Goldkuhl, 2012). Thus, the knowledge character of pragmatism is not restricted to either objective or subjective explanations and understandings, being the key forms of positivism and interpretivism, but rather recognized as forms of information that can provide new constructive knowledge (Ibid.). The epistemology further dictates that knowledge must not be perceived as a 'copy' of reality, but rather be constructed with the purpose of better manage existence in the world (Dewey 1931; Rorty, 1980). Therefore, from the thoughts of Goldkuhl (2012), the function of knowledge can eventually be summed up to the aim for constructive knowledge that is appreciated for being useful in action, generated through assessment and intervention, that takes into account in which more efficient and profitable relationship with objects may be established in the future.

Based on this worldview and the given problem statement, the selected focus on the practical usefulness of big data as a theoretical phenomenon that is able to create new organizational knowledge and support decision making, can be categorized under the so called functional pragmatism (Ibid.). The functional pragmatism is concerned with the idea of instrumentality, thus creating meaningful knowledge that may be useful for local as well as general practices in the Danish public sector. The practical usefulness of big data is in the given case study subjectively assessed by the various participants within. Lastly, concrete proposals for future applications of big data are to be given to the extent of which is possible with the goal of achieving higher internal efficiency and effectiveness with the use of big data. Thus, the data collection will assess the phenomenon of big data, which has the potential to create new knowledge that can support decision making and thus serves to promote change and efficacy of the future use of big data in the public sector.

3.3 Research Design

This thesis constitutes a single case study where the unit of investigation is a municipality. The case study as research strategy comes from the Latin concept of *casus* and means ‘what is happening’, by focusing on one instance of a ‘thing’ that is to be investigated, such as an organization, a department or an information system, seen within the boundaries of reality (Oates, 2006, p. 141). A case study is characterized by (1) focusing on depth rather than breadth, in order to obtain as much detail as possible about the one instance, (2) studying the instance in its natural setting without disturbing the scene too much, (3) aiming on the complexity of relationships and processes and how they are inter-related, and (4) using multiple data generation methods to achieve as much knowledge about life and work around the instance (Oates, 2006, p. 142). The characteristics of a case study relates very well to the definition by Yin (2003b), who describes the research strategy as “[...] *an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident*”. Further, this case study goes under the category of being an exploratory study, seeing that we are trying to fulfill the gap in research that currently exist on the actual outcome of the practical use of big data in the public sector (Ibid.). Identifying topics to be covered in future research projects is well-aligned with this study's paradigmatic underpinnings, seeing that pragmatism works within the thought of being more effective and obtain profitable relations with objects that may be established in the future.

Thus, exploring the effects of big data as a source for new organizational knowledge creation and support in decision making within the boundaries of reality, the practical utility for both now and in the future is presented.

Despite the fact that this study constitutes a single case study, we draw from the idea of cross-case displays (Miles & Huberman, 1994, p. 186). The aim of cross-case displays, i.e. studying multiple cases, is to primarily increase the generalizability, which is generally open for discussion in qualitative research. As Silverstein (1988) puts it, *"We are faced with the tension between the particular and the universal: reconciling an individual case's uniqueness with the need for more general understanding of generic processes that occur across cases"*. Thus, reassuring that the events and processes at one well-described setting are not wholly idiosyncratic, helps develop more sophisticated descriptions and more powerful explanations. In this study, we engage Randers Municipality as a second case display to enhance the generalizability by verifying whether the findings make sense beyond Svendborg Municipality.

3.4 Geospatial Big Data Analytics

A report from McKinsey Consulting points out that the use of geographic information systems (GIS) and geospatial big data analytics is an innovative technology that in the public sector has the potential of reducing costs, improving services, and creating a better community for citizens (Grant et al., 2014). GIS include the areas of hardware, software, data, management, and analysis of geographically referenced data, commonly referred to as geospatial analytics. Despite the fact that GIS have existed since the 1960s, Grant et al. (2014) summit how the last few years have brought remarkable advances and trends in relation to the technology: (1) large amount of geospatial data generated through particularly smart phones, credit cards, social media, GPS devices, Google, etc., (2) an increase in the accuracy of geospatial data to pinpoint locations, and lastly, (3) more advanced methods available to analyze geospatial information, partly due to greater standardization of data and databases. It is further stated by Lee & Kang (2015) that a significant portion of big data is actually geospatial data and the size of such data is growing rapidly at least by 20% every year. This is in relation to what Grant et al. (2014) explain that it is possible to innovate daily life and businesses by exploiting the power of location embedded in geospatial big data. The advances and trends in GIS make it possible to use geospatial data in new and innovative ways as it offer to integrate and analyze large disparate datasets that involve both geospatial- and non-geospatial data, which makes it possible to easily detect

patterns and trends in society that might otherwise may be overlooked, thus providing new knowledge that can support decision making (Grant et al., 2014; Lee & Kang, 2015).

From a conducted qualitative research with over 30 cities, Grant et al. (2014) refer to three general areas in the public sector where geospatial analytics offer significant benefits. These areas are information dissemination, urban planning, and service delivery. Across these areas, GIS technology and geospatial analytics can “[...] *increase speed, accuracy, and cost effectiveness related to a wide range of government priorities, including those related to crime prevention, emergency management, disaster recovery, social services, health care, transportation, urban planning, environmental initiatives, and facility planning and management*” (Grant et al., 2014). Thus, as Lee & Kang (2015) put it, GIS and geospatial analytics can help with a number of decisions by its mapping capabilities to visualize information in a more interactive way, creating new knowledge that can support decision making and help optimizing resource deployment and service delivery by determine which locations are in need for what services.

3.5 Case: Residential Social Monitoring

Residential Social Monitoring is a project and tool developed by Svendborg-, Randers- and Helsingør Municipality, with Kommunernes Landsforening (KL) as facilitator. The purpose of the project was to improve the strategic work with exposed residential areas in Danish society, a tool that, according to a written guide on the subject by KL (2015b), is sought by many to follow the developments in the municipality's residential areas and document the unique social housing challenges. It is recognized that data analytics and data-based knowledge within the municipality's residential areas have the potential to contribute to (1) to create a better foundation for decision making, (2) to improve the strategic work with the municipality's exposed residential areas, and (3) to strengthen the targeted efforts in practice (Ibid.).

According to Jensen et al. (2015), vulnerable children and young people take up a lot of both social- and economic resources and is an area of increased focus in Danish municipalities and have been so for many years. The challenges are many; how do we control the economic costs, how do we get an early eye on the children at risk, how do we create the needed motivation among them, how do we coordinate efforts, family cooperation, etc. Further, municipalities must through their casework and variety of offers contribute to content and quality in children and young peoples' lives.

The focus is on strengthening the efforts for vulnerable children and young people and their families to ensure coherence between the different efforts that is initiated across administrations, professional groups and areas in the municipality (Ibid.).

3.5.1 Technical Components

The case of Residential Social Monitoring consists of four technical components. The components are geographic information system (GIS), Fælleskommunalt Ledelsesinformationssystem (FLIS), health data from school nurse and the Danish National Statistics (DST), as visualized in the figure below (KL, 2015b).

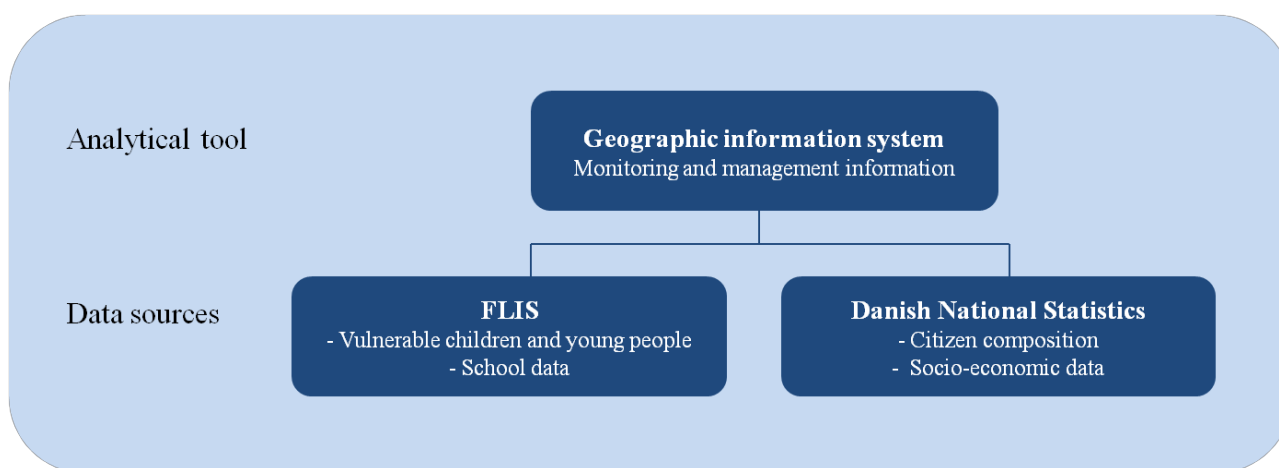


Figure 10 – Technical components in Residential Social Monitoring

GIS is the analytical tool in residential social monitoring, where one produce maps and analyzes geographical data (Ibid.). Citizens' addresses are used to encode the geographic location and CPR-data enables a coupling between addresses and relevant data on individual level.

FLIS is a large data warehouse, purchased and administrated by KOMBIT ("About FLIS", n.d.). FLIS offers access to calculations of ratios and raw data on the lowest registration level. With FLIS, municipalities can benchmark internally and externally against other municipalities, geographic regions and the national average. It allows the possibility to select the most comparable municipalities, i.e. municipalities that are the most similar in e.g. size and demographic distribution. However, FLIS also offers data that can be broken down to a single accounting or individual level, so municipalities are able to connect its strategic objectives and the actual efforts.

Further, FLIS also contains historical data that may describe developments over time. FLIS has a comprehensive data foundation, consisting of seven data fields that are visualized in the figure below (“About FLIS”, n.d.).

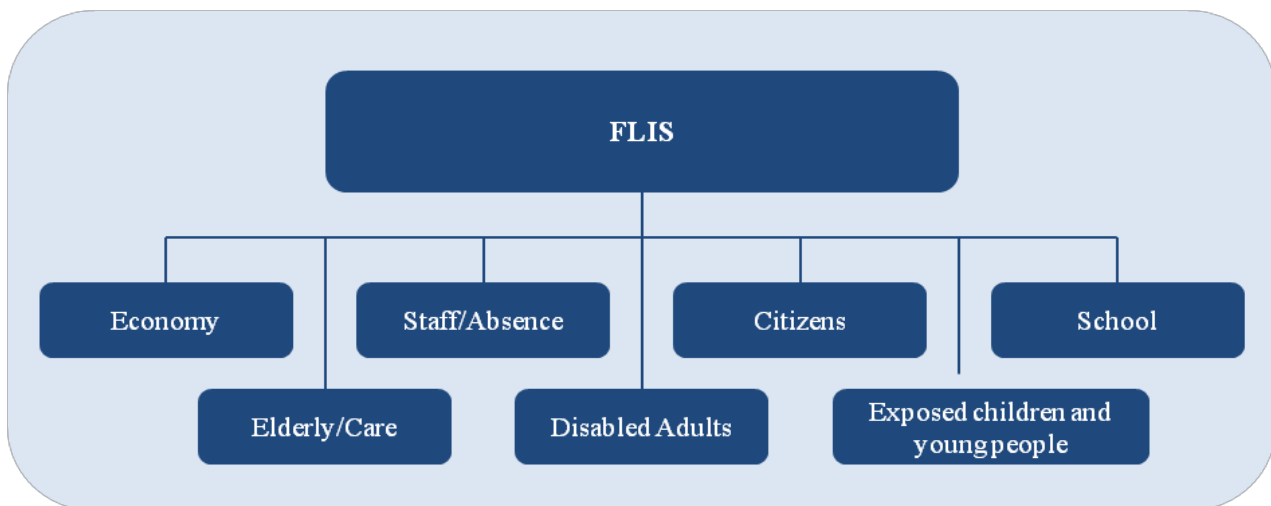


Figure 11 – Data fields in FLIS

The use of data on the areas of vulnerable children and young people, and school is of great relevance in Residential Social Monitoring as they tell something about the social heritage, which usually is of negative character in exposed residential areas. The views of data extraction on vulnerable children and young people, and school data are institutionalization, preventive measures, grades at primary school, inclusion and segregation, number of citizens and movements (KL, 2015b).

Lastly, data from the DST is also used in residential social monitoring. The data involves, among other, citizen composition and socio-economic data, broken down on residential areas. These data may reflect labor market attachment, level of education and sole family provider status (Ibid.).

3.5.2 Access to the Case Study

The initial contact to access the case of Residential Social Monitoring was carried out by writing to the project manager of FLIS at KOMBIT, Hedvig Hvarregaard Andersen. This was done by sending an e-mail in which the purpose of the thesis activities, what we wanted to study, the benefits for KOMBIT, as well as the time consumption, were described. Firstly, the intention was, in relation to pragmatism, to create an understanding of the practical use of data from FLIS within big data projects in the Danish public sector.

Secondly, what KOMBIT could expect from the project, including focused knowledge for the contemporary use and for future changes or improvements. Our enquiry was met with an invitation from KOMBIT for an initial interview, in which the further circumstances of the project were discussed in relation to its duration and requirements for access to participants. After an agreement on cooperation was made, an introduction of FLIS was presented by Andersen, followed by questions by us as researchers in regards to the concept of big data and current data projects in the municipalities where data from FLIS was used.

3.5.3 Selection of Participants

The selection of participants is based on the desire to get a thorough understanding of the subject on how big data is able to create new organizational knowledge and support decision making. From the work of Palys (2008), the process of selecting the right participants is based on the idea of purposive sampling, which is a superordinate concept that signifies a series of strategic choices tied to the research objectives. In line with pragmatism, we are interested in the actual outcome of the practical use of big data, which from the theoretical grounding can be separated into two groups of individuals. Firstly, knowledge creation with big data refers to the individuals dealing with big data management, i.e. the ones performing data analytics. Secondly, support in decision making refers to the individuals who participate in decision making processes and use the analyses performed by the first group. Below is an overview of the involved participants with name, title and workplace, all of whom are believed to provide detailed information that is relevant to the enquiry.

#	Participant	Role	Date and location for interview
1	Hedvig Hvarregaard Andersen (HHA) KOMBIT	Project Manager of FLIS	6 th of Marts At KOMBIT
2	Sara Thestrup Hansen (STH) Kommunernes Landsforening (KL)	Facilitator of 'Residential Social Monitoring'	9 th of Marts At KL
3	Jan Carlsson (JC) Svendborg Municipality	Urban and housing policy development consultant	21 th of Marts At Svendborg Municipality
4	Mikkel Næsvang (MN) Svendborg Municipality	Data analyst	21 th of Marts At Svendborg Municipality
5	Astrid Krarup Andersen (AKA) Svendborg Municipality	GIS-Coordinator	21 th of Marts At Svendborg Municipality
6	Troels Rasmussen (TR) Randers Municipality	AC-Consultant	29 th of Marts Frederiksberg, phone interview

Table 6 – The study's participants

3.6 Data Collection

This study takes stand in the field of qualitative research. The choice of method on the qualitative research interview is believed to generate the most detailed and comprehensive insight on how big data to create new organizational knowledge as support in decision making. Schwandt (2001) defines the purpose of qualitative, interview-based research to describe and clarify people's experiential life, "[...] as it is lived, felt, undergone, made sense of and accomplished by human beings". The purpose of qualitative study by Schwandt (2001) is closely related to what Miles & Huberman (1994) and Strauss & Corbin (1990) describe it. With the researcher's role to gain a 'holistic' overview of the context under investigation, the purpose is to attempt capturing data on the perceptions of local actors 'from the inside', through a process of deep attentiveness of empathetic understanding, also referred to as 'verstehen' (Miles & Huberman, 1994).

Strauss & Corbin (1990) define it as, “[...] *research about persons’ lives, lived experiences, behaviors, emotions, and feelings as well as about organizational functioning, social movements, cultural phenomena, and interactions between nations*”.

Qualitative interviews are a flexible, yet powerful tool to capture the voices and the ways people make meaning of their experiences (Rabionet, 2011; Andersen, 2010). Accordingly, with the qualitative interview as method, it provides us with unique opportunities to capture and discover new perspectives on the issues, which can be very useful for the further work process. This is particular important in the study of ours, seeing that the field may have relatively little practical knowledge. Though, through the qualitative interviews with our participants, whom are acknowledged as experts in the subject field, we get to access their vision and understanding of the use of big data as well as the possibilities and limitations it encompasses.

The design of the qualitative study is based on the framework of Brinkmann & Kvale (2009) for a systematic approach to structure, arrange and carry out the different parts of a qualitative research study. The seven stages of the framework are thematizing, designing, interviewing, transcribing, analyzing, verifying and reporting. Firstly, a description of the research’s purpose and topic was written, including answering the questions of what and why. Secondly, a plan of what components the study consisted of was described. This meant considering all the stages to ensure that we could obtain the knowledge that is desired to answer the research question, which also is reflected in the order of the interviews held. Thirdly, the interviews were carried out with the use of an interview guide and thoughtful perspectives on the knowledge that we were seeking. During this part, we also familiarized ourselves prior to the interviews with the various interpersonal skills to exploit the profound knowledge needed. Fourthly, all the interviews were recorded concerning the analysis with a complete transcription of the entire oral interview to a text. Fifthly, data coding of the entire empirical material were done for analytical purposes, seeing that this provides us with the best overview and results. Sixthly, a discussion on the verification in terms of generalizability, reliability and validity were discussed. Reliability, as a measurement of the consistency of the results, and with the concept of validity, it is analyzed and discussed how the qualitative study actually examined what is in stage one was set out to investigate. Lastly, a well-defined way of reporting the results and methods of inquiry that meets the scientific criteria is discussed for future research applications.

3.6.1 Partially-Structured Interviews

From the work of Newman & McNeil (1998), the adopted interview procedure can be characterized as being partially-structured. The reason for this choice is that we have a relatively wide and open theoretical framework where there are no unambiguous answers, but instead serves as an overall framework. The relatively wide and open theoretical framework is reflected in the interview guide as open-ended questions with the purpose of letting our participants to think and reflect before answering the questions. This provides us with the opportunity to be receptive and open to new perspectives and information that the participant can contribute with and therefore we must listen carefully. As our theoretical framework involves a number of items that should be clarified to answer the research question, an interview guide is prepared to illustrate the topics to be affected (Andersen, 2010). The interview guide has thus acted as a guideline for the interviews. The items are not necessarily needed to be affected in the order they are listed in, why the items are to be crossed out on a running basis.

The variety of interviews held with KOMBIT, Kommunernes Landsforening, Svendborg- and Randers Municipality, have had an average duration of 30 minutes. Besides a number of key items, which all participants have been interviewed on, specific questions related to the participants' area of responsibility and big data's practical benefit here, have also been asked. The interview guide, which also reflects the operationalization of the theories, is attached as appendix B. The table, which is divided into two main theoretical areas, shows a concrete set of questions regarding each area that is under investigation. From the table below, an abstract of the interview guide is shown. The left column indicates the theoretical areas in which the questions are operationalized and the right column shows the specific questions.

Theoretical frame	Interview questions
Concept of big data	- How do you understand big data? - For how long have you been familiar with big data?
Organizational Knowledge Creation	- What new knowledge has it created combining the various datasets in Residential Social Monitoring? - Has the use of big data provided you with answers to questions that were previously unknown? How?
Improvements in Effectiveness	- How do you experience that big data and knowledge from it contribute to decision making processes? - Allows data analytics you to make better decisions? How?
Improvements in Efficiency	- How do you experience that big data can be used in early social intervention? - How does evidence-based practice play in?

