NTNU
Norwegian University of Science and Technology

The Role of Big Data in Addressing Societal Challenges

Farzana Quayyum

Master of Science in Information Systems
Submission Date: 07 July 2019

Supervisors:
Letizia Jaccheri, IDI
Ilias O. Pappas, IDI

Norwegian University of Science and Technology
Department of Computer Science
Trondheim, Norway

Acknowledgement

I would like to express my special thanks of gratitude to my supervisors Professor Letizia Jaccheri and Post Doctoral Fellow Ilias O. Pappas for their supervision and support during the whole Master thesis period. Along with the researcher, the supervisors have contributed a lot in this research with their expertise and experience within Information Systems research. I would also like to thank NTNU, especially IDI for their services and for facilitating a great learning experience.

Abstract

In the last decade, the use of big data analytics has earned a lot of attention. But the potential use of big data is yet to be figured out by social sector organizations, from a social perspective. Harnessing the power of big data to increase collective and individual awareness about societal problems, and ultimately to create the needed intelligence and innovative solutions for societal problems is still challenging. From this master thesis, we aim to investigate the role of big data analytics in addressing societal challenges and social good, leading to generating social value; and also we aim to explore the associated challenges and drivers of employing big data analytics. Finally, the objective of this thesis is to examine how the big data challenges and enablers influence the intentions of social entrepreneurs and innovators to use big data analytics in their work. From this study, we aim to answer the following research questions:

RQ1: What is the role of big data and analytics in helping social innovators and entrepreneurs generate social value?

RQ2: What are the challenges and benefits of using big data and analytics to address societal challenges?

RQ3: What is the relation between challenges and enablers of big data and analytics with social entrepreneurial intentions?

The research direction was formulated on the basis of The role of big data in addressing societal challenges: A systematic mapping study"(Appendix A.1). We have collected a wide range of data to answer our research questions; we have interviewed entrepreneurs and innovators from the social sector who work to address various social issues and challenges. We have also conducted a questionnaire to validate and complement our research findings. Finally, this research provides an overview of the social areas where big data analytics is being used currently; how they are being used and also what future potentials it has. This research also identifies the relationship of big data analytics with social entrepreneurship and social innovation.

Sammendrag

Det siste tiåret har analyse av Big Data fått mye oppmerksomhet. Derimot er den potensielle bruken av Big Data ennå ikke oppdaget av sosiale organisasjoner basert på ett sosialt perspektiv. Det er fortsatt utfordrende å kunne utnytte styrken fra Big Data for å øke kollektiv og individuell bevissthet om samfunnsproblemer, og dermed til slutt kunne skape nødvendige intelligens og innovative løsninger for samfunnsproblemer. I denne masteroppgaven tar vi sikte på å undersøke rollen av Big Data og analyse i det å møte samfunnsmessige utfordringer og samfunnsmessig goder, noe som fører til generering av samfunnsverdier. I tillegg tar vi sikte på å utforske de utfordringene og assosierte endringer og drivere av Big Data og analyse. Denne oppgaven sikter etter å undersøke hvordan utfordringer ved Big Data og hjelpemidlene påvirker intensjonene til sosialentreprenører og innovatører i det å bruke stordataanalyse i deres arbeid. Basert på denne studien tar vi sikte på å svare på følgende forskningsspørsmål:

RQ1: Hva er Big Data og analyse sin rolle i å hjelpe sosiale innovatører og gründere å generere sosial verdi?

RQ2: Hva er utfordringene og fordelene ved å bruke Big Data og analyse for å løse samfunnsutfordringer?

RQ3: Hva er forholdet mellom utfordringer og hjelpemidlene av Big Data og analyse med sosialentreprenørintensjoner?

Forskningsretningen ble formulert basert på The role of big data in addressing societal challenges: A systematic mapping study"(Vedlegg A.1). Vi har samlet et bredt spekter av data for å svare på forskningsspørsmålene. Vi har intervjuet gründere og innovatører fra sosialsektoren som jobber med å møte ulike sosiale problemer og utfordringer. Vi har også tilføyet et spørreskjema for å validere og utfylle våre forskningsresultater. Denne undersøkelsen gir en oversikt over de sosiale områdene hvor Big Data og analyse brukes i dag, hvordan de blir brukt og hvilke fremtidige potensial det har. Denne undersøkelsen identifiserer også forholdet mellom utfordringer i Big Data analyse og hjelpemidlene innen sosialentreprenørskap og sosial innovasjon.

Preface

This thesis is submitted to the Norwegian University of Science and Technology (NTNU) as part of the course TDT4900 Computer Science, Master's Thesis. The thesis has been performed under the supervision of Professor Letizia Jaccheri and Post Doctoral Fellow Ilias O. Pappas at the Department of Computer Science, NTNU, Trondheim.

The Master thesis builds on the systematic mapping study from the course TDT4501 Computer Science, Specialization Project. The mapping study of this thesis has been accepted in the 18th IFIP Conference on e-Business, e-Services and e-Society (I3E2019) and will be published in a special volume of Lecture Notes in Computer Science by Springer. The paper can be found in Appendix A.1.

Contents

List of Tables

List of Figures

Abbreviations

ANOVA = Analysis of Variance
EI = Entrepreneurial Intention
GDPR = General Data Protection Regulation
GQM = Goal Question Metrics
SE = Social Entrepreneurship
SEI = Social Entrepreneurial Intentions
SPSS = Statistical Package for the Social Sciences
NSD = Norwegian Centre for Research Data

Chapter 1

Introduction

Ubiquitous data and services offer powerful storage and intelligent computation capability where only imagination is the limit for what software can achieve to capture the rich information that exist in the data. Society may be viewed as a digital ecosystem on which data, information, and knowledge is shared and transferred among its stakeholders, to achieve transformation and create sustainable societies. There is a growing need for findable, accessible, reusable, and interoperable infrastructures and data management standards that provide greater access to the information in the society (Wilkinson et al. (2016)). Such infrastructures enable thriving innovation in the society; motivating social innovators and entrepreneurs to use this data and information in their innovative solutions to tackle our complex societal challenges. Big data applications can be used for social good, like the ones already being done; for example - predicting epidemics from mobile data (Wesolowski et al. (2015)), or for transport planning and traffic engineering (Morner (2016)) etc.

In the last decade, the use of big data and data analytics has earned a lot of attention. In various fields of science, technology, and business, the merit of big data is undeniable. But from the social perspective, the potential use of big data is yet to be figured out by social sector organizations(Pappas et al. (2018)). To harness the power of big data, to increase collective and individual awareness about societal problems, and ultimately to create the needed intelligence that will lead to innovative solutions for societal problems is still a challenge. The aim of this master thesis is to investigate the role of big data in generating social value and also exploring the challenges that social innovators and social entrepreneurs face in employing big data analytics; along with the challenges we also investigate the drivers that can help and motivate them in using big data analytics. Finally, the objective of this thesis is to examine how the big data challenges and enablers influence the intentions of social entrepreneurs and innovators about using big data analytics in their work.

The remainder of this chapter proceeds as follows: Section 1.1 presents the motivation for this research. Section 1.2 presents the research questions. Section 1.3 defines the boundaries of the research. Section 1.4 explains the chosen research method and research process. Section 1.5 presents the outline of this thesis.

1.1 Motivation

The value of big data is clear for addressing technical and business problems (Chen et al. (2012)). However, research on the social value of big data is not at the same level (Agarwal and Dhar (2014); Zicari (2014)), raising the question on how well big data can be used and is being used to address complex social problems. Akoka et al. (2015) performed a literature review on big data research, which showed at that time the usages of big data did not attract researchers except for two application domains: Marketing and retail, and Healthcare. In 2017, the authors again conducted a literature review on big data research which showed that domains like earth, energy, medicine, ecology, finance, government, education have started getting attention also along with marketing and health from researchers (Akoka et al. (2017)). But still using big data to solve social problems or for social good has not been well acknowledged. So, in this research, we have explored the social side of big data and how it can be used for the society. Since Big data is a recent upcoming technology in the market which can bring huge benefits for organizations, it becomes necessary that various challenges and issues associated in bringing and adapting to this technology are brought into the light (Katal et al. (2013)). So with this master thesis, we aim to understand the role of big data and analytics in helping social entrepreneurs and innovators to address societal challenges for social good and thus helping to generate social value. Social value is defined differently by many different people. There is no single authoritative definition of "social value" but we can say that it refers to wider non-financial impacts of programs, organizations and interventions, including the wellbeing of individuals and communities, social capital and the environment (Mulgan (2010)). We investigate what conditions can enable successful solutions for creating social values and what are the challenges in harnessing the potential of big data for them. We also study how these challenges and enablers influence the intentions of the social innovators and social entrepreneurs to adapt big data applications that can lead to social innovation and thus societal transformation.

The specialization project "The Role of Big Data in Addressing Societal Challenges: A Systematic Mapping Study" undertaken in course TDT4501 - "Computer Science, Specialization Project" serves as the foundation for this Master thesis. The summary of the literature review was submitted to the *18th IFIP Conference on e-Business, e-Services and e-Society (I3E 2019)*[1] and has been accepted to publish in a special volume of *Lecture Notes in Computer Science by Springer*. The paper can be found in Appendix A.1. The objective of the systematic mapping was to offer a map of the research that is being done on this topic so far, thus offering the basis for a reflection process among the researchers in this field. We have used the findings of the mapping to develop our research agenda

[1] https://www.i3e2019.com

for this thesis and investigated the role of big data and their applications to generate social value.

1.2 Research Questions

Big Data is a significant area of study now, for both practitioners and researchers. But the social impact of big data has not been explored as it has been for marketing or technological innovation. In this research we aim to address this gap. Despite the success of big data applications, some obstacles and challenges in the development of big data applications remain (Akoka et al. (2017)). We want to explore those obstacles also in this research and want to understand how the obstacles and also the drivers impact the social entrepreneurs and innovators. So, to fulfill our research interests, this research aims to answer the following research questions:

RQ1: What is the role of big data analytics in helping social innovators and entrepreneurs to generate social value?

RQ2: What are the challenges and benefits of using big data analytics to address societal challenges?

RQ3: What is the relation between challenges and enablers of big data analytics with social entrepreneurial intentions?

1.3 Research Scope

From this thesis, we aim to understand how social innovators and social entrepreneurs are addressing societal challenges using big data analytics in their work; their experience of using big data and how well big data can be used to address complex social problems. We have also tried to identify how big data challenges influence their work and what opportunities the social entrepreneurs get when they use data and data driven technologies for their solutions. Social innovation is defined as "innovative activities and services that are motivated by the goal of meeting a social need" (Mulgan (2006)) and as change agents, social entrepreneurs harness innovation at a systemic level to bring about a change in social equilibrium (de Bruin and Ferrante (2011); Lehner and Kansikas (2012); Zahra et al. (2009)). Social entrepreneurs focus on bringing about improved social outcomes for a particular community or group of stakeholders (Phillips et al. (2015)). The terms social innovation, social good, social change, and societal transformation are related to each other. All these concepts concern about ideas and actions that have positive change and impact on the society. During this study our focus is on the applications of big data that have social impact and address social problems or challenges; so, to keep a broad and wide scope in this research we use all these terms.

The participants of this research are individuals who work in the social sector; they are involved with social innovation and/or social entrepreneurship individually or as part of an organization. The participants also had experience of using big data analytics in their work either in a small scale or in a larger scale.

We have collected data with semi-structured interviews and a questionnaire. We have not considered any geographical preferences regarding the participants. So the participants of both interview and questionnaire are from a diverse background. Though the participants are from all around the world but the majority of the participants and their organizations are based in Europe.

1.4 Research Process

This research started with a systematic mapping which will be presented in detail in chapter 3. The mapping study gave us an overview of the field and helped us to determine a conceptual framework for further research and then to come up with the research questions. Figure 1.1 presents an overview of the different research methods used in this research.

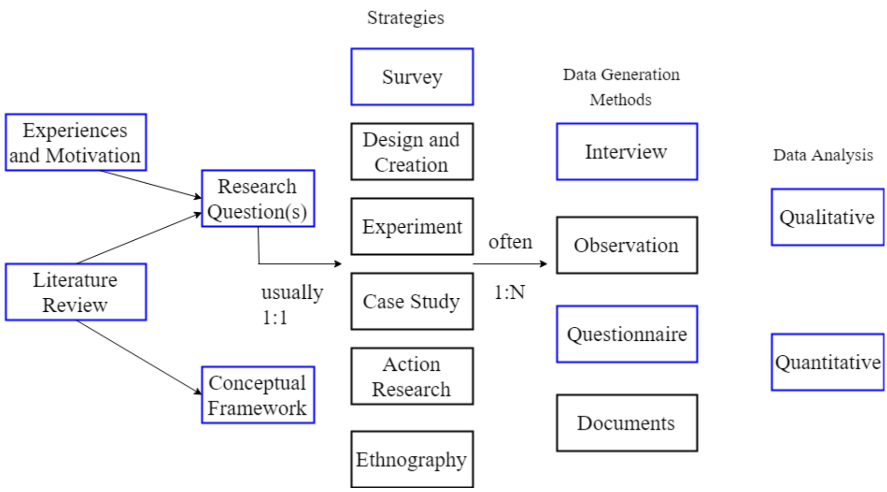

Figure 1.1: Research method overview, adapted from Oates (2005)

To address the research questions, we have performed both qualitative and quantitative methods in this research. We have collected qualitative data with semi-structured interviews and quantitative data with a questionnaire.

A survey approach was utilized for data generation. We have performed semi-structured interviews with 10 individual professionals who are working in the social sector. All the interviews were recorded and transcribed, and all participants gave oral consent for this. As the interview participants were from different countries, so we conducted the interviews using multiple online platforms (Skype and different online conference systems).

To develop the interview questions we used the GQM method (Basili (1992)); appropriate for answering our research questions, and also enabling us to focus on our predefined research topics in an exploratory approach. To analyze the interview data, thematic synthesis process by Cruzes and Dyba (2011) has been used and NVivo was used to code the transcribed interviews. The initial coding process resulted in 65 codes with 124 references; and finally gave us 8 themes. We finally classified the themes into two higher order themes.

Quantitative data were collected through an electronic questionnaire. We have performed statistical analysis (factor analysis, Spearman's rank correlation coefficient and analysis of variance) with the questionnaire data. We have used SPSS (Statistical package for the social sciences) to perform the analysis. The detailed data analysis process of both interview and questionnaire will be described in Chapter 4.

Multiple data generation methods were used so that we can collect a wide range of data; and also can support and complement the findings of each data generation method. By utilizing both data generation methods, a lot of qualitative and quantitative data was generated and analyzed. Thus a mixed method study and a method triangulation was used, which made it possible to provide additional evidence and support for the findings of this research.

1.5 Outline of the Thesis

The rest of the thesis proceeds as follows. Chapter 2 introduces the background of the study and then proceeds with the systematic mapping study undertaken in advance of the Master thesis, and its main findings in Chapter 3. Chapter 4 presents the research methodology undertaken, and the ethical considerations undertaken when conducting this empirical research. Chapter 5 presents the results of this research, including synthesized results obtained from interviews and questionnaire. Chapter 6 discusses the results in relation to the research questions. Finally in Chapter 7, the thesis concludes by answering the research questions and proposes directions for future work.

Chapter 2

Background

2.1 Big Data and Data Analytics

The digital and connected nature of modern-day life has resulted in vast amounts of data being generated by both people and organizations. This phenomenon of an unprecedented growth of information and our ability to collect, process, protect, and exploit it has been described with the catchall term of Big Data (Cuquet et al. (2017)). The extant literature identifies 'big data' as the 'next big thing in innovation' (Gobble (2013)), the next frontier for innovation, competition, and productivity" (Manyika et al. (2011)). The rationale behind such statements are that the 'big data' is capable of changing competition by "transforming processes, altering corporate ecosystems, and facilitating innovation" (Brown et al. (2011)).

The definition of big data has different dimensions and there is no universally accepted definition. The term is used to describe a wide range of concepts: from the technological ability to store, aggregate, and process data, to the cultural shift that is pervasively invading business and society, both drowning in information overload (Mauro (2015)). From the literature, we also see that many scholars use the term 'Big data' not only to refer large sets of data but also refers to analytical tools used to manage these large sets of data and turn them into useful and productive information. A couple of examples are Siemens and Long (2011), who define big data as "data sets whose size is beyond the ability of typical database software tools to capture, store, manage and analyze" and Chen et al. (2012), who call it "data sets … that is so large (from terabytes to exabytes) and complex (from the sensor to social media data) that they require advanced and unique storage, management, analysis, and visualization technologies". In the study of Wamba et al. (2015), the authors also stated that "We need to think about 'big data' not only in terms of analytics but more in terms developing high-level skills that allow the use of new generation of IT tools and architectures to collect data from various sources, store, organize, extract, analyze, generate valuable insights and share them with key firm stakeholders for competitive advantage co-creation and realization. These definitions and statements show that the authors think of big data in terms of how it gets analyzed, not how many specific terabytes of space it fills (Maltby (2011)). Based on that, we use the terms 'big data' and 'data analytics' interchangeably in this research.

2.2 Big Data and Social Innovation

Big Data contains a wealth of societal information. Analyzing big data and further summarizing and finding clues and laws it implicitly contains can help us better perceive the present (Jin et al. (2015)). To initiate any innovative steps or to address any social challenge, data are essential. A deliberate and systematic approach towards social innovation through big data is needed as it will offer social value and competitive advantage for social entrepreneurs (Pappas et al. (2017)) allowing the transformation of current business models leading to sustainable societies. New business models will go beyond economic needs, and address societal challenges generating shared value that impacts the companies, organizations, consumers, and the public at large (Porter and Kramer (2019)). Big data play a key role in this transformation and combining them from multiple sources, sharing them with various stakeholders, and analyzing them in different ways can lead to the achievement of digital transformation and creation of sustainable societies (Pappas et al. (2018)).

The authors of Pappas et al. (2017) define a Social Innovation Ecosystem (SIE) as Figure 2.1 that shows how all the stakeholders need to collaborate and cooperate together to enable the use of big data for societal transformation towards the achievement of social change. The main actors of a Social Innovation Ecosystem are individuals/social entrepreneurs, social enterprises/industry, government, civil society, and academia. Big data can empower the stakeholders like organizations, academia, entrepreneurs or citizens who want to follow data-driven decision making to solve various social problems and to achieve digital transformation to make society more sustainable. If we can identify how the stakeholders can take advantage of big data to enhance their impact on society, and how they can benefit from social innovation using big data, we can empower them. We aim to identify the ways to empower them with new empirically supported and data-driven practices and applications, including recommendations and descriptors of how to stimulate innovation through big data analytics, aimed at solving the societal problems with a greater impact.

2.3 Social Innovation and Social Entrepreneurship

In this research, we particularly focus on the role of social innovators and social entrepreneurs among other actors in the social innovation ecosystem (Figure 2.1) in creating innovative solutions for the society with data analytics. We will explore the concept of

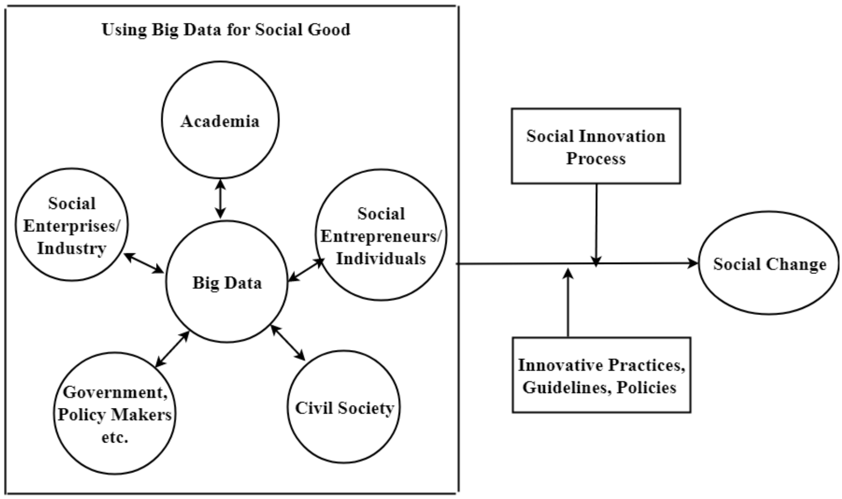

Figure 2.1: Using big data in the social innovation ecosystem by Pappas et al. (2017)

social entrepreneurship along with social innovation as the actors of the same ecosystem and how they connect to each other in addressing societal problems or challenges.

Phillips et al. (2015) conducted a mapping study on social innovation and social entrepreneurship, which provides collective insights into research linking social innovation with social entrepreneurship. The authors write that, social innovation is not undertaken in isolation by lone entrepreneurs, but is an interactive process shaped by the collective sharing of knowledge between a wide range of organizations and institutions that influence developments in certain areas to meet a social need or to promote social development (Phillips et al. (2015)).

Ashoka, a foundation, which has been promoting social entrepreneurship for many years, defines social entrepreneurs as "Social entrepreneurs are individuals with innovative solutions to society's most pressing social problems" (Ashoka (2019)). Social entrepreneurs focus on bringing about improved social outcomes for a particular community or group of stakeholders adopting a business approach (Phillips et al. (2015)). The underlying drive for social entrepreneurship is the creation of social value as opposed to personal or shareholder wealth (Noruzi et al. (2010); Thake and Zadek (1997)) and the activity of such social creation is characterized by pattern-breaking change or innovation (Munshi (2010); Noruzi et al. (2010)), through the creation of new combinations of, for example, products, services, organization, or production (Defourny and Nyssens (2010)). Phills et al. (2008) define social innovation as "a novel solution to a social problem that is more effective, efficient, or just than existing solutions and for which the value created accrues primarily to society as a whole rather than private individuals". And, Zahra et al. (2009) define social entrepreneurship as "the activities and processes undertaken to discover, define, and exploit opportunities in order to enhance social wealth by creating new ventures or managing existing organizations in an innovative manner". By aligning this definition of social entrepreneurship by Zahra et al. (2009) with the definition of social innovation by Phills et al. (2008), the authors of Phillips et al. (2015) suggested that social entrepreneurship and social innovation are both about identifying a problem-solving opportunity to meet a social need. So, researches (Phillips et al. (2015)) shows that both social entrepreneurship and social innovation share common understandings and overlaps, significantly in the process of identifying problem-solving opportunities for social needs.

From all this literature we can say social innovation and social entrepreneurship is very well connected to each other and also shares the same characteristics. They also have similarities in their works and outcomes towards society. So, in this research, we not only focus on social innovation but also on social entrepreneurship in creating social value and addressing societal challenges with big data analytics.

2.3.1 Social Entrepreneurship Scale

Similar to traditional or conventional entrepreneurs, social entrepreneurs also pursue success by creating and adding value to their companies. Social entrepreneurship has been labelled ‘caring capitalism’ because the achievement of relevant social goals relies on competitiveness in the marketplace (Hibbert et al. (2005)). Dees (1998) states that "Similar to conventional entrepreneurship, social entrepreneurs must first identify their competition and then develop strategies for competing effectively. Furthermore, they must realize that competition is not limited to non-profit or social-oriented competition". So, value creation, market competition, business models, business strategies these are some common concerns for social entrepreneurs also, just like traditional entrepreneurs. Thus, sometimes differentiating social entrepreneurs from conventional entrepreneurs is not easy.

As mentioned by Dees (1998), social entrepreneurs serve as change agents (change agents or world changers) who: (1) adopt a mission to create and sustain social value (have a mission), (2) recognize and pursue new opportunities to serve their mission (take action), (3) engage in a process of adaptation, innovation, and learning (active education), (4) are not limited by the resources they currently possess (resources), and (5) have a sense of accountability to those served and the outcomes created (results/outcomes). In the process of social entrepreneurship, social entrepreneurs also share similar characteristics like conventional entrepreneurs, but the their difference lies in their missions. Using these five facets of a unidimensional social entrepreneurship concept from Dees (1998), Carraher (2013) developed a 11-item social entrepreneurship scale to measure social entrepreneurship.

In our research, we have used this 11-item scale by Carraher (2013) to understand our questionnaire participants’ mission, their way of involvements with the society, innovation thinking etc. We asked them to rate (with a seven point likert scale) how much each of the statements from 11-item social entrepreneurship scale describes them and their position in the social sector. In Table 2.1 we present the 11-item social entrepreneurship scale statements that we have asked to the participants of our questionnaire.

Table 2.1: Social Entrepreneurship Scale by Carraher (2013)

Based on my experience and/or current role in the organization, the following are descriptive of me:
1. I am adopting a mission to create social value (not just private value).
2. I am recognizing new opportunities to serve my mission.
3. I am engaging in a process of continuous adaptation related to my mission.
4. I am acting boldly without being limited by resources currently in hand in the fulfillment of my mission.
5. I am relentlessly pursuing new opportunities to serve my mission.
6. I am caring deeply about the outcomes created by the fulfillment of my mission.
7. I seek to be a 'world changer' through the accomplishment of my mission.
8. I am adopting a mission to sustain social value (not just private value).
9. I am engaging in a process of continuous innovation related to my mission.
10. I am exhibiting a heightened sense of accountability to the constituencies served by my mission.
11. I am engaging in a process of continuous learning related to my mission.

2.3.2 Social Entrepreneurship Intentions (SEI)

In today's society, the role of social entrepreneurship is vastly increasing through both non-profit and for-profit businesses (Letaifa (2015)). The success of social entrepreneurs can create both private and public wealth. For example, the Bill and Melinda Gates Foundation; though Bill Gates is not a social entrepreneur, the foundation has made a major difference worldwide in advancing the health and education for all, an inspiring social entrepreneurship mission (Certo and Miller (2008); Selladurai and Carraher (2014); Stallworthy et al. (2014)). All though social entrepreneurs pursue goals and objectives that concerns social problems and needs; they have similar characteristics to conventional entrepreneurs; as we discussed in the previous section. Both social entrepreneurs and conventional entrepreneurs are innovative, creative, and motivated to pursue their venture. As an actor of the social innovation ecosystem, innovation is a key focus for social entrepreneurs. But social entrepreneurs will differ in the extent to which they prefer to include innovative elements in their business models and for which purpose specifically. With the context of this study, social entrepreneurs can use big data analytics to achieve different missions, for example- big data can be used for their own management purpose, for building innovative solutions for the society, for addressing and identifying problems or people who are in need; or even for profit seeking (though it is not the primary mission of social entrepreneurs). So, big data can have multiple dimensions of usage that directly or indirectly helps the social entrepreneurs and their organizations to achieve their mission.