Table 7 – An abstract of the interview guide (translated in English)

3.6.2 Participant Observation

The second part of the case study's primary data is based on participant observation. The qualitative method of participant observation is, as well as interviews, incorporated under the umbrella term of ethnographic methods (Kawulich, 2005). From the work of DeWalt & DeWalt (2002), participant observation is a strong method to develop a holistic understanding of the phenomena under investigation. It is suggested that the method may as well increase the validity of the study, as observations may help the researcher have a better understanding of the context and phenomenon under study (Ibid.).

The adopted process of conducting observations at Svendborg Municipality is by Angrosino & DePerez (2000) characterized as 'focused observation'. Focused observation is exemplified as observation supported by interviews, in which the participants' insights guide the researcher's decisions about what to observe (Ibid.).

The choice of carrying out focused observations was in need to achieve a practical understanding of the 'data value chain' (Flyverbom & Madsen, 2016). From the thoughts of Flyverbom & Madsen (2016), big data analysis can be conceptualized as consisting of distinguishable steps and sorting processes. As a form of 'data value chain', data, like other materials, are sourced, refined and turned into valuable products. The data value chain is closely related to what is implied by Strathern (2005) and Leach (2012) on the well-established distinctions between data, information, and knowledge, i.e. DIKW-pyramid, as explained in the theoretical grounding. This distinction suggests that data is unprocessed material that comes into senses, i.e. machines. Information is then that data organized or grouped in some logic or other. Lastly, knowledge is more than information, it is data organized in a way that has an effect.

Accordingly, in the case of Residential Social Monitoring at Svendborg Municipality, the focused observations were to provide us with insights on how data in practice are selected and gathered from the data warehouses and transferred to the municipality's local geographical information system for further analytical study. From the work of Whyte (1979), the strategy on acknowledging and viewing the participants as collaborators to conduct the most effective observations, was adopted. Whyte (1979) attaches great importance to the relationship between the researcher and participants as collaborative researchers who, through building solid relationship, improve the research process and improve the skills of the researcher to conduct research.

3.6.3 Reports and Articles

The second part of the thesis's empirical foundation constitutes of data in the form of reports and articles. From the work of Andersen (2010), such data represents secondary data techniques and is in this thesis brought in from particularly management consultancies and technology companies, seeing that big data in the public sector is driven first and foremost by these organizations (Thapa, 2016; Pollitt, 2014). The purpose of bringing such data is to support the learned image on big data as a source for organizational knowledge creation and support in decision making. The complete list of reports and articles is attached as appendix C.

Area	Reports and articles
Big data in the public sector	Economic Co-operation and Development (OECD, 2015) <i>Data-Driven Innovation: Big Data for Growth and Well-Being</i> Oracle (2012) <i>Big Data: A Big Deal for Public Sector Organizations</i>
A source for knowledge creation	McKinsey Global Institute (Manyika et al., 2011) <i>Big Data: The next frontier for innovation, competition, and productivity</i> Google (Andrade et al., 2014) <i>From Big Data to Big Social and Economic Opportunities: Which Politics Will Lead to Leveraging Data-Driven Innovation's Potential?</i>
Decision making	Ernst & Young (EY, 2016) <i>How can big data lead to more rational decision-making?</i> Capgemini (2012) <i>The Deciding Factor: Big Data & Decision Making</i>
Geospatial big data analytics	McKinsey Consulting (Grant et al. 2014) <i>Coordinates for Change: How GIS Technology and Geospatial Analytics Can Improve City Services</i> Deloitte Consulting (2014) <i>The Impact of Open Geographical Data</i>

Table 8 – An abstract of the secondary material used

3.7 Data Analysis

Having accounted for how we collect the primary material through qualitative methods, this section reflects upon the process of empirical analysis. The aim of the analysis phase is to draw valid meaning from the qualitative data that we can potentially theorize (Miles & Huberman, 1994). One of the key features of qualitative data is their richness and holism, with strong potential for revealing the complexity nested in real context (Miles & Huberman, 1994; Kvale & Brinkman, 2009). Considering qualitative data is based on 'raw' experiences, the analysis of it is based upon the transcribed interviews as the outcome of our interviews. The process of transcribing the interviews is needed since the 'raw' experiences are not immediately accessible for analysis why it requires processing (Miles & Huberman, 1994). This is a time consuming, yet necessary process, as it requires a holistic understanding in order to select the appropriate parts for further investigation (Andersen, 2010).

The process of transcribing interviews is a study in itself and imposes challenges of reliability and validity, as well as the distinction between oral and written language (Kvale & Brinkmann, 2009).

The initial analysis of the transcribed interviews relies on the process of open coding as a group to identify patterns within the statements of the participants. This process helps to build categories within knowledge creation and decision making with big data and is beneficial as it allows for individually identifications of differences and discrepancies within the group. After having defined the categories together, we coded the entire data individually in order to compare our individual results. Despite the fact that this process it is rather time consuming, it meets the concern in qualitative analysis of reliability of the results (Miles & Huberman, 1994).

After having completed the initial data analysis, we systematically copy the individual elements of empirical material to detailed textual analysis. This process is carried out electronically and is based on a data matrix, which helps to maintain an overview and assist with finding patterns, relationships and causality (Andersen, 2010). For review of the complete data matrix, please see appendix D.

Theme / Theory	Mikkel Næsvang (MN)	Astrid Krarup Andersen (AKA)	...
Concept of big data	<i>[...] the monitoring perspectives or at least the good case that I usually refer to regarding big data in general, is how to empty a stadium as fastest as possible, which is becoming aware of people's movement patterns and the pressure points and in relation to traffic and police [...]. (Line 15)</i>	<i>I would define it as a really good concept for me to come out with data because it is something that has gained obedience throughout the entire managerial area [...]. (Line 12)</i>	
Knowledge creation	<i>[...] it can be fairly easy to see if there is an area that sticks out of which one had not believed that ran. (Line 99)</i>	<i>[...] it has clearly contributed to the visibility of a lot of conditions that we might have had an idea of how things were, if nothing else have been made it more obvious and articulated with the use of more data in the analytics processes (line 96)</i>	
Support in decision making	<i>[...] it strengthens our ability and capacity to make decisions, better decisions. When we have it visually we can better and more easily get an overview than looking at data tables, which sometimes can be difficult (Line 257)</i>	<i>[...] the use of GIS provides us with more knowledge and the more knowledge we have about the decision that is to be taken, the better decision can be made by the politicians, whom ultimately are the ones making the final choices. (Line 104)</i>	

Table 9 – An abstract of the data matrix.

Despite the fact that the study is generally deductively driven, seeing that we are looking for support for a predefined proposition, i.e. big data is able to produce new organizational knowledge and support decision making, we have adopted techniques from inductive content analysis with the initial approach of analyzing the qualitative data through open-coding (Elo & Kyngäs, 2007). The reason for this is the lack of research about the actual outcome of the practical use of big data in the public sector. Thus, elements of both deductive and inductive reasoning are used throughout the data analysis, yet also reflected in the data collection with open-ended questions to provide us with the opportunity to be receptive and open to new perspectives that may be relevant.

4.0 Empirical Findings

4.1 Introduction

The following analysis is built on the previous examined theories and empirical data, which is collected by having followed the thesis's research design. The analysis presents the results that are found in relation to the aims and questions of the research. It is worth mentioning that the participants in the empirical material mention big data as either 'coupled figures' or 'management information', which generally seems to be the practical term used at the particular organization. In case of any doubt, big data is, as previously mentioned in the literature review, characterized as relatively 'large' and complex datasets, whose size excludes the use of ordinary software to collect, manage and analyze.

4.2 The Concept of Big Data

Before moving into the area of how Svendborg Municipality use big data in relation to knowledge creation and decision making with the specific case of Residential Social Monitoring, a further investigation of what acquaintance the participants possess of big data and how they would define this phenomenon was commenced.

In sum, the findings show that a broad understanding of what big data is exists. From a national point of view, the representative from Kommunernes Landsforening (KL) argues that big data, "*[...] is the possibility of being able to collect a vast amount of data and then see some patterns*" and with help from data their main focus is to benchmark and compare the municipalities with each other and easily present some differences between them (Hansen. S., personal interview, March 9, 2017, appendix F, line 18-19). This point of view is supported by the FLIS representative, who says that the essence of big data in the Danish public sector is to look across the population. However, this is only possible when the massive amount of data that exists in the public sphere, is actually used (Andersen, H., personal interview, March 6, 2017, appendix E, line 10). However, Hansen also argues that big data is not just something you easily define, "*[...] whether something is big data or not is a difficult question to answer*" (personal interview, March 9, 2017, appendix F, line 27), but at the same time she tries to explain it by using examples from her professional career:

Yes [that she has worked with big data], but you can say that in the municipality of Aarhus we collected all the data that we could possible get hold of from the case workers and cases and that Aarhus is a very large municipality so it was a lot of data that we collected and tried to create patterns and do something with red, yellow or green (e.g. indicators) depending on when an action was necessary in different settings. (Hansen. S., personal interview, March 9, 2017, appendix F, line 32)

Therefore, a clear conviction that big data should be used to look across borders of municipalities with the possibilities of benchmarking and comparing them from another, exists from the two national representatives, i.e. KL and FLIS. Moving closer to the participants from Svendborg Municipality we see that different opinions about big data continuous. When asked about how they would describe big data, one participant returned with an answer, which clearly states that big data is in such a character, that it is overwhelming for one to understand and easily portray with words, “[...] *it is so great that I almost can't [tell] [...] 'Oh boy'*” (Carlsson, J., personal interview, March 21, 2017, appendix G, line 44).

It is interesting that the national representatives had more ease in defining big data, where the participants from Svendborg shows difficulties, but they share the same approach, i.e. they can easily present examples of where they believe big data is at place. Surveillance is an example of such, which is presented by several participants, “*One can almost say that to some extent surveillance is a designation of it [big data]*” (Carlsson, J., personal interview, March 21, 2017, appendix G, line 44). This is also supported by another who believe that a good case of big data is the monitoring or surveillance perspectives regarding to how fast you can empty a stadium:

Those surveillance perspectives or rather the good case that I can think of is the one about how you can empty a stadium fastest, thus you get attentive to peoples movement patterns and where is the bottlenecks and regarding to the traffic and the police and all that sorts of things [...] that's the classical understanding that I have of it”. (Næsvang, M., personal interview, March 21, 2017, appendix H, line 15)

Furthermore, examples of behavioral tracking of the population in traffic, city centers and shopping malls are given, but most importantly, big data shall be seen as a context specific concept. One can change the circumstances and understanding, depending on what area that data is integrated into the work tasks (Næsvang, M., personal interview, March 21, 2017, appendix H, line 20).

Thus, it is interesting to see how the participants draw from their experiences when trying to define big data. From a national point of view, big data is more understandable and seen in how you monitor municipalities overall. In relation to Svendborg and Randers Municipality, the representatives use their technical skill and knowledge to draw conclusions about big data with specific examples. Lastly, one representative from Svendborg municipality is not convinced that big data exists in Denmark:

I would define it as a really good concept for me to promote data usage because it has won responsiveness in management perspectives and it's in that path that I can use data more. But if I scroll back I wouldn't say that we have enough data in Denmark to call it big data, if that's the original definition. So I let people talk about big data, and then I work with data. (Andersen, A., personal interview, March 21, 2017, appendix I, line 12)

That big data is difficult to conceptualize is supported by the representative from Randers Municipality, who claims that big data is hard for him to define, since they do not use it that much (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 9). Having too much data is not something that Randers Municipality worries about, due to the fact that they do not have larger datasets that exceed their population of 100.000 and which they, “[...] *data management wise are able to process and handle*” without problem (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 15, 19). Big data recur as a popular term, when asking the participants. As explained by one participant from Randers, the concept has become so extensive popular following that, “[...] *it relates to the massive datasets who Facebook, Google or similar companies are generating*” (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 27). Further, big data is not simply something you make a choice to use or not, it is not just a method, but rather an overall concept regarding to analysis of large data amounts (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 32). This argument relates well with Knapp (2013) who explains that big data, is “*Tools, processes, and procedures*” (p. 215) for organizations to create, manage and manipulate data (Fredriksson, 2016, p. 3). However, the participant does not see big data as being just one method. To this, Jordan (2014) describes big data as “*The analytical tools*” (p. 378). Thus, Big data is not only one way of doing things, but rather a concept which surrounds many different aspects of interpreting excessive amounts of data.

Thus, a clear skepticism about whether enough data is captured throughout the Danish public sector for it to be called big data exists, when asking the data coordinator from Svendborg and the AC-

consultant from Randers Municipality. However, it is interesting to see that big data is used as a marketing concept, or as explained by Rasmussen, “[...] a buzzword for a lot of these developments that happens when it comes to data” (personal interview, March 29, 2017, appendix J, line 9). Accordingly, this description does not view big data in terms of defining the volume, variety and velocity as Laney (2001) argued. Only two of the participants clearly mention the vast amount of data several times, as an important factor in what characterizes big data. Further, Rasmussen is still skeptic about the data volume dimension and shares similar thoughts of Andersen, that, “*It is not something which matters that much*” (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 9). A cohesive agreement exist, that big data is such a popular concept that it allows the municipalities to work with and conceptualize data more, than previously.

Regarding to how much experience the participants have with big data, the findings draw very different angles of the understanding and acquaintance that one has. Again, big data seems difficult for the participants to grasp when asked for how many years they have been aware of it or for how long it has existed. One says, “*That is a really good question*” (Hansen. S., personal interview, March 9, 2017, appendix F, line 24). Another draws attention to us as researchers and asks, “*How many years ago is it gradually [...] I don't know?*” (Andersen, A., personal interview, March 21, 2017, appendix I, line 12). Thus, it is clear that big data is not something that the participants can define. Further, the participants present different answers to how long they have been familiar with it. Where some participants have been familiar with big data for a shorter period of time, “*Hmm, 2,5 years*” (Næsvang, M., personal interview, March 21, 2017, appendix H, line 26) and “*then it is properly 3 years ago*” (Andersen, A., personal interview, March 21, 2017, appendix I, line 21). One find himself in the middle and argues that he has been familiar with it for 3 to 5 years (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 23). Finally, one have worked a lot with data and management information and familiarized with it 9 years ago during school (Hansen. S., personal interview, March 9, 2017, appendix F, line 24-25). Thus, it is interesting, that certain differences of big data acquaintance exist, when having the participants' role and age in perspective. Where the eldest participant has the least acquaintance or knowledge of big data, only knowing it for 2,5 years, the youngest one has 9 years of experience and familiarity with it. Where Hansen became familiar with it during education, Næsvang, Rasmussen and Andersen have become familiar during work. Hence, a margin of 6 years exists between the two groups of participants and it is interesting to highlight the fact that the phenomenon has existed since 2001, where Laney presented his thoughts on big data. That there exist many different views on big data is also confirmed by Fredriksson (2016),

who mentions that, “*Big data has no overall definition that's agreed upon by the industry and academic scholars*” (p. 2). That big data has no overall definition is further supported by Schönberger & Cukier (2014), whom argue that big data is a phenomenon without any exact definition (p. 6). Thus, big data as being a ‘buzzword’ is what PWC presented in 2014, with argumentation that big data has changed the way we do business today. Further, they highlighted that the concept itself has changed from simply being a buzzword to a trend which is lasting in the future (p. 2). However, there seems to be an existing belief that big data is still a buzzword, a marketing concept and a trend that allows people to promote and work easily with data, when asking participants from Svendborg and Randers municipalities. This is supported by Davenport et al. (2012) who also argue that big data is simply a buzzword “*for smarter, more insightful data analysis*”, used primarily by information technology providers (p. 22).

Viewing the findings in perspective of Laney's three big data dimensions, i.e. volume, variety and velocity, thoughts of similarity recurs, when participants interpret what big data is for them. Firstly, an agreement exists, that the volume perspective is something that big data encompasses. However, two participants are rather skeptic about the quantity of big data available for them. Thus, in context of the Danish municipalities, big data seems to deviate in minor degree, from the idea of being datasets, which cannot be easily understandable by people. As further explained by Manyika et al. (2011), big data is, “[...] *beyond the ability of typical database software tools to capture, store, manage, and analyze*” (p. 1). Secondly, the collection and usage of varieties of data (Laney, 2001), is something that the participants' answers address. Thirdly, data velocity seems not to be an area that is directly quoted by participants. However, several answers and statements are pointing in the direction, that big data has to do with this dimension.

Another perspective which recurs, especially from the national point of view, is that big data allows one to draw patterns and perform benchmarking between municipalities. This perspective is addressed by the report of Manyika et al. (2011). Firstly, data mining allows users, “[...] *to extract patterns from large datasets by combining methods from statistics and machine learning with database management*” (p. 28). Secondly, machine learning and neural networks, can form the foundation for finding patterns and insights by automatically learning, recognize and draw intelligent decisions of the data that is available (Manyika et al., 2011, p. 29). Thirdly, Power (2014) clarifies, that big data can be used in ‘retro perspective analyses’ that enhances the understanding of historical data and then make decisions based on the future (p. 225). Hence, letting municipalities and national

representatives work with big data, allows them to better draw patterns and compare themselves with others, creating new knowledge and thereby make better decisions.

To sum up, the findings show that big data supports the concept being a buzzword that in this study is explained by the use of examples. At Svendborg Municipality, the participants presented examples of big data being 'surveillance'. This is further addressed as something that big data can assist. Surveillance relates to what is also known as 'monitoring', but also to clarify the habits of people, as Knapp (2013) presents it, "[...] *discover your shopping habits*" (p. 215). That one can discover behavioral patterns using big data, is among others due to new innovations, such as footpath detectors, mobile applications and image analysis (Manyika et al., 2011, p. 68). The findings also show that the experience with big data is different from whom you ask. Finally, it is reasonable to say that no general definition of big data exists.

4.3 Big Data in the Danish Public Sector

This section seeks to explain and investigate how big data is, and to be applied, in the Danish public sector, drawing from both national- and municipal viewpoints. Hence, we seek to clarify the building blocks for the next section where the case of Residential Social Monitoring is presented.

Following the strategy by KL, 'Fælleskommunal digitaliseringsstrategi og handlingsplan' for 2016-2020, a clear goal was to promote and push for more innovation with the use of data. However, the common strategy has not put an overall pressure on the municipalities, to work more with data. As Carlsson explains, "*I haven't felt an increased pressure regarding the overall decision [...] instead I feel that it is our self who works with [data] and knows that we can benefit from it*" (personal interview, March 21, 2017, appendix G, line 38). Further, a more general agreement exists that it is the municipalities on their own, who can harvest the potential of including data in their work processes. Thus, a gap between the purpose of the strategy, plan presented and the actual effect it has made on the municipalities, seems existing. Despite the fact that no general urgency seems to exist, a clear belief is that big data holds great potential for the Danish public sector (Andersen, A., personal interview, March 21, 2017, appendix I, line 25; Hansen, S., personal interview, March 9, 2017, appendix F, line 40; Rasmussen, T., personal interview, March 29, 2017, appendix J, line 39). The future perspective is something that all participants happily share their opinion about and as Næsvang from Svendborg puts it, "*I believe that the soft areas in the municipality are very exciting*" (personal interview, March 21, 2017, appendix H, line 31). Here, he elaborates on how some social areas are

already characterized as being data-driven practices, where others, such as information regarding abuse of children, dentistry, vulnerability of citizens and the children and teenagers with special needs, adult disability and even in some elements of the job center and health perspectives, are lacking behind. Although some areas are lacking behind, it is important that, “[...] *we have some data that supports and tells us about what the world looks like, instead of just listening to stories within the authority departments*” (Næsvang, M., personal interview, March 21, 2017, appendix H, line 35). Even though a trust of other departments in the municipalities exists, data provides an overview and possibility of seeing new things that one can contribute extra time to (Næsvang, M., personal interview, March 21, 2017, appendix H, line 36). When asking Rasmussen from Randers, he puts on a ‘spokesman role’ for all the municipalities with a positive future belief that big data, “[...] *it’s getting more and more widespread because of the possibilities of coupling and combining new data sources than before*” (Rasmussen, T., personal interview, March 21, 2017, appendix J, line 40). These possibilities are something that Aarhus municipality completely master. Thus, based on his former job in the municipality, they are the key leader in the big data public sector race:

Aarhus municipality is following each and every residential area and thus they can categorize them after which one of them is the most vulnerable or not. I believe that this is the way to do things, but it certainly demands that you use the necessary resources for it to happen. (Næsvang, M., personal interview, March 21, 2017, appendix H, line 131)

However, several obstacles, including getting the resources needed, have to be confronted before utilizing the potential that can be archived. From a national point of view, Andersen elaborates, “*We have access to so much data in the public sphere, but no one actually uses it*” (personal interview, March 21, 2017, appendix I, line 10). One of the main reasons that data is not currently utilized is due to the lack in data quality, which is also recognized by KL as one of the main obstacles. (Hansen, S., personal interview, March 9, 2017, appendix F, line 51). To overcome this problem, Svendborg Municipality has started to implement new and more innovative information systems, rather than using their old electronic case and document management systems that do not allow data analytics. As Næsvang explains, “*We have just recently changed our old legacy system at the child and youth area [...] to DUBU which have a whole new set of data dimensions*” (personal interview, March 21, 2017, appendix H, line 46). Thus, this allows them to generate stronger management information and be more proficient and data driven, which is also something that the legislation and authorities

demand. Yet, lack of both technical and human resources seems to be a general challenges. From a national point of view, Hansen states that the field of data analytics might not be the one that receives the most resources from a central administration perspective (Hansen. S., personal interview, March 9, 2017, appendix F, line 50). This is also recognized by Andersen from Svendborg Municipality, “[...] *maybe it's not an area in which receives the major resources for it [big data] to be carried out with success*” (Andersen, H., personal interview, March 6, 2017, appendix E, line 115).

In spite the fact that more resources is needed, KL have been granted money for their initiative of developing the coupled KPIs, promoting the use of data (Hansen. S., personal interview, March 9, 2017, appendix F, line 63). They further help facilitate the connection between the municipality and government, by listening to the needs of the municipalities and then presenting several proposals for the government to evaluate in terms of distribution of resources (Hansen. S., personal interview, March 9, 2017, appendix F, line 169). According to Hansen, data is a key force in getting the necessary resources for the municipalities to do their work and continue to perform better, “*We would never be able to argument about anything towards the government if our numbers was not correct [...] it's crucial and something that you use in politics nowadays, that's numbers and data*” (Hansen. S., personal interview, March 9, 2017, appendix F, line 186). Yet, for now it seems that the municipalities must draw from their own existing workforce when applying data management and wish for more resources to be allocated in the future. As Andersen describes, “*I hope for it but I don't think we will get it [...] almost no municipalities adds more resources in that area*” (personal interview, March 6, 2017, appendix E, line 116). Thus, there is an existing need for economic funding and personal qualifications to assess the real potential of big data.

Even though a lack of both financial and personnel resources are present, there is a belief that big data as a buzzword or concept can still help getting closer the goal of becoming more data-driven, because it is popular. As Andersen explains, “[...] *you won't meet a leader who doesn't say “of course we should use data”* it rather seems to be a question of “[...] *how it should be used*” (personal interview, March 21, 2017, appendix I, line 89). Thus, for the benefits to arise, clear goals have to be developed in terms of big data's practical usage. That is why KL has moved towards the municipalities with the right help and guidance, for them to use data more:

[...] some of the things that we really work towards here in KL are to give the municipalities the best possible guidance to how they can easily make data available so that they can analyze and manage it. Some of the initiatives have been these publications

and websites we have made, but in general it's more important that they open their eyes to see how great the potential of using big data is. (Hansen. S., personal interview, March 9, 2017, appendix F, line 40)

The guidance that KL has facilitated, is something that KOMBIT has helped with, by endorsing municipalities and administrative units of Denmark to use even more data in their decisions, by inventing new strategies (Andersen, H., personal interview, March 6, 2017, appendix E, line 315).

Besides promoting the use of data analytics, KL uses it themselves to benchmark the municipalities and compare them to each other. By doing this, value is created in terms of showing for example the average expenses for each municipality on a national perspective, which enables a more efficient approach of making budgets and follow-ups on the performance. Big data can be viewed as a management tool, for KL to use in the broader perspective of facilitating the best interests for Danish municipalities. But also show the municipalities, where to improve interventions by letting data show patterns and developments in specific areas (Andersen, H., personal interview, March 6, 2017, appendix E, line 64). As stated by Andersen, *"KL uses all these key performance indicators (KPIs) across borders of the municipalities [...] typically 10 or 20 of them are chosen and a thorough examination and comparison of all the Danish municipalities are made each year"* (personal interview, March 6, 2017, appendix E, line 64). A clarification of how much money and resources have been used in an area such as vulnerable children and young people can thereby be shown. Yet, an 'efficiency measurement' and rating of municipalities are also made possible (Andersen, H., personal interview, March 6, 2017, appendix E, line 69).

One of the tools that allow the municipalities to work with big data is FLIS. The system is able to generate KPIs and also enables KL to gather information of each municipality. Even though FLIS is mostly seen as a data warehouse or business intelligence (BI) solution, for the single municipality to use, it is made clear that, *"[...] as such FLIS is both a solution made for KL as it is for the municipalities"* (Andersen, H., personal interview, March 6, 2017, appendix E, line 77). FLIS holds great potentials for future use as a big data tool, as it enables functionalities on data storage and extraction, making it possible to compare sector areas with the coupled KPIs (Hansen. S., personal interview, March 9, 2017, appendix F, line 71). There is an agreement between participants, that FLIS has the purpose of delivering increased cooperation between municipalities. However, Andersen argues, that even though 95% of the Danish population base is now covered and that most of the municipalities has implemented and bought the solution, *"[...] a lot of these municipalities are not*

using it to the extent it is meant for and some haven't used it at all" (personal interview, March 6, 2017, appendix E, line 280). Only a few minor municipalities are still not working with FLIS, but in general all others are starting to work against a more cooperative arrangement, so that everyone can benefit from the data that exists:

There are only a few minor municipalities who do not have FLIS. That's okay, we have allowed them to continue without, but all others are now in on the project and then they are all very interested because they now certainly know that they have to use it themselves since others are using their data and they are using others data. This also promotes a cleaning process of data. (Andersen, H., personal interview, March 6, 2017, appendix E, line 283)

This positive progression can be accredited to the new strategies that KOMBIT has developed the past years in pursuing the goal of promoting a more data-driven municipality and society (Andersen, H., personal interview, March 6, 2017, appendix E, line 315).

Despite the fact that the analytical systems are available, data is used very differently and in various degrees within the municipalities, where some are very creative and others are lacking behind. This is lately discovered by KL, who just recently in 2016 has made an effort on clarifying how data, management information and FLIS are used around Denmark (Hansen. S., personal interview, March 9, 2017, appendix F, line 207). The various reasons for how data is currently used, depends on which municipality you look at but a general rule of thumb implies that, *"[...] some of the smaller municipalities might not use it because they don't have the necessary resources an example could be an island municipality"* (Hansen. S., personal interview, March 9, 2017, appendix F, line 215). This is also something that Andersen is aware of and elaborates, *"Some new municipalities are only scratching the surface, since they only have a few responsible for the implementation"* (Andersen, H., personal interview, March 6, 2017, appendix E, line 25). Another reason for lower data commitment is also found in the lack of a data mind-set within the executive board, politicians and workers of the municipality, *"One has a reserved attitude towards data and big data that you have to be careful using it because you cannot trust it and that you are skeptical towards the registration and accounting practices is it done correctly and such"* (Hansen. S., personal interview, March 9, 2017, appendix F, line 217). Some municipalities are disbelieving the ability of comparing across borders of the municipalities due to the different registration practices and that one municipality are not identical. When looking at two municipality's placement costs and that they are 1 million apart,

one would say, “[...] *is this the reality or is it rather because we have different ways of doing things?*” (Hansen. S., personal interview, March 9, 2017, appendix F, line 228). To prevent this problem and to promote the usage of big data within the Danish public sector, some municipalities have started creating what Hansen calls, *“Very smart solutions for Ipads, so that politicians easily can open an app and then all the necessary information and data and other sources are shown easily”* (personal interview, March 9, 2017, appendix F, line 246). Thus, these innovative solutions, complements the belief that big data needs to be easy accessible for a politician to understand (Hansen. S., personal interview, March 9, 2017, appendix F, line 249). Examples are also given that municipalities across Denmark forms networks of similar municipalities to help in pursuing more efficient use of big data and to benchmark against each other’s (Andersen, H., personal interview, March 6, 2017, appendix E, line 128). GeoFyn is given as an example of where creativity is widely incorporated when working with data (Hansen. S., personal interview, March 9, 2017, appendix F, line 234).

Where small municipalities do not use data analytics in a broad extent, larger municipalities are very positive and committed about the usage of data in their work, *“Copenhagen and Aarhus are going crazy [...] they have a management information department in each of the magistrates exclusively working with data, presenting, collecting and extracting data”* (Hansen. S., personal interview, March 9, 2017, appendix F, line 220). Yet, it is not only the larger municipalities whom have implemented data analytics. Others, such as Svendborg and Randers, also agree upon the fact that data is everywhere and they work with it in each and every work situation and decision. As Andersen explains, *“I don’t think I can mention one area in which we do not include data”* (Andersen, A., personal interview, March 21, 2017, appendix I, line 83). That data permeates, is also recognized by Rasmussen, *“Yes, we use it all the time”* (personal interview, March 29, 2017, appendix J, line 38). Indeed, this is also what KL experiences:

Some municipalities are using it in everything they do. For example, when they present something for the local council numbers are always attached. And they follow it both on a daily and monthly basis, how is it going with our different areas and the goals thereto. (Hansen. S., personal interview, March 9, 2017, appendix F, line 209)

Even though data permeates the work processes of the municipalities in general, some areas are still more data surrounded than others. Svendborg is for example very experienced with data registrations and handling within the more technical areas, such as nature, climate/environment, property management and construction (Andersen, A., personal interview, March 21, 2017, appendix I, line

95). The way that the municipalities work with data has changed. According to Andersen, “*One has always worked with data as management information*” (Andersen, H., personal interview, March 6, 2017, appendix E, line 11). However, municipalities used to work with data in ‘pillars’ and now data is particularly merged together with other data, in new ways to present new knowledge and make decisions easier for example in residential social monitoring (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 45). This is furthermore supported by KL, who argues that one can now see,

Which kinds of residential areas are vulnerable and where is it our elders live so that we can plan where to place the next nursing home? Where is it our children live, so where can be shut down schools if needed? Where is it our refugees and immigrants live, in regards to planning the needed effort towards them? Where do the children do worst in the education environment, so that we know where to place the most effort? (Hansen, S., personal interview, March 9, 2017, appendix F, line 241)

That one can gather a full perspective of the municipality, including economical and goal oriented perspectives, is indeed what that Andersen believes the most experienced municipalities typically use big data to. Accordingly, FLIS functions as a great tool with over 1000 KPIs to generate an overview of the situation in the municipality regarding different areas such as social, residential and education (Andersen, H., personal interview, March 6, 2017, appendix E, line 23-30). Yet, also to reach own goals (Andersen, H., personal interview, March 6, 2017, appendix E, line 10). Furthermore, FLIS can be used in the context of, “[...] *showing what works in different areas across administrations [...] it is not just the economical part but also something as how much truancy exists within the primary school*” (Andersen, H., personal interview, March 6, 2017, appendix E, line 31). Information and data are thus extracted and used in different settings with other data from their data warehouses or business intelligence solutions that the municipalities might already have. Allowing them to compare and contrast upon the data from the various solutions, before making decisions (Andersen, H., personal interview, March 6, 2017, appendix E, line 49). Thus FLIS and big data have been widely used in making economic decisions, looking at future specific efforts and generating a ‘situation overview’ for the municipality to succeed.

To sum up, the findings show that great potentials exist within the future utilization of big data in the Danish public sector. That great potentials exist within data analytics in public sector organizations, is supported by several authors, whom claim that big data is indeed important for the future e-

government. According to Pollitt (2014), big data is to be considered the most important tool for future public administrations. Further, OECD (2015) explains how the public sector is one of the most data intensive areas and therefore has great potential, in deriving value from big data. A clear agreement exists that big data shall be seen as an important and valued strategic asset (O'Brien, 2012), which benefits the citizens of the municipalities and society (Thompson et al., 2014). Even though great potentials exist within the use of data analytics, several challenges need to be addressed before leveraging the full potential in the Danish public sector. That general challenges exist is what generally is found in literature. Firstly, according to (Klievink et al., 2016), readiness of big data in knowledge creation and decision making is a challenge. Secondly, challenges exist within the privacy and data quality aspects (Cheng, 2014). Thirdly, information and data is not yet completely shared between borders (Gamage, 2016). Fourthly, a shared belief that data standardization has to be addressed even more, before the real potential is in reach, which is also stated by Bays (2015), who focus on 'data inconsistency' within public sector administrations. Fifthly, agreement among participants exists that the necessary resources are not present to harvest the full potential of big data, even though literature argues that countries are massively investing in big data (Gamage, 2016). Yet, it is rather a matter of the human resources needed, for the release of the full potential of big data within the Danish public sector. This is due to the fact that the municipalities lack important professional, analytical and managerial talent (Manyika et al., 2011). Finally, challenges regarding to the organizational change that the municipalities might encounter, is also important to emphasize. As the findings show, some municipalities are very creative where others are lacking behind. This is further supported by Corbin (2012) who states that the reason for this difference should be found in the competitive environment that the governments operate in. Since the pressure from the environment is not existing in the same way as in the private sector, public sector organizations and administrations might lack behind simply because of the motivation to use data (Ibid.). However, the representatives still argue that big data is a buzzword (Jordan, 2014), which helps towards the harnessing of the full potential.

4.4 Residential Social Monitoring

The next part is concerned with the specific case of Residential Social Monitoring, a tool and method that is implemented at Randers, Svendborg and Helsingør Municipality. Overall, the findings show a clear conviction that the case is a big data project, since it uses FLIS as together with other data sources and systems. As clarified by Andersen, "*Yes, I believe so*" (Andersen, H., personal interview,

March 6, 2017, appendix E, line 17). Being a big data project is furthermore supported by KL, who highlights the terminology of big data being the key factor that determines that the case is a big data project. Large datasets are combined and arranged in an understandable manner with the result of taking decisions (Hansen. S., personal interview, March 9, 2017, appendix F, line 362). That the case consists of massive datasets is agreed upon by Svendborg, where one participant explains, *“It’s a project where we combine massive amounts and various data, that’s how I would characterize it”* (Andersen, A., personal interview, March 21, 2017, appendix I, line 30). Further, the combination of data sources is important in the context of residential social monitoring, since it helps the municipalities raise the quality and the safety of having newly updated data. Thus, the combination allows that the necessary responsibility can be taken in the right place (Andersen, A., personal interview, March 21, 2017, appendix I, line 143). In the study, different examples of data combination from both internal and external sources are presented. However, where FLIS is primarily used as the foundation, some ‘OPUS’ municipalities are using their specific systems to deliver data that they combine with the one FLIS provides (Andersen, H., personal interview, March 6, 2017, appendix E, line 119). The reason for this is because FLIS do not deliver all of the data needed for each and every sector area. Hence, municipalities then include data from various registers, such as ‘Indenrigsministeret’ data, combined with certain KPIs, and Danmarks Statistik (DST). Furthermore, they are also using data from their other legacy systems, which they usually combine it with their Business intelligence systems (Hansen. S., personal interview, March 9, 2017, appendix F, line 310). External data plays an important role in making the map based monitoring, as elaborated by Andersen:

We pull a lot of external data in, the whole BBR is external, our Geo Denmark data makes the foundation for the mappings that we have in KRAK and so forth. The addresses are also an external data source and the whole ‘grund’ data system. So it is something that we really work with. (personal interview, March 21, 2017, appendix I, line 140)

Thus, the focus is to combine the existing knowledge and wisdom to generate new value and knowledge (Andersen, A., personal interview, March 21, 2017, appendix I, line 31). Even though all of the municipalities’ systems are used to some extent, the combination of data still presents, especially when it comes to the extraction of data from several systems also including FLIS (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 80)

Residential Social Monitoring is a project facilitated and initiated by KL, who as mentioned in the latter section, works as a supportive role for the Danish municipalities. KL has taken on the role to foster more cooperation between the municipalities, since they do not talk together:

Well, we have been facilitating it. [...] since the municipalities are sitting separately, working on things, how do you precisely start the conversation between them. And how is it that we exactly [...] create this project and how do we gather the municipalities? We facilitate and make sure that it happens and connect the municipalities. (Hansen. S., personal interview, March 9, 2017, appendix F, line 326)

Thus, with the facilitation, KL has worked intensely to demolish the silos that the municipalities have consisted of, concerning only their own errand to becoming a more cooperative environment. However, whether the project has been successful and turned into real practice and results seems rather unclear, *“Well I am not sure. But some municipalities like Hedensted and Svendborg have come really far with it and they have used it and it looks good”* (Hansen. S., personal interview, March 9, 2017, appendix F, line 335). Although KL has facilitated the project and made efforts to help the municipalities, they are still cautious in making decisions based on the method and the big data project itself (Hansen. S., personal interview, March 9, 2017, appendix F, line 336). According to KL, this is a shame because that is what, *“[...] really matters, to make some decisions based on it [...] that is something that I did not feel that they had done yet”* (Hansen. S., personal interview, March 9, 2017, appendix F, line 341). Thus, it seems reasonable to say that the municipalities are still not data minded enough, to only draw conclusions based on numbers. Further, they still use their expertise as Andersen from Svendborg points out, *“We have been four different professionals involved in the facilitating process where others have also been associated when needed. We have used their views, expertise and comments towards it”* (personal interview, March 21, 2017, appendix I, line 36). The main reason for joining the project itself is explained by Carlsson, *“[...] it's simply because we need more tools to work with”* (personal interview, March 21, 2017, appendix G, line 52).

Despite having facilitated data analytics, Residential Social Monitoring has also worked as an idea generator. Randers Municipality now pursues a different approach, where innovative initiatives are developed based upon the original method and idea introduced (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 51). Namely, the idea of using map-based information is something that Randers has used in new innovative ways:

We actually don't use the map based part that much in sense of the map itself, the visualization as a map. In reality we use address localization of people in regards to that we know where they live. Thus, we use the geographical areas that the residential areas are when it comes to monitoring. (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 53)

Thus, Randers are still using the main concept of address monitoring i.e. the method developed by KL, but has moved away from the technical part of it in the use of GIS and other geographical based information systems (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 71). The overall purpose of the project is still essential, as Rasmussen explains, “[...] *a lot of the things in the guide is something that we are doing and of course we have gone through a process with selection between different indicators that we monitor in the case of residential and social areas*” (personal interview, March 29, 2017, appendix J, line 77). Even though Randers Municipality has changed direction, they clearly believe that the project is extremely relevant and important for them to continue with, since they are able to spot ‘something’ in advance or predict a certain outcome (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 125). When it comes to the future aspect, Rasmussen believes that additional potential can be released in the area of general monitoring, if the municipality, “[...] *followed all areas within it*”. Thus, they can be even more effective in spotting areas that have a bad development (personal interview, March 29, 2017, appendix J, line 126). What initiatively started the residential social monitoring in Randers is to be found in the master's plan, developed and financed by ‘Landsbyfonden’ (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 63). Presently, Randers is closely following three specific residential areas, which they have been granted extra financial support to make focused efforts within (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 56). Monitoring is done in the residential areas, which have shown to be the most vulnerable, where they are already making efforts and where they are aware of the ‘things’ that is happening (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 133).