In this research, we have investigated how social entrepreneurs and innovators are using big data to address societal challenges and to generate social value. We also tried to understand what are the challenges and drivers that entrepreneurs or innovators experience when they try to employ big data in their innovative solutions; and also, how these challenges or drivers influence their different entrepreneurial intentions. The motivation to innovate is an important issue for social entrepreneurship (Phillips et al. (2015)), because innovative new products and services, and/or innovative production and distribution processes, may be necessary to solve the "market failure" and government "failure" problems (Santos (2012)) that precipitate or allow social problems to persist. While innovation may be instrumental to the achievement of the pro-social outcomes, the achievement of successful innovation is also expected to be instrumental to the generation of psychic income for the social entrepreneur, arising due to self-satisfaction, self-esteem, and from recognition by others of one's social responsibility (Bacq and Alt (2018)). So, facing any challenge or facilitating by any driver can change the intentions of entrepreneurs for innovating new products or solutions to address social issues. In this research, we want to understand the relationship between the challenges and drivers of big data analytics with different social entrepreneurial intentions. To measure the influence and effect of big data challenges and drivers on social entrepreneurs, we have used the SEI scale developed by Douglas and Prentice (2019).

So, we have asked the following questions to our questionnaire participants as presented in Table 2.2 to measure their social entrepreneurial intentions, based on Douglas and Prentice (2019).

Table 2.2: Social Entrepreneurial Intentions by Douglas and Prentice (2019)

Please rate how likely it is that you would want to use big data and data analytics in your organization to:
1. Pursue a high-risk opportunity that has the possibility of very high profits
2. Grow the firm to be very large and profitable
3. Pursue profit maximization above all other objectives
4. Become a major, globally-recognized corporation
5. Generate high profits over many years
6. Locate the business at a place that suits your personal preferences
7. Enjoy the lifestyle and benefits of an independent business owner
8. Create a business around your personal hobbies or special interests
9. Have great flexibility to decide your work hours, your product lines, and so on
10. Be your own boss and make all the important decisions yourself
11. Gain great satisfaction because you are helping others who are in need
12. Solve social and economic problems that cause others to suffer
13. Help poor people get enough food, clothing, shelter, and medical assistance
14. Serve as a volunteer to help people who have social and/or economic problems
15. Help underprivileged people achieve what they are unable to achieve on their own

Chapter 3

Systematic Mapping Study

Prior to this master's thesis, a systematic mapping study of this research topic was undertaken as the specialization project, to provide an overview of the research available in the field of big data analytics and social innovation leading to social good. This chapter describes the mapping process and presents the finding of the mapping study. For the mapping, we followed several steps of the standardized process for systematic mapping studies. The study provides a mapping of total 156 papers; a list of which can be found in Appendix A.2. We focused on the finding of the mapping study and built our research agenda for this thesis. As mentioned in section 1.1, this chapter is a detailed version of our paper *"The Role of Big Data in Addressing Societal Challenges: A Systematic Mapping Study"* and proceeds as follows: section 3.1 explains the research process, section 3.2 presents the main findings of the mapping, and section 3.3 finally provides a discussion on the findings and section 3.4 concludes with the directions for future work.

3.1 Mapping Process

Systematic mapping studies are good in research areas with few relevant primary studies of high quality, as it provides a coarse-grained overview of the publications within the topic area (Petersen et al. (2008)). This systematic mapping study followed guidelines from Kitchenham (2004), and several steps of the standardized process for systematic mapping studies as Petersen et al. (2008), illustrated in figure 3.1. The main steps of our process are explained in this section and include the research questions, search and study selection strategies, manual search, and data extraction and synthesis method.

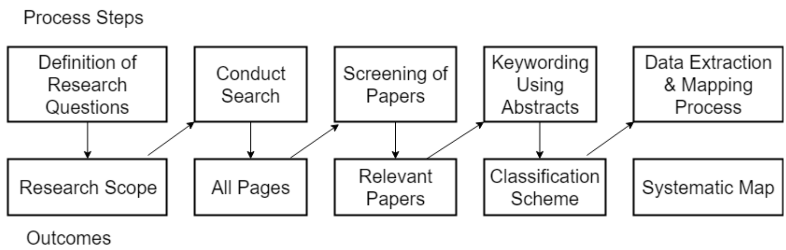

Figure 3.1: The systematic mapping study process by Petersen et al. (2008)

3.1.1 Research Questions

Using technology for addressing societal challenges is not new. But the term 'social innovation' is comparatively new and introducing data for innovative solutions to make an impact in the society can be seen mostly in the past few years. We have not found any other systematic mapping study published on this topic so far. But the newness and growing interest in the research field, argues the need for a mapping study to identify the focus and quality of research in using big data analytics for social challenges.

To provide an up to date overview of the research results within the field, we came up with the following research questions:

RQ1: How the research about 'big data and social innovation' has changed over time (in the last decade)?

RQ2: How much of the research is done based on empirical studies and what type of empirical studies?

RQ3: What are the challenges or barriers for successful implementation of big data for societal challenges?

3.1.2 Data Sources and Search Strategy

In our primary search, we have collected papers from all kind of sources including journals, conference papers, books, reports etc. The systematic search strategy consisted of searches in seven online bibliographic databases which were selected based on their relevance with our search topic and these databases are also well known for good quality literature resources in the field. To obtain high-quality data, we have searched in the following databases – Scopus, ISI Web of Science, ACM Library, IEEE Xplore, SAGE, Emerald and Taylor & Francis.

Then initial searches in the databases were conducted based on identified keywords related to this topic. The used search strings were:
("Big Data" OR "Data Analytics") AND "Social innovation"
("Big Data" OR "Data Analytics") AND "Societal Transformation"

("Big Data" OR "Data Analytics") AND "Social Change"
("Big Data" OR "Data Analytics") AND "Social Good"

3.1.3 Study Selection

The study selection process is illustrated in Figure 3.2, along with the number of papers at each stage. Searching the databases using the search string returned 593 papers, resulting in 465 unduplicated papers. These were imported into EndNote X8.

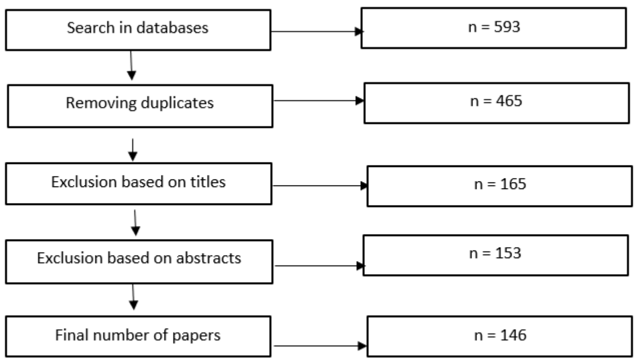

Figure 3.2: Study selection process of this mapping

Due to the importance of the selection phase in determining the overall validity of the literature review, a number of inclusion and exclusion criteria were applied. Studies were eligible for inclusion if they were focused on the topic of big data and data analytics, and their applications to foster social innovation, and lead to social impact, change and transformation.

This review was conducted in August 2018 and publications were searched from 2008 and onwards. We selected this timeframe as it is the time when these terms like big data and analytics, social innovation got the momentum. The systematic review included research papers published in academic outlets, such as journal articles and conference proceedings, as well as reports targeted at business executives and a broader audience, such as scientific magazines. In progress research and dissertations were excluded from this review, as well as studies that were not written in English. Given that our focus was on the social innovation and societal transformation that big data entails, we included quantitative, qualitative, and case studies. Since the topic of interest is of an interdisciplinary nature, a diversity of epistemological approaches was opted for.

We have not included books or book section in the mapping finally, but we have considered reports (e.g., Hitachi reviews, Fujitsu reports) because a lot of evidence is published by companies and a lot of work on social innovation and big data is done by companies as well.

3.1.4 Manual Search

Following the systematic search, a manual search was also conducted. Google Scholar was used to searching for papers manually. At this stage total 10 papers from Google scholar was added to our EndNote library and the final number of papers became 156.

3.1.5 Data Extraction and Synthesis

After the mapping, we finally ended up with 156 papers. We performed a systematic analysis and extracted data from the abstracts of the papers that we need to answer our research questions. We extracted data regarding the - publication frequency, publication source, research area, research type, empirical evidence and contribution type.

3.1.6 Threats to Validity

For the validity of this review, threats of retrieval of papers need to be considered. As following the standard procedure of systematic mapping, we used the keyword search in title and abstract of the papers, the papers which do not include these terms in the title or abstract are not included in this mapping. We are aware that some papers might be relevant to this topic but are not included in the mapping for this reason. Selection of databases might have affected the number of relevant papers retrieved. But we believe that we have searched the major databases and we have selected the databases which we considered relevant with our topic and also known for good quality papers.

3.2 Synthesized Results

3.2.1 RQ1: How the research about 'big data and social innovation' has changed over time (in the last decade)?

Publication Frequency
After analyzing the final 156 papers, we have found that the relevant papers are published from 2012 or later, with their frequency increasing yearly. The study was conducted in August 2018, so the finding of the year 2018 is until August. The findings (Figure 3.3) verify that the momentum or applications of big data are becoming increasingly popular.

Research Areas
Next, we examined the research sectors of the published articles, trying to give an overview of the general categories. The findings are shown in Figure 3.4.

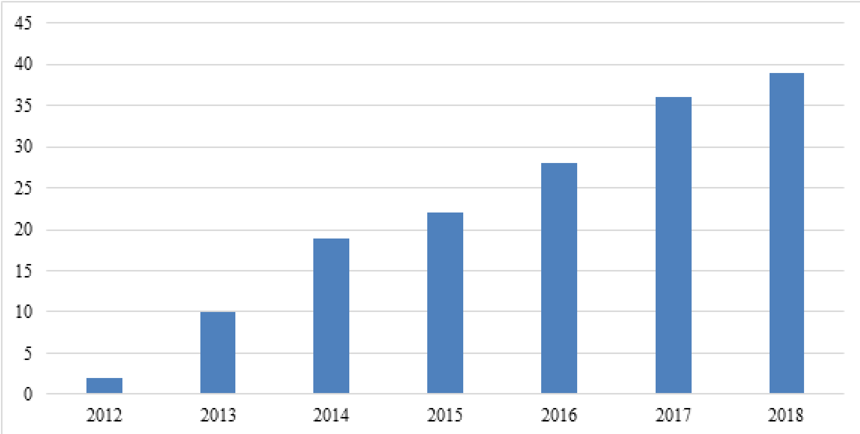

Figure 3.3: The publication frequency from 2012-2018

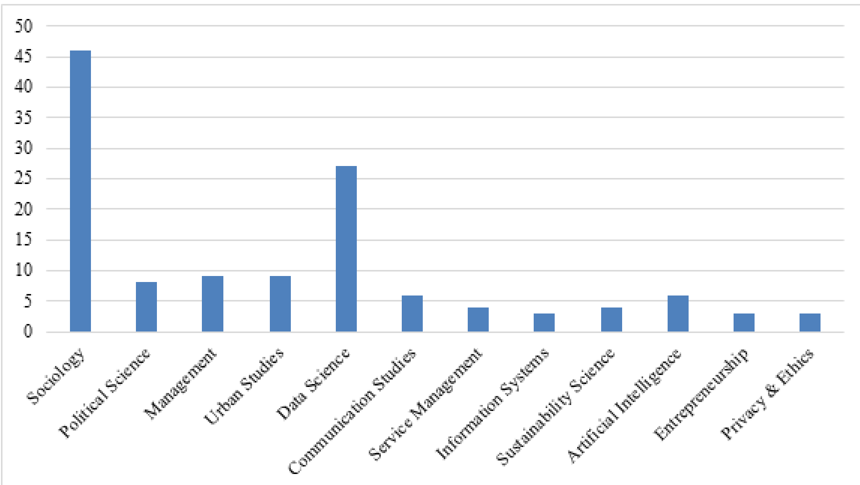

Figure 3.4: Distribution of articles by research areas

Publication Sources

This systematic mapping includes research papers published in academic outlets, such as journal articles and conference proceedings, as well as reports targeted at business executives and a broader audience, such as scientific magazines. We have kept reports (e.g., Hitachi reviews) because a lot of evidence is published by companies and a lot of work on social innovation and big data is done by companies as well.

We have tried to figure out how many of the relevant scientific papers are published in journals, how many as conference papers and from other sources. The statistic is given in

Figure 3.5.

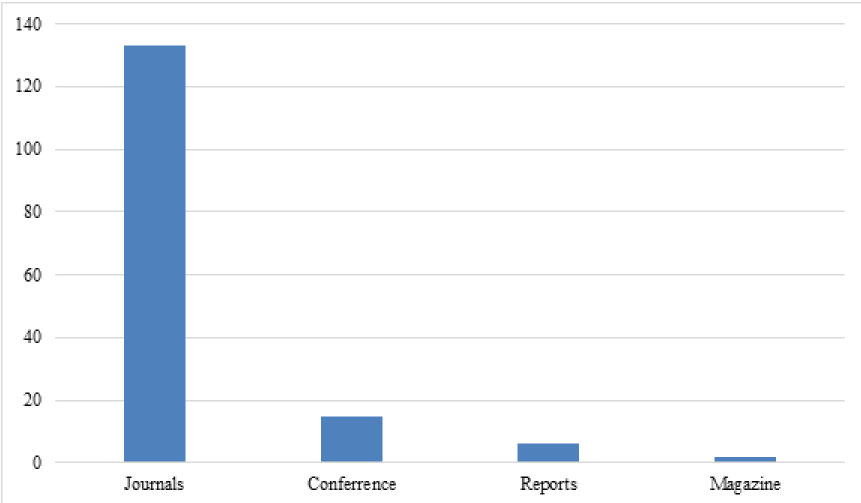

Figure 3.5: Publication sources of the articles

Along with the sources of the relevant papers, we have also searched for the top journals who published most papers about our research topic i.e. big data and social innovation. Here in Figure 3.6, we present the top journals according to the number of published papers from our review.

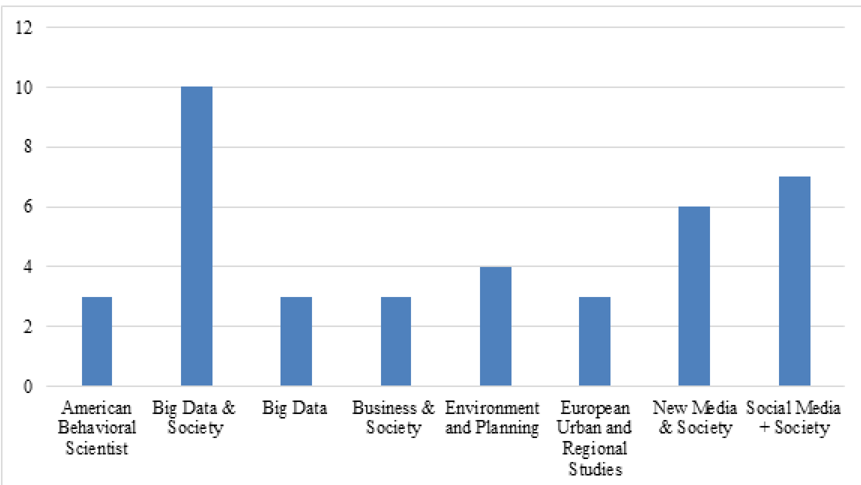

Figure 3.6: Distribution of articles per journals

3.2.2 RQ2: How much of the research is done based on empirical studies and what type of empirical studies?

Empirical Evidence
We primarily classified our reviewed papers as empirical and non-empirical papers. Non-empirical papers are conceptual papers. From the study, we see that majority (59%) of the papers are based on empirical evidence. With this finding (Figure 3.7), we also get the answer to our second research question of this mapping study.

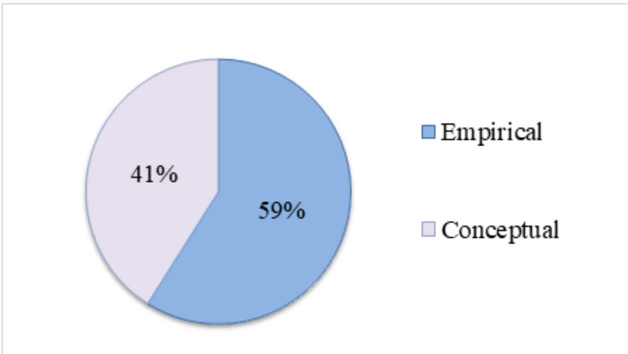

Figure 3.7: Empirical and non-empirical ratio

We then classified the empirical papers based on the type of study. The research types that have been assessed followed the guidelines from Oates (2005) include (1) survey, (2) design and creation, (3) experiment, (4) case study, (5) action research, and (6) ethnography. We have also included 'Discussion' as a research type, inspired by Zannier et al. (2006). We have added this last method as we felt that some papers are more suitable to categorize as a discussion paper. Discussion papers are also known as 'Expert opinion'.

After deciding about the research types, we counted the numbers for each type. The following figure shows which research types of the studies found from our map-ping. Only the papers providing empirical evidence (92 papers) were included in Figure 3.8, covering a total of 7 research methods.

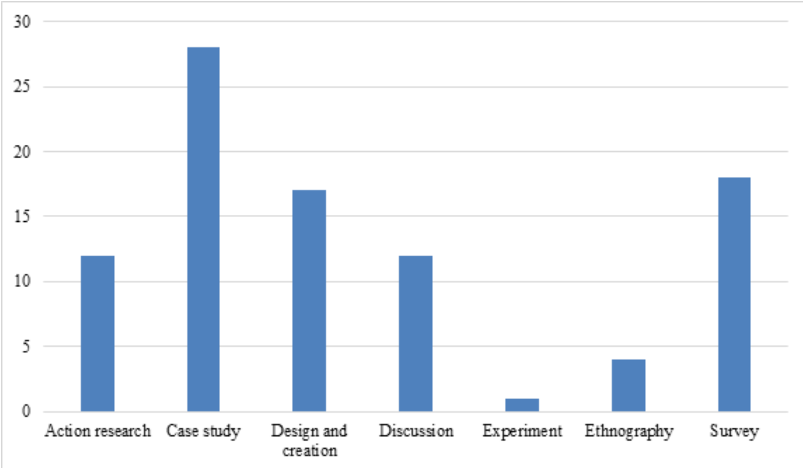

Figure 3.8: Type of empirical research

Contribution Type

Every research paper has some contribution to the advancement of research in the relevant field by providing something new. To illustrate which types of contributions that have been made within the research area between, Figure 3.9 was made. The figure shows the contribution type of papers. All 156 primary papers selected finally in our mapping study are considered in this figure.

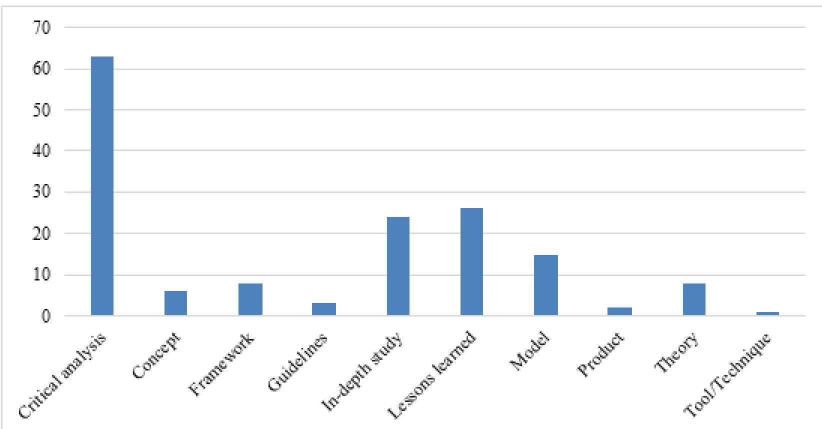

Figure 3.9: Type of research contribution

We differ between 10 contribution types. Based on Oates (2005), we define six different knowledge outcomes including (1) product, (2) theory, (3) tool/technique, (4) model, (5) in-depth study and (6) critical analysis. We also adapt some more knowledge outcomes or contribution types since some contribution types from Mathiassen et al. (2012) can describe the contribution of some papers more precisely; including (1) framework, (2) lessons learned, (3) tool/guidelines and (4) concept.

3.2.3 RQ3: What are the challenges or barriers to the successful implementation of big data for societal challenges?

Studying the title and abstract of all 156 papers, it has been found that only 3 papers mentioned challenges regarding employing big data in their studies. The challenges we find from this study are mentioned below:

1. Open data and privacy concern. (Taylor et al. (2014))
2. Challenge around obtaining data. (Ferreri and Sanyal (2018))
3. The prominence of marketing-driven software. (Dencik et al. (2017))
4. The interpretation of unpredictability. (Dencik et al. (2017))

So little evidence is not enough to generalize a fact for all and answer a research question like what the challenges or barriers for the successful implementation of big data for societal challenges are. So, to get a better understanding about specific challenges related to employing big data for societal change, we searched into literature outside our mapping, as we did not get enough references and evidence from our mapping. The focus of this literature search was only to find out the overall challenges of employing big data. We studied research papers from multiple fields including but not limited to business, healthcare, e-commerce, ICT, development etc. After this search, we have found more about the challenges related to big data when an organization wants to employ it. All these challenges are mentioned below:

1. Lack of understanding of how to leverage data analytics for business value (LaValle et al. (2011); Sivarajah et al. (2017))
2. Lack of management bandwidth due to competing priorities (LaValle et al. (2011))
3. Lack of skills ans staffs within line of business (LaValle et al. (2011); Katal et al. (2013); Wamba et al. (2015); Hilbert (2016))
4. Accessibility of data is difficult (LaValle et al. (2011); Sivarajah et al. (2017))
5. Perceived costs outweigh projected benefits (LaValle et al. (2011))
6. Existing culture does not encourage sharing of information (Katal et al. (2013); Wu et al. (2014))
7. Big data needs huge data storage capacity (Katal et al. (2013); Chen and Chun-Yang (2014))
8. Maintaining data privacy and security is challenging (Wu et al. (2014); Sivarajah et al. (2017); Hilbert (2016))

9. Analytical challenges with big unstructured data (Chen et al. (2012); Katal et al. (2013))
10. Scalability problem with big data (Chen and Chun-Yang (2014); Katal et al. (2013))
11. Current database software lacks in-database analytics (Russom (2011))
12. The difficulty of architecting big data analytic system (Russom (2011))
13. Quality of data (Katal et al. (2013); Wamba et al. (2015))
14. Ownership of data is not clear, or governance is ineffective (LaValle et al. (2011))
15. Proper interpretation of data (Dencik et al. (2017))
16. Lack of sponsorship (LaValle et al. (2011); Russom (2011))
17. Lack of compelling business case (Russom (2011))
18. No need for change (companies' perspective) (LaValle et al. (2011))
19. Don't know where and how to start with big data (LaValle et al. (2011))
20. Cannot make big data usable for end users (Russom (2011))

After searching for challenges from other fields of study we can see that the challenges we found from our mapping are similar to the challenges experienced in other fields of research. This finding inspired us to research further if all the challenges found from other fields are also applicable for big data analytics adoption when using for social causes; which we have addressed in this thesis.

3.3 Discussion

3.3.1 Findings from the Mapping

RQ1: How the research about 'big data and social innovation' has changed over time (in the last decade)?
Our study proves that terms like big data and social innovation gained the attention of academic and business communities later than in 2010. It can be also seen that the number of researches and publications are increasing every year since then, which proves the importance and increasing attention big data and social innovation is getting day by day. The study of Wamba et al. (2015) also verifies this finding where the authors stated that, "With regard to the literature review, 'big data' relevant journal articles have started appearing frequently in 2011. Prior to these years, the number of publications on the topic was very low. Publications on 'big data' related topics started only in 2008 (with 1 article) and then a steady increase in the number of publications in the following years."

From this mapping, we can see that many fields like social science, political science, management, service management, Sustainability science, information systems, urban management, communication, health care sector adapted big data for their applications. In the results section, we have presented the topmost fields, but other than these, there are also research fields we have found form the mapping like- education, ICT, journalism, tourism, etc. Here notable that all these papers with applications of big data in different fields are directly or indirectly related to various social issues.

RQ2: How much of the research is done based on empirical studies and what type of empirical studies?
In our systematic mapping, more than half of the papers (59%) provide empirical evidence. As there was no previous mapping on this topic, we cannot say how much empirical work was done before. But when 59% of the studies are empirical it proves that the researchers of this field are contributing much. With their contributions, the quality of research is also improving. The major contribution of the research papers from our mapping was a critical analysis, both empirical and non-empirical. When analyzing different topics, the authors also presented their insights, research agenda, guidelines for future research, what lessons they learned and their opinions. The empirical studies also presented models, frameworks and tools that can be used in future research.

RQ3: What are the challenges or barriers to the successful implementation of big data for societal challenges?
Among the few articles we found from our systematic mapping regarding big data challenges, the authors of Taylor et al. (2014) reflected on various cases related to big data challenges, including the challenge of maintaining data privacy and ethics when using all forms of big data for positive social change. The authors recommended exploring new formats for educating people about privacy/data protection risks to overcome data privacy challenges and to use templates to evaluate open data sources. Ferreri and Sanyal (2018) examines how the challenges around obtaining data to enforce new regulations are addressed by local councils to balance corporate interests with the public good . The authors stated that triangulating different sources of information is not always straightforward as the publicly available data might be partially obscured. In their case study, the authors recommend about platform economy to overcome the challenges regarding data collection. Dencik et al. (2017) focuses on the dominance of marketing-driven commercial tools for predictive analytics of data and their effectiveness to analyze data for completely different purposes such as law enforcement. Another challenge that Dencik et al. (2017) mentioned is, the notions of predictability and probability remain contentious in the use of social media big data. The authors reflected upon the challenges and points to a crucial research agenda in an increasingly datafied environment.

We have not found any study that addresses big data challenges when specifically applied for social problems or from the perspective of the society. We believe that there is a scope and necessity of future research regarding this; if there is any social challenge for adapting big data; and if there is, then how can we address that. So, we developed our plan for this thesis to focus on this research agenda.

3.3.2 Use of Keywords

As we have mentioned earlier (section 3.1.6), we found some research papers relevant to our study, but they have not been included in the mapping as they do not use the keywords we searched with. For example, Wesolowski et al. (2015) uses mobile call data to predict the geographic spread and timing of epidemics, and indeed they address a social challenge and has a significant societal impact. However, they do not use keywords regarding data analytics and societal impact, maybe because their focus is mainly on modeling and technical aspects. Instead their keywords include human mobility, mobile phones, epidemiology, dengue etc. Considering the importance of social implications of big data research as well as the inter-est of publication venues in contributing to societies (Pappas et al. (2018)), we suggest that future papers should take into account and report such implications in their abstract and keywords. We should note that indeed many papers discuss social implications, however they do not mention them in their abstracts, raising the need for a systematic literature review in the area. Thus, a more detailed analysis of the research articles can lead, among other things, to new combinations of keywords that will be able to better capture the current status regarding the impact of big data and analytics on societal challenges.

3.3.3 Limitation of the Study

A limitation of this study is, we have used only titles and abstracts to extract data. So, the categorizing and data extraction process depends on the quality of the abstracts. ICT-related research publications often do not use structured abstract (Kitchenham et al. (2008)) which results in poor accuracy when classifying papers based solely on abstracts. Also, following the standard procedure of systematic mapping, we could not include research papers in our study which do not have the keywords we searched with; even though some papers might be relevant to our topic.

3.4 Conclusion

This chapter presents findings of a systematic mapping study that researchers, social innovators, social entrepreneurs and all other stakeholders can use to unlock the power of big data for the benefit of the society. We have presented the current status that shows how research into big data and social innovation has increased over the last decade, attracting significant attention from across a wide array of disciplines. We have identified the major research areas where big data is getting significant attention; so future researchers can explore more about the impact of big data in those areas. This mapping study also proves that the empirical ground of research in this field is strong; research is not only limited to case studies, but also other forms of research is being done like action research, critical analysis, designing and creating new products, etc. The key contribution this mapping has made is offering the basis for a reflection process among the researchers in this field.

Chapter 4

Research Method

Data are an important element of social innovation. To initiate any innovative steps or to address any social challenge, data are needed. But comparing to technological and business perspective, using data or data analytics from a social perspective is not common yet. The objective of this study is to create a better understanding of big data analytics usage in the social sector. In particular, we will investigate how big data analytics has been used from a social perspective and what are the major challenges and drivers social entrepreneurs experience when they employ big data analytics in their work. We will finally examine the relationship between these challenges and drivers/enablers with different social entrepreneurial intentions.

This study is of exploratory nature as we seek to create knowledge by investigating the events and actions of those who experience them (Oates (2005)). For the investigation, we have considered multiple data generation methods and finally to answer the research questions we have used a both qualitative and quantitative method in this study. We have considered semi-structured interviews, as it gives a detailed understanding of the participant's work and experience and interviewees can also express themselves more freely and share their own perspectives on personal experiences related to the research topic. Then we have also conducted a survey to quantify and validate the findings from the interviews and also the finding from our mapping study which was done prior to this thesis. Considering the time constraints, the number of survey participants was aimed at 300 and finally ended up with 49 responses. Other data generation method like - observations were considered not feasible for this study as it would be too time-consuming to undertake. An overview of the research process is shown in Figure 4.1.

The rest of this chapter describes the research methods and followed steps further, and proceeds as follows: Section 4.1 presents and justifies the research questions of this empirical research, Section 4.2 explains the qualitative method of performing interviews, Section 4.3 explains the quantitative approach of the questionnaire and finally, section 4.4 explains how we managed the ethical considerations.

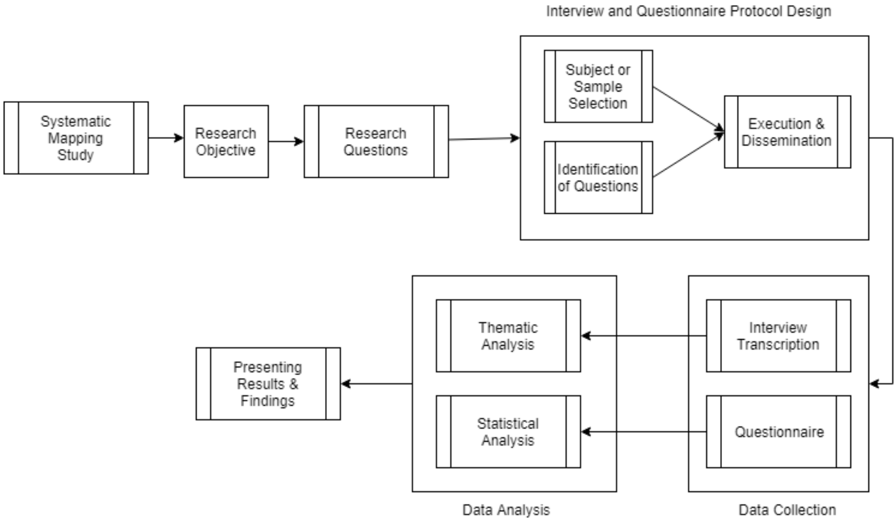

Figure 4.1: Overview of the research process

4.1 Research Questions

The aim of this research is to understand the role of big data analytics for social good. Big data analytics has a huge potential for many diverse sectors from many different perspectives. But in this research, we explore big data potential and contribution from a social perspective. We want to investigate how big data analytics is being used by social sector organizations, social entrepreneurs and innovators to address societal challenges and how big data can be used to generate social value.

After exploring the role and contribution of big data, we would explore the associated challenges, enablers and benefits of big data analytics. We want to identify the major contributions or benefits that big data analytics can offer to the organizations and entrepreneurs. From the systematic mapping study, we tried to find out the challenges of employing big data in social sector applications. We even looked into literature in addition to the mapping, to find out big data challenges from all sectors in general. From that literature search, we found a wide range of challenges of employing big data. In this research, we want to investigate if all those big data challenges which we found from different fields are also applicable when applied to innovation solutions for social good. Along with the challenges, we also focused on identifying big data drivers or enablers for successful solutions on societal challenges. We have found some drivers or enablers from literature that researchers find helpful for using big data and data analytics. We also want to investigate how much these drivers actually help social innovators and social entrepreneurs to employ big data analytics and how significant each of the drivers is to them.

After investigating all the challenges, benefits and drivers, we will finally try to know how these challenges and drivers impact the intentions of social innovators and social entrepreneurs to employ big data analytics in their work. We want to examine the relation of big data challenges and enablers with social entrepreneurial intentions.

So, finally from this research we aim to answer the following research questions:

RQ1: What is the role of big data analytics in helping social innovators and entrepreneurs to generate social value?

RQ2: What are the challenges and benefits of using big data analytics to address societal challenges?

RQ3: What is the relation between challenges and enablers of big data analytics with social entrepreneurial intentions?

4.2 Interview

Qualitative method is used to gain an understanding of the underlying reasons, opinions, and motivations. It can provide better insights into the problem and also helps to develop ideas or hypotheses for potential quantitative research. We have collected qualitative data using semi-structured individual interviews. An interview is a widely used data generation method in qualitative research and has been used extensively in multiple disciplines, including Information Systems (IS). We have done a total of 10 individual interviews and the duration of each interview was around 25-35 minutes. In the following sections, we describe the process of identifying interview questions, participant selection and detailed analysis procedure of the interview data.

4.2.1 Identification of Interview Questions

To identify the research questions, we have followed the Goal Question Metric (GQM) paradigm (Basili (1992)). Although the GQM approach was developed to define and evaluate goals for a project in a particular environment, its use has been expanded to a larger context (Basili (1992)). The author also stated that this paradigm can be used not just for management, engineering and quality assurance interests, but also for interpreting questions and the metrics. So, in this research we have used GQM from a generalized perspective to set our interview goals and to formulate our interview questions and the metrics. From the interviews, our goal was to answer our first two research questions (RQ1 and RQ2) mainly. As this part of our study is a qualitative study through semi-structured interviews, we have considered the interview questions as the metrics.

The overall objective of this study is to create a better understanding of how big data analytics is being used for social good strategies and what is the perception of social entrepreneurs and social innovators about using big data analytics to address societal challenges. We have defined our interview questions, so that we can achieve the objective of this research and can gather better insights about the problem. Since the interviews had an exploratory nature, the interview questions were not created to get a yes/no answer. The questions were designed to target the experiences that the interviewees had acquired through working in the social sector. Therefore, the interview guideline contains open questions like - about their missions, work practices, their experiences, challenges etc. The complete interview guide can be found in Appendix B.

In the following Figure 4.2, we present a mapping of the goals and our interview questions.

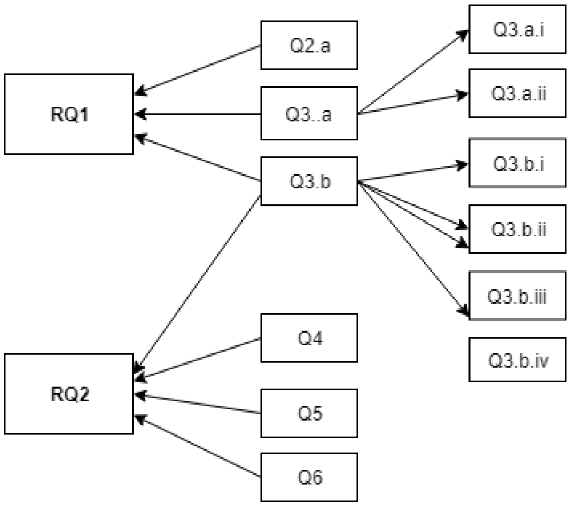

Figure 4.2: Mapping between research questions and the metrics

4.2.2 Selection of Participants

For the interviews, we have selected people who work in social sector organizations and have experiences of working with big data analytics in their organizations. We primarily considered the participants of two conferences which were focused on social innovation. These two conferences are the *5th annual Data on Purpose conference 2019*[1] organized by Stanford Social Innovation Review and *Nesta Education 2019: Shaping the Future, Shifting the System*[2] organized by NESTA, an innovation foundation. The focus of Data on Purpose 2019 was to help nonprofit leaders identify the best ways to build data and technology capacity and the focus of the NESTA event was to explore how the society can use data based evidence to make decisions about the future education system, considering the role of technology, including artificial intelligence, in solving some of the most pressing social issues, and looking at some of the most exciting education innovations. Other than these two events, we have used Linkedin to search for organizations who work

[1] http://www.ssirdata.org

[2] https://events.nesta.org.uk/nestaeducation2019?_ga=2.49983078.1783782477.1558538855-1907512832.1551172439

in the social sector and employ any form of data analytics in their work. After considering the organizations' work and relevance to our research, we contacted and invited one representative from each organizations through email and finally interviewed who agreed.

4.2.3 Analysis Procedure of Interview Data

For analyzing the interview data, we have applied Thematic analysis process; which is a codes-to-theory model for qualitative research (Cruzes and Dyba (2011)). Thematic analysis is defined as "a method for identifying, analyzing, and reporting patterns (themes) within data" (Braun and Clarke (2006)). In the interviews, we have tried to understand the works of the organization, how they are involved with social innovation, what is their working process, their experience etc. So, the interview data has given us a more detailed understanding of our research topic and also a background to perform the questionnaire later on this research. The objective of the thematic analysis process is to identify and understand the patterns or themes in the interview data about the work practices and experiences of the organizations and thus to answer our research questions.

The main steps of thematic analysis process are illustrated in figure 4.3 along with our finding from the analysis. In the following section we explain our data analysis procedure.

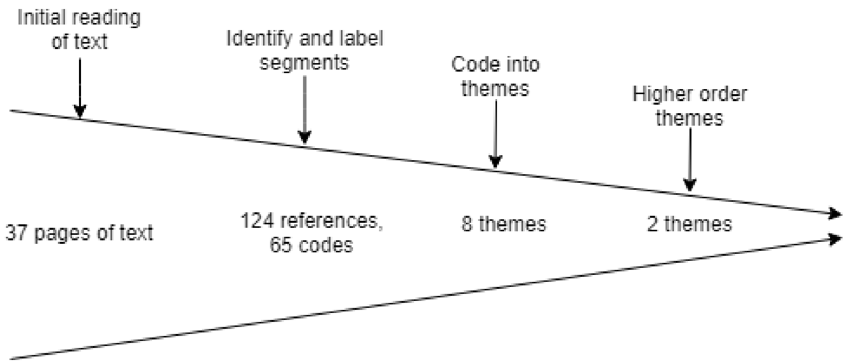

Figure 4.3: Thematic analysis process based on Cruzes and Dyba (2011)

Initial Reading We started the thematic analysis with reading the transcribed interview texts to generate initial ideas and to identify common patterns in the data. The interviews were transcribed shortly after they were taken, to ensure that the actual meaning of what interviewee answers were written correctly. After reading each transcribed interview text, initially some highlighted points were identified which helped in the further coding process.

We followed the GQM method to identify the interview questions, so it helped to ensure that we asked interviewees the questions that will help us to best address our research questions. This facilitated us an early and efficient analysis, as we could easily connect respondents' answers directly to our research objective.

Coding Process To codify is to arrange things in a systematic order, to make something a part of a system or classification, or to categorize (Saldana (2015)). In this research we have used descriptive coding. In descriptive coding, we summarize the basic topics of the data in a word or in short phrases to identify interesting topics, important findings from the data. According to Saldana (2015), descriptive coding techniques is appropriate for beginning qualitative researchers learning how to code data.

The coding process is an integrated approach and is a mix of inductive and deductive approaches. In the inductive coding process, data is reviewed line by line and a code is assigned when a concept appears. In deductive coding, we start with a list of codes based on theories or other key concepts in which we categorize data. For our research, we followed inductive approach. When analyzing the data, we had some clear ideas and thoughts of what we expect to find from the data but we did not have any predefined list of codes; we created the codes as concepts appeared in the data. This approach allowed us to avoid limiting the codes within our expectations and also prevented us from coding data out of context; we identified what the text was saying rather than what we wanted to see.

To make the coding process and the overall thematic analysis process organized and efficient, we have used NVivo[3]. NVivo is a qualitative data analysis computer software package which gives a place to organize, store and retrieve data so we can work more efficiently, save time and rigorously back up findings with evidence. This tool is very helpful for new researchers as it allows for a better understanding and exploration of unstructured data by facilitating the quick discovery of key topics and themes. Coding with NVivo initially resulted in 112 codes with 124 references. After getting the initial codes, we reviewed the codes again removed duplicates and merged similar codes. So finally this resulted in 65 codes. NVivo facilitated independent coding of each interview transcript, while at the same time allowed us to classify data from each interview into similar codes from others.

Translate Codes into Themes After finalizing the codes, we have reviewed all the codes and categorized them based on themes. A theme can be seen as a way of grouping initial codes into a smaller number of sets, to create a meaningful whole of unstructured codes (Cruzes and Dyba (2011)). We also reviewed the themes to ensure that the themes are not overlapping and are in line with the research context. Having some clear ideas and thoughts of what we expect to find from the data also helped us to define the themes. After the initial theming process, we got a total of 8 themes. We categorized the themes again and divided them into 2 higher order themes.

4.3 Questionnaire

Questionnaires are usually designed so that answers to questions can be scored and based on the scores we can obtain an overall measure of the attitudes and opinions of the respondents. The goal of our questionnaire was to validate and quantify some of the findings of this research. In the questionnaire we have asked the participants questions about big

[3]https://www.qsrinternational.com/nvivo/what-is-nvivo

data challenges, drivers and their influence; as the aim of this research is to find out the challenges and drivers of big data in the social sector, their significance; and to investigate their influence on social entrepreneurs' intention.

4.3.1 Identification of Questions

Designing a questionnaire plays an important role in the questionnaire performance. This is because the information gained from a questionnaire is proportional to the quality of the questionnaire which leads back to the design (Peterson (2000)). We have identified and designed questions so that the answers can directly help us to measure and validate the findings of this research and finally can lead us to answer our research questions. Using the questionnaire, we mainly answered our RQ3 and validated other research findings. For the questionnaire, we have designed the main questions based on the 11-items social entrepreneurship scales and social entrepreneurial intention scale as discussed in chapter 2; along with the big data challenges and enablers found from the literature search of this research.

Keeping the design concern in mind, we have kept the questions brief and specific. We have gathered both factual and opinion based data. The factual data was gathered to get some general demographic data to better understand the participants, and the opinion data was gathered to measure and validate the findings. The factual data was designed to be nominal data, and the opinion data was designed to be ordinal. Except the demographic questions, all other questions used closed Likert scale data. The questionnaire can be found in Appendix C.

Before disseminating the questionnaire, it was discussed with the supervisor. The questionnaire was disseminated through the internet, by utilizing collaborative online tools (like email and Linkedin) that were easy to reach to the right population. Due to time limitation, no pilot test was done with the questionnaire.

4.3.2 Sample

For selecting the sample of this questionnaire, we have applied the same criteria as our interview. We invited people to participate in the questionnaire who work in social sector organizations that use some sort of big data analytics in their works. We have primarily used the same platforms i.e. the two conference events (5th annual Data on Purpose conference 2019 and Nesta Education 2019: Shaping the Future2, Shifting the System) for finding out the organization who work in this sector. For the interviews we invited specific individuals but for the questionnaire we invited all members or members of specific teams of that organization. There was no geographical or ethnographic consideration regarding the choice of the sample. Our sample comprises both male and female, aged between 25 - 54+. We aimed and invited around 300 individuals, finally 49 of whom responded. In the next chapter we will present a more detailed overview of our sample.

4.3.3 Constructs

In the questionnaire, we have used four constructs. The first construct is *Social Entrepreneurship Scale* by Carraher (2013); which is used to better understand the status and mission of the participants. The next construct is the *Big Data Challenges*; with this instruments we have asked the participants to scale of the challenges that we found from the literature review according to their experience and opinion. The third construct was the *Big Data Enablers*. Similar as the second instrument, we have asked the participants to scale the enables that we found from the literature. Finally we have used the *Social Entrepreneurial Intentions* by Douglas and Prentice (2019) as our fourth construct. With this construct, we have examined that how likely our questionnaire participants will use big data analytics to achieve their mission.

4.3.4 Analysis of Questionnaire

To analyze the questionnaire data, some quantitative analysis has been done. The nominal data of the questionnaire will be presented with simple tables to provide an overview in general. For the ordinal data, we will initially perform *Factor analysis*, which is a statistical data reduction and analysis technique. We have used factor analysis to reduce our large number of variables into fewer numbers of factors; these factors will be used to explain correlations among multiple outcomes as the result of the underlying factors. Using factor analysis, we will categorize the variables of our constructs i.e. social entrepreneurship scales, big data challenges, big data enablers, and social entrepreneurial intentions. After factor analysis, we will present the correlation of all the factors and later we will perform ANOVA (analysis of variance). With ANOVA, we will examine the influence of big data challenges and enables on different social entrepreneurial intentions.

The questionnaire was conducted electronically, so techniques like eyeballing, spot checking were not required to verify the results. These methods are usually used to verify if the data is valid and if any mistakes were made during the transition from paper to a digital medium (Oates (2005)).

4.4 Ethics and Intellectual Property Rights

Ethical issues are important to consider when planning and performing empirical research (Oates (2005); Runeson and Host (2009); Singer and Vinson (2002)). Norwegian Centre for Research Data (NSD) is a resource centre, which assists researchers with regard to data gathering, data analysis, and issues of methodology, privacy and research ethics. The main objective of NSD is to improve possibilities and working conditions for empirical research that is primarily dependent on the access to data. NSD is the data protection official for NTNU also, so all NTNU students and researchers are obligated to notify NSD about their project if they are going to process personal data.

In this research, all ethical and privacy concerns with our research data have been considered. All interview participants were asked to explicitly agree to participate through oral consent before conducting the interviews. As in this research project we will not handle any personal information related to the participants, and as we will only register anonymous information (from both interview and questionnaire), meaning that the data contains no information that can enable the information to be traced back to an individual, this project has not been notified to NSD. None of the organizations or interviewees and questionnaire respondents are registered or presented by name, or other personal information that may be used to identify them.

Moreover, this master thesis is connected to EU-H2020 project INITIATE[4]. The purpose of project INITIATE is to institutionalize innovation through big data and social entrepreneurship. Project INITIATE is already registered in NSD, so this was another reason that this research has not been registered in NSD again. This thesis followed the same ethical and privacy standards as project INITIATE.

[4]http://initiate2020.eu

Chapter 5

Results

In the previous chapter, we have explained the methods of performing the interviews and questionnaire. So, in this chapter we present the results of them. In section 5.1 we present the interview results and in section 5.2 we present the results of the questionnaire.

5.1 The Interviews

The interviews were performed as explained in section 4.2 and here we are presenting the results will in a qualitative manner. First, we present an overview of the the participants of the interviews to gain a better insight of who the participants are. Afterward, the quantitative results of the interviews will be presented.

5.1.1 The participants

Total 10 interviews were performed with 10 different people. All the participants are involved in the social sector specifically with social innovation and/or social entrepreneurship. In the following table 5.1, we present an summary of the participants. In this research we have treated all data anonymously, so we also present the data in an anonymous manner. In the table we present the participants' role in the organization and an summary of the organizations missions and working area.

Table 5.1: Overview of the interview participants

Organization	Participant's Role	Organization's Description
Organization 1 (O1)	Director	Facilitates diverse organizations to tackle complex data and technology challenges, and to build technologies in a way that are socially beneficial and also responsible. The firm also works with the private sector to build decision-making infrastructure for socially beneficial deployment of emerging technology.
Organization 2 (O2)	Co-founder	Brings cutting-edge practices in data science and crowd-sourcing to some of the world's biggest social challenges and the organizations taking them on. It hosts online challenges, where a global community of data scientists competes to come up with the best statistical model for difficult predictive problems that make a difference to the society.
Organization 3 (O3)	Director	This is a policy research organization. The interviewee is the director of a coalition of organizations of cities that collect and organize neighborhood level data. The local mission of the organization is to assemble and organize local data and then provide direct technical assistance and insights around that data also finally to build community level capacity to use data for social change in a broader level.
Organization 4 (O4)	Researcher & Senior Data Scientist	The organization is a telecommunication company. The interviewee works at the research unit of the company and leads a project that aims to utilize telecom data for solving societal problems; for example- quantifying and analyzing human mobility patterns to map the spread and timing of epidemic.
Organization 5 (O5)	Researcher & Senior Associate	The organization is a research and consultancy company with a focus on social impact and social innovation. The organization conduct studies considering new models for digital social innovation, entrepreneurship, boundary pushing models for sustainable and impact financing amongst other topics.

Organization 6 (O6)	Co-founder & CEO	A public benefit corporation that works on building data collaboratives. These collaboratives, or data trusts, provide the legal, technical, and governance framework that empowers a collective of organizations to share and integrate their respective data sources in order to respond to current and future data demands. The mission is to unlock the human potential with data, and how individuals can act more potentially for the society.
Organization 7 (O7)	Vice President	The organization funds social entrepreneurs and social innovators. It invests in strategies that address some of the society's biggest challenges like to ensure everyone has the food they need to lead an active, healthy life, nurture early brain development, ensure students are equipped to complete the higher education of their choice and unlock an array of future careers, and transform end of life care.
Organization 8 (O8)	Head of Program	This foundation is a global non-profit organization focused on realizing open data's value to society by helping civil society groups access and use data to take action on social problems.
Organization 9 (O9)	Data Scientist	The organization's vision is to use data in the service of humanity. It supports charities and social enterprises large and small across a variety of issue areas including enabling them to do their jobs easier, to use their resources more efficiently and to gain better insight into the work that they do.
Organization 10 (O10)	CEO	Aims to help society and other organizations with new innovative solutions and technologies; Specially focuses on using social data.

5.1.2 Interview results and findings

As we have done a "Thematic analysis" of the interview data, our focus was to identify different themes and patterns from the information that our interviewees provided us about their work practices and experiences. Primarily we have identified 8 themes. Then we classified these 8 themes into 2 higher order themes. An overview of the themes can be seen in Table 5.2 and a more detailed breakdown of the themes along with associated codes of those themes is shown in Figure 5.1. We present a discussion of the findings for each of the sub-themes below.