With the use of maps in monitoring perspectives of vulnerable and exposed areas, this enables the municipalities to increase the communication value of data by creating an overview of the situation (Andersen, H., personal interview, March 6, 2017, appendix E, line 105). Further, maps help decision makers to orientate themselves in the vast amount of information that exists, an area which will be elaborated more on later in this analysis. With maps, Svendborg municipality is now able to

differentiate themselves when it comes to delivering solutions regarding to residential social monitoring (Andersen, A., personal interview, March 21, 2017, appendix I, line 61). This is done by making interactive maps, which simply makes the world easier for everyone to understand as KL explains (Hansen. S., personal interview, March 9, 2017, appendix F, line 87). Among other innovative solutions, the interactive maps help the municipalities in presenting which area should be financed when it comes to vulnerable children for example. The main reason is that the maps allows the politicians to actually look at what is going on within an area and thereby make better decisions based on it (Hansen. S., personal interview, March 9, 2017, appendix F, line 260).

In the future the case of Residential Social Monitoring can lead to further positive things. According to Carlsson, it can deliver increased internal cooperation, but more importantly that one can prioritize between the resources given to certain areas that are monitored, thus delivering an increased understanding (personal interview, March 21, 2017, appendix G, line 186). However, before additional potential can be released several challenges exists. Firstly, data quality is something that Randers highlights as the main obstacle. Secondly, that more resources are needed i.e. financial and personnel. When these are addressed, the project can really become a 'live', 'day to day monitoring' project which show trends and enable that one can effectively do specific efforts in areas that you do not monitor already (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 143). Massive potential are locked and needs to be addressed to achieve the full potential:

I see a great potential, that we in reality use residential social monitoring and monitoring in general as a tool to "spot" the development before, thus as a predictive tool or forecasting tool. If you will, to show where some developments are occurring and that these are some you are interested in strengthening, or some you might want to brake before they evolve to something that you can't control. All in all, the fact that you can follow the development of the municipality is something I think that monitoring and also geographic monitoring will only be even more prominent in the future.
(Rasmussen, T., personal interview, March 29, 2017, appendix J, line 366)

A main part of Residential Social Monitoring is the use of data on vulnerable children and young people. According to Hansen, KL has worked closely in publishing material and other relevant information on the subject (personal interview, March 9, 2017, appendix F, line 58). The cooperation has resulted in many, "[...] exiting things, since there is an existence of so many different patterns" (Hansen. S., personal interview, March 9, 2017, appendix F, line 73). Furthermore, the project is also

focusing on measuring the effects of specific efforts made and which patterns and developments are showing in an early stage so that the necessary actions can be taken in time. As Andersen explains, *"It is possible to follow the child all the way up through the system"* (personal interview, March 6, 2017, appendix E, line 213). Thereby, with the use of big data, it is possible to improve the conditions for the vulnerable children and young people.

4.5 Knowledge Creation and Predictive Insights

Within the operationalization of knowledge creation theory, the decision to use big data in the Danish public sector rests on a clear conviction that it answers questions that were previously unknown. As stated by Hansen, this is an area where the municipalities have returned to Kommunernes Landsforening and said that the use of big data generally provides them with answers that could not have been answered without extensive use of data analytics. Further, data reflects how incredible large differences that currently exist between municipalities, differences that are to be investigated and addressed (Hansen, S., personal interview, March 9, 2017, appendix F, line 120). This is further explained by Andersen, who argues that the use of big data within the topic of vulnerable children and young people may help answer questions like how many vulnerable children and young people are forcibly removed in some form or another and how many continues to stay home with their parents and getting better (personal interview, March 6, 2017, appendix E, line 213). Also, this is something that you really want to measure the effect of, in spite of the fact that such information can be used for preventive and evident purposes, why Andersen argues that data from many years back is genuinely important to be included in FLIS.

In relation to the case of Residential Social Monitoring, Carlsson explains that the use of big data has provided Svendborg Municipality with knowledge on how bad the situation really is, *"[...] it has surprised me that it is so evident all around, how bad it really is"* (personal interview, March 21, 2017, appendix G, line 89). Accordingly, the use of big data, and yet the unpleasant insights from data analytics, has today affected the municipality's master's plan for 2018 – 2021 to be considerably expanded from two exposed residential areas to involve eight areas (Carlsson, J., personal interview, March 21, 2017, appendix G, line 63). The surprise is moreover experienced by Næsvang, who puts it like this, *"[...] yes, it can be fairly easy to see if there is an area that sticks out of which one had not believed that ran"* (personal interview, March 21, 2017, appendix H, line 99).

At Randers municipality, Rasmussen agrees that residential social monitoring have verified their immediate understanding of how things currently are:

[...] it has clearly contributed to the visibility of a lot of conditions that we might have had an idea of how things were, if nothing else have been made it more obvious and articulated with the use of more data in the analytics processes. (personal interview, March 29, 2017, appendix J, line 96)

At Randers Municipality, the use of residential social monitoring have made them more aware of the development in how many children in exposed areas that go to school in the same neighborhood, which surprisingly not that many do. The surprise lies in that they have not earlier been able to couple data between children and school, which obviously is a foundation for some of the social services that were initiated. Rasmussen brings up the example that if you want to measure the effect of social services to combat school absence, it is crucial to know whether the students actually go to the particular school (personal interview, March 29, 2017, appendix J, line 110).

Not only has the use of big data provided knowledge on what is currently happening at this very moment, but a clear conviction exists that social residential monitoring also have provided adequate awareness on developmental trends in society that is highly useful for predicting future outcomes with a high level of reliability, based on detected, recognized and accepted patterns in data. According to Hansen, FLIS offers a wide range of key figures for the past five years, which means that trends between these data easily can be viewed (personal interview, March 9, 2017, appendix F, line 142). Further, from the perspective of KL, data from FLIS is primarily used within the areas of economics and specifically budgeting. However, not only is spending patterns investigated across municipalities, but also certain activities related to vulnerable children and young people are studied, seeing that this group of people makes up a relatively large part of the budgets. That is for instance whether more or less children are attending school as they are supposed to.

In relation to residential social monitoring, the use of school data is, according to Næsvang, very valuable, as if such data over a longer period of time for example shows falling grades in primary school or a later school start may inform us about what is happening in these concrete areas that is needed to be taken action upon (personal interview, March 21, 2017, appendix H, line 113).

This is further explained by Carlsson, who mentions that the technical features of GIS makes it possible to 'boil' it down to a lot of small areas, which is extremely useful knowledge (personal interview, March 21, 2017, appendix G, line 74). Lately, detailed insight into the municipality's residential areas have proved to be particularly helpful in the decision of where to put refugee children and families, to be sure that the municipality can handle them properly as well to estimate and predict how the areas will evolve over the forthcoming years (Næsvang, M., personal interview, March 21, 2017, appendix H, line 113).

Despite the fact big data has provided knowledge on developmental trends that can be used for predicting certain outcomes, it is generally difficult within the area of vulnerable children and young people. The municipalities' strategies on how to best handle these children and young people changes over time and make it somewhat difficult to compare the evidence-based indications you work with (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 119). Further, if you want to predict certain residential outcomes, it is important to monitor the entire municipality, and not just certain selected areas, which is what currently is being done at Randers Municipality:

[...] if you want to spot something in advance, I would say that in a monitoring perspective, it is optimal that you follow the entire municipality and not just the public estates that knowingly are most at risk, with the purpose to investigate and intervene at an early time if an area might be in a bad development. (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 124)

That big data has provided new knowledge and answers to formerly unknown questions, relates well to what is considered in big data literature. According to Munné (2016), big data analytics can be used to analyze large datasets for problem solving that can reveal data-driven insights. The abilities of such can for example be comprehended through detected and recognized patterns and to produce forecasts, such as planning purposes and public services (Yiu, 2012). Particularly the statement of how big data influences the master's plan of Randers Municipality is an example of planning new purposes and public services within new areas. This can further be connected to the idea of creating clear goals (Jacobsen & Thorsvik, 2008), which then allows one to come nearer the ideal state of perfect rationality (Simon, 1997). That big data has been able to make predictive insights by looking at trends in data over years, is supported by for example Schmarzo (2013), who states that big data

allows organizations to move closer in being a 'predictive enterprise'. Further, from the thoughts of Knapp (2013), big data is considered to create new knowledge as large datasets offers a higher form of intelligence and knowledge that may be able to generate new insights and thereby also deliver new increased forms of value (Schönberger & Cukier, 2014, p. 6). To this, Elgendy & Elragal (2014) and Ohlhorst (2012) state that big data has the potential to transform information to new and valuable knowledge that can be used to solve problems, increase profits and productivity, find opportunities and develop working patterns for as well private and public sector organizations. The findings show that municipalities are using existing systems, internal data and external data with new smart big data tools such as visualization i.e. using maps. Thus, they follow what is said in literature, where Marr (2015) who argues that exactly the combination generates improved insights (p.10), which then leads to new knowledge and information (Fredriksson, 2016).

From the theory of organizational knowledge creation by Nonaka (1994), the essential feature is to create new knowledge, compared to processing already existing information, which without doubt, is being done in this case. Big data offers municipalities with knowledge on what happens around in different areas and transfer the newly created knowledge to the necessary stakeholders, divisions and people, whom can easily understand and interpret it. Thus, big data allows working with explicit or codified knowledge, such as visualization (LaValle et al., 2011), in new and better ways than before (Fredriksson, 2016). Big data compliments the idea of tacit knowledge formerly hold by experts and caseworkers, transfer into explicit knowledge. Hence, the knowledge derived from big data can be seen in the 'combination' phase of the SECI process (Nonaka et al., 2000), where knowledge is gathered from explicit sources and undergoes a process of being transformed into new explicit understandable knowledge (Nonaka, 1994). However, it seems reasonable to say that the newly created knowledge and insights are still a part of a tacit knowledge creation process. The main reason is that municipalities have analysts working on the data itself and then experts using the provided knowledge (e.g. visualization in residential social monitoring) to make decisions. Furthermore, it is argued by the municipalities that an overall idea of the situation, due to tacit knowledge has always been present. Yet, big data has created a more "clear picture" of the reality, hence explicit knowledge. This have for example resulted in better allocation of immigrants, because the knowledge derived from big data has laid the foundation of decisions (Grant, 1996). As explained by Marr (2015), big data and info graphics is something you allocate to the data scientists, who process it. However, there is still a need for collaboration, between the analysts and the decision makers to generate value (p.

185). Thus, the case of residential social monitoring and big data in general within the Danish municipalities, paves the way towards explicit knowledge creation, but it is still in the category of a never ending process (Nonaka et al., 2000).

4.6 Decision Making and Strategic Planning

Within the operationalization of decision making theory, the use of geospatial big data analytics rests on a clear conviction that it supports decision making processes. As stated by Andersen, it is really useful to integrate and utilize geo codes with data from FLIS as they have done in Svendborg, Randers and Helsingør Municipality (personal interview, March 6, 2017, appendix E, line 105). Thus, it is possible to view small areas in a municipality to see whether the development is in a good or bad direction:

Instead of just pulling raw data out themselves, they can now see it on a map. You may say that it could also just be seen in a spreadsheet, but having it on a map makes it more visual and for those who actually make the decisions higher in the hierarchy is typically more visual brought. Thus, you may look at the map and easily figure out of whether Svendborg is green, yellow or red, relatively compared to other municipality on Funen. (Andersen, H., personal interview, March 6, 2017, appendix E, line 110)

This is further explained by Hansen, who argues that maps contain larger promotion value that helps to get a better overview of the current situation that is investigated and thus make better decisions, “[...] it strengthens our ability and capacity to make decisions, better decisions. When we have it visually we can better and more easily get an overview than looking at data tables, which sometimes can be difficult” (personal interview, March 9, 2017, appendix F, line 257). Furthermore, Hansen agrees that decision makers are, without doubt, more visually brought, especially when it comes to politicians, whom might have a hard time to understand certain lower-level issues (personal interview, March 9, 2017, appendix F, line 245). The use of geospatial big data analytics is not only taking place in the local municipalities, but also at Kommunernes Landsforening, it is becoming much more prevalent, “[...] we actually have a unit that only works with data analytics in the building, including geo analytics, which are highly relevant we make new common strategies or new objectives for the Danish municipalities” (Hansen, S., personal interview, March 9, 2017, appendix F, line 167). It is among other used when developing national goals where you look into areas all over Denmark to

view and compare e.g. how many readmissions that take place at hospitals. Readmissions account for a relatively large part of the budgets and if there is tendency that more citizens are sent to the hospital, the municipalities must make greater health efforts with home nursing or similar social services (Hansen, S., personal interview, March 9, 2017, appendix F, line 197). Geospatial big data analytics can help view such development in an easily and interactive way and thus provide information and knowledge on what to believe is ahead and what social services to implement to comprehend the development. Hansen even goes so far and says that almost every decision and goals are in some degree or another is, “[...] *permeated by big data*” (personal interview, March 9, 2017, appendix F, line 204).

In the context of residential social monitoring at Svendborg Municipality, Næsvang states that the use of geospatial big data analytics in regards to vulnerable children and young people helps align and contribute to a common understanding of a given situation and thus be in agreement of what social services to be initiated:

[...] it can be really difficult to work within the area of children and young people and understand what adult caseworkers talk about when they say that there is a great drug abuse in certain areas. Here, map-based solutions make it easily accessible and possible to put multiple disciplines together to provide greater social services and achieve a common understanding of whether the world is red, yellow or green. (personal interview, March 21, 2017, appendix H, line 53)

Achieving a better understanding with the use of maps is further explained by Andersen, who argues that geospatial analytics provides a great form of communication that quickly provides a common understanding of what is studied. Further, the use geospatial analytics on lower-level areas helps to differentiate and view shades that may stand out (Andersen, A., personal interview, March 21, 2017, appendix I, line 61). This is further explained by Carlsson, who state that decision making processes have become a whole a lot easier with the use of geographic information systems, seeing that it provides non-experts a greater overview and common understanding, “[...] *it is so much easier to get an overview through GIS and also more straightforward when it comes to convince non-specialists on how the world look, that being red, yellow or green*” (personal interview, March 21, 2017, appendix G, line 95). Carlsson continues by explaining how easier it is with GIS to stage and work

towards a common understanding with decision makers, whom are perfectly clear on what is to be dealt with and often surprised on how incredible smart the systems are, “[...] *the systems have become so advanced that you are able to work with all these crazy many numbers. And everyone say so and when we talk to novice people they almost scream “wow, why haven’t we seen this before”* (personal interview, March 21, 2017, appendix G, line 95). At this point, the use of geospatial big data analytics is primarily used internally in dialogues with management in the municipality and especially when it comes to convince the allocation of economic resources to you, which sometimes can be a ‘fight’. Here, besides geospatial analytics helps you to create a common understanding by visualizing what it is about in an interactive way, the advanced systems make it possible to present hard facts so the decision making process is not just being pure politics with arguments without content. By pinpointing locations to being either green, yellow or red, based on a number of data inputs, is way to involve factual information, thus arguments become common understood and decisions easier to follow through (Carlsson, J., personal interview, March 21, 2017, appendix G, line 54). Accordingly, in a decision making process with funding on the agenda, it is obviously necessary to present good arguments or else you will be ‘run over’, “*You're got to have some good arguments and here the use GIS to present data-based arguments have proved effectively, seeing we have been allocated more money than earlier*” (Carlsson, J., personal interview, March 21, 2017, appendix G, line 175). This is further explained by Næsvang, who argues that they internally have been able to make better decisions within the group and a lot of goals have already been achieved with the continuous strategic use of management information (personal interview, March 21, 2017, appendix H, line 117). The comments by Carlsson and Næsvang agree well with what Andersen says, that being the use of GIS with different data input provides a higher level of knowledge that supports the decision making processes, primarily at local politicians:

[...] the use of GIS provides us with more knowledge and the more knowledge we have about the decision that is to be taken, the better decision can be made by the politicians, i.e. on an informed basis, whom ultimately are the ones making the final choices. GIS can help provide with more accurate analytics and thus offer an easier approach to get these decisions done in a correct way. (Andersen, A., personal interview, March 21, 2017, appendix I, line 104)

Also at Randers Municipality, the use of GIS is said to be naturally used in wide range of decisions, *“Primarily GIS, but also in general data-based knowledge, is a part of our decisions and we do not make any decisions without having worked with data analytics before it”* (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 200). The decisions who Rasmussen mainly talks about comprehends the establishment of goals and visions for the future, seeing that the use of integrating historical data on among other vulnerable children and young people in GIS provides adequate awareness on developmental trends in society that is highly useful for predicting future outcomes with a high level of reliability, based on detected, recognized and accepted patterns in data. With an extended use of big data analytics it has become possible to propose concrete and measureable goals and visions that eventually can be followed up on:

[...] yes, goals and objectives have earlier been dominated by being highly complex and diffuse that you cannot really follow up on, but now it is possible to create milestones that you then can substantiate performance if you like. Thus, you come one step closer to the more general goals and objectives and it helps that we can actually answer “how many students in what areas have over 20 days of school absence”, thus answering whether more or fewer students stay in school at a time after a social initiative has been carried out. So yes, it helps that we have data in most areas that we can connect and view in GIS. (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 209)

The use of data analytics within a strategic perspective is also presented by Næsvang, who argues that the most important factor when developing new strategies is to create a common understanding, which is especially important when having new knowledge to present. This is where geospatial big data analytics really comes in handy, *“When you sit 10-12 people around a table and looking at these maps, it is quite easy to go from there to overall strategic decisions”* (personal interview, March 21, 2017, appendix H, line 129). This is further explained by Carlsson, who mentions that Svendborg's new master's plan is built upon geospatial big data analytics, where the strategy now involves eighth residential areas. That geospatial big data analytics is useful within a strategic context, are also mentioned by Rasmussen. According to Rasmussen, Randers's new strategic visions and goals are closely related to the use of data analytics, which has provided them with new knowledge on where to insert and improve social services:

The plan is very much concentrated on children and young people alone. The old one was about employment rate in general and level of education in the entire population, but children and young people have proved to be much more urgent, most definitely because we have not been able to see the things as we can today. (personal interview, March 29, 2017, appendix J, line 171)

That big data has provided the municipalities with improvements in effectiveness relates greatly to what is considered in big data literature. According to Munné (2016), big data analytics can provide greater internal transparency, thus being able to take better decisions and be more effective, and even create new services, by letting data ‘speak’ between multiple decisions Fredriksson (2016). Further, as Power (2014) points out, big data allows anyone to make better decisions, which certainly takes place within the municipalities. Especially by using geospatial data, the municipalities have been able to make decisions based on the risks within the different residential social areas. This is furthermore supported by Marr (2015), who mentions that big data creates new possibilities of visualization with such as info graphics. Hence, it contributes to a higher level of understanding information and thereby makes it possible for decision makers to initiate interactions and conversations that can lead to better and improved decision making and performance (ibid.). As stated by several municipal representatives, it is now possible to ‘color the world’, which then makes the case of residential social monitoring suited in terms of a successive info graphic, since, “[...] *the person looking at the information can pull out the key messages that suit their agenda or confirm their preferred decision*” (Marr, 2015, p. 180). Further, visualization makes the information and data ‘come alive’ and thereby trumps intuition (LaValle et al., 2011), which certainly takes place in the case, whom are moving towards perfect rationality in decision making. Since the politicians and decision makers are more visually brought, it makes sense to use maps and geospatial data, which is also mentioned by Marr (2015), who argues that it is effective, because, “[...] *it appeals to the way human beings receive and process information*” (p. 180) and thereby allow important information and knowledge to flow within the borders of the municipality.

From a theoretical point of view, the case organizations use big data to form goals and objectives of upcoming strategies and decisions. It is especially interesting to see how big data allows KL, from a national point of view, to form overall goals for Danish municipalities, where the municipalities themselves use it to be more specific in their strategies towards residential social monitoring and

targeted social services. As mentioned, big data has allowed Svendborg to create specific, concrete and measureable milestones and goals before the introduction of GIS. Thus, the goals are following Hansen's statement, i.e. decisions are permeated by big data and it permits the development of clear goals, which is important in getting closer to the 'economic man', who acts with perfect rationality (Simon, 1997). In relation to what EY (2016) present, big data prevents the cognitive bias of the human to occur, which of what can be confirmed in the case of residential social monitoring, seeing that one creates new strategies, goals, objectives, milestones and targeted services based on data.

To sum up, big data is able to take one closer to the ideal state of perfect rationality in decision making, due to the ability of gathering complete information and creating clear goals, thus being able to base strategies and decisions on relevant acquired knowledge. Thus, as we have seen throughout this section, decision making and knowledge creation are fuel for each other, which are generally acknowledged in literature (Sorofman, 2013; Manyika et al., 2011, Raghupathi, 2014, Feldman & March, 1981).

4.7 Targeted Social Services and Evidence-Based Practice

A general belief exist that targeted social services is the primary goal and what is wishfully thought. According to Hansen, the municipalities can better determine the required social services with the use of data analytics, such as knowing if vulnerable children and young people perform below average at primary school's final examinations. Thus, having such knowledge can help with the work of what social interactions that is required, such as inserting more teachers and thus improve the services (Hansen. S., personal interview, March 9, 2017, appendix F, line 95). This is further explained by Næsvang, who argues that within the soft areas, that being vulnerable children and young people, a lot of data-driven improvements exist:

Well, it is actually a very simple question. I do not think that anyone would disagree on the fact that well placed resources and early social intervention is what we all dream about and within the soft areas such as vulnerable children and young people that is where there really is something to achieve. (personal interview, March 21, 2017, appendix H, line 155)

At Randers Municipality, Rasmussen agrees that the 'soft areas' are interesting and the use of data really has an effect. An example of such is school absence, where due to the fact that registrations now are improved, knowledge on how big a problem absence really is, is recognized. Accordingly, initiatives are now approved on social interventions to turn this development around, driven by insights from data analytics, *"It is a selected area which we now insert extraordinary support. The majority of our ideas are most definitely related to data and new knowledge from it"* (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 322). That action is taken to improve targeted social services by the use of data analytics is further explained by Carlsson. Accordingly, the use of geospatial big data analytics in residential social monitoring is most lately reflected in last year's budgets, where extra financial support was given to establish small affordable houses, primarily led by the pressure of refugees, covered in terms of the cuts in the financial benefit reform. Thus, in terms of urban planning, larger buildings have been knocked down and smaller buildings constructed that people could rent for less than 3000 DKK a month, which is the financial limit (Carlsson, J., personal interview, March 21, 2017, appendix G, line 128).

Targeted social services are closely related to evidence-based practice, which is the study of what works best. According to Andersen, it is an area of immense importance, seeing that the public sector spends quite a lot of money on social services, *"[...] we use an inconceivable amount of money in Denmark on all the social services to e.g. get them into work, but do the services really have the desired effect?"* (personal interview, March 6, 2017, appendix E, line 237). This is further explained by Hansen, who argues that the municipalities, with insights from big data, now are able to streamline their social services:

One can streamline services because you now know what social interventions work and which do not. As an example, you are now able to view whether or not preventive measures have the intended effect, such as if children perform better at school with the given intervention. Thus, by making impact assessments with coupling of figures, you may dismantle the particular service, if it does not work. (Hansen, S., personal interview, March 9, 2017, appendix F, line 99)

In relation to case of residential social monitoring, Carlsson explains that they have tried to include the evidence-based indications from the 'Nationale Forskningscenter for Velfærd' in their work on

exposed areas in the municipality, “[...] *the indications have helped us in identifying and pinpointing the individual subjects and areas we measure on, when comparing several different measurements*” (personal interview, March 21, 2017, appendix G, line 152). That the indications help provide valuable information on exposed residential areas is also experienced by Rasmussen at Randers Municipality. Rasmussen reveals that the indications are widely used and something they increasingly follow up upon to categorize how exposed the individual areas are, “[...] *indications such as student absence, placements and grades have provided knowledge on how vulnerable the areas really are*” (personal interview, March 29, 2017, appendix J, line 271). Rasmussen continues by saying that the large amounts of data they have available today have improved their evidence-based practice in spite of the fact they now are able to follow up upon whether or not a student's school absence is changed after a particular social service intervention is carried out. This is further agreed by Andersen, who mentions that they at Svendborg Municipality, due to their work with big data, have achieved higher internal efficiency and better daily routines. However, despite the fact that data tells a lot, it cannot stand alone, “[...] *the data cannot stand alone, we can monitor but it may not tell us the whole story about the effect of something. So it is important that we constantly have professionals included*” (Andersen, A., personal interview, March 21, 2017, appendix I, line 38).

That big data has provided the municipalities with improvements in efficiency relates great to what is considered in big data literature. According to Munné (2016), big data analytics can improve efficiency, such as the ability to provide better services and continues improvement based on the personalization of services and learnings from the performance of such services, that being evidenced-based practice. Hence, evidence-based practice is seen as a vital area within data analytics, seeing it allows individuals to make the right decisions, based on evidence and facts, rather than on assumptions, experiences or ‘gut feelings’ (Marr, 2015, p. 229). Further, the high-powered analytical tools that exist, provides the opportunity of identifying, “[...] *what works without having to worry about why it worked*” (Schmarzo, 2013, p. 55). Thus, being able to compare and contrast between initiatives and alternatives, evidence-based practice in social services relates very well with the idea of move closer to the idea state of perfect rationality, seeing you on an objective base chooses the right decision (Simon, 1997). Finally, the hype around the potentials that big data may offer, such as increased public value, better social services and efficiency (Manyika et al., 2011), is evident in residential social monitoring.

4.8 Data Challenges and Privacy Issues

Despite the fact that Svendborg and Randers Municipality have achieved many useful gains by their use of data analytics with GIS, the work with large and integrated datasets are still a new field with a lot of challenges. As stated by Carlsson, “[...] and I’m quite sure that working with large datasets is a relatively new phenomenon in the Danish municipalities, but it definitely become much more widespread over the forthcoming years” (personal interview, March 21, 2017, appendix G, line 127). He further explains that due to it being a relatively new field, they have at Svendborg Municipality been afraid to present how the world actually looks, because they have felt that their data have not been strong enough. This is further explained by Næsvang, who puts it like this, “[...] we’re trying to keep it as simple as possible and we’re mostly only working with high-validated data to begin with in trying to do it well” (personal interview, March 21, 2017, appendix H, line 81). Thus, they are trying to keep it as simple as possible and ‘harvest the lowest hanging fruits’ for eventually to go deeper. Despite the fact that working with large and integrated datasets is still a new phenomenon that contains many challenges, great potentials and ‘fruits to harvest’ are available and should be pursued. Communicating this message to the wider public sector is what Andersen works with:

Internally, we work really hard to communicate the benefits that exist with data analytics, so you may say that our role is a little more in the communicative aspect we’re dealing with here, trying to create awareness on how important it is to register the data in the first place to ensure transparency and validated data. (personal interview, March 21, 2017, appendix I, line 128)

The data challenges are also recognized at Kommunernes Landsforening and KOMBIT. According to Hansen, the challenges in data imply that the potential of data analytics cannot be fully harvested, “[...] then there is also the question about data quality, where there without doubt exist some challenges to why the full potential is not resolved” (personal interview, March 9, 2017, appendix F, line 50). One of the social areas that are particular challenging regarding the data quality is within vulnerable children and young people. According to Hansen, this is due to the fact that a lot of different social workers are involved in the individual cases and they register their work in dissimilar ways, which reduces the data quality and eventually results in that the data cannot be trusted in data analytics (personal interview, March 9, 2017, appendix F, line 290). This is further agreed by Andersen, who argues that how information is registered has been a general challenge throughout the

whole municipality and something they currently are working on to improve, *“It has created some challenges in combining data and this have primary regarded how we as municipality register information, but something we’re dealing with at the moment”* (personal interview, March 21, 2017, appendix I, line 47). The issues with registering information are also present at Randers Municipality. Rasmussen mentions that particularly school absence is not very good and it clearly depends on how good the records are, *“School absence is not very law regulated and because of that, no real process or control seems to exist”* (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 216). He further points to the fact that there exists opposing attitudes towards registering school absence, seeing that some people find it not being an optimal utilization of teacher's time related to their core work:

[...] yes, it is at least in some areas that it is hard and sometimes also contradictory trends. For example, you probably wish to make the most of teachers' effective time and thus not spend so much time on registering absence, or it may also be due to the fact that the teachers' time is already cut and optimized so additional time spent on registration is not appropriate. (personal interview, March 29, 2017, appendix J, line 232)

Despite the fact that management in local municipalities is beginning to become more data-minded and of interest promote the idea to register all findings and not just what duty says, Rasmussen points out that you will have to figure out a way to do it effectively; else it will never be very good. The problematic area of registration does not only take place in school absence, but also within the health area. According to Rasmussen, it is still really difficult to get reliable health data, because the registration part is not well developed. An example is introduced with certain statements of caries in childrens' teeth, which also is useful in residential social monitoring, but the problem is that it is not particularly systematized, so you cannot just integrate it with other datasets (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 280). Accordingly, it is further stated that one might think that something is in fact registered in one way, but later to find out it is not. Thus, as Rasmussen explained, *“[...] and it can be a job in itself to find out how it really is registered and evaluated, that being the case if you want to use it for analytics purposes, you have to absolutely sure that it is registered in the right way”* (personal interview, March 29, 2017, appendix J, line 290). However, despite the fact that the practice of registering information in a homogeneous way is a difficult task

that affects the rest of the data value chain, it is certainly under improvement. Yet, in relation to the case of residential social monitoring, the registration part has been a barrier to succeed with the project, “[...] *it has been a far greater barrier to try and get the data we actually wanted to use in this project, that being to find them and understand how they have been registered in the first place*” (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 347). Thus, it is said that a large part of the resources that have been involved in the project have been to find validated data as input to GIS.

However, the registration part is not the only data challenge. There seems to be a general belief that while the data related area of school is quite strong, vulnerable children and young people are very weak. According to Næsvang, vulnerable children and young people is characterized as a dynamic and soft social area with a lot of complicated systems, which makes data analytics difficult, “*Well, the data quality is highly reflected in the areas that are very dynamic, thus being the big soft areas with complicated systems and data. This is especially within vulnerable children and people and adult disability*” (personal interview, March 21, 2017, appendix H, line 187). He further points out that these social areas consist of several systems that often are changed, making it a challenge to trust the data. Due to the challenges and limitations of using the municipalities’ own IT-systems, they increasingly tend to use FLIS, because the data is much more validated and structured, thus easier to work with in analytic perspectives, “*Firstly, we do not have very good data on the school area. Secondly, it is difficult for us to extract the data from our systems, but we have very good opportunities in FLIS, both in terms of data quality as well as to extract it into our own analytic tools*” (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 82). That FLIS provides great opportunities is further recognized and explained by Næsvang, who mentions that, “[...] *it’s the areas of citizens, staff, school, disabled adults, vulnerable children and young people, and the elderly that are like the FLIS universe we’re working with. They are classified as being the most simple and yet the most validated*” (personal interview, March 21, 2017, appendix H, line 179). Despite the fact that FLIS now offers highly validated and structured data is not without a lot hard work being put into it. As stated by Andersen, trusting the data has earlier been a general issue:

[...] the biggest problem was that you could not trust the data, why we along with Kommunernes Landsforening have carried out a great effort to improve the data quality

and thus corrected the gaps and mistakes that have existed, so our data now matches what is showed by DST. (personal interview, March 6, 2017, appendix E, line 271)

Thus, the effort to get the data quality on track has succeeded and KOMBIT now works on a new strategy in connection with FLIS's EU Public Procurement, where the new solution will be updated on a daily basis, "[...] the problem with FLIS is that the data warehouse is only updated once a month and that is not enough for management, who makes decisions every week. Thus, selected areas will in the future be updated on a daily basis" (Andersen, H., personal interview, March 6, 2017, appendix E, line 158). Furthermore, with FLIS's EU Public Procurement, the data areas of health and employment are to be included, which evokes enthusiasm at particularly Randers Municipality. Despite the new data areas evoke eagerness; Rasmussen estimates that it may take up to five years before there is validated data on the employment area, "It takes to time to get the data in and then having them validated afterwards, which is quite an extensive process." (personal interview, March 29, 2017, appendix J, line 90).

When asked about the Danish personal data legislation in relation to carrying out data projects in the municipalities, a general belief exists that it can be seen as a barrier when data is, or potentially can, be broken down to single individuals. According to Rasmussen, it is a really difficult area to work within for non-lawyers with limited knowledge on what you can or cannot do with data. Thus, in the beginning of the project of residential social monitoring, a lot of resources were spent on gaining knowledge on what data you could gather, how you would do it and for what purposes,

[...] in connection with this whole startup of the project, we have spent quite a lot of time to identify what data you wanted to collect, on what level and for what purpose. There have been some descriptions and reports that we have filled out, so we were allowed to use the data. (Rasmussen, T., personal interview, March 29, 2017, appendix J, line 333)

Rasmussen further points out that privacy lately has become a quite complex area in spite of the fact that they now have all this data available, why they internally at Randers Municipality have begun to create more permanent procedures for when using it in analytic purposes. Also at Svendborg Municipality, the personal data legislation has caused challenges. According to Næsvang, they have

become increasingly aware when performing data analytics, “[...] *I always reflect upon it when I come across some data that potentially can be degraded to individuals. In these cases, we have become very attentive to look at anonymizations or data agreements*” (personal interview, March 21, 2017, appendix H, line 166). That the personal data legislation is seen as a barrier is further explained by Carlsson. According to Carlsson, the data legislation prevents them to monitor small residential areas due to too few citizens, thus the need to group smaller areas together is necessarily. Carlsson concludes that it somewhat limits the purposes it was meant to be (personal interview, March 21, 2017, appendix G, line 164).

The privacy challenges are also recognized by Hansen and Andersen. According to Hansen, it is an area of strong focus at the data analytics division at Kommunernes Landsforening, where they tend to buy help from external lawyers, “*We sometimes buy external layers to help evaluate what we are dealing with. If we are not following the rules, a bad case can blow up in the media and may ruin everything*” (personal interview, March 9, 2017, appendix F, line 407). In relation to Carlsson’s statement regarding the need to group smaller areas together in residential social monitoring at Svendborg Municipality, Hansen recognizes the issue and argues that you are not allowed to show coupled figures if it is formed of less than five individuals. Also at KOMBIT, Andersen considers privacy as a ‘dangerous’ area, both for KOMBIT, yet also the municipalities, “*That is where it becomes slightly dangerous due to the personal data legislation. You can break down the data to single individuals, yet you are not allowed to. Further, the municipalities may not actually look across administrations*” (personal interview, March 6, 2017, appendix E, line 212). On the subject of FLIS, due to the new personal data legislation, technical limitations for coupled figures have been added:

Right now we have to set some limits for how to couple figures in relation to benchmarking, in spite of the fact that you are not allowed to identify a vulnerable child in school and you can if it is a small area such as Læsø, where there only may be five vulnerable children in total. Thus, when the population is below a particular level, you simply cannot get data on it. (Hansen. S., personal interview, March 9, 2017, appendix F, line 252)

That big data contains both technical, legal and management challenges relate great to what is generally considered in big data literature. Despite big data is coupled with knowledge discovery,

difficulties in analyzing 'large' datasets exist and presents unique systems engineering and architectural challenges (UN Global Pulse, 2012). This is further explained by Jamiy et al. (2014), who mention that the development of big data requires an extensive understanding of the data foundation and poor data governance usually is a significant factor in limiting one's effort to pursue big data (Desouza & Jacob, 2014). Precisely to what the findings in this study show, understanding how the data is registered and stored is necessary to take advantage from the knowledge hidden behind these sources of information. Additionally, the data challenges and privacy issues relate very well to what is written and recognized in the Danish digitization strategy. The strategy, which was introduced in the beginning of thesis, points at four specific areas of where data analytics lag behind. In the table below, the four challs and issues are commented upon with the experiences achieved from this case study.

#	Danish digitization strategy	Experiences from this study
1	Data is not standardized, making it difficult to use in analytical settings. Data standardization is an essential condition to fulfill the potential of digitization so data can be shared across relevant authorities or combined with other data to produce new knowledge.	Data standardization is an immense problem. Data is generally registered and stored differently, making it rather hard to use in analytical settings. Further, the frequent shift in systems makes it challenging to share data internally.
2	Some IT-suppliers takes well paid for the municipalities to access their own data.	FLIS collects data from most systems through interfaces and organizes it in a structured manner. Thus, the use of FLIS, which is owned by the public themselves, is a cheaper and more flexible solution than going through the respective IT-suppliers.
3	It is important to ensure that the regulatory environment is lasting, so the possibilities of sharing data across service areas is not unnecessarily constrained by outdated laws that troubles to keep up with the technological developments.	Despite the fact that shared opinions exist within the legal perspective of data analytics, it generally seems to limit the possibility to utilize data and create new knowledge.
4	Not only is sharing data a technical, financial and legal challenge, but the ability to exploit new technology is inextricably linked with the ability to create change in the organization and culture. Only when that happens is it possible to create real added value for citizens, businesses and municipalities.	Despite the fact that a 'data-mindset' is developed by the participants in this study, an undertone of absent readiness from decision makers higher in the hierarchy seems to be reluctant in terms of relying and trusting data analytics.

Table 10 – Data challenges and privacy issues

5.0 Discussion

5.1 Introduction

In the following chapter, we present and discuss the implications for research and contributions on the conceptualization of big data, as a source for new organizational knowledge creation and support in decision making. Further, we present and discuss reflections on the research process on conducting qualitative research as well as evaluating design-science research. The implications for practice are incorporated and presented during evaluating design-science research.

5.2 Implications for Research and Contributions

Despite the fact that the findings show that big data is a rather unclear concept, it is commonly referred to being able to provide new organizational knowledge and a broader understanding, thus being able to contribute with an increase in quality of decision making and better argumentation. In the following three subsections, we discuss the findings of this study on the concept of big data, the effect of it on organizational knowledge creation and support in decision making. The findings are discussed in relation to how they respectively support and extend the literature presented in chapter 2.

5.2.1 Findings that Support and Extend the Conceptualization of Big Data

The findings in this study show that the concept of big data is rather unclear. Big data is generally considered as a popular concept, yet very difficult to define in own words. Further, it seems hard to grasp for how long one has been familiar with the concept, although the familiarity with it is said to be around a couple of years in average. In the empirical data, it is noticeably that big data has two perceived understandings. At respectively Kommunernes Landsforening, KOMBIT and Randers Municipality, big data is commonly acknowledged as 'large' datasets that among other can be used for internal and external benchmarking purposes. However, at Svendborg Municipality, big data is not recognized as the amount of data, but instead expressed through examples of surveillance perspectives. Considering that the concept of big data is fairly incomplete, it is surprising that it is generally agreed upon that it has potential to be used much more in the public sector.

Therefore, the findings of this study both support and extend what is found on how big data is seen and understood in literature. It supports the literature that big data has no overall definition that is agreed upon and thus further investigation and clarification of the concept is necessary to manage it

properly. Yet, the findings also somewhat exceed the literature by showing that big data is vastly impenetrable and thus only can be expressed through concrete examples, such as surveillance.

5.2.2 Findings that Support and Extend Big Data Literature and Knowledge Creation

The findings in this study show that a strong relation exists between big data literature and organizational knowledge creation. Big data is generally referred to a source of information creating new knowledge. In the empirical data, the use of big data is commonly said to provide a greater picture of an issue or situation, hence the creation of new organizational knowledge. However, providing a greater picture of an issue or situation is not just referred to the visual representation of big data, which in this case study uses geographic information systems as analytical tools, but in terms of that the combination of different data sources provides new insights that were not possible to discover using just a single source. This relates very well to the essential feature of organizational knowledge creation that is the relevancy of creating new knowledge, compared to processing already existing information (Nonaka, 1994).