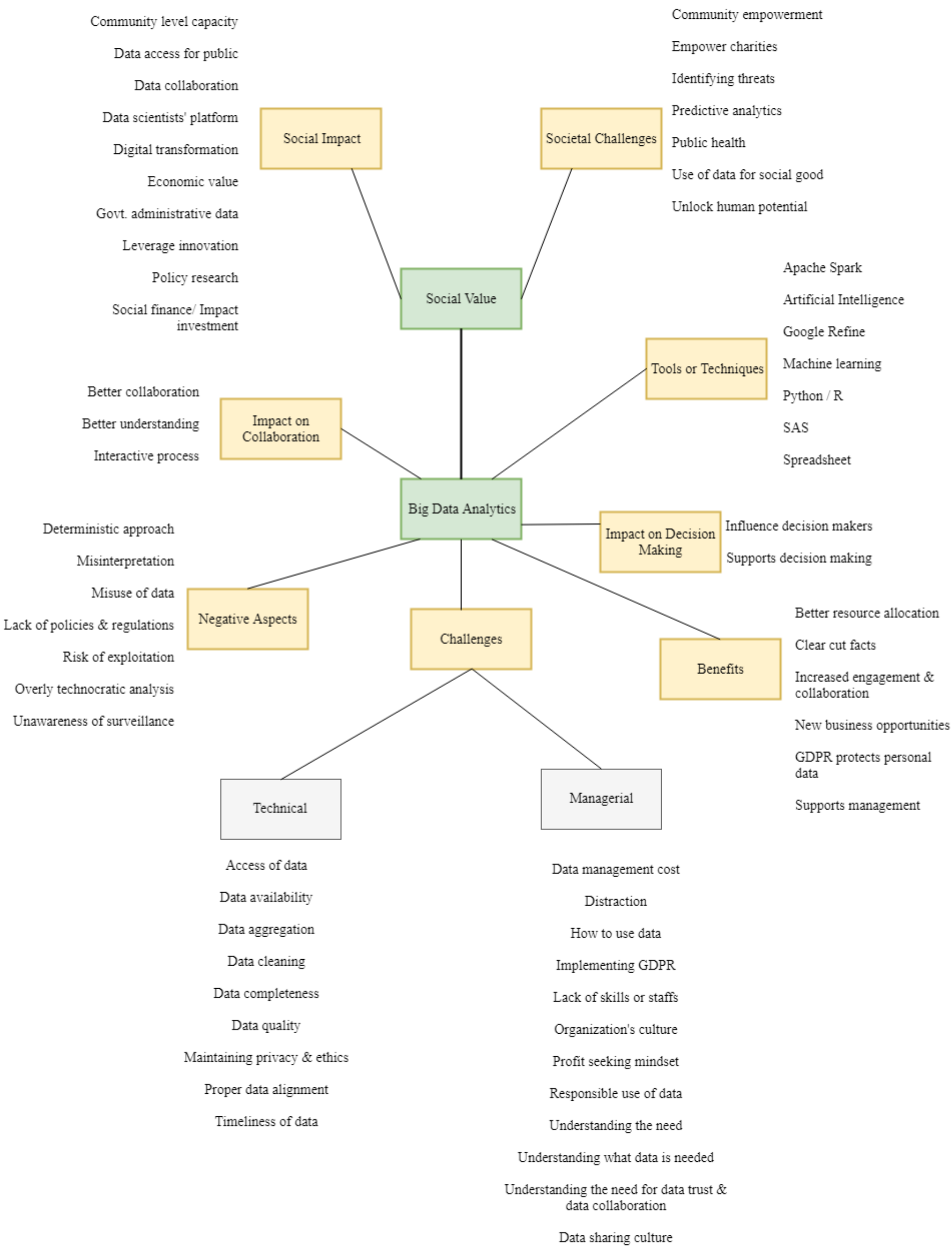

Figure 5.1: A detailed overview of the themes and associated codes

Table 5.2: Overview of the themes

Theme	Codes	References
Social impact	10	19
Societal challenges	7	12
Big data challenges	21	42
Big data benefits	6	12
Big data negative aspects	7	12
Big data tools	9	15
Impact of big data on collaboration	3	7
Impact on big data on decision making	2	5

Social impact

Many organizations in the social sector work with data to create or enhance a positive social impact on the society with their activities. There are different forms or ways of creating social impact with big data analytics. During the interviews, we asked our participants about their organization, background, how they are generating social value etc. to better understand them and their work. Here we present some examples that we have learned from our interviews about different ways of creating social impact using bid data and data analytics. For example, one of our interviewee mentioned:

> O3 - *"We collect and organize neighborhood level data for better decision making with a particular focus on building community level capacity to use data for social good in a broader level."*

Our another interviewee who is a Co-founder of the organizations said that, in 2013/2014 he and his team mates wanted to do an independent study with some data sets which had some social impact relationships; but he could not find similar kind of examples or sources. Because that time data science for social impact was a much less known quantity. With this experience they were motivated to do something about it. He mentioned that,

> O2 - *"We decided to built a platform where lots of people can work on social impact problems by having the data from non-profits or NGO; then having it, packaging it up in a competition format, and putting the competition out there for anybody to work on. "*

Another example of creating social impact with data is to make it available for people, so that they can use it to address any social problem or social good. Like our another interviewee mentioned he works in an organization that strive for a world that is more open, where open data and open knowledge can be used for social good.

> O8 - *"Our organization is a global non-profit organization focused on realizing open data's value to society by helping civil society groups access and use data to take action on social problems."*

There are organizations also who funds social entrepreneurs and social innovators. They invest in strategies that will address some of the societies biggest challenges ans these strategies also include using big data analytics. Like our interviewee said:

> O7 - *"We generate social value by investing in strategies to ensure that everyone has the food they need to lead an active, healthy life, nurture early brain development, ensure students are equipped to complete the higher education of their choice and unlock an array of future careers, and transform end of life care."*

They fund strategies to better utilize data, to use predictive analytics for socio-economic development. One particular example he mentioned is:

> O7 - *"Funding universities to use data for predictive analytics; to identify students who need help, to ensure that unprivileged students enroll, persist through, and graduate with the skills they need to find jobs and pursue their career goals."*

There are many examples of using big data analytics to create positive social impact and for social good. So, with all these examples we made up this theme.

Societal challenges

Along with creating social impact big data and data analytics can be also used to address social challenges, to solve social problems. Some of our interviewees also belong to organizations who use big data analytics to solve various social various problems. Here we can quote some examples of what our interviewees mentioned. The first example can be, using big data and data analytics to predict spread of epidemics, so that proper prevention strategies and measures can be taken ahead of the situation occurs. One of our interviewee mentioned his work as:

> O4 - *"Understanding how we can forecast the spread of epidemic diseases, that's a problem of the society and the ministry of health would like to understand and solve the problem".*

Another interviewee explains her work as a consultant, like this:

> O9 - *"Supporting all types of organizations that use technology strategically in what they do; which is partly about working with civil society who maybe trying to solve very complex problems and data driven technologies might be helpful for them. But it is also thinking about they can do that in a way that are socially beneficial, mitigate harm and also responsible."*

There are many different ways of facilitating the society to solve social problems with the help of data. But to use data for social purpose, data availability is an important issue. One of our interviewee works in an organization that recognizes this issue and is trying to make data available for public who want to use it for any social cause. He mentioned that:

> O8 - *"We are primarily supporting other civil society organizations to be more open. We also work with government. We develop software tools which can be used to publish data. And govt. around the world are using our software to make data available to anyone who wants to use it, so the data can be used for the common good."*

Data are also used to find out threats or to build predictive models to better address problems that may occur in future. Our one interviewee mentioned that their mission is:

> O6 - *"to unlock the human potential with data, how individuals can act more potentially for the society."*

And they are doing this in different ways. Some examples he mentioned includes:

> O6 - *"Chicago is a city with a large inventory of aging homes that still contain lead paint and pipes. So old building in Chicago have exposure to lead and children can be exposed to this threat. We use data to find the building with threat and make people aware of the threat."*

> O6 - *"In Africa, we used predictive modeling to find out which children may need immediate assistance from the teachers, we used these model to better allocate resources. With data they predicted student dropout rate, identified students who need more assistance etc."*

These are some examples of using data to address societal challenges which helped us to better understand how big data is being used today to solve social problems and also how it can be used in future.

Big data benefits

The benefits of using big data have a long several dimensions. It can be the corporate social responsibility dimension, it can be long term benefits of exploring new venues for possibly earning new revenue for a company. Here we mention the benefits our participants talked about.

> O1 - *"It helps alignment and digital transformation process in the company to increase collaboration, increase engagement of more people and then also changes the way they learn and think about the problems they are tackling."*

> O3 - *"The promise of our whole network is community empowerment; and improvement can happen better with the data. It can change the pro-section relationship and get everyone on the same set of knowledge and that helps to move into the important conversation far ahead of time and to understand what actions need to be taken."*

Another important benefit that many of our interviewees mentioned is, looking for new opportunities by analyzing data. Like our interviewees said:

> O4 - *"We are always looking towards that how we can possibly explore new revenue potential. So its about learning, about how we can use the data, how useful it is and in what setting are the data valuable. So all these learning are important for a company that is exploring the data driven side of business."*

> O5 - *"You can find opportunities that you would not expect in some cases, by using and analyzing data. It can open up new business opportunities and also social business; you can realize what is needed from a social point of view."*

> O10 - *"Organizations can get new business models from analyzing data, you can find what your customers, employees, people are saying. You can make new strategies based on that for the future and for next years. You can also predict the future."*

Analyzing data helps organizations to know and understand facts with evidence. As our interviewees said:

> O5 - *"The main benefit is to have the ability to show clear cut facts to support your theory of change. So when you go to your clients you can say "look, we have a theory in our mind and if we do X then Y will follow". So using data is way to put clear cut, straight and undeniable truth behind your functions, behind your theories of change and behind the whole attempt of changing the mindset."*

> O2 - *"Well, if you not using data then you are just guessing. And I think if you have data then you can try to understand what it can tell you."*

Big data challenges
One of our main goal from this research was finding out the challenges that social entrepreneurs and innovators face when they employ big data analytics. During the mapping study (prior to this thesis), we found some challenges from literature. So, this time we wanted to see if the challenges our interviews face in the organizations are similar to all those challenges we found from literature. And yes, after analyzing the interviews we have seen that the challenges our interviewees face who work in the social sector are similar those that we found in literature. We have found many different challenges from the interviews, so to better understand the challenges and to present more clearly we have categorized them into two categories as presented in Figure 5.1.

Now we present some of our interviewees thoughts and comments about the challenges. The most important and common challenge we have found is the privacy and ethical concern of data; and all of our interviewees mentioned about this during the conversation. One of our interviewee said:

> O4 - *"I guess there are other challenges too, but the hardest thing is to sort out the privacy and if we are allowed to use the data."*

> O6 - *"Knowing the ethics of data use and using the data ethically and legally is the most important challenge."*

This privacy concern also triggers other challenges. The GDPR has been implemented to protect the privacy of people. But, implementing this GDPR has become a challenge especially for small or bottom up organizations. On this issue one of our interviewee mentioned that:

> O5 - *"The challenges rise from protecting the privacy of whoever produce the data. Though it has been a year that GDPR was approved but I would say the bottom up organizations are still struggling to do with that. So there is still a big potential that is still untapped at the moment because people are still understanding how to use data properly. That is affecting a lot to the grass root organizations. Corporations*

already know how to do that; but the social organizations and social enterprises are still struggling a bit with that."

Other than privacy and ethical concerns, understanding the potential of data is also an important challenge. On this issue, one of our interviewee said that:

> O9 - *"I would say the biggest challenge for the charity sector is making a persuasive case about why they should spend their time doing data analysis and analytics."*

Another interviewee also mentioned the similar kind of challenges, she said that:

> O3 - *"I think the biggest challenge is how the decision makers use the data. You can make the best report using fancy analysis and may be also fancy visualization of data but if the decision makers don't use the data in decision making than it can turn as a trash."*

She also added that:

> O3 - *"Knowing how to connect the data with our questions, for example how people can use administrative data is really challenging and I think we are still far away in doing that now."*

Another important challenge we have found is about the data sharing culture.

> O3 - *"I think our data sharing culture is also challenging; it is still in developing the norms and practices around that."*

> O6 - *"Making people understand why we need data collaboration is challenging."*

There are also many challenges related to technical issues like: data quality, data cleaning, timeliness of data etc. Some comments of the interviewees related the technical issues are as follows:

> O1 - *"Data is difficult. Different kind of challenges occur in different projects. Like - cleaning is a issue, having a data set that you can use, statistical significance of data, the confidence level you need to have to claim anything from the data."*

> O2 - *"Sometimes it is messy, sometimes it is hard to understand that how the values come into the system. Data cleaning is a large part of the problem."*

> O8 - *"All the data published are not of sufficient qualities, they have a lot of errors. And there is risk of misinterpretation when technical language is used. We are trying to address this kind of problems and we try to make our data easier for interpretation."*

He also added that:

> O8 - *"The timeliness of data is also a issue, if the data published very late then it is not up to date and the accuracy of data is not always there as I said. The completeness of data is also always not there, so these are some problems that are hard to resolve."*

Some organizations also face challenges regarding the cost and expenses of developing and maintaining data structures. Specially for small organizations they can not afford or allocate dedicated resources to maintain their own data structures. Like one of the interviewee said:

> O5 - *"Another challenge is the cost of producing, analyzing and managing data. The bigger the organization is, the easier it is to manage data. The small organizations cannot afford to do that and to have dedicated people to manage the data."*

Negative aspects of big data
Big data is becoming an important part of various industries like business, academia and also for social good. Though big Data has immense positive impact on our society in all aspects, but the rise of big data is also negative for society nearly in some respects. When talking about big data benefits and challenges, we also asked our interviewees about the negative aspects of bid data from their experience. In this sections we discuss and present the negative sides or risks that big data may cause for the society.

An important concern about using big data for social good is, big data should be used for the good of the whole society, not any specific group of people or individuals. Our interviewees also mentioned about this issue like this:

> O4 - *"I truly believe that if you really going to do something that is in the realm of social good then it has to be something that benefits the whole society not only a few people, because then the rest will be just exploited. I think one have to carefully consider the questions and the problems that you are actually aiming at solving. That is a very important dimension and it is like a tool, you can use it in a good way or in a bad way."*

> O5 - *"We live in a society where the economic power is the most impactful power. Being the weaker link in a chain means you need to make sure that you don't end up being exploited by bigger and more powerful organizations like Cambridge Analytica or Facebook. These challenges are especially dangerous for small organizations who do not know how to deal with those. So the disparity in opportunities between the smaller and bigger organizations is the biggest potential threat in using data."*

Another risk of big data use for social good is, before using the data we need to be clear why we are using them or why we need them. Using data without proper understanding might be a risk for the society.

> O6 - *"A negative aspect of big data can be the 'Distraction'. People may get distracted without knowing properly why we are using data or what data exactly we need. We should not just use data, we should use it well."*

> O9 - *"When using data, organizations need to be aware that they don't focus on vanity metrics but they're always focusing on the actual mission that they're trying to achieve and thinking about."*

> O3 - *"All data do not have social values and people put lot of money for that but it*

actually does not make any difference in the world. These ineffective data projects can also cause damage, sucking up resources but do not deliver anything."

Another negative aspect that one of the interviewee mentioned is lack of proper data alignment. He mentioned that:

O7 - *"Before implementing technology, everything in the organization should be aligned to use the data and technology properly. If the organization does not have proper alignment to use data than the data can not be used properly, people will not understand why they need the data or what is the benefit of those data."*

Some of the interviewees also mentioned about overly technocratic analysis of data, overestimation of the value of big data, messy and wrong interpretation of data; these can also cause negative impact on the society.

Big data tools or techniques

In interview, we asked our participants about the tools and techniques they use to analyze data in their organizations. Table. 5.3 shows the tools and techniques that were used by the participants in their work.

Table 5.3: Tools used by the participants

Tools	**References**	**Sources**
Artificial Intelligence	1	1
Apache Spark	1	1
Google Refine	1	1
Machine Learning	1	1
Python Ecosystem	3	3
R	1	1
SAAS	1	1
Spreadsheet	4	4
STATA	1	1

It has been seen that although the participants use a lot of different tools, but the most common tool used by them is Spreadsheet. It might be because, many of the interviewees think that for simple and small/medium scale data sets, Spreadsheet is sufficient. Some of our interviewees answers were like this:

O1 - *"I use so many different tools. But for data mostly Spreadsheets with thousands of rows sometimes. Moving with other tools is not necessary often."*

O8 - *"Our focus is on small data sets that can be analyzed with just normal Spreadsheets or Microsoft excel for example."*

Organizations who work with large data and complex data sets, they prefer more advance tools like Python, R, Apache Spark etc. Among these advanced tools, Python seems to be most commonly used by our interview participants.

> O5 - *"We are not an organization that uses many data analysis tools. The tools we use can go from normal Excel spreadsheets to analysis with STATA and few big data analyses on Python, with the libraries of Python."*

> O2 - *"For our use internally, we use Python Ecosystem, Numpy, Pandas, Scikit learn and so forth. If working with large datasets we sometimes use tools like Apache Spark."*

Other than Spreadsheet and Python, some organizations also use tools for specific purposes. For example: one of our interviewee mentioned that in their organization they use Google Refine for data cleaning purpose.

> O8 - *"Our another project or initiative which uses and also teaches Google refine for data cleaning and they use a suit of tools that they teaches other organizations to use for their specific needs."*

Overall, it is seen that there are a lot of tools and techniques that are available for data analysis purpose and organizations use a lot of different tools for data analysis based on their needs.

Impact on cooperation
The impact of big data on organization's cooperative practices was agreed upon by almost all participants of our interview. The impact on work practices seem to be similar in all organizations. Interviewees agreed that big data ensures better cooperation, better understanding in team and also make the processes interactive. Some comments of our interviewees are as follow:

> O2 - *"In data analysis everything is a question, nothing is certain. So, you have to ask questions to understand the data. Thus, while working with data it becomes a much more collaborative process to understand what you are doing."*

He also added that:

> O2 - *"Whenever you are doing data analysis with an organization that you are not part of, or even if from the same organization that is very big and you don't know where the data came from and how it was generated; it makes the process a lot more interactive."*

Another example of using data for cooperative purposes is, using it to connect with partners. As one of interviewee mentioned:

> O8 - *"We work with different partner organizations and we use the data to identify the priorities and how they shape our connection with our partners. "*

Data also helps to form a better understanding in a team. It help all team members to be on the same page, so all team members can discuss the same thing to agree or disagree. Our another interviewee phrased it like this:

O1 - *"It supplements the decision making and creates shared understanding about the problem and potentially the shape of solution and it makes it easier to discuss and debate."*

Another participant also said similar, he said that:

O6 - *"This helps the company and the team for better understanding and to take better decisions."*

Another important comment we found from one of our interviewee is, the way data helps to change peoples mindset. Data shows clear cut facts, so with data it is easier to show and convince people about any fact. He phrased it like this:

O5 - *"You really start to see the data having an impact on the organization for the cooperation when you did a good job in changing the mindset. So data is something that can be used one way or another. And if you are very good at changing the mindset, showing the value of creating the social opportunities, then yes the data can be very effective, helpful in changing the process the organization works."*

Impact on decision making

Just like collaboration, from interviews we have seen that big data has a great influence on decision making also. Organizations who work with data analytics, also use data for organizational decision making. All of our interviewee agreed that data influences both their decisions and also the decision making process. One of our interviewee said that their whole purpose of using data is to influence external community decision makers. She mentioned that:

O3 - *"The whole purpose of the data is to influence the decision making processes of external groups. So our whole purpose with data is to influence the decision makers of other community actors which are generally nonprofits, govt agencies and foundations."*

O1 - *"Using data and more information to make more informed decision not determent decisions; but at least complementing the regular decision-making process is really important."*

One of our interviewee also mentioned that they use data to convince and influence their stakeholders about their projects. Talking about their current project, he said that:

O4 - *"You know there are many stakeholders here. So getting their attention and getting them interested and teaching them what this is about and finally seeing them realize that this is something that the society needs is the ultimate goal."*

Some other answers we found form our interviewees regarding use of data analytics for their decision making purpose includes:

O2 - *"We do not take decisions without analyzing data. It impacts both our decision making process and the decisions."*

O9 - *"When we're producing all these analytics, there is a set of results that can then be used by the charity and we measure our impacts when we have initial feedback from the charities about what is useful and then three months down the line we actually call them up and ask them to fill out a survey about what the impact has actually been. It is actually taken up and used within decision making."*

5.2 The Questionnaire

The goal of the questionnaire was to understand what are the major challenges and drivers for the social entrepreneurs and social innovator to employ big data; and how these challenges and drivers influence their intentions. There were a total of 49 respondents from different countries. All the questions were required to answer to submit the questionnaire; so that the respondents do not skip any question by mistake or intentionally. This was because if any respondent skip any question, then the data may not to present a full contextual view of the respondents' opinion.

5.2.1 The respondents

The participants were fairly diverse in terms of geographical areas and a nearly even gender split. Though the respondents live in different countries, but most of them are from Europe and a few from Asia and North America. The majority of the respondents belong to age group 25-44 and a few respondents belong to age group 44 and above. Except 3 respondents, all other respondents said that they have previous experience of social innovation or social entrepreneurship to some extent. And, all participants had experience with big data analytics to some extent. In the following Table 5.4, we present an overview of the respondents in terms of age, gender and background.

Table 5.4: Overview of the respondents

Total number or respondents, n = 49	**Frequency**
Gender	Male = 57% (n = 28) Female = 41% (n = 20) Prefer not to say = 2%(n = 1)
Age	18-24 = 0% (n = 0) 25-34 = 41% (n = 20) 35-44 = 41% (n = 20) 45-54+ = 18% (n = 9)
Social Innovation and/or social entrepreneurship experience	Yes (in different scales) = 94% (n = 46) Not at All = 6% (n = 3)
Experience with big data analytics	Yes (in different scales) = 100% (n = 49) Not at All = 0% (n = 0)

In our questionnaire, we have used the 11 social entrepreneurship scales by Carraher (2013) to better understand the respondents. We asked the respondents to answer the 11 SE scale questions based on their experience and/or current role in the organization; we asked them to scale how these statements describe them with a 1 to 7 Likert-type scale. The questions and the overview of the responses can be seen at the following Table 5.5.

Table 5.5: Overview of the respondents' social entrepreneurship scales

Question	Scale	Frequency(%)
1. I am adopting a mission to create social value (not just private value).	Not at All	0%
	Moderately disagree	2%
	Slightly disagree	4.1%
	Neutral	10.2%
	Slightly agree	20.4%
	Moderately agree	18.4%
	Strongly agree	44.9%
2. I am recognizing new opportunities to serve my mission.	Not at All	2%
	Moderately disagree	2%
	Slightly disagree	4.1
	Neutral	14.3%
	Slightly agree	22.4%
	Moderately agree	18.4%
	Strongly agree	36.7%
3. I am engaging in a process of continuous adaptation related to my mission.	Not at All	0%
	Moderately disagree	2%
	Slightly disagree	8.2%
	Neutral	22.4%
	Slightly agree	24.5%
	Moderately agree	16.3%
	Strongly agree	26.5%
4. I am acting boldly without being limited by resources currently in hand in the fulfillment of my mission.	Not at All	0%
	Moderately disagree	12.2%
	Slightly disagree	20.4%
	Neutral	16.3%
	Slightly agree	24.5%
	Moderately agree	16.3%
	Strongly agree	10.2%
5. I am relentlessly pursuing new opportunities to serve my mission.	Not at All	0%
	Moderately disagree	2%
	Slightly disagree	10.2%
	Neutral	20.4%
	Slightly agree	24.5%
	Moderately agree	28.6%
	Strongly agree	14.3%

6. I am caring deeply about the outcomes created by the fulfillment of my mission.	Not at All	0%
	Moderately disagree	2%
	Slightly disagree	2%
	Neutral	12.2%
	Slightly agree	22.4%
	Moderately agree	30.6%
	Strongly agree	30.6%
7. I seek to be a 'world changer' through the accomplishment of my mission.	Not at All	2%
	Moderately disagree	2%
	Slightly disagree	4.1%
	Neutral	32.7%
	Slightly agree	26.5%
	Moderately agree	12.2%
	Strongly agree	20.4%
8. I am adopting a mission to sustain social value (not just private value).	Not at All	0%
	Moderately disagree	2%
	Slightly disagree	6.1%
	Neutral	8.2%
	Slightly agree	16.3%
	Moderately agree	22.4%
	Strongly agree	44.9%
9. I am engaging in a process of continuous innovation related to my mission.	Not at All	2%
	Moderately disagree	0%
	Slightly disagree	4.1%
	Neutral	14.3%
	Slightly agree	30.6%
	Moderately agree	20.4%
	Strongly agree	28.6%
10. I am exhibiting a heightened sense of accountability to the constituencies served by my mission.	Not at All	2%
	Moderately disagree	0%
	Slightly disagree	6.1%
	Neutral	24.5%
	Slightly agree	30.6%
	Moderately agree	22.4%
	Strongly agree	14.3%
11. I am engaging in a process of continuous learning related to my mission.	Not at All	0%
	Moderately disagree	0%
	Slightly disagree	6.1%
	Neutral	16.3%
	Slightly agree	26.5%
	Moderately agree	40.8%
	Strongly agree	10.2%

From table 5.5, it is notable that all of our respondents more or less are working to create and sustain social value; 44.4% strongly agree that they are working to create social value (scale 1) and 44.9% respondents strongly agreed that they are working to sustain the social value (scale 9). In scale 6, we can see that almost all the respondents said that they care deeply, moderately or even slightly about the outcomes of their works and missions. Most of the respondents also agreed (slight, moderately or strongly) that they exhibit intense accountability to the constituencies served by their my mission. As discussed in section 2.3.1, social entrepreneurs differ from conventional entrepreneurs in their mission. So, overall from the data we can see, our respondents also engage in innovation; pursue new opportunities and have missions to benefit the society, not just for individual. Although they possess similar characteristics to conventional entrepreneurs, they are more concerned with satisfying social needs rather than commercial needs (Roberts and Woods, 2005).

5.2.2 Questionnaire results and findings

Here we are presenting the results of the questionnaire. In this study, we aimed at around 300 participants and finally had total 49 respondents. So our response rate was 16%. With the current number of respondents and a confidence level of 95%, the results give 13% margin of error, which we have calculated with a proportion of 50%, z-score of 1.96 and a population of 300). Statistically these calculations will not provide any robust results. Considering these facts, the results might have big deviations and might not be accurate to generalize the results for all. Despite having all these limitations, the results can still give implications towards the present practical situation.

We first carried out Factorial analysis for all constructs, with principal components and varimax rotation. We have exhibited items in factor loading that were higher than 0,5. We have conducted factor analysis to categorize the variables of our instruments; so that we can examine the characteristics and influence of each category of variables on other variables more clearly.

For the big data challenges, we have done a forced three-factor analysis and identified three distinct factors; 1. Data challenges (challenges related to the data itself), 2. Process challenges (challenges related to the data processing) and 3. Management challenges (challenges related to the data and organization management). We have identified these categories based on Sivarajah et al. (2017). The factor analyzed big data challenges are presented in Table 5.6

Table 5.6: Factor analysis of big data challenges

Items	Data Challenges	Process Challenges	Management Challenges
Lack of skills within the line of business	.775		
Cannot make big data usable for end users	.699		
Don't know how and where to start with big data and data analytics	.661		
Lack of sponsorship for big data and data analytics projects	.630		
Lack of management bandwidth or inadequate staffing due to competing priorities	.607		
No need to use big data or data analytics to change the organization	.596		
Lack of compelling business case to use big data	.558		
Analytical challenges with big unstructured data.		.851	
Big data needs huge data storage capacity		.821	
Scalability problem with big data		.777	
Architecting big data analytic system is difficult.		.678	
Perceived costs may outweigh projected benefits		.579	
Accessibility of data is difficult			.772
Quality of available data			.707
Maintaining data privacy and security is challenging			.544
Ownership of data is not clear, or governance is ineffective			.531

After the challenges, we have factor analyzed the big data enablers. Table 5.7 presents the factor analyzed big data enablers. For enablers also, we have conducted a forced two-factor analysis, which resulted in two factors; 1. Growth of Data and Tools and 2. Commitment. We have categorized the enablers into these two categories inspired from Russom (2011).