Despite the fact that big data is generally referred to as explicit, or codified knowledge in literature, seeing it is information that can be shared in formal and systematic language in the form of data and thus can be processed, transmitted and stored, tacit knowledge has in this study proved to be very important in big data management. Tacit knowledge, which is associated with action and involvement (Nonaka, 1994), usually clarified as personal 'know-how' (Choo, 1996), is estimated to be of great need in order to fully understand how big data creates new organizational knowledge. As explained in the conceptualization of big data, it is not just seen as 'large' datasets, but also a process of different analytical steps that is carried out by individuals. The combination of both tacit and explicit knowledge is especially evident within the data challenges presented in the empirical findings. Here, the process of finding and understanding how data is registered by different groups of primarily caseworkers is necessarily to be familiar with in order carry out the analytical steps and thus to create new knowledge.

Therefore, the findings of this study both support and extend what is found in the big data literature and its effects on organizational knowledge creation. It supports the literature that big data is able to generate new explicit or codified knowledge, through the combination of analyzing different data sources in analytical tools, such as geographic information systems. Yet, the findings also exceed the literature by showing that tacit knowledge is highly important to fully understand the knowledge

creation process. Tacit knowledge plays a significant role in big data management, seeing it is a premise to understand how data is registered in the first place to perform data analytics.

5.2.3 Findings that Support and Extend Big Data Literature and Decision Making

The findings in this study show that a strong relation exists between big data literature and its effect on decision making. Big data is generally acknowledged to provide new organizational knowledge that ultimately contributes with an increase in quality of decision making and better argumentation. In the empirical material, big data is commonly said to strengthen the ability and capacity to make better decisions, seeing it now can be made on an informed basis, thus decreasing uncertainty. Further, big data and the extensive amount of data that now can be processed in information systems is seen to have the effect that decisions are and expected to be made on a higher degree of 'facts' rather than 'gut feeling'. From the thoughts of organizational knowledge creation, 'facts' and 'gut feeling' are referred to respectively explicit and tacit knowledge. The empirical findings all point towards the fact that you with the use of big data is able to create clear goals and objectives, prioritize between alternatives and choose the best solution, thus becoming closer to the ideal state of perfect rationality.

However, what is noticeably is that big data is not only used for informational purposes, but also for persuading managers. This is especially evident within negotiations on budgeting allocation. Here, big data is seen useful to make one's case why funding should be allocated to you by drawing attention to certain issues in the society that requires action. Thus, big data is in this context positioned as a tool that is able to create shared understanding of choice situations, which then can be thoroughly discussed. Creating a shared understanding is widely found in the empirical data where the use of geographic information systems as analytical tool, i.e. geospatial big data analytics, is seen particular useful with its visual methods, in spite of the fact that the managers in question are recognized as being more visually brought.

Therefore, the findings of this study both support and extend what is found in the big data literature and its effects on decision making. It supports the literature that big data is able to take one closer towards the ideal state of perfect rationality with the extensive amount of information and knowledge available in decision making processes. Yet, the findings also exceed the literature by showing that big data is not only used for informational purposes and thus to relieve uncertainty, but as well to reduce equivocality by creating shared understandings of choice situations. Thus, the findings show that big data has two functions, both of which are told to be of great value and commonly said have the potential to be used much more in the future.

5.3 Reflections on the Research Process

In the following two subsections, we reflect on the common issue of ensuring trustworthiness in qualitative research as well as evaluating design-science research that is this study's paradigmatic underpinnings. Moreover, the implications for practice that are found in this study are incorporated and presented in line with the guidelines for evaluating design-science research, seeing that they refer and apply to the future and practical applications of big data in the public sector.

5.3.1 Conducting Qualitative Research

In this section, we reflect on the study in relation to the principles of conducting qualitative research. In literature, a wide range of different evaluation criterias for qualitative research exists (Lincoln & Cuba, 1994; Klein and Myers, 1999; Creswell, 2007). While these sets of criteria differ in some degree, it is generally acknowledged that qualitative research, due to its study of, *"People's experiential life as it is lived, felt, undergone, made sense of and accomplished by human beings"* (Schwandt, 2001), calls for a judgment that exceeds the traditional application of validity, generalizability and reliability, that being the criterias for carrying out positivist research.

Despite the fact that many critics are reluctant to accept the trustworthiness of qualitative research, frameworks, from example of Klein & Myers (1999) and Lincoln & Cuba (1994), for ensuring rigor in this field of work have been in existence for many years (Shenton, 2004). Where Klein & Myers's (1999) set of principles for conducting fields studies in information systems is made specifically within the interpretative paradigm perspective, Lincoln & Cuba's (1994) four evaluation criterias refer to general qualitative research projects (Shenton, 2004). The four evaluation criterias by Lincoln & Cuba (1994), that is credibility, transferability, dependability and confirmability, have won considerable favor by various scholars (Shenton, 2004). Thus, these form the focus on the reflections of the research process for ensuring trustworthiness for undertaking this inquiry.

The four criterias to ensure trustworthiness in qualitative research by Lincoln & Cuba (1994) correspond to the criteria employed by the positivist investigator (Shenton, 2004).

- | | |
|---------------------|---|
| (1) Credibility | (in preference to internal validity) |
| (2) Transferability | (in preference to external validity/generalizability) |
| (3) Dependability | (in preference to reliability) |
| (4) Confirmability | (in preference to objectivity) |

Within each of the criteria, Shenton (2004) establishes several strategies that can improve and ensure the trustworthiness of the qualitative research. Firstly, within the first criteria of credibility, that is the confidence in the 'truth' of the findings (Lincoln & Cuba, 1994), Shenton (2004) indorses fourteen strategies that may be applied by researchers to promote confidence. The strategies are (a) the adoption of research methods well established both in qualitative investigation in general and in information science particular, (b) the development of an early familiarity with the culture of participating organizations before the first data collection dialogues take place, (c) random sampling of individuals to serve as participants, (d) triangulation, (e) tactics to help ensure honesty in participants when contributing data, (f) iterative questioning, (g) negative case analysis, (h) frequent debriefing sessions, (i) peer scrutiny of the research project, (j) the researcher's 'reflective commentary', (k) background, qualifications and experience of the investigator, (l) member checks, (m) thick description of the phenomenon under scrutiny, and (n) examination of previous research findings to assess the degree to which the project's results are congruent with those of past studies.

Secondly, within the criteria of transferability, that is to show that the findings have applicability in other contexts (Lincoln & Cuba, 1994), sufficient thick descriptions of the phenomenon under investigated are mentioned, to allow readers to have a proper understanding of it, thus being able to compare the described instances in the report to those they have seen in other situations. Further, the use of multiple case studies, which is carrying out a similar project with the same methods in a different environment, is estimated to be of great value to promote transferability (Shenton, 2004).

Thirdly, within the criteria of dependability, that is to show that the findings are consistent and could be repeated (Lincoln & Cuba, 1994), a thorough description of the processes carried out within the study should be reported in detail, enabling future researchers to repeat the work, if not necessarily to gain the same results (Shenton, 2004).

Lastly, within the criteria of confirmability, that is the degree of neutrality or the extent to which the findings of a study are shaped by the participants and not researcher bias, motivation, or interest (Lincoln & Guba, 1994), 'audit trails' are giving to comprehend the issues of subjectivity, seeing it allows the reader and observer to trace the decisions made and procedures followed throughout the study (Shenton, 2004).

Criteria	Reflections and techniques for ensuring trustworthiness in this study
Credibility	We conducted interviews with key participants at KOMBIT and Kommunernes Landsforening (KL) to achieve and gather an early familiarity and information about the geospatial big data analytics project of Residential Social Monitoring before carrying out the central interviews with Svendborg- and Randers Municipality. The conducted interviews with participants at Svendborg- and Randers Municipality make up somewhat random sampling, based on the ones who had the time and were willing to participate. Besides interviews, we carried out observations at Svendborg Municipality to help us gain a better understanding of the context and the phenomenon of big data under investigation. To help ensure honesty in participants when contributing data, we offered full anonymity at any time.
Transferability	We have tried to provide thick descriptions of the complex phenomenon of big data to enable transparency and the possibility to transfer it to other times, settings, situations and people.
Dependability	We have tried to provide thorough descriptions of the various processes and decisions carried out in this study to enable future researchers to repeat the work.
Confirmability	To help ensure confirmability, all interviews have been transcribed and the presented knowledge can be traced back to the origin. Thus, we have taken cover in the written documentation.

Table 11 – Reflections on the qualitative evaluation criterias by Lincoln & Cuba (1994)

5.3.2 Evaluating Design-Science Research

Within design-science, as this study's paradigmatic underpinnings, the worldview is concentrated on whether the study's results have potential to drive concrete actions or behavioral changes (Goldkuhl, 2012; Kvale & Brinkmann, 2009; Hevner et al., 2004). Hevner et al. (2004) describe the practical outcome of design-science in information systems research as,

[...] information systems are implemented within an organization for the purpose of improving the effectiveness and efficiency of that organization. Capabilities of the information system and characteristics of the organization, its work systems, its people, and its development and implementation methodologies together determine the extent to which that purpose is achieved. (Hevner et al., 2004)

Thus, with design-science being inherently a problem solving process, Hevner et al. (2004) present a framework for understanding information system research with a set of guidelines for conducting and evaluating good design-science research. The framework consists of seven guidelines all of which are focused on the idea of creating and evaluating the IT artifacts intended to solve the identified organizational problems.

Guideline	Description
Guideline 1: Design as an Artifact	Design-science research must produce a viable artifact in the form of a construct, a model, a method, or an instantiation.
Guideline 2: Problem Relevance	The objective of design-science research is to develop technology-based solutions to important and relevant business problems.
Guideline 3: Design Evaluation	The utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods.
Guideline 4: Research Contributions	Effective design-science research must provide clear and verifiable contributions in the areas of the design artifact, design foundations, and/or design methodologies.
Guideline 5: Research Rigor	Design-science research relies upon the application of rigorous methods in both the construction and evaluation of the design artifact.
Guideline 6: Design as a Search Process	The search for an effective artifact requires utilizing available means to reach desired ends while satisfying laws in the problem environment.
Guideline 7: Communication of Research	Design-science research must be presented effectively both to technology-oriented as well as management-oriented audiences.

Table 12 – Design-science research guidelines (Hevner et al. 2004)

While it is generally acknowledged that design-science evaluation methods is difficult to apply in practice (Hevner et al., 2004), we have for this study selected to draw inspiration from the two guidelines of ‘Problem Relevance’ and ‘Research Contributions’, which we find useful and doable in this study. Firstly, within the guideline of ‘Problem Relevance’, the purpose of this study, and in line with design science, was to examine the actual outcome of the practical use of big data in the public sector, based on the theoretical framework of organizational knowledge creation and decision making. Big data is in this thesis considered as a source with great potential to create new organizational knowledge and support decision making, thus providing a higher degree of rationality in decision making processes, which together make up great relevance to research on big data and also serves as inspiration to the rest of the Danish public sector, whom wishes to work more

effectively with their data. For the case organization's own concern, i.e. implications for practice, it is estimated that the greatest relevance lie in the future applications, identified as the thesis's artifact, which can provide concrete proposals for the direction of the public sector's use of big data in the future and being able to drive change and improvements in effectiveness and efficiency. In regards to our findings presented in section 5.2.1, a general setting of the concept of big data in the Danish public sector is appropriate and advised. Big data is still in an early stage, why a common understanding of the concept is needed to manage it properly in the future. Secondly, we draw attention to the role of tacit knowledge when carrying out big data projects. Due to the many data challenges that we experienced throughout the study, a shift towards the consideration and prioritization of tacit knowledge to understand and handle how data is registered and stored are highly important for individuals in order to perform data analytics. Thirdly, big data is only not limited to be used for informational purposes and thus to relieve uncertainty, but also to reduce equivocality by creating shared understanding of choice situations and thus being able to facilitate collaboration, dialogue and reflection. It is advised that this point of view is conceived in big data projects.

Within the guideline of 'Research Contributions', this study focuses on big data, a relatively new phenomenon in the public sector. Accordingly, a number of learning points in relation to the organizational utility of big data can be presented. The study has proved that the use of big data allows public sector entities to create new organizational knowledge and support decision making and thus makes a profound example of the actual outcome of the practical use, which currently makes up a gap in literature. Furthermore, in spite of the fact that the findings at Svendborg Municipality also are present at Randers Municipality, improves the generalization and thus support that the results are transferable to other public sector entities. Lastly, seeing that public sector entities have different datasets than private sector entities, it is worth mentioning that the research contributions of this study may not be immediate transferable to other than the sector they have been explored in.

6.0 Concluding Remarks

The final chapter presents the concluding remarks of this study. It elaborates on how the empirical findings accommodate the research aims and answers the research questions that steered this study. The chapter ends by suggesting future research opportunities that we have become familiar with through the project.

6.1 Conclusion

The primary aim of this case study was to investigate the actual effects of big data applications to enable a better understanding of the complex concept and the possibilities and limitations it encompasses. Big data has received a great deal of attention in recent years, yet little is known about the actual outcome of the practical use of it in the public sector. In this study, big data is considered a source for new organizational knowledge creation and support in decision making. The phenomenon is investigated through the geospatial big data project of Residential Social Monitoring in Danish local government; Svendborg Municipality, as one level of the public sector. The technical structure consists of an analytical tool and several data sources. As analytical tool, advanced geographic information systems (GIS) are used for data management and visualization. Fælleskommunalt Ledelsesinformationssystem (FLIS) and Danish National Statistics (DST) are external data warehouses used as input in GIS. Data fields of vulnerable children and young people, school, citizen composition and socio-economic data are retrieved and used as they together tell something about the social heritage, which usually is of negative character in exposed residential areas.

Looking at the concept of big data, the empirical analysis indicates that a clear perception of the concept seems yet to be absent. Big data is popular, yet rather unclear. A variety of meanings of big data exists and reflects a wide range of what the concept encompasses. Big data is generally referred to as 'large' datasets, but also characterized as being surveillance. Despite the fact that big data is fairly incomplete, it is generally agreed upon that it holds great potentials to be used much more in the Danish public sector.

Big data is generally said to be able to create new explicit knowledge as the analytical process of combining 'large' datasets provides a greater picture of an issue or situation. At Svendborg

Municipality, the use of big data has provided correct answers to most previously unknown questions. Big data has contributed with knowledge on how bad the situation currently is within the municipality's residential social areas. The visibility of such conditions could not have been achieved without extensive data analytics. Besides having provided knowledge on what is currently happening, big data has raised adequate awareness on developmental trends in society that is highly useful for predicting future outcomes with a high level of reliability, based on detected, recognized and accepted patterns in data.

Big data and the newly created knowledge is commonly said to strengthen the ability and capacity to make better decisions at Svendborg Municipality. Big data is seen to have the effect that decisions are and expected to be made on a higher degree of 'facts' rather than 'gut feelings'. Accordingly, big data permeates every decision seeing that the municipality does not do anything before having worked with data analytics. In relation to strategy, the use of big data has made it possible to create clearer measureable goals and objectives that eventually can be followed up on. Big data has affected the municipality's master's plan for 2018-2021 to be considerably expanded from two exposed residential areas to involve eight areas. Besides having leaved one's stamp on strategy, the use of big data helps optimizing internal resources and target social services. This takes place within the ability to provide better services, such as early social interventions, and continues improvement based on personalization of services and learnings from the performances of such services.

Despite the great opportunities that follows big data, a number of organizational challenges and IT-related requirements exist. Firstly, managing the knowledge creation process is carried out by employees with analytical competences and thus is an important factor to address for problem solving and reveal data-driven insights. Secondly, the data foundation has to be standardized and analytical systems allow integrating and analyzing 'large' datasets.

Finally, we can conclude that big data shows great practical utility for Svendborg Municipality and the rest of the Danish public sector. The proposed theoretical framework in this thesis brings forward big data as a source for new organizational knowledge creation and support in decision making. At Svendborg Municipality, the use of geographic information systems to analyze 'large' datasets creates new knowledge and breaks the limitations that are considered in the theory of bounded rationality. With these opportunities, big data paves the way towards the ideal state of perfect rationality.

6.2 Suggestions for Future Research

The research presented in this thesis addressed and answered the research question posed in the introduction. However, it is evident that the subject has proven to be far more extensive than the scope of the project. The findings of this study point to a number of interesting possibilities that could be pursued in future research projects. Some of these aspects are described in this section.

Firstly, in spite of the fact that big data is a relatively new phenomenon in the public sector and thus very little is known about the actual outcome of the practical use of it, the findings in this study are highly interesting and demonstrate a new paradigm and era of how public sector entities can work effectively with data. There is no doubt that the potential of working creatively with datasets is great and it is expected that the field will evolve rapidly over the forthcoming years. Thus, further practical research in this field is estimated to be of great value and may work as inspiration to how municipalities can invest in the technical and human resources needed to harvest the low hanging fruits that seem to be present, with respect to the legal environment.

Secondly, we observed that data challenges permeated throughout the study. Initially, it was only planned to lightly touch upon the area, yet it came to represent a significant part of the primary data collection. We were somewhat aware of the issues within the subject from among other the Danish digitization strategy, but it showed to be far more extensive than originally imagined. In terms of this, we believe that further investigation into data challenges are appropriate and rather needed, seeing it is necessary to cope with these to generate true value from big data.

Thirdly, we observed that big data is not just a technical or legal challenge, but the ability to exploit new technology is inextricably linked with the capability to create change in the organization and culture. Despite the fact that a 'data-mindset' seemed to be developed by the participants in this study, an undertone of absent readiness from decision makers higher in the hierarchy seems to be reluctant in terms of relying and trusting data analytics. Thus, we believe that investigation from a change management perspective is important to fully utilize big data in decision making processes. Such a study would make it possible to observe 'what is missing' to get the decision makers on board and participating in the change to make a difference.

Overall, we would like to highlight the fact that the conclusion of this study was intended for two purposes. Firstly, with pragmatism as the paradigmatic underpinnings, a discovery of the actual outcome of the practical use of big data in the public sector, seeing that gaps in research exists within this field. Secondly, contributing to a deeper understanding of the 'fuzzy' and complex concept of big data. Thus, the study has further added to the appreciation of empirical variability. Although the findings are primarily based on a single organization, the study contributes by providing new practical insights on how big data creates new organizational knowledge and support decision making, which are an important starting point for future empirical studies on the subject.

7.0 References

“About FLIS” (n.d.). Retrieved from <http://www.kombit.dk/indhold/om-flis>

Ackoff, R. (1989). From Data to Wisdom. *Journal of Applied Systems Analysis.*, 16, pp. 3-9.

Amunts, K., Ebell, C., Muller, J., Telefont, M., Knoll, A., & Lippert, T. (2016). *The Human Brain Project: Creating a European Research Infrastructure to Decode the Human Brain*. Neuron. Elsevier Inc.

Andersen, I. (2010). *Den Skinbarlige Virkelighed (Studieteknik) – Vidensproduktion I samfundsvidenskaberne* (4th ed.). Samfundslitteratur.

Andrade, P. L., Hemerly, J., Recalde, G., & Ryan, P. (2014). From Big Data to Big Social and Economic Opportunities: Which Policies Will Lead to Leveraging Data-Driven Innovation's Potential? *The Global Information Technology Report 2014: Rewards and Risks of Big Data, Chapter 1.8*. Cornell University and the World Economic Forum. Retrieved from https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2429942

Andrews, R. & Entwistle, T. (2010). Does Cross-Sectoral Partnership Deliver? An Empirical Exploration of Public Service Effectiveness, Efficiency, and Equity. *Journal of Public Administration Research and Theory*, 20 (3), 679,701.

Angrosino, M. V. & DePerez, K. A. (2000). Rethinking observation: From method to context. *Handbook of Qualitative Research*, 2, 673-702. Thousand Oaks, CA: Sage.

Aronson, E. (1995). The Rationalizing Animal. In B. M. Staw, *Psychological dimensions of organizational behavior*. Englewood Cliffs: Prentice Hall.

Ashforth, B. E., & Mael, F. (1989, Jan). Social Identity Theory and the Organization. *The Academy of Management Review*, 14(1), pp. 20-39.

Bassi E., Leoni D., Leucci S., Pane J. & Vaccari, L. (2014). Opening Public Deliberations: Transparency, Privacy, Anonymisation. In: Casanovas P., Pagallo U., Palmirani M., Sartor G. (eds) *AI Approaches to the Complexity of Legal Systems*. Lecture Notes in Computer Science, vol 8929. Springer, Berlin, Heidelberg

Bays, J. (2016). *Harnessing big data to address the world's problems*. Retrieved from <http://voices.mckinseyonsociety.com/harnessing-big-data-to-address-the-worlds-problems/>

Blumer, H. (1969). *Symbolic Interactionism: Perspective and Method*. University of California Press, Berkeley, CA.

Boyd, D. & Crawford, K. (2012). Critical Questions For Big Data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15, 662-679.

Braa, K. and Vidgen, R. (1999). Interpretation, intervention, and reduction in the organizational laboratory:

a framework for in-context information system research. *Accounting, Management & Information Technology*, 9, 25–47.

Brinkmann, S., & Kvale, S. (2009). *Interview: Introduktion til et håndværk* (2nd ed.). Hans Reitzel.

Brunsson, N. (1989). *The organization of hypocrisy: talk, decisions and actions in organizations*. Chichester: John Wiley & Sons.

Capgemini (2012). *The Deciding Factor: Big Data & Decision Making*. Retrieved from <https://www.capgemini.com/resources/the-deciding-factor-big-data-decision-making>

Carter, P. (2011). Big data analytics: Future architectures, skills and roadmaps for the CIO. Retrieved from <http://triangleinformationmanagement.com/wp-content/uploads/2013/12/bigdata-idc-wp.pdf>

Chen, M., Mao, S. & Liu, Y. (2014). Big Data: A Survey. *Mobile Networks and Applications*, 19 (2), (171-209).

Cheng, J. H.-W. (2014). *Big Data for Development in China*. UNDP. UNDP.

Christensen, P. H. (2010). *Mere videndeling*. Copenhagen: Hans Reitzels Forlag.

Cohen, W. M., & Levinthal, D. A. (1990, Mar). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 35(1), pp. 128-152.

Corbin, K. (2012). *Federal Government's Big Data Efforts Lagging*. Retrieved 2017, from CIO.com: <http://www.cio.com/article/2390059/government/federal-government-s-big-data-efforts-lagging.html>

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches*. Thousand Oaks, CA. Sage.

Cronen, V. (2001). Practical theory, practical art, and the pragmaticsystemic account of inquiry. *Communication Theory*, 11 (1), 14–35.

Daft, L. R., Bettenhausen, K. R., & Tyler, B. B. (1993). Implications of Top Managers' Communication Choices for Strategic Decisions. In G. P. Huber, & W. H. Glick, *Organizational Change and Redesign*. New York: Oxford University Press.

Davenport, T. H., Barth, P., & Bean, R. (2012). How 'Big Data' is Different. *MIT SLOAN MANAGEMENT REVIEW*.

Deloitte Consulting (2014). *The impact of the open geographical data*. Retrieved from <http://eng.gst.dk/media/gst/2626131/GST%20The%20impact%20of%20the%20open%20geographical%20data.pdf>

Demchenko, Y., Ngo, C., de Laat, C., Membrey, P., & Gordijenko, D. (2014). Big Security for Big Data: Addressing Security Challenges for the Big Data Infrastructure. In W. Jonker & M. Petkovic (Eds.), *Secure*

Data Management (Vol. 8425, 76-94): Springer International Publishing.

Desouza, K. C., & Jacob, B. (2014). Big Data in the Public Sector: Lessons for Practitioners and Scholars. *Administration & Society*, 38, 309-334

DeWalt, K. M. & DeWalt, B. R. (2002). *Participant observation: a guide for fieldworkers*. Walnut Creek, CA: AltaMira Press.

Dewey, J. (1938). *Logic: The Theory of Inquiry*. Henry Holt, New York.

Eilon, S. (1969, December). What Is a Decision? *Management Science*, 16(4), pp. B172-B189.

Elgendy, N., & Elragal, A. (2014). Big Data Analytics: A Literature Review Paper. In P. Perner (Ed.), *Advances in Data Mining. Applications and Theoretical Aspects* (Vol. 8557, pp. 214-227): Springer International Publishing.

Elo, S., & Kyngäs, H. (2007). The qualitative content analysis process. *Journal of Advanced Nursing*. 62 (1), 107-115.

Espedal, B. (1997). Perspectives on Leadership Development. *NHH-rapport*.

European Commission. (2012). *The Human Brain Project*. European Commission. Lausanne: The HBP-PS Consortium.

EY (2016). *How can big data lead to more rational decision-making?* Retrieved from <https://advisory.ey.com/analytics/rational-decision-making>

EY (2016). *How is big data analytics transforming corporate decision-making?* Retrieved from <https://advisory.ey.com/analytics/big-data-analytics>

Feldman, M. S., & March, J. G. (1981). Information in Organizations as Signal and Symbol. *Administrative Science Quarterly*, 26(2), 171-186.

Flyverbom, M. & Madsen, A. K. (2016). Sorting Data Out - Unpacking Big Data Value Chains and Algorithmic Knowledge Production. *Die Gesellschaft der Daten: Über die digitale Transformation der sozialen Ordnung*, 123-144. Bielefeld, Transcript Verlag.

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, 165, 234-246.

Fredriksson, C. (2015). *Knowledge Management with Big Data: Creating New Possibilities for Organizations*. Retrieved from http://spa.gu.se/digitalAssets/1552/1552434_fredriksson--knowledge-management-with-big-data_paper_norkom2015.pdf

Fredriksson, C. (2016). *Organizational knowledge creation with Big data: A case study of the concept and*

practical use of Big data in a local government context. Retrieved from <https://www.abo.fi/fakultet/media/22103/fredriksson.pdf>

Gamage, P. (2016, June). New development: Leveraging 'big data' analytics in the public sector. *Public Money & Management*, 36(5), pp. 385-390.

Gamage, P. (2016). New development: Leveraging 'big data' analytics in the public sector. *Public Money & Management*, 36 (5), 385-390.

Goldkuhl, G. (2012). Pragmatism vs interpretivism in qualitative information systems research. *European Journal of Information Systems*, 21 (2), 135-146.

Grant, A., Razdan, R., & Shang, T. (2016). *Coordinates for Change: How GIS Technology and Geospatial Analytics Can Improve City Services*. Retrieved from http://www.mckinsey.com/~media/mckinsey/dotcom/client_service/public%20sector/gdnt/gdnt_coordinates_for_change_gis_final.ashx

Grant, R. M. (1996). TOWARD A KNOWLEDGE-BASED THEORY OF THE FIRM. *Strategic Management Journal*, 17, pp. 109-122.

Hayek, F. A. (1945, Sep). The use of knowledge in society. *The American Economic Review*, 35(4), pp. 519-530.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, 28, (1), 75-105

Howarth, B. (2014, June 13). *Big data: how predictive analytics is taking over the public sector*. Retrieved from <https://www.theguardian.com/technology/2014/jun/13/big-data-how-predictive-analytics-is-taking-over-the-public-sector>

Huber, G. P. (1990). A Theory of the Effects of Advanced Information Technologies on Organizational Design, Intelligence, and Decision Making. *The Academy of Management Review*, 15(1), 47-71.

Huber, G. P. (2004). *The Necessary Nature of Future Firms: Attributes of Survivors in a Changing World*. Sage Publications, Inc.

Huber, G. P., & Glick, W. H. (1993). *Organizational Change and Redesign: Ideas and Insights for improving Performance*. New York: Oxford University Press.

Hyldborg, P. (2016). *Et digitalt drømmesamfund. Illusion eller realitet? - Fem forudsigelser om vores digitale fremtid*. Retrieved from <http://www.denoffentlige.dk/et-digitalt-droemmesamfund-illusion-eller-realitet-fem-forudsigelser-om-vores-digitale-fremtid>

Irisgroup (2013). *Big Data som vækstfaktor i dansk erhvervsliv – potentialer, barrierer og erhvervspolitiske konsekvenser*. Retrieved from <https://erhvervsstyrelsen.dk/sites/default/files/big-data-som-vaekstfaktor.pdf>

Jacobsen, D. I., & Thorsvik, J. (2008). *Hvordan organisationer fungerer: Indføring i organisation og ledelse* (Vol. 2. udgave). København: Hans Reitzels Forlag.

Jamiy, F., Daif, A., Azouazi, M., & Marzak, A. (2014). The potential and challenges of Big data - Recommendation systems next level application. *International Journal of Computer Science Issues*, 11 (5), 6.

Janis, I. L. (1983, Fall). Groupthink: Psychological Studies of Policy Decisions and Fiascos. *Presidential Studies Quarterly*, 13(4), pp. 654-656.

Jensen, N. R., Langager, S., Petersen, K. E., & Erlandsen, T. (2015). *Udsatte børn og unge*. Hans Reitzels.

Jordan, S. R. (2014). Beneficence and the Expert Bureaucracy. *Public Integrity*, 16(4), 375- 394.

Kahneman, D. (1997). New Challenges to the Rationality Assumption. *Legal Theory*, 3(2), pp. 105-124.

Kahneman, D., & Tversky, A. (1984, Apr). Choices, values, and frames. *American Psychologist*, 39(4), pp. 341-350.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Kawulich, B. (2005). Participant Observation as a Data Collection Method. *Forum: Qualitative Social Research*, 6 (2), Art. 43.

KL (2015b). *Guide til Boligsocial Monitorering. Et tværgående kommunalt redskab til kortbaseret ledelsesinformation om udsatte by- og boligområder*. Retrieved from https://www.kl.dk/ImageVaultFiles/id_75343/cf_202/Guide_til_boligsocial_monitorering.PDF

KL (2015a). Offentlige data skal skabe effektivitet og merværdi. In *Lokal og Digital – Et Sammehængende Danmark; Fælleskommunal digitaliseringsstrategi 2016-2020* (15-17). Retrieved from http://www.kl.dk/ImageVaultFiles/id_72295/cf_202/F-lleskommunal_Digitaliseringstrategi_2016-2020.PDF

Klein, H. K., & Myers, M. D. (1999). A set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems. *MIS Quarterly*, 23 (1), 67-94.

Klievink, B., Romijn, B.-J., Cunningham, S., & Bruijn, H. d. (2016, August). Big data in the public sector: Uncertainties and readiness. *Inf Syst Front*, pp. 267-283.

Knapp, M. (2013). Big Data. *Journal of Electronic Resources in Medical Libraries*, 10(4), pp. 215-222.

Laney, D. (2001). 3D Data Management: Controlling Data Volume, Velocity, and Variety. *Application Delivery Strategies*.

Langley, A., Mintzberg, H., Pitcher, P., Posada, E., & Saint-Macary, J. (1995, May-Jun). Opening up Decision Making: The View from the Black Stool. *Organization Science*, 6(3), pp. 260-279.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big Data, Analytics and the Path From Insights to Value. *MIT SLOAN MANAGEMENT REVIEW*, 52(2), pp. 21-31.

Leach, J. (2012). Step Inside: Knowledge Freely Available. The Politics of (making) Knowledge objects. *The Politics of Knowledge*, 79-95. London: Routledge.

Lee, J., & Kang, M. (2015). Geospatial Big Data: Challenges and Opportunities. *Big Data Research*, 2 (2), 74-81.

Lincoln, Y. S., & Guba, E. G. (1994). Competing paradigms in qualitative research. *Handbook of qualitative research*, 105-117. Thousand Oaks, CA: Sage.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh & Byers, A. H. (2011). *Big Data: The next frontier for innovation, competition and productivity*. McKinsey Global Institute.

March, J. G. (1994). *A primer on decision making: How decisions happen*. New York: The Free Press.

March, J. G., & Olsen, J. P. (1976). *Ambiguity and choice in organizations*. Bergen: Universitetsforlaget.

Marieta, D., Opreana, A., & Christescu, M. P. (2010). Efficiency, Effectiveness and Performance of the Public Sector. *Romanian Journal of Economic Forecasting*, 4, 132-147.

Marr, B. (2015). *Big Data: Using SMART Big Data, Analytics and Metrics To Make Better Decisions and Improve Performance*. John Wiley & Sons Inc.

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work and think*. London: John Murray.

McAfee, A., & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review*, October.

McNeely, C. L., & Hahm, J.-o. (2014). The Big (Data) Bang: Policy, Prospects, and Challenges. *Review of Policy Research*, 31(4), pp. 304-310.

Michael, M., & Lupton, D. (2016). Toward a manifesto for the 'public understanding of big data'. *Public Understanding of Science*, 25(1), 104-116.

Miles, M. B., & Huberman, M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks, CA: Sage.

Munné, R. (2016). Big Data in the Public Sector. In *New Horizons for a Data-Driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe* (195-209). Springer Open.

- Newman, I., & McNeil, K. A. (1998). *Conducting Survey Research in the Social Sciences*. University Press of America.
- Nonaka, I. (1994, Feb). A Dynamic Theory of Organizational Knowledge Creation. *Organization Science*, 5(1), pp. 14-37.
- Nonaka, I., Toyama, R., & Konno, N. (2000). SECI, Ba and Leadership: a Unified Model of Dynamic Knowledge Creation. *Long Range Planning*, 33, pp. 5-34.
- Nonaka, I., Umemoto, K., & Senoo, D. (1996). From Information Processing to Knowledge Creation: A Paradigm Shift in Business Management. *Technology In Society*, 18(2), pp. 203-218.
- O'Brien, A. (2012). *Big Data is Big Business in Smart Government*. Retrieved from https://idc-community.com/government/smart_government/bigdataisbigbusinessinsmartgovernment
- Oates, B. J. (2006). *Researching Information Systems and Computing*. University of Teesside, Middlesbrough: SAGE Publications Ltd.
- OECD (2015). *Data-Driven Innovation: Big Data for Growth and Well-Being*. Paris: OECD Publishing.
- Ohlhorst, F. J. (2012). Wiley and SAS Business Series : Big Data Analytics : Turning Big Data into Big Money. Somerset, NJ, USA: John Wiley & Sons.
- Oracle (2012). *Big Data: A Big Deal for Public Sector Organizations*. Retrieved from <http://www.oracle.com/us/industries/public-sector/public-sector-big-data-br-1676649.pdf>
- Palys, T. (2008). Purposive sampling. *The Sage Encyclopedia of Qualitative Research Methods*, 2, 697-698. Los Angeles: Sage.
- Phillips-Wren, G., & Hoskisson, A. (2015). An analytical journey towards big data. *Journal of Decision Systems*, 87-102.
- Polanyi, M. (1958). *Personal Knowledge*. Chicago, IL: The University of Chicago Press.
- Polanyi, M. (1967). *The Tacit Dimension*. New York: Anchor Books.
- Pollitt, C. (2014). Future Trends in European Public Administration and Management: An Outside-in Perspective. Retrieved from <http://www.cocops.eu/wp-content/uploads/2014/04/FutureTrendsInEuropeanPublicAdministrationAndManagement.pdf>
- Power, D. J. (2014). Using "Big Data" for analytics and decision support. *Journal of Decision Systems*, 222-228.
- Provost, F. & Fawcett, T. (2013a) *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. O'Reilly Media.

Provost, F., & Fawcett, T. (2013b). Data science and its relationship to big data and data-driven decision-making. *Big Data*, 1(1), 51-59.

PWC (2014). *An introduction to Big Data – how a buzzword morphed into a lasting trend that will transform the way you do business*. Retrieved from <http://www.pwc.co.uk/industries/technology/insights/an-introduction-to-big-data-how-a-buzzword-morphed-into-a-lasting-trend-that-will-transform-the-way-you-do-business.html>

Rabionet, S. E. (2011). How I Learned to Design and Conduct Semi-structured Interviews: An Ongoing and Continuous Journey. *The Qualitative Report*, 16 (2), 563-566.

Raghupathi, W., & Raghupathi, V. (2014). Big Data Analytics - Architectures, Implementation Methodology, and Tools. In S. Kudyba (Ed.), *Big Data, Mining, and Analytics. Components of Strategic Decision Making* (pp. 49-70). Boca Raton: CRC Press Taylor & Francis Group.

Regalado, A. (2014). The Power to Decide. *MIT TECHNOLOGY REVIEW*, 117(2).

Rorty, R. (1980). Pragmatism, relativism and irrationalism. *Proceedings and Addresses of the American Philosophical Association*, 53 (6), 719-738.

Rowe, F. (2014). What Literature Review is Not: Diversity, Boundaries and Recommendations. *European Journal of Information Systems*, 23, 241-255.

Rowley, J. (2007, April). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, 33(2), pp. 163-180.

Schmarzo, B. (2013). *Big data: understanding how data powers big business*. Indianapolis, IN: John Wiley & Sons.

Schmidt, E. & Cohen, J. (2014). *The New Digital Age - Reshaping the Future of People, Nations and Business*. John Murray.

Schon, D. A. (1987). *Educating the reflective practitioner*. Jossey-Bass.

Schott, R. L. (1991, May 1). Administrative and Organization Behavior: Some Insights from Cognitive Psychology. *Administration & Society*, 23(1), pp. 54-73.

Schultze, U., & Avital, M. (2011). Designing interviews to generate rich data for information systems research. *Information and Organization*, 21 (1), 1-16.

Schwandt, T. A. (2001). *Dictionary of Qualitative Inquiry*. SAGE Publications.

Schwandt, T. A. (2001). *Qualitative inquiry: A dictionary of terms* (2nd ed.). Thousand Oaks: Sage.

Schwarz, A., Mehta, M., Johnson, N., & Chin, W. (2006). Understanding frameworks and reviews: a

commentary to assist us in moving our field forward by analyzing our past. *Database* 38 (3), 29–50.

Schönberger, V. M. & Cukier, K. (2014) *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Eamon Dolan/Mariner Books.

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

Shollo, A. (2013). *The Role of Business Intelligence in Organizational Decision-making* (Vol. 1). Copenhagen: Copenhagen Business School.

Silverstein, A. (1988). An Aristotelian resolution of the ideographic versus nomothetic tension. *American Psychologist*, 43 (6), 425-430.

Simon, H. A. (1957). A Behavioral Model of Rational Choice. In H. A. Simon, *Models of man: social and rational: mathematical essays on rational human behavior in a social setting*. New York: John Wiley & Sons.

Simon, H. A. (1960). *The new science of management decision*. New York: Harper & Brothers.

Simon, H. A. (1973, May-Jun). Applying Information Technology to Organization Design. *Public Administration Review*, 33(3), pp. 268-278.

Simon, H. A. (1997). *Administrative behavior: a study of decision-making processes in administrative organisations* (Vol. 4). New York: Free Press.

Sorofman, J. (2013, February 28). Data, Data Everywhere, Gartner Blog Network. Retrieved from <http://blogs.gartner.com/jake-sorofman/data-data-everywhere/>

Staw, B.M. (1976). Knee-Deep in Big Muddy: A Study of Escalating Commitment to a Chosen Course of Action. *Organizational Behaviours and Human Performance* 16: 27-44.

Strathern, M. (2005). Robust Knowledge and Fragile Future. *Global Assemblages: Technology, Politics and Ethics as Anthropological Problems*, 464-481. Malden/Oxford: Blackwell Publishing.

Strauss, A. L., & Corbin, J. M. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Thousand Oaks: Sage.

Szulanski, G. (1996). EXPLORING INTERNAL STICKINESS: IMPEDIMENTSTO THE TRANSFER OF BEST PRACTICE WITHINTHE FIRM. *Strategic Management Journal*, 17, pp. 27-43.

Szulanski, G. (2003). Sticky knowledge: barriers to knowing in the firm. *SAGE strategy series*.

Thapa, B. (2016). *Foreshadowing Data-driven Government: Citizen-Government Interaction in Visions of the Data-powered State*. Retrieved from <http://basantathapa.de/academia-3/talks-presentations/>

Thompson, N., Ravindran, R., & Nicosia, S. (2014, November 18). Government data does not mean data governance: Lessons learned from a public sector application audit. *Government Information Quarterly*(32), pp. 316-322.