Table 5.7: Factor analysis of big data enablers

Items	**Growth of Data and Tools**	**Commitment**
Increased data storage capacity	.841	
Data availability from various sources	.821	
Rapid growth of data from data-intensive sensor network applications	.753	
New set of analytics tools designed specifically to analyze large amounts of data	.625	
General Data Protection Regulation (GDPR) as a pathway to ensure better processing of data and the rights of the data subject		.908
Adaptability of a culture that brings together the power of technology and culture of embracing changes to improve the organization		.696

Douglas and Prentice (2019) mentions three different kinds of social entrepreneurial intentions with different perspectives, i.e. in terms of their social-purpose intention, psychic income or profit-seeking intentions. The psychic-income and profit-seeking intentions are based on Douglas (2013). After factor analyzed the intentions, we have divided and named the factors same as Douglas and Prentice (2019); which are: 1) Social purpose EI, 2) Psychic income EI and 3) Profit seeking EI. Table 5.8 presents the factor analyzed social entrepreneurial intentions.

Table 5.8: Factor analysis of Social Entrepreneurial Intentions (SEI)

Items	**Social Purpose EI**	**Psychic Income EI**	**Profit Seeking EI**
Help underprivileged people achieve what they are unable to achieve on their own	.930		
Help poor people get enough food, clothing, shelter, and medical assistance	.896		
Gain great satisfaction because you are helping others who are in need	.848		
Solve social and economic problems that cause others to suffer	.679		

Serve as a volunteer to help people who have social and/or economic problems	.541		
Have great flexibility to decide your work hours, your product lines, and so on		.930	
Locate the business at a place that suits your personal preferences		.901	
Be your own boss and make all the important decisions yourself		.885	
Enjoy the lifestyle and benefits of an independent business owner		.855	
Create a business around your personal hobbies or special interests		.812	
Grow the firm to be very large and profitable			.847
Pursue a high-risk opportunity that has the possibility of very high profits			.825
Pursue profit maximization above all other objectives			.808
Generate high profits over many years			.785
Become a major, globally recognized corporation			.691

In Table 5.9, we have correlated all categories of big data challenges and enablers with different social entrepreneurial intentions. Here, we have performed Spearman's rank correlation coefficient.

We can see, social purpose EI has a correlation of .129 with data challenges; correlation of .377 with process challenges and has correlation of .288 with management challenges. So, we can see all these categories of big data challenges are correlated with the social purpose intentions. Same way, big data enablers i.e growth of data & tools (.485) and commitment(.333) also has strong correlation with the social purpose intentions of social entrepreneurs. Notable that, the management related challenges of big data has a negative correlation (-.035) i.e strongly correlates with the profit seeking intentions of social entrepreneurs. The correlation matrix also exhibits negative correlation of data related challenges on the psychic income (-.439) and profit seeking (-.166) intentions. Overall from this correlation matrix, it is visible that all categories of challenges correlates to all categories of social entrepreneurial intentions. And also, the big data enablers are seem to be most influential for social purpose entrepreneurial intentions.

Table 5.9: Correlation of big data challenges and enablers with Social Entrepreneurial Intentions

	Social Purpose EI	Psychic Income EI	Profit Seeking EI	Management challenges	Process challenges	Data challenges	Growth of data & tools	Commitment
Social purpose SEI	1.000	.039	-.056	**.129**	.377**	.288*	**.485****	**.333***
Psychic income SEI	.039	1.000	.517**	**.154**	.436**	**-.439****	-.037	-.215
Profit seeking SEI	-.056	.517**	1.000	**-.035**	.333*	**-.166**	.000	-.051
Management challenges	**.129**	**.154**	**-.035**	1.000	.234	.013	-.134	-.005
Process challenges	.377**	.436**	.333*	.234	1.000	.194	.139	.029
Data challenges	.288*	**-.439****	**-.166**	.013	.194	1.000	.283*	.359*
Growth of data & tools	**.485****	-.037	.000	-.134	.139	.283*	1.000	.563**
Commitment	**.333***	-.215	-.051	-.005	.029	.359*	.563**	1.000

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

After performing factor analysis and examining the correlation of the constructs, we have performed the one-way analysis of variance (ANOVA) to determine whether there are any statistically significant differences between the effects of different categories of challenges and enablers on each social entrepreneurial intention. In Table 5.10, we present the ANOVA for social purpose EI. We can see, the big data process related challenges has a significant effect ($F=2.436$; $p<0.05$) on the intention; where other two categories of challenges has less significance. The growth of big data and tools also has a significant effect (F=2.255; p<0.05 on social purpose intentions; similarly the other enabler category i.e commitment also has similar kind of significance (F=2.263; p<0.05).

Table 5.10: ANOVA for Social Purpose EI

		Sum of Squares	**df**	**Mean Square**	**F**	**Sig.**
Management challenges	Between Groups	14.895	17	.876	.865	.615
	Within Groups	31.413	31	1.013		
	Total	46.307	48			
Process challenges	Between Groups	32.073	17	1.887	2.436	.015
	Within Groups	24.007	31	.774		
	Total	56.080	48			
Data challenges	Between Groups	16.431	17	.967	1.178	.336
	Within Groups	25.444	31	.821		
	Total	41.875	48			
Growth of data and tools	Between Groups	32.168	17	1.892	2.255	.024
	Within Groups	26.008	31	.839		
	Total	58.176	48			
Commitment	Between Groups	35.996	17	2.117	2.263	.024
	Within Groups	29.004	31	.936		
	Total	65.000	48			

In Table 5.11, we present the ANOVA for profit seeking EI. We see that here, the data related challenges has the most significant effect on profit seeking EI (F=1.885; p<0.5). Data processing related challenges and also the commitments as enabler are also significant (p<0.5); where the management challenges and growth of data & tools as enablers has less effect on the profit seeking EI.

Table 5.11: ANOVA for Profit seeking EI

		Sum of Squares	df	Mean Square	F	Sig.
Management challenges	Between Groups	15.438	20	.772	.700	.793
	Within Groups	30.869	28	1.102		
	Total	46.307	48			
Process challenges	Between Groups	26.842	20	1.342	1.285	.266
	Within Groups	29.238	28	1.044		
	Total	56.080	48			
Data challenges	Between Groups	24.028	20	1.201	1.885	.060
	Within Groups	17.847	28	.637		
	Total	41.875	48			
Growth of data and tools	Between Groups	16.181	20	.809	.539	.922
	Within Groups	41.995	28	1.500		
	Total	58.176	48			
Commitment	Between Groups	34.642	20	1.732	1.598	.125
	Within Groups	30.358	28	1.084		
	Total	65.000	48			

For psychic income EI, we have found that both management and process little effect on the psychic income intentions. But other items like data challenges and both categories of enablers (growth of data & tools and commitment) have less effect on psychic income EI. The details are presented in the following Table 5.12.

Table 5.12: ANOVA for Psychic income EI

		Sum of Squares	df	Mean Square	F	Sig.
Management challenges	Between Groups	21.114	20	1.056	1.173	.342
	Within Groups	25.194	28	.900		
	Total	46.307	48			
Process challenges	Between Groups	27.060	20	1.353	1.305	.253
	Within Groups	29.020	28	1.036		
	Total	56.080	48			
Data challenges	Between Groups	15.554	20	.778	.827	.665
	Within Groups	26.321	28	.940		
	Total	41.875	48			
Growth of data and tools	Between Groups	14.202	20	.710	.452	.965
	Within Groups	43.974	28	1.571		
	Total	58.176	48			
Commitment	Between Groups	20.923	20	1.046	.665	.826
	Within Groups	44.077	28	1.574		
	Total	65.000	48			

Chapter 6

Discussion

In this chapter we discuss the findings presented in chapter 5. Section 6.1 reflects the role of big data in helping social innovators and entrepreneurs to generate social value. Section 6.2 discusses the challenges and benefits of using big data to address societal challenges; and Section 6.3 discusses the relation between challenges and enablers of big data with social entrepreneurial intentions (SEI).

6.1 Role of big data in generating social value

In this study, we have seen different ways of generating social value. Mulgan (2010) stated that social value does not have any fixed definition and boundary; it is something that refers to non-financial impacts and aimed at the wellbeing of people and society. After analyzing the interview data, we have seen organizations are working for the wellbeing of the society in many different ways. Some organizations work to create social impact (for example - organization O5, who works to embed social impact in every business through experimental research, policy, and consultancy) and some organizations work to address social problems or challenges directly, for example organization O7, who works with strategies that address societies challenges like food waste, human health, Lead contamination and early age brain development etc. So, in all cases the organization's mission is the same, that is to offer wellbeing for the society and its people; but their strategies or programs are different. When organizations are offering wellbeing to the society (in any way), they are generating social value. So, in our thematic analysis we divided our *Social Value* theme into two categories - social impact and societal challenges. In the following paragraphs, we discuss these two themes in more details.

Social Impact
Some organizations are generating social value by creating or increasing the social impact with the help of big data and analytics. For example, organization O3 is an organization

which is a coalition of organizations of cities that collect and organize neighborhood level data for better decision making with a particular focus on helping stakeholders use the data and promoting actuatable access and outcome with the data. This organization creates value by empowering the communities and its citizens. Our case organization O1 supports people to do socially beneficial work using data and data driven technology and supports other organizations that are building technology to build in a way that is conscious about the social impact of what they are building and the value they create are not undermined by the harm they cause. Another example is organization O5, which carries out consulting and researches in the area of social finance and impact investing. Impact investing tries to find solution how to finance projects in a way that can create social impact along with financial profit. They try to show their customers how social impact can be also a value proposition and in their research they use data to better understand their projects and facts. Organization O10 focuses more on innovation. They use data to analyze how innovation can be made for products and services that better addresses the needs of the society. So in all these examples we see how these social organizations are using big data and analytics to create an positive impact and thus offering social value to the society.

Societal Challenges

Another form of generating social value is solving or addressing social problems with the help of big data and analytics. Organization O2 is an organization that funds social entrepreneurs and social innovators. They invest in strategies that address various social problems like - ensuring food safely for everyone, ensuring education for underprivileged students, early brain development problem of children etc. The organization does not use big data and analytics itself; instead it funds strategies of other organizations to use big data and data analytics for social causes. For example, this organization funds universities to use data for predictive analytics to identify students who need help and assistance; to ensure that unprivileged students enroll, persist through, and graduate with the skills they need to find jobs and pursue their career goals. Another organization O4 is a telecommunication company, that uses their customers' mobile data to forecast the spread of epidemic diseases to help the society and the ministry of health to understand and solve the problem. Organization O6 generates social value by analyzing data for better addressing various social issues. For example: Chicago is a city with a large inventory of aging homes that still contain lead paint and pipes. So old building in Chicago have exposure to lead and children can be exposed to this threat. They use data to find the building with threat and make the people aware of the threat. In another project in Africa, O6 used predictive modeling to find out which children may need immediate assistance from the teachers, to better allocate resources. With data they predicted student dropout rate, and also identified students who need more assistance from their teachers. Organization O8 is a global non-profit organization that works to make data available to people who wants to use it for social good; they help other civil society organizations to learn how to work with data; primarily open data but sometimes also closed data to take action on social problems. Organization O9 supports charities and non profit organizations by providing resources to use and analyze data more efficiently and to gain better insights into the work the charities do.

From all these examples, we answered our first research question (RQ1) and understood how social entrepreneurs, social innovators and also social organizations are using big data and analytics to generate social value. As Pappas et al. (2018) mentioned, analyzing big data in different ways can lead to sustainable societies; here in this research, we could see how these social organizations are using and analyzing big data in different ways to benefit the society. In every organization we present here, big data plays a great role to perform its tasks, leading to the achievement of its mission; and creating economic value in a way that also creates value for society by addressing its needs and challenges, as Porter and Kramer (2019) discussed. So, there are different ways of using big data to address social needs facilitating societal transformation. These examples can also help future social entrepreneurs and innovators to get inspired to employ big data analytics in their work.

6.2 Benefits and challenges of using big data

Benefits

Big data analytics is a fast-growing and influential practice (Russom (2011)) and it has a huge potential of benefits that are driving the adoption of it in organizations. From literature and interviews, we have seen all organizations get many kinds of benefits by using big data in their works. Gupta et al. (2018) presents a framework for the societal applications of big data. In this research, we have found that along with the big data societal application areas mentioned by Gupta et al. (2018), there are some additional areas also where organizations are using big data and getting benefit in creating social impact. The new areas where big data is used and benefits includes: exploring new working opportunities or identifying social problems in the society, changing peoples' mindset by showing clear cut facts with data, increased engagement and collaboration among the stakeholders in a society; and supporting civil society decision making.

So, organizations are getting benefit from big data in various forms. There are some common benefits we have found from our interviews that organizations get when they use big data analytics for management and decision making purposes. From the interviews, we have known that most often big data analytics help organizations to get a base line understanding of a given social issue. So it can help organizations to get all stakeholders on the same page about what is happening, to change the conversation and to discuss the implications and what should they be doing differently. Based on our interviewees' opinion, it improves the relationship among stakeholders and get everyone on the same set of knowledge; and that helps to move into the important conversation far ahead of time and to understand what actions need to be taken. From the interviews, we have also seen that he usage of data and analytics in organizations has many other different dimensions. Our interviewees mentioned that their organizations explore opportunities or social problems to work with that they would not expect in some cases, by using and analyzing data. It opens up new possibilities and opportunities for the social organizations; they can realize what is needed from a social point of view. It is also about learning about how they can use the data, how useful it is and in what setting are the data valuable. This data driven side of the organizations has the potential not just to improve itself but also to improve the society and peoples' lives. So, employing big data analytics in innovative solutions and works can help future entrepreneurs and innovators to get all these benefits in their organizations too.

Challenges
While the potential benefits of big data are real and significant, there still remain many different challenges associated with this; and that need to be addressed to fully realize this potential. From the mapping study we performed in this research, we found some challenges related to big data adoption in general, from different fields of studies. In this thesis, we have examined if all those challenges also exist when big data is used for social causes and from a social perspective. After talking with our interview participants and conducting the questionnaire, we have seen that they also experience the similar kind of challenges we found from literature. Some of these challenges are technical, some are managerial and some are even related to organizational culture or the mindset. The data sharing culture in our society is still in developing the norms and practices around that. Some most common challenges that we have known from the interviews are related to maintaining privacy and ethics of the data. The challenges rise from protecting the privacy of whoever produce the data; ensuring ethical guidelines to work with that data. Though it is been a year that GDPR was approved but we have seen that some bottom up organizations we interviewed, are still struggling to adapt the rules completely. So there is still a big potential that is untapped at the moment because social entrepreneurs, innovators and also the social organizations are still understanding how to use data properly. Big corporations already know how to do that; but the social organizations and social enterprises are still struggling a bit with that. Another common challenge we have seen is having the understanding about the potential of data and what kind of data is needed. One of our interviewee mentioned that people may get distracted without knowing properly why they are using data; and this kind of distraction can lead to exploitation of people in the society without achieving any common social good. These are some common issues we have found from this study that we believe need to focus on, but we should not consider these as obstacles that we can not overcome.

There are still challenges in employing big data for social good, even a lot of negative examples of using big data for evil, for profit or for exploitation. But along with some negative aspects, new frameworks for data ethics, advanced technologies like Blockchains and algorithms are also developing; we have seen this this study how coalitions are working on making data collaboration, developing ethical frameworks to ensure better use of data for the society. We have also seen that our interviewees who are facing these kind of challenges are also concentrating and working to address the challenges to make the big data adoption process easier. Overall from this research we see, along with technical and business organizations, social organizations are also adopting big data analytics and its applications; our social entrepreneurs and innovators are getting more concerned about the big data potential; also about the issues and debates around it. So we believe, our researchers, entrepreneurs and social enterprises and all other stakeholders of the social innovation ecosystem as a whole are making good progress on using big data application for social good.

Enablers or Drivers
Along with challenges, in this research we have also examined the drivers or enablers that can boost the big data adoption among the social entrepreneurs and innovators. Along with the advancement of data technologies and data generation, commitments from the industry also helps social organizations to adapt big data applications. Organizations are moving towards a culture that embraces the power of technology and associated changes to improve the organization. Organizations are also working with governments and authorities to develop data collaboratives, ethical frameworks, policies and regulations, data governance body etc. to ensure better use of data for the society and social good. All these developments and initiatives are helping and effecting the stakeholders of the social innovation ecosystem to leverage the employment of big data applications.

6.3 Relationship of big data challenges and enablers with SEI

Santos (2012) stated that innovation is necessary as it can be a potential solution for the society to solve problems and needs, like government or market failures that allows social problems to exist. If innovation is necessary then having the motivation and intention to do something innovative and achieve successful innovation is also important, as stated by Phillips et al. (2015). If social entrepreneurs are working with big data analytics then associated big data challenges and drivers are some factors that can influence their innovation motivation and intentions also. In this study, we have found many different challenges and drivers associated with big data adoption, but all of them may not be equally significant to social entrepreneurs and innovators. We examined this relationship of big data challenges and drivers with social entrepreneurial intentions by analyzing the questionnaire data. The analysis of the questionnaire where we categorized the challenges and enablers depending on their factor loading, shows that all different categories of challenges and enablers affect social entrepreneurial intentions differently.

Our analysis shows that for social purpose intentions, data processing related challenges like - analytical challenges, architecting and scaling big data applications, cost of storing and processing big data etc. effects the most. Challenges related to the data itself also has a strong effect on social purpose intentions. The most common and important data related challenges is maintaining data privacy and security. All our interviewees and questionnaire respondents agreed that maintaining data privacy and data ethics is a big concern for them. Data related other challenges includes data accessibility, data quality, data governance etc. And comparing to data processing and data characteristics, data and organization management related challenges seem to have less effect on social purpose intentions. This management related challenges includes lack of skills and staffs, sponsorship, lack of understanding of the big data potential etc. For psychic income intentions also, we have seen data processing challenges has the most significant effect. But unlike social purpose intentions, here management challenges are more important than the data characteristics related challenges. And for the third category of intentions i.e. the profit seeking intentions, data related challenges has the most significant effect. And here, management related challenges has the least significance.

Similarly in terms of enablers, all social entrepreneurial intentions have different implications. For social purpose intentions both categories of enablers have similar significance. But for profit seeking and psychic income intentions, commitment related drivers seem to have more influence than enablers related to growth and development of data and technologies. Though data and technological advancement may have less significance, but still they have enough influence on the intentions regarding the big data adoption.

So, overall we see big data analytics related challenges and the driving forces do have influence on social entrepreneurs' intentions. This research also shows the differences in significance of all the challenges and drivers on intentions.

Chapter 7

Conclusion

This thesis aimed to explore and understand the role of big data analytics in creating social value and social good; and how the challenges and drivers associated with big data analytics influence the social entrepreneurs and innovators about employing big data. To achieve our goal we investigated social organizations and individuals working in this sector, with both qualitative and quantitative manner. Our research results show that entrepreneurs and social organizations are mainly using big data in two different ways; one is for solving or addressing social problems and the other is for creating social impact. In both cases, the use of big data analytics leads to creating and sustaining social value and social good. Our study also indicates that the challenges and drivers have different influence on social entrepreneurs. The challenges one organization experiences may differ from the experience of another organization, for example the issues related to implementing GDPR and ensuring data privacy. We have seen from interviews that small organizations struggle to implement GDPR than any bigger organizations. Also, from the interviews, we have seen that drivers like data access, increased data storage capability matters more for a small or grass-roots level organization than any big corporations. So, not all the enablers or challenges are equally important to social entrepreneurs or innovators when it comes to employing big data; the impact may differ from organization to organization based on the context, organization's resources, status, etc.

This study contributes to the social perspective that emphasizes the importance of adoption and applications of big data; by identifying the trends and practices, associated challenges, driving forces along with their significance and influences. The findings of this research will act as a guideline for social entrepreneurs, innovators and researchers to better understand the big data practices and needs; and will help to develop future agenda and roadmaps. For other researchers of this field, this study will provide a detailed understanding of the context and also about the big data applications in the social sector. It will also provide an outlining of the potential areas for future research. More investigations can be undertaken to understand the challenges in a tailored way to develop and propose specific solutions for them. Future researches can also focus on making the enablers more effective, so they can help to increase the organizational benefits from big data adaption even more.

Another contribution of this research is the systematic mapping study. This systematic mapping study also extends the big data research in several ways. To best of our knowledge, there was no systematic mapping or literature review done before this that focuses solely on how big data and their analytics can lead to social good and social value. The findings of the mapping show the publication frequency, big data research areas, type of contributions big data research is making in the industry overall; based on that future researchers can think what type of contributions we are lacking, what areas we are missing, etc. and make their research agenda.

There are some limitations to this study that can be identified. First of all regarding the questionnaire sample. We did not have enough respondents to be statistically robust. A small sample size increases the margin of errors skewing the results, which decreases the power of the study. Another limitation of the questionnaire was, the participants. We invited all individuals of the social organizations to participate in the questionnaire. But not necessarily all employees of a social organization have enough understanding and the same level of motivation towards social innovation and social good. This issue was also reflected in the questionnaire analysis, where few people mentioned that they had less experience about social innovation and/or social entrepreneurship than expected.

Another shortcoming to the study can be the diversity and number of the investigated organizations. A wider collection of qualitative data could be helpful for this research, both to discover more relevant themes and to ensure credible conclusions. Further investigations of more social sector organizations operating in different geographical locations can improve the reliability of the research results.

Bibliography

Agarwal, R., Dhar, V., 2014. Editorial—big data, data science, and analytics: The opportunity and challenge for is research. Information Systems Research 25, 443–448.

Akoka, J., Comyn-Wattiau, I., Laoufia, N., 2015. Research on big data – characterizing the field and its dimensions, advances in conceptual modeling. MOBID Workshop, Proceedings of ER2015 Workshops. 9382, 173–183.

Akoka, J., Comyn-Wattiau, I., Laoufia, N., 2017. Research on big data – a systematic mapping study. Computer Standards & Interfaces 54(2), 105–115.

Ashoka, 2019. Social entrepreneurship. Accessed 09-02-2019.
URL `https://www.ashoka.org/en-SE/focus/social-entrepreneurship`

Bacq, S., Alt, E., 2018. Feeling capable and valued: A prosocial perspective on the link between empathy and social entrepreneurial intentions. Journal of Business Venturing, 33, 333–350.

Basili, V. R., 1992. Software modeling and measurement: the goal/question/metric paradigm. Tech. rep.

Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. Qualitative research in psychology 3(2), 77–101.

Brown, B., Chul, M., Manyika, J., 2011. Are you ready for the era of big data? McKinsey Quarterly 4, 24–27, 30–35.

Carraher, S., 2013. Isi, social entrepreneurship, and research. Journal of Management History 19 (1).

Certo, S., Miller, T., 2008. Social entrepreneurship: key issues and concepts. Business Horizons 51, 267–271.

Chen, C. L. P., Chun-Yang, Z., 2014. Data-intensive applications, challenges, techniques, and technologies: A survey on big data. Information Sciences 275(10).

Chen, H., Chiang, R. H., Storey, V. C., 2012. Business intelligence and analytics: From big data to big impact. MIS Quarterly 36, 1165–1188.

Cruzes, D. S., Dyba, T., 2011. Recommended steps for thematic synthesis in software engineering. book: Empirical software engineering and measurement (esem). International Symposium on 2011, 275–284.

Cuquet, M., Vega-Gorgojo, G., Lammerant, H., Finn, R., ul Hassan, U., 2017. Societal impacts of big data: challenges and opportunities in europe. Cornell University Library.

de Bruin, A. M., Ferrante, F. M., 2011. Bounded opportunity: A knowledge based approach to opportunity recognition and development. Entrepreneurship Research Journal 1(4), 1–21.

Dees, J., 1998. The meaning of social entrepreneurship. Unpublished paper funded by the Kauffman Center for Entrepreneurial Leadership, Kansas City, Missouri.

Defourny, J., Nyssens, M., 2010. Conceptions of social enterprise and social entrepreneurship in europe and the united states: Convergences and divergences. Journal of Social Entrepreneurship 1, 32–53.

Dencik, L., Hintz, A., Carey, Z., 2017. Prediction, pre-emption, and limits to dissent: Social media and big data use for policing pro-tests in the united kingdom. New Media & Society 20.

Douglas, E., 2013. Reconstructing entrepreneurial intentions to identify predisposition for growth. Journal of Business Venturing 28, 633–651.

Douglas, E., Prentice, C., 2019. Innovation and profit motivations for social entrepreneurship: A fuzzy-set analysis. Journal of Business Research 99, 69–79.

Ferreri, M., Sanyal, R., 2018. Platform economies and urban planning: Airbnb and regulated deregulation in london. Urban Studies 55(15).

Gobble, M. M., 2013. Big data: the next big thing in innovation,. Res. Technol. Manag., 56(1), 64–66.

Gupta, A., Deokar, A., Iyer, L., Sharda, R., Schrader, D., 2018. Big data & analytics for societal impact: Recent research and trends. Information Systems Frontiers 20, 185–194.

Hibbert, S., Hogg, G., Quinn, T., 2005. Social entrepreneurship: understanding consumer motives for buying the big issue. Journal of Consumer Behavior 4(3), 159–172.

Hilbert, M., 2016. Big data for development: A review of promises and challenges. Development Policy Review.

Jin, X., Wah, B. W., Cheng, X., Wang, Y., 2015. Significance and challenges of big data research. Big Data Research 2, 59–64.

Katal, A., Wazid, M., Goudar, R. H., 2013. Big data: Issues, challenges, tools and good practices. International Conference on Contemporary Computing (IC3).

Kitchenham, B., 2004. Procedures for performing systematic reviews. Keele, UK, Keele University 33, 1–26.

Kitchenham, B. A., Pearl Brereton, O., Owen, S., 2008. Length and readability of structured software engineering abstracts. IET Softw 2(37).

LaValle, S., Lesser, E., Shockley, R., Kruschwitz, N., Hopkins, M., 2011. Big data analytics and the path from insights to value. MIT Sloan Management Review.

Lehner, O. M., Kansikas, J., 2012. Opportunity recognition in social entrepreneurship: A thematic meta analysis. Journal of Entrepreneurship 21, 25–58.

Letaifa, S., 2015. How social entrepreneurship emerges, develops and internationalizes during political and economic transitions. European Journal of International Management.

Maltby, D., 2011. Big data analytics. 74th Annual Meeting of the Association for Information Science and Technology.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Byers, A., 2011. Big data: The next frontier for innovation, competition and productivity. McKinsey Global Institute.

Mathiassen, L., Chiasson, M., Germonprez, M., 2012. Style composition in action research publication. MIS Quarterly 32(2), 347–363.

Mauro, A. D., 2015. What is big data? a consensual definition and a review of key research topics. AIP Conference Proceedings. 1644(97).

Morner, M. V., 2016. Application of call data records – chances and obstacles. World Conference on Transport Research, WCTR.

Mulgan, G., 2006. The process of social innovation. Innovations: Technology, Governance, Globalization 1, 145–162.

Mulgan, G., 2010. Measuring social value. Stanford Soc Innov Rev.

Munshi, N. V., 2010. Value creation, social innovation and entrepreneurship in global economies. Journal of Asia-Pacific Business 11, 160–165.

Noruzi, M. R., Westover, J. R., , Rahimi, G. R., 2010. An exploration of social entrepreneurship in the entrepreneurship era. Asian Social Science 6(6), 4–10.

Oates, B. J., 2005. Researching Information Systems and Computing. Sage Publications Ltd.

Pappas, I. O., Jaccheri, L., Mikalef, P., Giannakos, M. N., 2017. Social innovation and social entrepreneurship through big data: Developing a research agenda. 11th Mediterranean Conference on Information Systems (MCIS).

Pappas, I. O., Mikalef, P., Giannakos, M., 2018. Big data and business analytics ecosystems: paving the way towards digital transformation and sustainable societies. Inf Syst E-Bus Manage 16, 479.

Petersen, K., Feldt, R., Mujtaba, S., Mattsson, M., 2008. Systematic mapping studies in software engineering. . In EASE 8, 68–77.

Peterson, R. A., 2000. Constructing effective questionnaires. Chronicle Books.

Phillips, W., Lee, H., Ghobadian, A., N., O., James, P., 2015. Social innovation and social entrepreneurship: A systematic review. Group & Organization Management 40(3), 428–461.

Phills, J. A., Deiglmeier, K., Miller, D. T., 2008. Rediscovering social innovation. Stanford Social Innovation Review, 6, 34–43.

Porter, M., Kramer, M., 2019. Creating shared value. In: Lenssen G., Smith N. (eds) Managing Sustainable Business Managing Sustainable Business, 327–350.

Runeson, P., Host, M., 2009. Guidelines for conducting and reporting case study research in software engineering. Empirical software engineering 14(2), 131.

Russom, P., 2011. Big data analytics. TDWI best practices report Fourth quarter.

Saldana, J., 2015. The coding manual for qualitative researchers. Sage.

Santos, F. M., 2012. A positive theory of social entrepreneurship. Journal of Business Ethics 111, 335–351.

Selladurai, R., Carraher, S., 2014. Servant leadership: Research and practice. IGI Global, Hersey, PA.

Siemens, G., Long, P., 2011. Penetrating the fog: Analytics in learning and education. EDUCAUSE Review 46(5), 30–32.

Singer, J., Vinson, N. G., 2002. Ethical issues in empirical studies of software engineering. IEEE Transactions on Software Engineering 28(12), 1171–1180.

Sivarajah, U., Kamal, M., Irani, Z., Weerakkody, V., 2017. Critical analysis of big data challenges and analytical methods. Journal of Business Research 70, 263–286.

Stallworthy, G., Boahene, K., Ohiri, K., Pamba, A., Knezovich, J., 2014. Roundtable discussion: what is the future role of the private sector in health. Globalization & Health (serial online) 10(1), 1–9.

Taylor, L., Cowls, J., Schroeder, R., Meyer, E., 2014. Big data and positive change in the developing world. Policy and Internet 6, 418–444.

Thake, S., Zadek, S., 1997. Practical people, noble causes: How to support community social entrepreneurs. London, England: New Economics Foundation.

Wamba, S. F., 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.

Wesolowski, A., Qureshi, T., Boni, M. F., Sundsoy, P. R., Johansson, M. A., Rashed, S. B., Engo-Monsen, K., Buckee, C., 2015. Impact of human mobility on the emergence of dengue epidemics in pakistan. PNAS September.

Wilkinson, M., Dumontier, M., Aalbersberg, I., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J., da Silva Santos, L., Bourne, P., 2016. The fair guiding principles for scientific data management and stewardship. Scientific data 3.

Wu, X., Zhu, X., Wu, G.-Q., Ding, W., 2014. Data mining with big data. IEEE Transactions on Knowledge and Data Engineering 26(1).

Zahra, S. A., Gedajlovic, E., Neubaum, D. O., Shulman, J. M., 2009. A typology of social entrepreneurs: Motives, search processes and ethical challenges. Journal of Business Venturing 24, 519–532.

Zannier, C., Melnik, G., Maurer, F., 2006. On the success of empirical studies in the international conference on software engineering. ICSE '06 Proceedings of the 28th international conference on Software engineering, 341–350.

Zicari, R. V., 2014. Big data: Challenges and opportunities. AKERKAR, R. (ed.) Big data computing. Boca Raton, FL: Chapman and Hall/CRC.

Appendix A1

The Role of Big Data in Addressing Societal Challenges: A Systematic Mapping Study

Farzana Quayyum[1], Ilias O. Pappas[1,2] [0000-0001-7528-3488] and Letizia Jaccheri[1] [0000-0002-5547-2270]

[1] Norwegian University of Science and Technology (NTNU), Trondheim, Norway
[2] University of Agder, Kristiansand, Norway
farzanaq@stud.ntnu.no; ilias.pappas@uia.no; letizia.jaccheri@ntnu.no

Abstract. Big data has recently become the focus of academic and corporate investigation due to its high potential in generating business and social value. We have done a systematic mapping of the literature related to big data and its applications leading to social change through the lens of social innovation. The search strategy initially resulted in 593 papers, and after applying inclusion exclusion criteria a total of 156 papers were mapped; 59% of which were identified as empirical studies. This mapping investigated the publication frequency of the studies, research approach and contributions, research areas and article distribution per journal. We also address some challenges found from the mapping associated with the research topic. This mapping study will offer the basis for a reflection process among the researchers in this field and will allow us to develop a research agenda and roadmap of big data and its applications leading to social change.

Keywords: Big Data, Data Analytics, Social Innovation, Social Good, Social Change, Societal Transformation, Systematic Mapping Study.

1 Introduction

The evolution of Information and Communication Technology drives the digitalization process in many aspects of peoples' daily lives, which generates huge amount of data every moment from a growing number of sources. In the last decade, the use of big data and their analytics has earned a lot of attention. In various fields of science, technology, and business, the merit of big data is undeniable. But from the social perspective, the potential use of big data is yet to be figured out by social sector organizations [1]. Several definitions of big data exist, and they typically refer to the 'three Vs'' that characterize big data: volume, velocity, and variety, which have been extended including more characteristics of big data as explained a recent literature review [2]. Furthermore, several definitions of big data analytics exist [2], and they roughly refer to the combination of the data itself, the analytics applied to the data, and the presentation of results to generate value [3]. Thus, here we are interested in big data and their analytics.

The potential of big data and analytics to generate social value seems clear [1], however the main focus of big data analytics research has been on business value [4].

Combining big data analytics with social innovation can be the solution to address this gap [5,6]. Social innovation is defined as a novel solution to a social problem that is more effective, efficient, sustainable, or just than existing solutions and for which the value created accrues primarily to society as a whole rather than private individuals [7]. Social innovation can generate social good and lead to social change. Social good is typically defined as an action that provides some sort of benefit to the people of the society. The concept of social change refers to addressing the root causes of societal problems and changing them.

The terms social innovation, social good, social change, and societal transformation are related to each other. During this study our focus was on the applications of big data that have social impact and address social problems or challenges; so, to keep a broad and wide scope in this mapping review study we use all these terms.

Systematic literature reviews on big data applications have been conducted and investigate, among other things, big data dynamic capabilities [2], the operation and strategic value of big data for business [8], the impact of big data on business growth [9], the social economic value of big data [10]. Furthermore, there is an increasing number of studies that address both the business and social impact of big data as well as ways on how big data analytics can solve societal challenges, with evident examples the following recent special issues [1,11]. However, to best of our knowledge there is no systematic mapping or literature review that focuses solely on how big data and their analytics can lead to societal transformation and social good.

A systematic mapping can help us to understand what conditions can enable successful solutions, combined with strategies, tactics, and theories of change that lead to lasting impact [5,12,13]. Furthermore, this mapping will allow capturing the needed capabilities, resources, and conditions that the big data actors need to develop or acquire in order to manage big data applications, increase social value and solve societal challenges and create a sustainable society. To contribute to the creation of sustainable societies, we have done this systematic mapping of the literature related to big data and their applications leading to social innovation and thus societal transformation.

The objective of this study is to offer a map of the research that has being done, thus offering the basis to develop a research agenda and roadmap of big data and analytics and their applications leading to societal transformation and change. We have followed the standardized process for systematic mapping studies [14]. Based on the primary search with search strings, a total of 593 unduplicated papers was retrieved. After applying some exclusion criteria, the number was reduced to 165 (based on titles), then 153 (based on abstracts) and finally 146 were selected from the search and later 10 more papers were added manually from Google scholar.

The relative newness and growing interest in the research field, argues the need for a mapping study to identify the focus and quality of research in using big data analytics for social challenges. To provide an up to date overview of the research results within the field, we came up with the following research questions:

RQ1: How the research about ‘big data and social innovation' has changed over time (in the last decade)?
RQ2: How much of the research is done based on empirical studies and what type of empirical studies?
RQ3: What are the challenges or barriers for successful implementation of big data for societal challenges?

The paper proceeds as follows: Section 2 introduces the background of this study, Then, section 3 explains the detailed procedure of the research method, section 4 presents the results and findings of the mapping study, section 5 discusses the findings in relation to the research questions, section 6 concludes the paper presenting the implications of this study.

2 Background

2.1 Big Data

The digital and connected nature of modern-day life has resulted in vast amounts of data being generated by people and organizations alike. This phenomenon of an unprecedented growth of information and our ability to collect, process, protect, and exploit it has been described with the catchall term of Big Data [15]. Literature identifies ‘big data’ as the ‘next big thing in innovation’ [16], the next frontier for innovation, competition, and productivity” [17]. The rationale behind such statements is that the ‘big data’ is capable of changing competition by “transforming processes, altering corporate ecosystems, and facilitating innovation” [18]. It can be acknowledged as a key source of value creation. Beyond improving data-driven decision making, it also is crucial to identify the social value of big data [4], and what are the role of big data and potential impact of it in the society.

2.2 Social Innovation

The term social innovation has largely emerged in the last few years and there is much discussion about it now. The field of social innovation has grown up primarily as a field of practice, made up of people doing things and then, sometimes, reflecting on what they do [19]. The term social innovation has not any fixed boundaries, as it cuts across many different sectors like public sector, the social benefit sector, technology sector, and many others. The social innovation process has been described by scholars in multiple contexts as it needs to be multidisciplinary and cross social boundaries, for its impact to reach more people [20,21,22]. Social innovations are ideas that address various social challenges and needs.

2.3 Big Data and Social Innovation

Big data contains a wealth of societal information and can thus be viewed as a network mapped to society; analyzing big data and further summarizing and finding clues and laws it implicitly contains can help us better perceive the present [23]. Data are an important element of social innovation. To initiate any innovative steps or to address any social challenge, data are needed. A deliberate and systematic approach towards social innovation through big data is needed as it will offer social value [5]. Since more data become available at a smaller cost, big data can be used as actionable information to identify needs and offer services for the benefit of the society and ensure aid to the individuals and society that generate them [13].

Following the importance of big data and social innovation, further work is needed to better define and understand how well society can benefit from big data to increase social value and lead to social good [4,6]. From this mapping, we can contribute to the field of big data research from a social perspective. While presenting an overview of the present research status, we also want to identify if there are any obstacles and challenges in using big data analytics that the stakeholder might face in their way to employ big data for their social innovative solutions. Big data can empower policymakers and entrepreneurs to provide solutions for social problems [6]. Identifying possible challenges and having a clear picture of the big data research in the social sector can also help stakeholders to prepare beforehand, to take advantage of the big data that are available, filter them and proceed to decisions that will help them innovate for social good and change.

3 Research Methodology

A systematic mapping study was undertaken to provide an overview of the research available in the field of big data analytics and social innovation leading to societal transformation, following the standardized process for systematic mapping studies [14] as illustrated in Fig. 1; along with guidelines from [24].

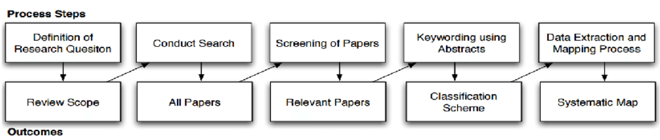

Fig. 1. The Systematic Mapping Study Process [14]

3.1 Data Sources and Search Strategy

In our primary search, we collected papers from all kind of sources including journals, conference papers, books, reports etc. This review was conducted in August 2018 and publications were searched from 2008 and onwards. We selected this timeframe as it is the time when these terms like big data and analytics, social innovation got the momentum. The systematic search strategy consisted of searches in seven online bibliographic databases which were selected based on their relevance with our search topic and these databases are also well known for good quality literature resources in the field. To obtain high-quality data, we searched in the following databases – Scopus, ISI Web of Science, ACM Library, IEEE Xplore, SAGE, Emerald and Taylor & Francis. Then initial searches in the databases were conducted based on identified keywords (Table 1) related to this topic. The used search strings were:

Table 1. The keyword combination for initial search

"Big Data" OR "Data Analytics"	AND	"Social Innovation"
		"Societal Transformation"
		"Social Good"
		"Social Change"

3.2 Study Selection

The study selection process is illustrated in Fig. 2, along with the number of papers at each stage. Searching the databases using the search string returned 593 papers, resulting in 465 unduplicated papers. These were imported into EndNote X8. Due to the importance of the selection phase in determining the overall validity of the literature review, a number of inclusion and exclusion criteria were applied. Studies were eligible for inclusion if they were focused on the topic of big data and data analytics, and their applications to foster social innovation, and lead to social impact, change and transformation. We used "big data" and "data analytics" separate to broader our search as several studies employ big data analytics techniques but do not use the term big data.

The mapping included research papers published in journals, conference proceedings, reports targeted at business executives and a broader audience, and scientific magazines. In progress research and dissertations were excluded from this mapping, as well as studies that were not written in English. Given that our focus was on the social innovation and societal transformation that big data entails, we included quantitative, qualitative, and case studies. Since the topic of interest is of an interdisciplinary nature, a diversity of epistemological approaches was opted for.

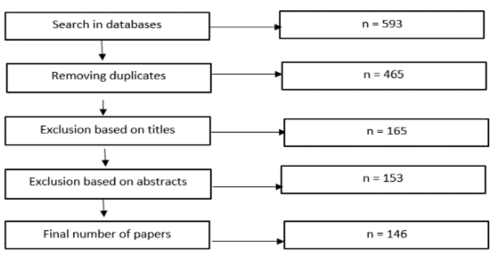

Fig. 2. The study selection process

3.3 Manual Search

Following the systematic search, a manual search was also conducted. Google Scholar was used to searching for papers manually. At this stage total 10 papers from Google scholar was added to our EndNote library and the final number of papers became 156.

3.4 Data Extraction

After the mapping, we finally ended up with 156 papers. We performed a systematic analysis and extracted data from the abstracts of the papers that we need to answer our research questions. We extracted data regarding the - publication frequency, publication source, research area, research type, empirical evidence and contribution type.

4 Results and Findings

RQ1: How the research about ‘big data and social innovation' has changed over time (in the last decade)?

Publication Frequency. The analysis shows that relevant papers are published from 2012 or later, with their frequency increasing yearly. The study was conducted in August 2018, so the year 2018 is not complete. The findings (Fig. 3) verify that the momentum or applications of big data are becoming increasingly popular.

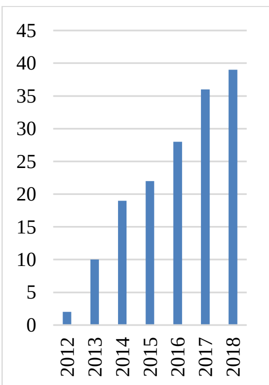

Fig. 3. Publication frequency

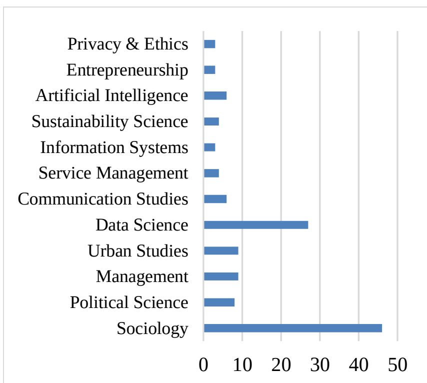

Fig. 4. Research areas

Research Areas. Next, we examined the research sectors of the published articles, to give an overview of the general categories. The findings are shown in Fig. 4.

Publication Sources. As mentioned in section 3, our mapping includes research papers published in academic outlets; but we have considered reports also (e.g., Hitachi reviews) because a lot of evidence is published by companies and a lot of work on social innovation and big data is done by companies as well.

We have tried to figure out how many of the relevant scientific papers are published in journals, how many as conference papers and from other sources. The statistic is given in Fig. 5.

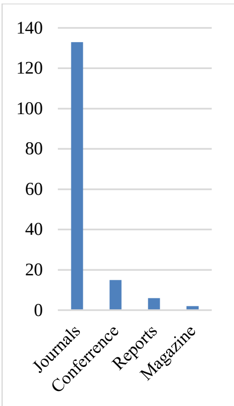

Fig. 5. Sources of publication

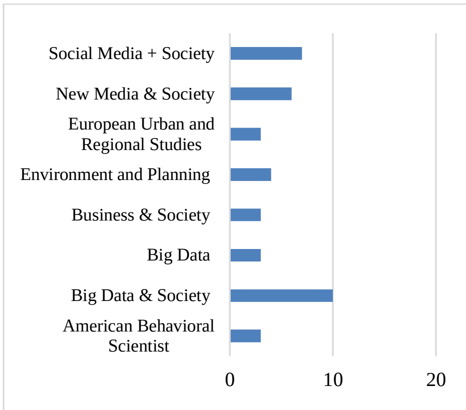

Fig. 6. Journals with a higher number of relevant publications

Along with the sources of the relevant papers, we have also searched for the journals who published maximum number of papers about our research topic i.e. big data and social innovation. Here in Fig. 6, we mention a few journals with maximum number of published papers from our review.

RQ2: How much of the research is done based on empirical studies and what type of empirical studies?

Empirical Evidence. We primarily classified our reviewed papers as empirical and non-empirical papers. Non-empirical papers are conceptual papers. From the study, we see that majority (59%) of the paper s are based on empirical evidence. With this finding (Fig. 7), we also get the answer to our second research question.

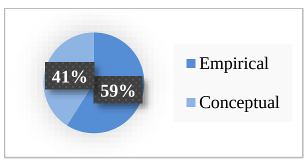

Fig. 7. Empirical and non-empirical ratio

We then classified the empirical papers based on the type of study. The research types that have been assessed followed the guidelines from [25] include: (1) survey, (2) design and creation, (3) experiment, (4) case study, (5) action research, and (6) ethnography. We have also included 'Discussion' as a research type, inspired by [26]. We have added this last method as we felt that some papers are more suitable to categorize as a discussion paper. Discussion papers are also known as 'Expert opinion'.

After deciding about the research types, we counted the numbers for each type. The following figure shows which research types of the studies found from our mapping. Only the papers providing empirical evidence (92 papers) were included in Fig. 8, covering a total of 7 research methods.

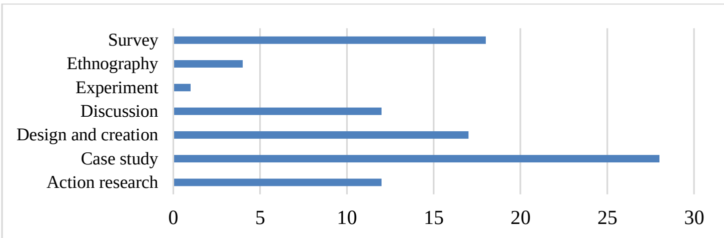

Fig. 8. Empirical evidence (research type)

Contribution Type. Every research paper has some contribution to the advancement of research in the relevant field by providing something new. To illustrate which types of contributions that have been made within the research area between, Fig. 9 was made. The figure shows the contribution type of papers. All 156 primary papers selected finally, in our mapping study are considered in this figure.

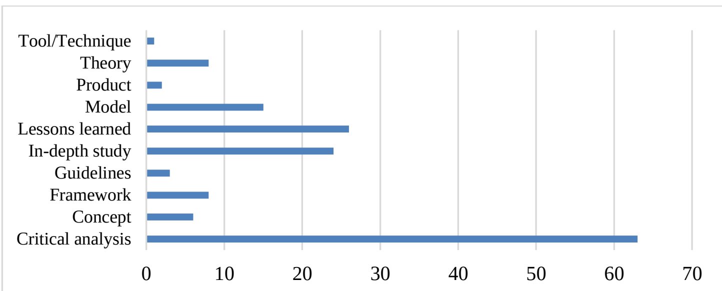

Fig. 9. Contribution types

We differ between 10 contribution types. Based on [25], we define six different knowledge outcomes including (1) product, (2) theory, (3) tool/technique, (4) model, (5) in-depth study and (6) critical analysis. We also adapt some more knowledge outcomes or contribution types since some contribution types from [27] can describe the contribution of some papers more precisely; including (1) framework, (2) lessons learned, (3) tool/guidelines and (4) concept.

RQ3: What are the challenges or barriers to the successful implementation of big data for societal challenges?

Studying the title and abstract of all 156 papers, it has been found that only 3 papers mentioned challenges regarding employing big data in their studies. The challenges we find from this study are mentioned below:

- Open data and privacy concern [28]
- Challenge around obtaining data [29]
- The prominence of marketing-driven software [30]
- The interpretation of unpredictability [30]

So little evidence is not enough to generalize a fact for all and answer a research question like what the challenges or barriers for the successful implementation of big data for societal challenges are. So, we believe there is further research scope on this issue.

5 Discussion

RQ1: How the research about 'big data and social innovation' has changed over time (in the last decade)?

From this mapping we have presented an overview of the research status on big data and social innovation that has been done in the last decade. As we could not find any prior systematic study on this topic, we cannot compare the results. But this mapping will now help other researchers to understand to social potential of big data research and applications. Our study proves that terms like big data and social innovation gained the attention of academic and business communities later than in 2010. It can be also seen that the number of researches and publications are increasing every year since then, which proves the importance and increasing attention big data and social innovation is getting day by day. Another study on big data [8] also stated that, "With regard to the literature review, 'big data' relevant journal articles have started appearing frequently in 2011. Prior to these years, the number of publications on the topic was very low. Publications on 'big data' related topics started only in 2008 (with 1 article) and then a steady increase in the number of publications in the following years".

From this mapping, we can see that many fields including social science, political science, information systems, urban management, communication, healthcare sector adapted big data for their applications. In the results section, we have presented the fields with major number of research studies, but there are also research fields we have found form the mapping where big data is being used; like- education, journalism, tourism, etc. Here notable that all these papers with applications of big data in different fields are directly or indirectly related to various social issues; which proves that big data applications have a big potential to be used for the good of the society not only for business or technology.

RQ2: How much of the research is done based on empirical studies and what type of empirical studies?
In our systematic mapping, more than half of the papers (59%) provide empirical evidence. As there was no previous mapping on this topic, we cannot say how much empirical work was done before. But when 59% of the studies are empirical it proves that the researchers of this field are contributing much. With their contributions, the quality of research is also improving. The major contribution of the research papers from our mapping was a critical analysis, both empirical and non-empirical. When analyzing different topics, the authors also presented their insights, research agenda, guidelines for future research, what lessons they learned and their opinions. The empirical studies also presented models, frameworks and tools that can be used in future research.

RQ3: What are the challenges or barriers to the successful implementation of big data for societal challenges?
In article [28], the authors reflected on various case related to big data challenges, including the challenge of maintaining data privacy and ethics when using all forms of big data for positive social change. The authors recommended exploring new formats for educating people about privacy/data protection risks to overcome data privacy challenges and to use templates to evaluate open data sources. In [29] authors investigate how the challenges around obtaining data to enforce new regulations are addressed by local councils to balance corporate interests with the public good. The authors stated that triangulating different sources of information is not always straightforward as the publicly available data might be partially obscured. In their case study, the authors recommend about platform economy to overcome the challenges regarding data collection. In [30], the authors examine the dominance of marketing-driven commercial tools for predictive analytics of data and their effectiveness to analyze data for completely different purposes such as law enforcement. Another challenge that [30] mentions is, the notions of predictability and probability remain contentious in the use of social media big data. The authors reflected upon the challenges and points to a crucial research agenda in an increasingly datafied environment.

5.1 Use of Keywords

We found some research papers relevant to our study, but they have not been included in the mapping as they do not use the keywords we searched with. For example, [31] use mobile call data to predict the geographic spread and timing of epidemics, and indeed they address a social challenge and has a significant societal impact. However, they do not use keywords regarding data analytics and societal impact, maybe because their focus is mainly on modeling and technical aspects. Instead their keywords include human mobility, mobile phones, epidemiology, dengue etc. Considering the importance of social implications of big data research as well as the interest of publication venues in contributing to societies [1], we suggest that future papers should take into account and report such implications in their abstract and keywords. We should note that indeed many papers discuss social implications, however they do not mention them in their abstracts, raising the need for a systematic literature review in the area. Thus, a more detailed analysis of the research articles can lead, among other things, to new combinations of keywords that will be able to better capture the current status regarding the impact of big data and analytics on societal challenges.

5.2 Limitation of the Study

Systematic mapping study approach is not without limitations [32]. For the validity of this review, threats of retrieval of papers need to be considered. Even though a systematic approach was used during this mapping, the selection of papers dealing with "big data" that we have included was based on our subjective judgment. Another limitation is, we have used only titles and abstracts to extract data. So, the categorizing and data extraction process depends on the quality of the abstracts. ICT-related research publications often do not use structured abstract [33] which results in poor accuracy when classifying papers based solely on abstracts. Also, following the standard procedure of systematic mapping, we could not include research papers in our study which do not have the keywords we searched with; even though some papers might be relevant to our topic.

6 Implication for Research and Practice

This systematic mapping study extends the big data research in several ways. Our work contributes to the social perspective that emphasizes the importance of adoption and applications of big data. This study can guide other researchers of this field to develop their agenda and roadmap for future research. The findings of this research show the type of contributions big data research is making in the industry; based on that future researchers can think what type of contributions we are lacking and make their research agenda on that. In this research, we have also identified the challenges of big data adoption in the social sector. Future researchers can explore more about these challenges and can investigate if there are other challenges. There is possible future research potential to address and propose solutions for these challenges; so that employing big data can be easier and more efficient for the stakeholders.

7 Conclusion

This paper presents findings of a systematic mapping study that researchers, social innovators, social entrepreneurs and all other stakeholders can use to unlock the power of big data for the benefit of the society. We have presented the current status that shows how research into big data and social innovation has increased over the last decade, attracting significant attention from across a wide array of disciplines. We have identified the major research areas where big data is getting significant attention; so future researchers can explore more about the impact of big data in those areas. This study also proves that the empirical ground of research in this field is strong; research is not only limited to case studies, but also other forms of research is being done like action research, critical analysis, designing and creating new products, etc. The key contribution this paper has made is offering the basis for a reflection process among the researchers in this field.

References

1. Pappas I.O., Mikalef P., Giannakos M.N., and Krogstie J.: Big data and business analytics ecosystems: paving the way towards digital transformation and sustainable societies. Inf Syst E-Bus Manage 16: 479, (2018).
2. Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: a systematic literature review and research agenda. *Information Systems and e-Business Management*, *16*(3), 547-578.
3. Gantz J, Reinsel D (2012) The digital universe in 2020: big data, bigger digital shadows, and biggest growth in the far east. IDC iView IDC Anal Future 2012:1–16.
4. Agarwal, R. & Dhar, V.: Editorial—Big data, data science, and analytics: The opportunity and challenge for IS research. Information Systems Research, 25, 443-448 (2014).
5. Pappas I. O., Jaccheri L., Mikalef P. and Giannakos M. N.: Social Innovation and Social Entrepreneurship through Big Data: Developing a Research Agenda. 11th Mediterranean Conference on Information Systems (MCIS) (2017).
6. Desouza, K. C., and Smith, K. L.: Big data for social innovation. Stanford Social Innovation Review, Summer (2014).
7. Phills J. A., Deiglmeier K. and Miller D. T.: Rediscovering social Innovation, Stanford Social Innovation Review 6(4), 34-43, (2008).
8. Wamba S. F., Akter S., Edwards A., Chopin G. and Gnanzou D.: How ‘big data’ can make big impact: Findings from a systematic review and a longitudinal case study, International Journal of Production Economics, Volume 165, Pages 234-246, July (2015).
9. de Camaro Fiorini, P., B.M. Roman Pais Seles, C.J. Chiappetta Jabbour, E. Barberio Mariano and AB. L. de Sousa Jabbour: Management theory and big data literature: From a review to a research agenda. International Journal of Information Management 43: 112-129 (2018).
10. Gunther, W. A., M. H. Rezazade Mehrizi, M. Huysman and F. Feldberg: Debating big data: A literature review on realizing value from big data. The Journal of Strategic Information Systems 26(3): 191-209 (2017).
11. Tseng F. M. and Harmon R. R.: The impact of big data analytics on the dynamics of social change, Technological Forecasting and Social Change, Vol 130, (2018).

12. Cajaiba-Santana, G.: Social innovation: Moving the field forward. A conceptual framework. Technological Forecasting and Social Change, 82, 42-51, (2014).
13. Zicari R. V.: Big data: Challenges and opportunities. AKERKAR, R. (ed.) Big data computing. Boca Raton, FL: Chapman and Hall/CRC, (2014).
14. Petersen K., Feldt R., Mujtaba S., and Mattsson M.: Systematic mapping studies in software engineering. In EASE, volume 8, pages 68–77, (2008).
15. Cuquet M., Vega-Gorgojo G., Lammerant H, Finn R., Umair ul Hassan: Societal impacts of big data: challenges and opportunities in Europe, Cornell University Library, (2017).
16. Gobble M. M.: Big data: the next big thing in innovation, Res. Technol. Manag., 56 (1), pp. 64-66, (2013).
17. Manyika J., Chui M., Brown B., Bughin J., Dobbs R., Roxburgh C., and Byers A. H.: Big Data: The Next Frontier for Innovation, Competition and Productivity, McKinsey Global Institute (2011).
18. Brown B., Chul M., Manyika J.: Are you ready for the era of ′big data′?, McKinsey Q., 4 (2011), pp. 24-27+30-35, (2011).
19. Mulgan G.: The theoretical foundation of social innovation, Social Innovation, Springer, (2012).
20. Mulgan G.: The process of social innovation. Innovations, 1(2), 145-162, (2006).
21. Herrera M. E. B.: Creating competitive advantage by institutionalizing corporate social innovation. Journal of Business Research, 68(7), 1468-1474, (2015).
22. Murray R., Caulier-Grice, J., and Mulgan, G.: The open book of social innovation. London, UK. National endowment for science, technology and the art (NESTA), (2010).
23. Jin X., Wah B. W., Cheng X., Wang Y., Significance and challenges of Big Data Research, Big Data Research, Volume 2, Pg. 59-64 (2015).
24. Kitchenham B.: Procedures for performing systematic reviews. Keele, UK, Keele University, 33:1–26 (2004).
25. Oates B. J.: Researching Information Systems and Computing: Sage Publications Ltd., (2006).
26. Zannier C., Melnik G., and Maurer F.: On the success of empirical studies in the international conference on software engineering, ICSE '06 Proceedings of the 28th international conference on Software engineering, Pages 341-350, (2006).
27. Mathiassen L., Chiasson M., and Germonprez M.: Style Composition in Action Research Publication. MIS Quarterly, 32(2), pp. 347-363, (2012).
28. Taylor L., Cowls J., Schroeder R., Meyer E. T.: Big data and positive change in the developing world. Policy and Internet 6, 418-444, (2014).
29. Ferreri, M. and Sanyal, R.: Platform economies and urban planning: Airbnb and regulated deregulation in London, Urban Studies, Volume: 55 issue: 15 (2018).
30. Dencik L., Hintz A., Carey Z.: Prediction, pre-emption, and limits to dissent: Social media and big data use for policing protests in the United Kingdom. New Media & Society 20, (2017).
31. Wesolowski A., Qureshi T., Boni M. F., Sundsoy P. R., Johansson M. A., Rashed S. B., Engo-Monsen K. and Buckee C.O.: Impact of human mobility on the emergence of dengue epidemics in Pakistan, PNAS September 22, (2015).
32. Kitchenham BA, Budgen D, Pearl Brereton O (2011) Using mapping studies as the basis for further research—a participant observer case study. Inf Softw Technol 53:638–651.
33. Kitchenham BA, Pearl Brereton O, and Owen S.: Length and readability of structured software engineering abstracts. IET Softw 2:37 (2008).

Appendix A2

List of papers from the Systematic Mapping Study

Paper	Year	Published in	Contribution	Method
A. Williams, J., Edgeman, R., 2014. Enterprise self-assessment analytics for sustainability, resilience and robustness. The TQM Journal 26, 368-381.	2014	The TQM journal	Design and creation	Model
Adolf, M.T., Stehr, N., 2018. Information, Knowledge, and the Return of Social Physics. Administration & Society 50, 1238-1258.	2018	Administration & Society	Conceptual	Critical analysis
Agarwal, R., Dhar, V., 2014. Big data, data science, and analytics: The opportunity and challenge for IS research. INFORMS.	2014	Information Systems Research	Conceptual	Critical analysis
Agrawal, D., Budak, C., El Abbadi, A., Georgiou, T., Yan, X., 2014. Big data in online social networks: User interaction analysis to model user behavior in social networks, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 1-16.	2014	Conference paper	Design and creation	Framework
Almirall, E., Wareham, J., Ratti, C., Conesa, P., Bria, F., Gaviria, A., Edmondson, A., 2016. Smart Cities at the Crossroads: New Tensions in City Transformation. California Management Review 59, 141-152.	2016	California Management Review	Discussion/Expert opinion	Critical analysis
Angeler, D.G., Alvarez-Cobelas, M., Sánchez-Carrillo, S., 2018. Sonifying social-ecological change: A wetland laments agricultural transformation. Ecology and Society 23.	2018	Ecology and Society	Action research	Product
Arafah, Y., Winarso, H., 2017. Redefining smart city concept with resilience approach, IOP Conference Series: Earth and Environmental Science, 1 ed.	2017	IOP **Conference** Series: Earth and Environmental Science	Survey	Theory

Arnold, J.D.M., Lafreniere, D., 2017. Creating a longitudinal, data-driven 3D model of change over time in a postindustrial landscape using GIS and CityEngine. Journal of Cultural Heritage Management and Sustainable Development.	2017	Journal of Cultural Heritage Management and Sustainable Development	Design and creation	Model
Arora, P., 2016. The Bottom of the Data Pyramid: Big Data and the Global South. International Journal of Communication 10, 1681-1699.	2016	International Journal of Communication	Conceptual	Critical analysis
Austin, M.J., 2017. Mack Center on Nonprofit and Public Sector Management in Human Service Organizations. Research on Social Work Practice 28, 386-391.	2017	Research on Social Work Practice	Case study	In-depth study
Bakir, V., Feilzer, M., McStay, A., 2017. Introduction to Special Theme Veillance and transparency: A critical examination of mutual watching in the post-Snowden, Big Data era. Big Data & Society 4, 2053951717698996.	2017	Big Data & Society	Conceptual	Critical analysis
Barassi, V., 2016. Contested visions: Digital discourses as empty signifiers from the 'network' to 'big data'. Communication and the Public 1, 423-435.	2016	Communication and the Public	Conceptual	Lessons learned
Bassett, C., 2015. Plenty as a response to austerity? Big Data expertise, cultures and communities. European Journal of Cultural Studies 18, 548-563.	2015	European Journal of Cultural Studies	Conceptual	Critical analysis
Benyounes, M., Vanessa, F.-M., Enrique, F.-M., 2013. Characterizing social response to urban earthquakes using cell-phone network data: the 2012 oaxaca earthquake, Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication %@ 978-1-4503-2215-7. ACM, Zurich, Switzerland, pp. 1199-1208.	2013	ACM **conference** on Pervasive and ubiquitous computing adjunct publication	Case study	In-depth study
Berg, M., 2017. Making sense with sensors: Self-tracking and the temporalities of wellbeing. DIGITAL HEALTH 3, 2055207617699767.	2017	Digital health	Case study	In-depth study

Biahmou, A., Emmer, C., Pfouga, A., Stjepandić, J., 2016. Digital master as an enabler for industry 4.0, Advances in Transdisciplinary Engineering, pp. 672-681.	2016	**Conference** - Advances in Transdisciplinary Engineering	Case study	In-depth study
Blauer, B., 2017. Building the Data City of the Future. The ANNALS of the American Academy of Political and Social Science 675, 151-165.	2017	The ANNALS of the American Academy of Political and Social Science	Discussion/Expert opinion	Critical analysis
Bogomolov, A., Lepri, B., Staiano, J., Letouzé, E., Oliver, N., Pianesi, F., Pentland, A., 2015. Moves on the street: Classifying crime hotspots using aggregated anonymized data on people dynamics. Big Data 3, 148-158.	2015	Journal - Big Data	Design and creation	Model
Boyd, R., Holton, R.J., 2017. Technology, innovation, employment and power: Does robotics and artificial intelligence really mean social transformation? Journal of Sociology 54, 331-345.	2017	Journal of Sociology	Discussion	Critical analysis
Brooker, P., Dutton, W., Greiffenhagen, C., 2017. What would Wittgenstein say about social media? Qualitative Research 17, 610-626.	2017	Journal - Qualitative Research	Conceptual	Critical analysis
Burns, R., 2015. Rethinking big data in digital humanitarianism: practices, epistemologies, and social relations. Geojournal 80, 477-490.	2015	Geojournal	Conceptual	Critical analysis
Calder, B.J., Malthouse, E.C., Maslowska, E., 2016. Brand marketing, big data and social innovation as future research directions for engagement. Journal of Marketing Management 32, 579-585.	2016	Journal of Marketing Management	Discussion/Expert opinion (Commentary)	Critical analysis
Candelieri, A., Archetti, F., Giordani, I., Arosio, G., Sormani, R., 2013. Smart cities management by integrating sensors, models and user generated contents. WIT Transactions on Ecology and the Environment 179 VOLUME 1, 719-730.	2013	Journal - WIT Transactions on Ecology and the Environment	Design and creation	Product

Castelnovo, W., Misuraca, G., Savoldelli, A., 2015. Smart Cities Governance: The Need for a Holistic Approach to Assessing Urban Participatory Policy Making. Social Science Computer Review 34, 724-739.	2015	Social Science Computer Review	Design and creation	Framework
Castillo de Mesa, J., Palma García, M.d.l.O., Gómez Jacinto, L., 2018. Analysis of social innovation on social networking services. European Journal of Social Work, 1-14.	2018	European Journal of Social Work	Case study	In-depth study
Cesario, E., Catlett, C., Talia, D., 2016. Forecasting Crimes Using Autoregressive Models, Proceedings - 2016 IEEE 14th International Conference on Dependable, Autonomic and Secure Computing, DASC 2016, 2016 IEEE 14th International Conference on Pervasive Intelligence and Computing, PICom 2016, 2016 IEEE 2nd International Conference on Big Data Intelligence and Computing, DataCom 2016 and 2016 IEEE Cyber Science and Technology Congress, CyberSciTech 2016, DASC-PICom-DataCom-CyberSciTech 2016, pp. 795-802.	2016	IEEE 14th International **Conference**	Design and creation	Model
Chandler, D., 2015. A World without Causation: Big Data and the Coming of Age of Posthumanism. Millennium 43, 833-851.	2015	Journal - Millennium	Conceptual	Critical analysis
Chandler, D., 2016. How the World Learned to Stop Worrying and Love Failure: Big Data, Resilience and Emergent Causality. Millennium 44, 391-410.	2016	Journal - Millennium	Conceptual	Critical analysis
Chang, N., 2015. Marrying IoT and big data: Are you ready? Hitachi Review 64, 255-258.	2015	Hitachi Review	case study	In-depth study
Chatfield, A.T., Reddick, C.G., 2015. Smart City Implementation Through Shared Vision of Social Innovation for Environmental Sustainability: A Case Study of Kitakyushu, Japan. Social Science Computer Review 34, 757-773.	2015	Journal - Social Science Computer Review	(Theory-building) Case study	Framework

Chavalarias, D., 2016. The unlikely encounter between von Foerster and Snowden: When second-order cybernetics sheds light on societal impacts of Big Data. Big Data & Society 3, 2053951715621086.	2016	Journal - Big Data & Society	Conceptual	Critical analysis
Cohen, M., 2017. Fake news and manipulated data, the new GDPR, and the future of information. Business Information Review 34, 81-85.	2017	Business Information Review	Discussion/Expert opinion	Critical analysis
Couldry, N., Rodriguez, C., Bolin, G., Cohen, J., Volkmer, I., Goggin, G., Kraidy, M., Iwabuchi, K., Qiu, J.L., Wasserman, H., Zhao, Y., Rincón, O., Magallanes-Blanco, C., Thomas, P.N., Koltsova, O., Rakhmani, I., Lee, K.-S., 2018. Media, communication and the struggle for social progress. Global Media and Communication 14, 173-191.	2018	Global Media and Communication	Conceptual	Critical analysis
Cowan, D., Alencar, P., McGarry, F., 2014. Perspectives on open data: Issues and opportunties, Software Science, Technology and Engineering (SWSTE), 2014 IEEE International Conference on. IEEE, pp. 24-33.	2014	IEEE International **Conference**	Conceptual	Guidelines
Criado, J.I., Sandoval-Almazan, R., Gil-Garcia, J.R., 2013. Government innovation through social media. Elsevier.	2013	Government Information Quarterly	Conceptual	Lessons learned
Culhane, D., Fantuzzo, J., Hill, M., Burnett, T.C., 2017. Maximizing the Use of Integrated Data Systems: Understanding the Challenges and Advancing Solutions. The ANNALS of the American Academy of Political and Social Science 675, 221-239.	2017	The ANNALS of the American Academy of Political and Social Science	Action research	Lessons learned
Dalton, C.M., Stallmann, T., 2018. Counter-mapping data science. Canadian Geographer-Geographe Canadien 62, 93-101.	2018	Canadian Geographer-Geographe Canadien	Conceptual	Theory
Davies, W., 2015. The return of social government: From 'socialist calculation' to 'social analytics'. European Journal of Social Theory 18, 431-450.	2015	European Journal of Social Theory	Conceptual	Critical analysis

Dayal, U., Akatsu, M., Gupta, C., Vennelakanti, R., Lenardi, M., 2014. Expanding global big data solutions with innovative analytics. Hitachi Review 63, 333-339.	2014	Hitachi Review	Case study	In-depth study
de Falco, S., Angelidou, M., Addie, J.-P.D., 2018. From the “smart city” to the “smart metropolis”? Building resilience in the urban periphery. European Urban and Regional Studies, 0969776418783813.	2018	European Urban and Regional Studies	Action research	Critical analysis
Dencik, L., Hintz, A., Cable, J., 2016. Towards data justice? The ambiguity of anti-surveillance resistance in political activism. Big Data & Society 3, 2053951716679678.	2016	Big Data & Society	case study	In-depth study
Dencik, L., Hintz, A., Carey, Z., 2017. Prediction, pre-emption and limits to dissent: Social media and big data uses for policing protests in the United Kingdom. New Media & Society 20, 1433-1450.	2017	New Media & Society	Action research	Critical analysis
Deng, H., Bai, W., Chao, L., An, X., 2014. Knowledge management in supporting collaborative innovation community capacity building. Journal of Knowledge Management 18, 574-590.	2014	Journal of Knowledge Management	Survey -- you have the other literature reviews as surveys	Lessons learned
Desouza, K.C., Smith, K.L., 2014. Big data for social innovation. Stanf Soc Innov Rev 2014, 39-43.	2014	Magazine (Stanford Social Innovation Review)	Conceptual	Critical analysis
Duke, S.A., 2017. Classical sociology meets technology: Doing independent large-scope research. Current Sociology, 0011392117702428.	2017	Current Sociology	Conceptual	Lessons learned
Edvardsson, B., Frow, P., Jaakkola, E., Keiningham, T.L., Koskela-Huotari, K., Mele, C., Tombs, A., 2018. Examining how context change foster service innovation. Journal of Service Management.	2018	Journal of Service Management	Conceptual	Framework

Ferreri, M., Sanyal, R., 2018. Platform economies and urban planning: Airbnb and regulated deregulation in London. Urban Studies, 0042098017751982.	2018	Urban Studies	Case study	In-depth study
Field, J.M., Victorino, L., Buell, R.W., Dixon, M.J., Meyer Goldstein, S., Menor, L.J., Pullman, M.E., Roth, A.V., Secchi, E., Zhang, J.J., 2018. Service operations: what's next? Journal of Service Management 29, 55-97.	2018	Journal of Service Management	Conceptual	Guidelines
Fink, A., 2018. Bigger data, less wisdom: the need for more inclusive collective intelligence in social service provision. AI and Society 33, 61-70.	2018	AI and Society	Action research	Critical analysis
Fiorini, R.A., 2017. Would the big government approach increasingly fail to lead to good decision?: A solution proposal. Kybernetes 46, 1735-1752.	2017	Kybernetes	Design and creation	Model
Frey, W.R., Patton, D.U., Gaskell, M.B., McGregor, K.A., 2018. Artificial Intelligence and Inclusion: Formerly Gang-Involved Youth as Domain Experts for Analyzing Unstructured Twitter Data. Social Science Computer Review, 0894439318788314.	2018	Social Science Computer Review	Ethnography	Lessons learned
Gill, C., Rendon, H., Rodriguez, J., 2017. Problem framing in the age of data analytics, Proceedings of the 19th International Conference on Engineering and Product Design Education: Building Community: Design Education for a Sustainable Future, E and PDE 2017, pp. 543-548.	2017	19th International **Conference**, E and PDE 2017	case study	In-depth study
Gillespie, M., Osseiran, S., Cheesman, M., 2018. Syrian Refugees and the Digital Passage to Europe: Smartphone Infrastructures and Affordances. Social Media + Society 4, 2056305118764440.	2018	Social Media + Society	Case study	In-depth study

Givoni, M., 2016. Between micro mappers and missing maps: Digital humanitarianism and the politics of material participation in disaster response. Environment and Planning D: Society and Space 34, 1025-1043.	2016	Environment and Planning D: Society and Space	case study	In-depth study
Gombocz, E.A., 2013. Changing the model in Pharma and Healthcare - Can we afford to wait any longer?, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 1-22.	2013	International Conference on Data Integration in the Life Sciences	Conceptual	Critical analysis
González-Bailón, S., 2013. Social science in the era of big data. Policy & Internet 5, 147-160.	2013	Policy & Internet	Conceptual	Critical analysis
Gulson, K.N., Webb, P.T., 2017. Mapping an emergent field of 'computational education policy': Policy rationalities, prediction and data in the age of Artificial Intelligence. Research in Education 98, 14-26.	2017	Research in Education	Conceptual	Critical analysis
Guo, C., Saxton, G.D., 2017. Speaking and Being Heard: How Nonprofit Advocacy Organizations Gain Attention on Social Media. Nonprofit and Voluntary Sector Quarterly 47, 5-26.	2017	Nonprofit and Voluntary Sector Quarterly	Design and creation	Model
H. Dutton, W., 2014. Putting things to work: social and policy challenges for the Internet of things. info 16, 1-21.	2014	info	Discussion (Critical Assessment)	Critical analysis
Halavais, A., 2015. Bigger sociological imaginations: framing big social data theory and methods. Information, Communication & Society 18, 583-594.	2015	Information, Communication & Society	Conceptual	Critical analysis
Halford, S., Savage, M., 2017. Speaking Sociologically with Big Data: Symphonic Social Science and the Future for Big Data Research. Sociology 51, 1132-1148.	2017	Sociology	Conceptual	Critical analysis

Hara, H., Ishigaki, K., 2012. Overview of Research toward Realization of Intelligent Society. Fujitsu Scientific & Technical Journal 48, 105-109.	2012	Fujitsu Scientific & Technical Journal	Case study	In-depth study
Hayashi, H., Asahara, A., Sugaya, N., Ogawa, Y., Tomita, H., 2015. Spatio-temporal similarity search method for disaster estimation, 2015 IEEE International Conference on Big Data (Big Data), pp. 2462-2469.	2015	IEEE International **Conference** on Big Data (Big Data)	Design and creation	Model
Hilbert, M., 2016. Big data for development: A review of promises and challenges. Development Policy Review 34, 135-174.	2016	Development Policy Review	Survey -- you have the other literature reviews as surveys	Lessons learned
Hitomi, S., Muro, K., 2013. Social innovation through utilization of big data. Hitachi Review 62, 384-388.	2013	Hitachi Review	Case study	In-depth study
Holtzhausen, D., 2016. Datafication: threat or opportunity for communication in the public sphere? Journal of Communication Management 20, 21-36.	2016	Journal of Communication Management	Conceptual/Theoretical	Theory
Housley, W., Procter, R., Edwards, A., Burnap, P., Williams, M., Sloan, L., Rana, O., Morgan, J., Voss, A., Greenhill, A., 2014. Big and broad social data and the sociological imagination: A collaborative response. Big Data & Society 1, 2053951714545135.	2014	Big Data & Society	Conceptual	Critical analysis
Housley, W., Webb, H., Williams, M., Procter, R., Edwards, A., Jirotka, M., Burnap, P., Stahl, B.C., Rana, O., Williams, M., 2018. Interaction and Transformation on Social Media: The Case of Twitter Campaigns. Social Media + Society 4, 2056305117750721.	2018	Social Media + Society	Design and creation	Theory
Jacobson, J., Mascaro, C., 2016. Movember: Twitter Conversations of a Hairy Social Movement. Social Media + Society 2, 2056305116637103.	2016	Social Media + Society	Case study	In-depth study

Janssen, M., van den Hoven, J., 2015. Big and Open Linked Data (BOLD) in government: A challenge to transparency and privacy? Elsevier.	2015	Government Information Quarterly	Conceptual	Critical analysis
Jayachandran, J., 2018. Media Literacy and Education in India During Times of Communication Abundance. Journal of Creative Communications 13, 73-84.	2018	Journal of Creative Communications	Discussion (Commentary)	Critical analysis
Jeffrey, J., Peter, D., David, S.-R., Kemal, A.D., 2017. Big Data, Digitization, and Social Change: Big Data (Ubiquity symposium), Ubiquity %@ 1530-2180, pp. 1-8.	2017	Magazine (Ubiquity)	Conceptual	Critical analysis
Jiang, P., Leng, J., Ding, K., Gu, P., Koren, Y., 2016. Social manufacturing as a sustainable paradigm for mass individualization. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 230, 1961-1968.	2016	Journal of Engineering Manufacture	Conceptual	Concept
Kaika, M., 2017. 'Don't call me resilient again!': the New Urban Agenda as immunology … or … what happens when communities refuse to be vaccinated with 'smart cities' and indicators. Environment and Urbanization 29, 89-102.	2017	Environment and Urbanization	Conceptual	Critical analysis
Kayser, V., Blind, K., 2017. Extending the knowledge base of foresight: The contribution of text mining. Technological Forecasting and Social Change 116, 208-215.	2017	Technological Forecasting and Social Change	Conceptual	Guidelines
Kejriwal, M., Szekely, P., 2018. Knowledge Graphs for Social Good: An Entity-centric Search Engine for the Human Trafficking Domain. IEEE Transactions on Big Data, 1-1.	2018	IEEE Transactions on Big Data	Design and creation	Tool/Technique
Khosla, R., Nguyen, K., Chu, M., 2015. Service personalisation of assistive robot for autism care, IECON 2015 - 41st Annual Conference of the IEEE Industrial Electronics Society, pp. 002088-002093.	2015	IECON 2015 - 41st Annual **Conference** of the IEEE Industrial Electronics Society	Action research	Lessons learned

Krivý, M., 2016. Towards a critique of cybernetic urbanism: The smart city and the society of control. Planning Theory 17, 8-30.	2016	Planning Theory	Conceptual	Critical analysis
Leitch, S., Warren, M., 2016. The Syrian Electronic Army – a hacktivist group. Journal of Information, Communication and Ethics in Society 14, 200-212.	2016	Journal of Information, Communication and Ethics in Society	Ethnography	Lessons learned
Leurs, K., Smets, K., 2018. Five Questions for Digital Migration Studies: Learning From Digital Connectivity and Forced Migration In(to) Europe. Social Media + Society 4, 2056305118764425.	2018	Social Media + Society	Conceptual	Critical analysis
Lin, Z., Yang, L., Zhang, Z.a., 2018. To include, or not to include, that is the question: Disability digital inclusion and exclusion in China. New Media & Society, 1461444818774866.	2018	New Media & Society	Ethnography	Lessons learned
Liu, C.-H., Wang, J.S., Lin, C.-W., 2017. The concepts of big data applied in personal knowledge management. Journal of Knowledge Management 21, 213-230.	2017	Journal of Knowledge Management	Conceptual	Concept
Loader, B.D., Dutton, W.H., 2012. A decade in internet time: The dynamics of the Internet and society. Information Communication and Society 15, 609-615.	2012	Information Communication and Society	Discussion (Critical assessment)	Critical analysis
Loebbecke, C., Picot, A., 2015. Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. Journal of Strategic Information Systems 24, 149-157.	2015	Journal of Strategic Information Systems	Discussion (Viewpoint paper)	Concept
Luka, M.E., Millette, M., 2018. (Re)framing Big Data: Activating Situated Knowledges and a Feminist Ethics of Care in Social Media Research. Social Media + Society 4, 2056305118768297.	2018	Social Media + Society	Conceptual	Critical analysis

Lupton, D., 2017. Digital health now and in the future: Findings from a participatory design stakeholder workshop. DIGITAL HEALTH 3, 2055207617740018.	2017	Digital Health	Conceptual	Lessons learned
Madsen, A.K., 2015. Between technical features and analytic capabilities: Charting a relational affordance space for digital social analytics. Big Data & Society 2, 2053951714568727.	2015	Big Data & Society	Survey (analysis of 8 projects)	Lessons learned
Malthouse, E.C., Haenlein, M., Skiera, B., Wege, E., Zhang, M., 2013. Managing customer relationships in the social media era: Introducing the social CRM house. Journal of Interactive Marketing 27, 270-280.	2013	Journal of Interactive Marketing	Conceptual	Critical analysis
Mannheimer, S., Pienta, A., Kirilova, D., Elman, C., Wutich, A., 2018. Qualitative Data Sharing: Data Repositories and Academic Libraries as Key Partners in Addressing Challenges. American Behavioral Scientist, 0002764218784991.	2018	American Behavioral Scientist	Conceptual	Critical analysis
March, H., Ribera-Fumaz, R., 2014. Smart contradictions: The politics of making Barcelona a Self-sufficient city. European Urban and Regional Studies 23, 816-830.	2014	European Urban and Regional Studies	Conceptual	Critical analysis
Marti, P., Megens, C., Hummels, C., 2016. Data-Enabled Design for Social Change: Two Case Studies. Future Internet 8.	2016	Future Internet	Design and creation	Model
Meng, A., DiSalvo, C., 2018. Grassroots resource mobilization through counter-data action. Big Data & Society 5, 2053951718796862.	2018	Big Data & Society	Action research	Lessons learned
Milan, S., 2015. When Algorithms Shape Collective Action: Social Media and the Dynamics of Cloud Protesting. Social Media + Society 1, 2056305115622481.	2015	Social Media + Society	Conceptual	Critical analysis
Mor Barak, M.E., 2018. The Practice and Science of Social Good: Emerging Paths to Positive Social Impact. Research on Social Work	2018	Research on Social Work Practice	Survey	Model

Practice, 1049731517745600.				
Moumni, B., Frias-Martinez, V., Frias-Martinez, E., 2013. Characterizing social response to urban earthquakes using cell-phone network data: The 2012 oaxaca earthquake, UbiComp 2013 Adjunct - Adjunct Publication of the 2013 ACM Conference on Ubiquitous Computing, pp. 1199-1207.	2013	UbiComp 2013 Adjunct - Adjunct Publication of the 2013 ACM **Conference** on Ubiquitous Computing	Case study	In-depth study
Nambisan, S., 2017. Digital Entrepreneurship: Toward a Digital Technology Perspective of Entrepreneurship. Entrepreneurship Theory and Practice 41, 1029-1055.	2017	Entrepreneurship Theory and Practice	Conceptual	Critical analysis
Nathan, E., 2014. Big data for social good, Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining %@ 978-1-4503-2956-9. ACM, New York, New York, USA, pp. 1522-1522.	2014	Proceedings of the 20th ACM SIGKDD international **conference** on Knowledge discovery and data mining	Conceptual	Lessons learned
Nunan, D., Di Domenico, M., 2013. Market Research and the Ethics of Big Data. International Journal of Market Research 55, 505-520.	2013	International Journal of Market Research	Conceptual	Critical analysis
Öberg, C., Graham, G., Hennelly, P., 2014. Smart cities: A literature review and business network approach discussion on the management of organisations. IMP Journal 11, 468-484.	2014	IMP Journal	Survey	Lessons learned
Obeysekare, E., Marucci, A., Mehta, K., 2016. Developing a lean data management system for an emerging social enterprise, 2016 IEEE Global Humanitarian Technology Conference (GHTC), pp. 54-62.	2016	2016 IEEE Global Humanitarian Technology **Conference** (GHTC)	Case study	In-depth study

Obschonka, M., 2017. The quest for the entrepreneurial culture: psychological Big Data in entrepreneurship research. Current Opinion in Behavioral Sciences 18, 69-74.	2017	Current Opinion in Behavioral Sciences	Conceptual	Lessons learned
Oztemel, E., Gursev, S., 2018. Literature review of Industry 4.0 and related technologies. Journal of Intelligent Manufacturing.	2018	Journal of Intelligent Manufacturing	Survey	Theory
Patrício, L., Gustafsson, A., Fisk, R., 2017. Upframing Service Design and Innovation for Research Impact. Journal of Service Research 21, 3-16.	2017	Journal of Service Research	Design and creation	Concept
Pedersen, J.S., Wilkinson, A., 2018. The digital society and provision of welfare services. International Journal of Sociology and Social Policy 38, 194-209.	2018	International Journal of Sociology and Social Policy	Survey	Critical analysis
Pink, S., Lanzeni, D., Horst, H., 2018. Data anxieties: Finding trust in everyday digital mess. Big Data & Society 5, 2053951718756685.	2018	Big Data & Society	Conceptual	Critical analysis
Pink, S., Sumartojo, S., Lupton, D., Heyes La Bond, C., 2017. Mundane data: The routines, contingencies and accomplishments of digital living. Big Data & Society 4, 2053951717700924.	2017	Big Data & Society	ethnography	Concept
Prasad, S., Shankar, R., Gupta, R., Roy, S., 2014. A TISM modeling of critical success factors of blockchain based cloud services. Journal of Advances in Management Research 15, 434-456.	2014	Journal of Advances in Management Research	Survey	Theory
Priante, A., Ehrenhard, M.L., van den Broek, T., Need, A., 2017. Identity and collective action via computer-mediated communication: A review and agenda for future research. New Media & Society 20, 2647-2669.	2017	New Media & Society	Survey	Lessons learned
Purdam, K., 2014. Citizen social science and citizen data? Methodological and ethical challenges for social research. Current Sociology 62, 374-392.	2014	Current Sociology	Action research	Lessons learned

Ranganathan, S., Nicolis, S.C., Spaiser, V., Sumpter, D.J.T., 2015. Understanding Democracy and Development Traps Using a Data-Driven Approach. Big Data 3, 22-33.	2015	Big Data	Design and creation	Model
Ribes, D., 2018. STS, Meet Data Science, Once Again. Science, Technology, & Human Values, 0162243918798899.	2018	Science, Technology, & Human Values	Conceptual	Critical analysis
Rodríguez, C., Ferron, B., Shamas, K., 2014. Four challenges in the field of alternative, radical and citizens' media research. Media, Culture & Society 36, 150-166.	2014	Media, Culture & Society	Survey	Critical analysis
Rossi, U., Di Bella, A., 2017. Start-up urbanism: New York, Rio de Janeiro and the global urbanization of technology-based economies. Environment and Planning A: Economy and Space 49, 999-1018.	2017	Environment and Planning A: Economy and Space	Conceptual	Critical analysis
Ruppert, E., Law, J., Savage, M., 2013. Reassembling Social Science Methods: The Challenge of Digital Devices. Theory, Culture & Society 30, 22-46.	2013	Theory, Culture & Society	Conceptual	Model
Salvatore, S., Mannarini, T., Avdi, E., Battaglia, F., Cremaschi, M., Fini, V., Forges Davanzati, G., Kadianaki, I., Krasteva, A., Kullasepp, K., Matsopoulos, A., Mølholm, M., Redd, R., Rochira, A., Russo, F., Santarpia, A., Sammut, G., Valmorbida, A., Veltri, G.A., 2018. Globalization, demand of sense and enemization of the other: A psychocultural analysis of European societies' sociopolitical crisis. Culture & Psychology, 1354067X18779056.	2018	Culture & Psychology	Conceptual	Critical analysis
Sarikakis, K., Korbiel, I., Piassaroli Mantovaneli, W., 2018. Social control and the institutionalization of human rights as an ethical framework for media and ICT corporations. Journal of Information, Communication and Ethics in Society.	2018	Journal of Information, Communication and Ethics in Society	Conceptual	Theory

Saxena, S., 2017. Open Linked Statistical Data (OLSD): prospects and issues. The Bottom Line 30, 195-200.	2017	The Bottom Line	Discussion (viewpoint)	Critical analysis
Saxena, S., 2017. Prospects of open government data (OGD) in facilitating the economic diversification of GCC region. Information and Learning Science 118, 214-234.	2017	Information and Learning Science	Survey	Lessons learned
Schebesch, K.B., 2017. Some facilitators and inhibitors of knowledge-based socio-technological transformations, Proceedings of the European Conference on Knowledge Management, ECKM, pp. 872-880.	2017	Proceedings of the European **Conference** on Knowledge Management, ECKM	Conceptual	Critical analysis
Schrock, A.R., 2016. Civic hacking as data activism and advocacy: A history from publicity to open government data. New Media & Society 18, 581-599.	2016	New Media & Society	survey	Concept
Schroeder, R., 2016. Rethinking digital media and political change. Convergence 24, 168-183.	2016	Convergence	Design and creation	Framework
Selinger, E., Hartzog, W., 2015. Facebook's emotional contagion study and the ethical problem of co-opted identity in mediated environments where users lack control. Research Ethics 12, 35-43.	2015	Research Ethics	Conceptual	Critical analysis
Semcow, K., Morrison, J.K., 2018. Lean Startup for social impact: Refining the National Science Foundation's Innovation Corps model to spur social science innovation. Social Enterprise Journal 14, 248-267.	2018	Social Enterprise Journal	Case study	In-depth study
Shimizu, T., Shomura Dr, Y., Masukawa, H., Takeda, Y., 2014. Traffic management solutions for social innovation business. Hitachi Review 63, 51-56.	2014	Hitachi Review	Case study	In-depth study
Shin, D.-H., Jin Park, Y., 2016. Understanding the Internet of Things ecosystem: multi-level analysis of users, society, and ecology. Digital Policy, Regulation and Governance	2016	Digital Policy, Regulation and Governance	case study	Critical analysis

19, 77-100.				
Siapera, E., Boudourides, M., Lenis, S., Suiter, J., 2018. Refugees and Network Publics on Twitter: Networked Framing, Affect, and Capture. Social Media + Society 4, 2056305118764437.	2018	Social Media + Society	Survey	Lessons learned
Smigiel, C., 2018. Urban political strategies in times of crisis: A multiscalar perspective on smart cities in Italy. European Urban and Regional Studies, 0969776418792049.	2018	European Urban and Regional Studies	Conceptual	Framework
Snow, C.C., Håkonsson, D.D., Obel, B., 2016. A Smart City Is a Collaborative Community: Lessons from Smart Aarhus. California Management Review 59, 92-108.	2016	California Management Review	Action research	Framework
Solman, P., Henderson, L., 2018. Flood disasters in the United Kingdom and India: A critical discourse analysis of media reporting. Journalism, 1464884918762363.	2018	Journalism	Survey	Lessons learned
Solovyev, V.D., Bochkarev, V.V., Kaveeva, A.D., 2015. Variations of social psychology of Russian society in last 100 years, Proceedings - 2015 IEEE International Conference on Smart City, SmartCity 2015, Held Jointly with 8th IEEE International Conference on Social Computing and Networking, SocialCom 2015, 5th IEEE International Conference on Sustainable Computing and Communications, SustainCom 2015, 2015 International Conference on Big Data Intelligence and Computing, DataCom 2015, 5th International Symposium on Cloud and Service Computing, SC2 2015, pp. 519-523.	2015	Proceedings - 2015 IEEE International **Conference** on Smart City	Case study	Critical analysis

Stephens, S.H., 2018. A narrative approach to interactive information visualization in the digital humanities classroom. Arts and Humanities in Higher Education, 1474022218759632.	2018	Arts and Humanities in Higher Education	Action research	Lessons learned
Stilgoe, J., 2017. Machine learning, social learning and the governance of self-driving cars. Social Studies of Science 48, 25-56.	2017	Social Studies of Science	Conceptual	Critical analysis
Sung, B., Park, S.D., 2018. Who Drives the Transition to a Renewable-Energy Economy? Multi-Actor Perspective on Social Innovation. Sustainability 10.	2018	Sustainability	Experiment	Model
Swist, T., Collin, P., 2017. Platforms, data and children's rights: Introducing a 'networked capability approach'. New Media & Society 19, 671-685.	2017	New Media & Society	Conceptual	Theory
Tani, S., Fermin, I., Maruyama, Y., Ito, A., 2015. Solution for improving hospital management in Denmark. Hitachi Review 64, 455-461.	2015	Hitachi Review	Case study	In-depth study
Taylor, L., Cowls, J., Schroeder, R., Meyer, E.T., 2014. Big data and positive change in the developing world. Policy and Internet 6, 418-444.	2014	Policy and Internet	Discussion	Lessons learned
Thayyil, N., 2018. Constructing global data: Automated techniques in ecological monitoring, precaution and reification of risk. Big Data & Society 5, 2053951718779407.	2018	Big Data & Society	Conceptual	Critical analysis
Till, C., Haverkamp, J., White, D., Bhaduri, B., 2016. Understanding climate-induced migration through computational modeling: A critical overview with guidance for future efforts. The Journal of Defense Modeling and Simulation 15, 415-435.	2016	The Journal of Defense Modeling and Simulation	Conceptual	Model

Udanor, C., Aneke, S., Ogbuokiri, B.O., 2016. Determining social media impact on the politics of developing countries using social network analytics. Program 50, 481-507.	2016	Program	Case study	In-depth study
Van Dijck, J., Poell, T., 2016. Understanding the promises and premises of online health platforms. Big Data & Society 3, 2053951716654173.	2016	Big Data & Society	case study	In-depth study
Viana Chaves, H., Maia Filho, O.N., 2018. Destinies of the subject in a society almost completely seduced by knowledge. Social Science Information 57, 344-356.	2018	Social Science Information	Conceptual	Critical analysis
Viitanen, J., Kingston, R., 2014. Smart Cities and Green Growth: Outsourcing Democratic and Environmental Resilience to the Global Technology Sector. Environment and Planning A: Economy and Space 46, 803-819.	2014	Environment and Planning A: Economy and Space	case study	In-depth study
Waddell, S., 2016. Societal Change Systems: A Framework to Address Wicked Problems. The Journal of Applied Behavioral Science 52, 422-449.	2016	The Journal of Applied Behavioral Science	Conceptual	Framework
Waddock, S., 2016. Taking Stock of SIM: Social Issues in Management Division of the Academy of Management. Business & Society	2016	Business & Society	Conceptual	Critical analysis
Wahl-Jorgensen, K., 2015. The Chicago School and Ecology: A Reappraisal for the Digital Era. American Behavioral Scientist 60, 8-23.	2015	American Behavioral Scientist	Conceptual	Critical analysis
Weerakkody, V., Kapoor, K., Balta, M.E., Irani, Z., Dwivedi, Y.K., 2017. Factors influencing user acceptance of public sector big open data. Production Planning & Control 28, 891-905.	2017	Production Planning & Control	Action research	Model

Wehn, U., Evers, J., 2015. The social innovation potential of ICT-enabled citizen observatories to increase eParticipation in local flood risk management. Technology in Society 42, 187-198.	2015	Technology in Society	case study	In-depth study
West, S.M., 2017. Data Capitalism: Redefining the Logics of Surveillance and Privacy. Business & Society, 0007650317718185.	2017	Business & Society	Survey	Critical analysis
Williamson, B., 2014. Knowing public services: Cross-sector intermediaries and algorithmic governance in public sector reform. Public Policy and Administration 29, 292-312.	2014	Public Policy and Administration	Survey	Critical analysis
Wilner, A.S., 2018. Cybersecurity and its discontents: Artificial intelligence, the Internet of Things, and digital misinformation. International Journal 73, 308-316.	2018	International Journal	Conceptual	Critical analysis
Wu, J., Guo, S., Huang, H., Liu, W., Xiang, Y., 2018. Information and Communications Technologies for Sustainable Development Goals: State-of-the-Art, Needs and Perspectives. IEEE Communications Surveys & Tutorials 20, 2389-2406.	2018	IEEE Communications Surveys & Tutorials	Survey	Lessons learned
Xu, Y.Y., Gonzalez, M.C., 2017. Collective benefits in traffic during mega events via the use of information technologies. Journal of the Royal Society Interface 14.	2017	Journal of the Royal Society Interface	Action research	Model
Yeung, D., 2018. Social Media as a Catalyst for Policy Action and Social Change for Health and Well-Being: Viewpoint. Journal of Medical Internet Research 20.	2018	Journal of Medical Internet Research	Discussion/Expert opinion (viewpoint)	Critical analysis
Zeng, J.Y., 2018. Fostering path of ecological sustainable entrepreneurship within big data network system. International Entrepreneurship and Management Journal 14, 79-95.	2018	International Entrepreneurship and Management Journal	Conceptual	Lessons learned

Zeng, R., Greenfield, P.M., 2015. Cultural evolution over the last 40 years in China: Using the Google Ngram Viewer to study implications of social and political change for cultural values. International Journal of Psychology 50, 47-55.	2015	International Journal of Psychology	Case study	Lessons learned
Zhu, G., Huang, C., Hu, B., Li, G., 2016. Autonomy in individual behavior under multimedia information. Multimedia Tools and Applications 75, 14433-14449.	2016	Multimedia Tools and Applications	Design and creation	Model

Appendix B

Interview Protocol

Q1. What is your relationship with social innovation? Or how are you involved with social innovation?

Q2. What is your mission?

   a. How do you generate social value?

Q3. Have you used big data or data analytics in your work/organization?

   a. If not, then are you planning to employ big data for your next project?

      i. How likely?
      ii. For what purpose – give an example maybe?

   b. If yes, would you please describe how you are using them?

      i. How this impact the process you follow to achieve your mission?
         1. What process do you follow?
            Give an example maybe of how you reach your mission and generate social value?
         2. Does this impact your decision making?
            a. Impacts the way/process decisions are made?
            b. Impact the decisions themselves?
      ii. How this impact the cooperation within your company/group/etc.
      iii. How this impact the cooperation with your customers/citizens?
      iv. What kind of tools have you used to support your work?
         1. How the tools you chose have influenced your work?

Q4. In your opinion, what challenges you face or might face if you use big data for your projects?

Q5. What benefits would ensue if your organization implement some form of big data or data analytics?

Q6. Do you foresee any negative aspects of using big data for generating social value?

   a. Any issue with privacy, trust or ethics?

Appendix C

Employing big data and data analytics in addressing societal challenges

Dear Participant,

You have been invited to share your knowledge and experience with us about the use of data and data analytics to address societal challenges and social good.

This survey is a part of a master's thesis. From this survey, we aim to collect information about the drivers, enablers, and challenges of using big data analytics in the intersection of social innovation and social entrepreneurship. Before starting the survey here we present a short description of the key terms -

Big data and data analytics: Big data refers to data sets that extend beyond single data repositories like databases or data warehouses. It has three main characteristics: the data itself, the analytics of the data, and the presentation of the results of the analytics.

Social innovation: Social innovation refers to innovative activities and services that are motivated by the goal of meeting a social need and that are predominantly diffused through organizations whose primary purposes are social.

Societal challenges: Societal challenges refer to problems that people have interacting with people in society or engaging in normal social behaviors. Examples of societal challenges may include climate change, public health, the aging population, energy security, food security, transportation problem, and many others. By solving societal challenges, we mean to help the community finding out ways to solve this kind of social challenges by using data analytics.

Thank you for participating in our research.
We will treat your data anonymously according to GDPR.

Declaration:
Hereby, the research team declares that the collected data will be stored safely, it will be handled anonymously and will not be given to any third parties not involved in the research project.

Contact info of researchers:
Farzana Quayyum (farzanaq@stud.ntnu.no (mailto:farzanaq@stud.ntnu.no))
Ilias O. Pappas (ilias.pappas@uia.no (mailto:ilias.pappas@uia.no))

* Required

1. What is your age group? *

- [ ] 18 - 24
- [ ] 25 - 34
- [ ] 35 - 44
- [ ] 45 - 54+

2. What is your gender? *

- ○ Male
- ○ Female
- ○ Other
- ○ Prefer not to say

3. What is your occupation or job title? *

Enter your answer

4. Which country do you live? *

Enter your answer

5. Do you have previous experience with the followings? *

	Far below	Moderately below	Slightly below	Meet expectations	Slightly above	Moderately above	Far above
Social innovation and/or social entrepreneurship							
Big data and/or data analytics							

6. Based on my experience and/or current role in the organization, the following are descriptive of me: *

	Not at All	Moderately disagree	Slightly disagree	Neutral	Slightly agree	Moderately agree	Strongly agree
I am adopting a mission to create social value (not just private value).							
I am recognizing new opportunities to serve my mission.							
I am engaging in a process of continuous adaptation related to my mission.							
I am acting boldly without being limited by resources currently in hand in the fulfillment of my mission.							

I am relentlessly pursuing new opportunities to serve my mission.							
I am caring deeply about the outcomes created by the fulfillment of my mission.							
I seek to be a 'world changer' through the accomplishment of my mission.							

7. Based on my experience and/or current role in the organization, the following are descriptive of me (question 6 continues): *

	Not at all	Moderately disagree	Slightly disagree	Neutral	Slightly agree	Moderately agree	Strongly agree
I am adopting a mission to sustain social value (not just private value).							
I am engaging in a process of continuous innovation related to my mission.							
I am exhibiting a heightened sense of accountability to the constituencies served by my mission.							

I am engaging in a process of continuous learning related to my mission.							

8. From a literature review study, we have found that organizations face multiple challenges when employing big data and data analytics in their organizations. Here we have listed down some challenges. In your opinion, how much challenging each of them is? *

	Not at all	Moderately below	Slightly below	Average	Slightly above	Moderately above	Far above
Lack of understanding of how to leverage data analytics for business value.							
Lack of management bandwidth or inadequate staffing due to competing priorities.							
Lack of skills within the line of business.							
Accessibility of data is difficult.							
Perceived costs may outweigh projected benefits.							
Existing culture does not encourage the sharing of information.							
Big data needs huge data storage capacity							

Maintaining data privacy and security is challenging.	◯	◯	◯	◯	◯	◯	◯
Analytical challenges with big unstructured data	◯	◯	◯	◯	◯	◯	◯
Scalability problem with big data	◯	◯	◯	◯	◯	◯	◯

9. From a literature review study, we have found that organizations face multiple challenges when employing big data and data analytics in their organizations. Here we have listed down some challenges. In your opinion, how much challenging each of them is? (question 8 continues) *

	Not at all	Moderately below	Slightly below	Average	Slightly above	Moderately above	Far above
Current database software lacks in-database analytics.	◯	◯	◯	◯	◯	◯	◯
Architecting big data analytic system is difficult.	◯	◯	◯	◯	◯	◯	◯
Quality of available data	◯	◯	◯	◯	◯	◯	◯
Ownership of data is not clear, or governance is ineffective	◯	◯	◯	◯	◯	◯	◯
Proper interpretation of data is challenging (especially social media data)	◯	◯	◯	◯	◯	◯	◯

Lack of sponsorship for big data and data analytics projects.							
Lack of compelling business case to use big data.							
Cannot make big data usable for end users.							
Don’t know how and where to start with big data and data analytics.							
No need to use big data or data analytics to change the organization.							

10. We have also found some enablers of big data and data analytics from the literature. In your opinion how much important each of these enablers is to help or motivate using big data and data analytics? *

	Not at all	Moderately below	Slightly above	Average	Slightly above	Moderately above	Far above
Increased data storage capacity							
Data availability from various sources							
New set of analytics tools designed specifically to analyze large amounts of data							

General Data Protection Regulation (GDPR) as a pathway to ensure better processing of data and the rights of the data subject							
Rapid growth of data from data-intensive sensor network applications							
Adaptability of a culture that brings together the power of technology and culture of embracing changes to improve the organization							

11. Please rate how likely it is that you would want to use big data and data analytics in your organization to: *

	Not at all	Moderately disagree	Slightly disagree	Neutral	Slightly agree	Moderately agree	Strongly agree
Pursue a high-risk opportunity that has the possibility of very high profits							
Grow the firm to be very large and profitable							
Pursue profit maximization above all other objectives							

Become a major, globally-recognized corporation							
Generate high profits over many years							
Locate the business at a place that suits your personal preferences							
Enjoy the lifestyle and benefits of an independent business owner							
Create a business around your personal hobbies or special interests							
Have great flexibility to decide your work hours, your product lines, and so on							
Be your own boss and make all the important decisions yourself							

12. Please rate how likely it is that you would want to use big data and data analytics in your organization to: (question 11 continues) *

	Not at all	Moderately agree	Slightly agree	Neutral	Slightly agree	Moderately agree	Strongly agree
Gain great satisfaction because you are helping others who are in need							

Solve social and economic problems that cause others to suffer

Help poor people get enough food, clothing, shelter, and medical assistance

Serve as a volunteer to help people who have social and/or economic problems

Help underprivileged people achieve what they are unable to achieve on their own

Submit