Toffler, A. (1991). *Powershift: knowledge, wealth and violence at the edge of the 21st century*. New York: Bantam.

Troester, M. (2012). Big Data Meets Big Data Analytics - Three Key Technologies for Extracting Real-Time Business Value from the Big Data That Threatens to Overwhelm Traditional Computing Architectures. Retrieved from http://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/big-data-meets-big-data-analytics-105777.pdf

UN Global Pulse (2012). *Big Data for Development: Challenges and Opportunities*. Retrieved from <http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGloabalPulseMay2012.pdf>

Vaidhyanathan, S., & Bullock, C. (2014). Knowledge and Dignity in the Era of “Big Data”. *The Serials Librarian*, 66(1-4), 49-64.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *Management Information Systems Quarterly*, 26 (2), xiii-xxiii.

Weick, K. E. (1979). *The social psychology of organizing* (Vol. 2. ed.). New York: McGraw-Hill.

Whyte, W. F. (1979). On Making the Most of Participant Observation. *The American Sociologist*, 14 (1), 56-66. American Sociological Association.

Woodie, A. (2015). *Why Gartner Dropped Big Data Off the Hype Curve*. Retrieved from <https://www.datanami.com/2015/08/26/why-gartner-dropped-big-data-off-the-hype-curve/>

Yin, R. K. (2003b). *Case study research. Design and methods* (3rd ed.). Thousand Oaks, CA: Sage.

Yiu, C. (2012). *The Big Data Opportunity. Making government faster, smarter and more personal*. London: Policy Exchange.

8.0 Appendices

8.A Concept matrix

Concept matrix	Conceptualization of big data										Data format			Opportunities																												
	3v's (volume, variety and velocity)	5v's	Digital information	Algorithms	Tools, processes and procedures	Analytical tool	Extremely large data sets	Cultural, technological and scholarly phenomenon	Change in data management	Big data technologies	Structured	Unstructured	Semi-structured	Predictive analytics	Open data	Actions	Reduce risks	Data-driven innovation	Real-time information	Improve efficiency	Improve effectiveness	New insights	New knowledge discovery	Social value of data	Economic value of data / savings	Improve policymaking	Improve competitiveness	Value creation	Improve strategic	Marketing	Merchandising	Supply chain	New business models	Improve operational	Build successful products and services	Optimize business processes	Improve decision making (support)	Improve firm performance	Improve productivity	Efficient deployment of resources		
Andrade et al. (2014)					x										x																											
Bell (2017)															x																											
Begoli & Henry (2012)																																										
Boyd & Crawford (2012)																																										
Capcemin (2012)																																										
Cartier (2011)	x																																									
Davenport et al. (2012)	x																																									
Flyverthom & Madsen (2016)																																										
Flyverthom (2015)																																										
Fosso Warba et al. (2015)																																										
Frederiksson (2015)																																										
Frederiksson (2016)																																										
Gillispie (2014)																																										
Howard (2012)																																										
Immy et al. (2014)																																										
Kabir & Caryannis (2013)	x																																									
Kadel (2010)																																										
Kasler et al. (2013)	x																																									
Kamthala et al. (2014)																																										
Kim et al. (2014)	x																																									
Kichin (2014)																																										
Kump (2013)																																										
Labridis & Jaganishi (2012)																																										
Laney (2001)	x																																									
Lohr (2012)																																										
Manjras et al. (2011)																																										
McAfee & Brynjolfsson (2012)	x																																									

Concept matrix	Implications/challenges																Sectors									
	Need for policies for the use of big data	Disclosure of personal data	Privacy issues	Demand for highly skilled workers	Technical challenges	Security	Surveillance	Management challenges	Data culture and change	Leaks and data breaches	Laws	Corporate big data profiling	Data silos	Difficult to interpret / data analysis	Ramifications for trust	Ethical issues	Data quality challenges	Private sector	Hype: Public sector	Case studies: Public sector	Healthcare	Financial services	Transportation	Social media	Smart cities	Human welfare
Avardale et al. (2014)	x	x	x	x															x							
Bafl (2017)			x																							
Bespal & Honey (2012)				x	x																					
Boyd & Crawford (2012)																x										
CapGemini (2012)				x								x		x												
Carer (2011)				x	x								x													
Davenport et al. (2012)																										
Flyverboom & Madsen (2016)																										
Flyverboom (2015)			x				x								x		x									
Fosso Wamba et al. (2015)																										
Fredenlissou (2015)			x	x	x			x						x	x	x	x									
Fredenlissou (2016)																										
Gillespie (2014)																										
Howard (2012)																										
Jamny et al. (2014)				x	x									x			x								x	
Kabir & Carayannis (2013)								x																		
Kadell (2010)																										
Kasler et al. (2013)					x	x		x																		x
Kamshila et al. (2014)			x		x												x									
Kim et al. (2014)		x	x	x		x									x				x							
Kishor (2014)																										
Kisapi (2013)																										
Ladomiris & Jureidat (2012)														x												
Laney (2001)																										
Lohr (2012)														x												
Manvika et al. (2011)																										
McAfee & Brynjolfsson (2012)						x																				
Michael & Lupton (2016)																										
Minne (2016)			x	x	x	x																				
OECD (2014)					x																					
Orsac (2012)																										
Polio et al. (2015)																										
Power (2013)					x	x																				
Provost & Fawcett (2013)						x																				
Purcell (2012)					x	x																				
Reynolds (2014)																										
Rutter (2014)			x	x	x	x																				
Saleh & Jurewski (2016)																										
Schomberger & Chaker (2013)			x		x	x										x										
Tronster (2012)																										
Zacharakis et al. (2013)			x	x	x	x																				
Zuboff (2015)			x				x																			
Total	42	1	2	11	17	15	6	2	5	6	3	1	1	3	5	4	7	5	0	4	3	5	1	1	1	1

8.B Interview guide

Emne / Teoretisk ramme	Spørgsmål
Indledning	I forlængelse af den fælleskommunale digitaliseringsstrategi 2016-2020, oplever du et øget pres for brugen af data i det daglige?
Big Data	Hvordan forstår du Big Data? Hvor længe har du været bekendt med Big Data?
Big Data i den offentlige sektor	Fænomenet Big Data er mest udbredt i den private sektor. Den offentlige sektor besidder andre typer data end den private sektor - er der potentiale i at benytte big data i den offentlige sektor? Hvordan?
Boligsocial Monitorering	Vil du karakterisere projektet Boligsocial Monitorering som et Big Data projekt? Hvorfor er kortbaseret ledelsesinformation interessant i boligsocial monitorering?
Videnskabelse	Hvilken ny viden har det skabt at kombinere data fra FLIS, Danmarks Statistik, samt interne datakilder i jeres GIS-system? Har brugen af denne databaserede viden givet jer svar på spørgsmål som tidligere var ukendte? Giver databaseret viden jer svar på at udpege aktuelle udviklingstendenser?
Beslutningsgrundlag	Hvordan oplever du, at kortbaseret ledelsesinformation og databaseret viden bidrager til beslutningsprocesser hos jer? Hvilke fordele ser du at databaseret viden giver jer ift. beslutningsgrundlag? Lader databaseret viden jer træffe bedre beslutninger? Hvordan? Hjælper databaseret viden jer med at definere mere klare mål i kommunen? Hvordan?
Strategisk arbejde	Hvordan oplever du, at databaseret viden kan bidrage til at løfte det strategiske arbejde med kommunens udsatte boligområder? Har brugen af databaseret viden og indsigten derfra betydning for jeres valg af strategiske arbejde i dag? Hvilke fordele ser du, at databaseret viden kan løfte det strategiske arbejde med kommunens udsatte boligområder? Det Nationale Forskningscenter for Velfærd (tidl. Socialforskningsinstituttet) har opstillet en række evidensbaserede indikatorer - hvordan benyttes disse i forbindelse med boligsocial

	monitorering?
Målrettede indsatser	Hvordan oplever du, at databaseret viden kan styrke de målrettede indsatser i praksis? Hvilke fordele ser du, at databaseret viden styrke de målrettede indsatser i praksis? Hvordan oplever du, at databaseret viden kan benyttes i en tidlig indsats? Hvilke fordele ser du, at databaseret viden kan benyttes i en tidlig indsats?
Persondatalovgivning	Hvordan oplever du, at persondatalovgivningen spiller ind ift. arbejdet med boligsocial monitorering? Er persondatalovgivningen en barriere for at indfri de fordele og potentialer, som data analyser giver?
Ressourceallokering	Hvordan oplever du, at databaseret viden kan bidrage til jeres fordeling samt udnyttelse af ressourcer i boligsocial monitorering? Hvilke fordele ser du, at databaseret viden kan bidrage jeres fordeling samt udnyttelse af ressourcer i boligsocial monitorering?
Dataudfordringer	Har du oplevet datamæssige udfordringer (eks. kvalitet) i forbindelse med Boligsocial monitorering? Har du oplevet ressourcemæssige udfordringer med projektet? Har du oplevet faglige udfordringer med projektet?
Afslutning	Hvilke muligheder ser du for hvordan databaseret viden kan blive benyttet ift. Boligsocial monitorering i fremtiden? Hvilke andre typer data ser du som meningsfulde i en boligsocial monitorering sammenhæng? Snart vil Sundheds- og Beskæftigelsesdata komme i FLIS, hvilke muligheder ser du med det i forhold til boligsocial monitorering?

8.C Reports and articles

Area	Reports and articles
Big data in the public sector	Economic Co-operation and Development (OECD, 2015) <i>Data-Driven Innovation: Big Data for Growth and Well-Being</i> McKinsey Global Institute (Manyika et al., 2011) <i>Big Data: The next frontier for innovation, competition, and productivity</i> Oracle (2012) <i>Big Data: A Big Deal for Public Sector Organizations</i> Pricewatercooperhouse (PWC, 2014) <i>An introduction to Big Data – how a buzzword morphed into a lasting trend that will transform the way you do business</i> McKinsey & Company (Bays, 2015) <i>Harnessing big data to address the world's problems</i> SAS Institute Danmark (Hyldborg, 2016) <i>Et digitalt drømmesamfund. Illusion eller realitet? – Fem forudsigelser om vores digital fremtid</i>
A source for knowledge creation	McKinsey Global Institute (Manyika et al., 2011) <i>Big Data: The next frontier for innovation, competition, and productivity</i> Google (Andrade et al., 2014) <i>From Big Data to Big Social and Economic Opportunities: Which Politics Will Lead to Leveraging Data-Driven Innovation's Potential?</i>
Decision making	Ernst & Young (EY, 2016) <i>How can big data lead to more rational decision-making?</i> Ernst & Young (EY, 2016) <i>How is big data analytics transforming corporate decision-making?</i> Capgemini (2012) <i>The Deciding Factor: Big Data & Decision Making</i>
Geospatial big data analytics	McKinsey Consulting (Grant et al. 2014) <i>Coordinates for Change: How GIS Technology and Geospatial Analytics Can Improve City Services</i> Deloitte Consulting (2014) <i>The Impact of Open Geographical Data</i>

8.D Data matrix

Concept of Big data									
H4A	ST	JG	NN	SA	AA	FR	GR	Grundidé	
[...] at vi har så meget data i det offentlige som man rent faktisk ikke bruger til noget (line 10) [...] kigger på værs af Danmarks befolkning (line 11)	Jamen det er jo mulighed for at kunne samle en hel masse data og kunne se nogle mønstre og i vores kontekst at Danmark er struktureret som det er så er det jo (line 16) der handler rigtig meget om benyttning, så det her med at kunne ved hjælp a en masse data kunne se nogle mønstre og nogle forskelle (line 19) spørgsmålet er hvorvidt man kan hælde noget for Big Data men der arbejder jeg (red. hendes røde hår) altså i og så med beslutningskraft, altså bare inden for kommunen hvor er det så på tværs af alle kommuner (line 27) Ja, men altså man kan sige Aarhus kommune der samlede vi al data, fra alle sagsbehandlere og alle sager og Århus kommune er en mega stor kommune så det var lidt noget data vi samlede og forsøgte at skabe mønstre og lave noget med det, gul grøn eller rød hvor der skulle handles på det for sålænge ting var (line 32)	Det er så stort, at det synes jeg næsten ikke jeg kan (line 16) / Altså, hvidt da lige op. Det er jo at fra. Man kan jo næsten sige en eller anden grad så er det jo det der med hvordan lærer man et sådan hulspil, overvejning en helles betegnelse ikke for det. (line 44)	[...] de der overvågningsperspektiver eller hvert fald det gode case jeg tit kommer i tanke om det er jo den her med hvordan lærer man et sådan hulspil, overvågningsperspektiver og hvor er press punkterne og i forhold til værkten og politiet og alle de der ting. Det kan jo også være at man under trafikale forhold eller i køkø gæren og laden i stocente eller bryndte osv. Det er sådan, synes jeg den klassiske forståelse jeg har at det. (line 13)	Så vi lag defineret det som et super godt begreb, til for mig at kunne ud med data fordi det er noget der har været iobryndet ud i hele det her med hvordan lærer man et sådan hulspil, overvågningsperspektiver og hvor er press punkterne og i forhold til værkten og politiet og alle de der ting. Det kan jo også være at man under trafikale forhold eller i køkø gæren og laden i stocente eller bryndte osv. Det er sådan, synes jeg den klassiske forståelse jeg har at det. (line 13)	Så vi lag defineret det som et super godt begreb, til for mig at kunne ud med data fordi det er noget der har været iobryndet ud i hele det her med hvordan lærer man et sådan hulspil, overvågningsperspektiver og hvor er press punkterne og i forhold til værkten og politiet og alle de der ting. Det kan jo også være at man under trafikale forhold eller i køkø gæren og laden i stocente eller bryndte osv. Det er sådan, synes jeg den klassiske forståelse jeg har at det. (line 13)	[...] fordi vi bruger det ikke særligt meget. Det er faktisk jeg har hørt om det som det til eksempel for en hel masse af de udbudsagere der står på det her med hvordan lærer man et sådan hulspil, overvågningsperspektiver og hvor er press punkterne og i forhold til værkten og politiet og alle de der ting. Det kan jo også være at man under trafikale forhold eller i køkø gæren og laden i stocente eller bryndte osv. Det er sådan, synes jeg den klassiske forståelse jeg har at det. (line 13)	Grundide opfattelse - ikke nogen enlyd definition (fraksæson under det samme) (3V's) han og Mikkel Overvejning Mikkel benytter eksempel på at kunne foretage begrebet. Benchmark Hedvig/Jas/Sara: Kigger på værs af kommuner. Hedvig/Jas/Sara: Mågt data. Se mønstre / (red. gul, grøn) Mikkel: Konkrethedsfangst begrebet Astrid: I forbindelse med eget arbejde, dvs. data koordinatør, sagsbehandler. Brugte Big Data som et markedsbegreb, der muliggør hendes arbejde med data fordi det er populært Torris: buzzword. Men det bruges ikke meget. Mængde af data. Ikke kempe meget grundet den begrænsede befolkningsmængde. Populært begreb ifm. google og facebook. Det er ikke noget man siger vi bruger det eller under værs. Mågt forståelse for det.	Der findes forskellige erfaringer og kendskab til Big Data Så, der er ung har stans kendskab. Stilet kendskab med det gennem stude. Mikkel, Torris & Astrid: Kort tid, kun 3 år måske. Disse har stilet kendskab gennem praksis. Det er spændende at se hvordan kendskabet varierer i forhold til alder og røle. Forsøgsbefragning: 2V (2011) Gælder:	

Udfordring	TR	AKA	MN	IC	SDH	HHA	Digitaliseringsstrategi
Der indses eget potentiale i data	Ja, altså det bruger vi jo hele tiden. Det tænker jeg man bruger, det tænker jeg at alle bruger i hvert fald hver dag. Så det tænker jeg det er meget udbredt og noget af det der bliver mere og mere forskelligt data kilder end i tidligere kunne. Til har man meget sidet og i søjler og kigget på data som, na så har vi data om hvilken offentlig ydelse du modtager og så et helt andet sted har vi nogle data om at du måske også ja modtager en eller anden ydelse fra et andet område social området eller andet. (line 39)	Men det er klart at selvom vi nu ved hvad Big Data er som begreb og det er nødvendigt at blive dataleverne og det er en dataverden vi ved en og i kender alle de her ting. Så er det måske ikke et område som endnu har tilført det store ressourcer for at få gennemført. Det er noget vi skal trække på eksisterende ressourcer her. (line 113)	[...] nu hvor vi for eksempel på Berne og unge lige har skrevet fagsystem, fra et ESDB ligeme fagsystem til at udbyde det nye som også var okay til DUBU som har nogle helt andre data dimensioner. Og det er også det er en nydel sammenhæng mellem de to systemer men det er svært fald i eksempel på de både kræves fra lovgivningen og myndighed ønsker at få nogle stærkere ledelsesinformation og være mere dataorientet. (line 46)	Der har ikke været et pres på egent brug på grund af det der er sket med den såkaldte besparelser i digitaliseringsstrategien 2016-2020. Jeg oplever det som at det er os selv der arbejder med det og godt kan se, at det kan vi få meget mere ud af. (line 38)	Ja, i høj grad (red. potentiale for big data i den offentlige sektor). Altså noget af det vi virkelig arbejder for i KL, det her med at give kommunerne, altså både det her med at hjælpe dem med at få stillet det her data til rådighed bl.a. ved hjælp af nogle publikationer og hjemmesider, men også simpelthen at de får øjnene op for hvad for noget potentiale der er i det. Fordi, vi kender det jo meget ledelsesinformation, men det her med at kunne se tingene lidt oppe fra og kunne sammenligne på værst både inden i en kommune og på værst af kommuner det er noget hvor man måske kommunerne er allerede langt med det og bl.a. via FLIS men der er stadig rigtig meget potentiale, ulovtst potentiale. (line 40)	[...] Vi har så meget data i det offentlige som man rent faktisk ikke bruger til noget på værst af kommunerne. (red. i spørgsmålet om FLIS brunt og opførelsen af potentialer i den offentlige sektor) (line 10)	
Der indses eget potentiale i data	Der indses eget potentiale i data	Der indses eget potentiale i data	Der indses eget potentiale i data	Der indses eget potentiale i data	Der indses eget potentiale i data	Der indses eget potentiale i data	Der indses eget potentiale i data

[illegible]

[illegible]

8.E Transcribed interview – Hedvig Hvarregaard Andersen (HHA)

Interviewer: Anders Mårtensson (AM) and Lasse Vinter (LV)

Date: 6th of March 2017

Length: 26:56 min.

Location: KOMBIT, Copenhagen

Found on attached USB

8.F Transcribed interview – Sara Thestrup Hansen (STH)

Interviewer: Anders Mårtensson (AM) and Lasse Vinter (LV)

Date: 9th of March 2017

Length: 33:47 min.

Location: Kommunernes Landsforening (KL), Copenhagen

Found on attached USB

8.G Transcribed interview – Jan Carlsson (JC)

Interviewer: Anders Mårtensson (AM) and Lasse Vinter (LV)

Date: 21st of March 2017

Length: 24:07 min.

Location: Svendborg Municipality

Found on attached USB

8.H Transcribed interview – Mikkel Næsvang (MN)

Interviewer: Anders Mårtensson (AM) and Lasse Vinter (LV)

Date: 21st of March 2017

Length: 20:02 min.

Location: Svendborg Municipality

Found on attached USB

8.I Transcribed interview – Astrid Krarup Andersen (AKA)

Interviewer: Anders Mårtensson (AM) and Lasse Vinter (LV)

Date: 21st of March 2017

Length: 18:36 min.

Location: Svendborg Municipality

Found on attached USB

8.J Transcribed interview – Troels Rasmussen (TR)

Interviewer: Anders Mårtensson (AM) and Lasse Vinter (LV)

Date: 29th of March 2017

Length: 46:03 min.

Location: Frederiksberg, Phone interview

Found on attached USB

