Big Data Challenges in Emergency Management

What challenges do emergency management organizations face when implementing Big Data technology in their operations during a crisis?

ØISTEIN SYVERSEN FONGAARD &
THOMAS FARSTAD NESTAAS

SUPERVISOR
Devinder Thapa

University of Agder, 2019
Faculty of Social Sciences
Department of Information Systems

Master

Preface

This study is conducted as the finalizing thesis in the Masters degree of Information Systems at the Department of Information Science at the University of Agder. It is conducted in the time period of January 2019 until June 2019.

The purpose of the study has been to study the challenges emergency management organisations are facing when implementing Big Data technology based tools in their operations during a crisis.

We would like to thank our advisor, Professor Devinder Thapa at the Department of Information Science, for constructive feedback and support. We really appreciate all the feedback and advices, without it the study would not have been compiled.

We would also like to thank the Centre for Integrated Emergency Management at the University of Agder for the support throughout the project, as well as all of the informants that have participated, allocated their schedules, and shared their knowledge and opinions. The study would have been impossible to conduct without these contributions.

Kristiansand

3th June 2019

Øistein Syversen Fongaard

Thomas Farstad Nestaas

Abstract

The use of Big Data technologies has been more common in recent years, spreading across other fields than just the business sector. Emergency management organisations are seeking to utilise technologies that can improve their response and situational understanding during crises and disasters. Big Data and Big Data Analysis are working its way into the field of emergency management as well. Innovation and new technology results in new opportunities that need to be explored, but it is just as important to be aware of its challenges and limitations. This study will address the challenges emergency management organisations that use Big Data technology are facing. The following research question was formulated for the study: *What challenges do emergency management organisations face when implementing Big Data technology in their operations during a crisis?* To answer this question, we have conducted a qualitative study where we have interviewed informants in emergency management organisations and relevant research facilities. These informants have insight in how Big Data is used in emergency management, with its subsequent challenges. A total of 16 informants were interviewed in semi-structured interviews.

The literature review reveals that the challenges with implementing Big Data in organisations can be divided into data challenges, process challenges and management challenges. The literature review provided us with a basis for our study, and the comparison of literature with the findings from the data collection resulted in a revised model presenting Big Data challenges in emergency management.

Findings from the study concludes with a conceptual classification of Big Data challenges in emergency management. The data and process challenges revealed in the literature review were kept. Data challenges include *volume, velocity, variety, variability, veracity, visualisation,* and *value*. Process challenges include *data acquisition & warehousing, data mining & cleansing, data aggregation & integration, analysis & modelling,* and *data interpretation*. The management challenge of *data ownership* is not considered as a challenge in emergency management, because the population are willing to sacrifice such rights in the course of a disaster where lives are at stake. This leaves the management challenges of *privacy, security, data governance, data & information sharing*, and *cost/operational expenditures*. In addition, the study results in three new challenges being added to the existing. The first, *data access,* is a data challenge which refers to problems relating to being unable to collect a necessary volume of information. The second, *data neutrality,* is a process challenge that refers to that no information is objective, it is always tainted and there is always someones data collection behind it. The third, *time pressure* is a management challenge that refers to the speed imperative that occurs by the nature of disasters and its connection the risk of losing lives.

We provide four recommendations on how practitioners can overcome some of the mentioned challenges. We recommend organisations using Twitter data as their data source to look into data quality assurance tools such as TweetCred to overcome challenges relating to data veracity. Challenges relating to data access caused by communication networks being knocked out, can be used as information in nature disasters like typhoons or earthquakes, in the way that the areas not producing any data probably are the most critical areas affected by the disaster. To handle challenges related to data and information sharing we recommend that emergency management organisations seek towards common data standards such as Humanitarian Exchange Language (HXL). In addition, we recommend emergency management organisations to develop techniques to provide fast and easy-to-read

visualisations of the analysed data that is suited to their personnel in the field, to maintain a data driven crisis response.

Suggestions for future research includes an empirical evaluation of the concluding conceptual classification of Big Data challenges in emergency management provided in this study. Further, an analysis of the Big Data challenges related to the different terms of emergency, disaster and crisis to see if there is any differences between the situations, as well as differences in other phases of a disaster, where this study only has examined the response phase of the disaster.

Table of Contents

List of Tables

List of Figures

1.0 Introduction

Disasters, terror, epidemics and refugee crises are situations with the need of accurate usage of resources, and coordination of emergency personnel and emergency actors. These are situations that threatens values and weakens one's ability to carry out important functionality (Eriksen, 2011, s. 13). Big Data technologies provide the possibility for emergency management organisations to act predictable, increase situational awareness, as well as administer social benefits in the event of a crisis. This can be achieved through analysis of large datasets from diverse sources, from geolocations to social media (Watson, Finn & Wadhwa, 2017). Existing research on Big Data challenges in general suggests that there should be conducted research on Big Data challenges in specific associated areas, as Sivarajah, Kamal, Irani, & Weerakkody, (2016) presents:

> It would be valuable to expand the scope of the subject area and to repeat this exercise to identify and draw links with established theoretical contributions in other different associated areas. A publication based on such analysis would provide an extremely valuable platform for the BD and BDA research and practitioners' community. (Sivarajah et al., 2016, p. 280)

The research endeavoured by Sivarajah et al. (2016) and the lack of research on Big Data challenges in emergency management revealed a gap which gave us the motivation for this study. To address the Big Data challenges in emergency management we formulated the following research question:

What challenges do emergency management organisations face when implementing Big Data technology in their operations during a crisis?

To answer the research question we have conducted a qualitative study, where the data collection has been conducted through semi-structured interviews with informants in the emergency management organisations Standby Task Force (SBTF), ACAPS, Direct Relief, Médicins Sans Frontiéres (MSF), the emergency medical communication centre in Southern Norway, Red Cross and Kristiansand Municipality, as well as the research facilities Centre for Integrated Emergency Management (CIEM), Western Norway Research Institute, Qatar Computing Research Institute (QCRI) and the College of Information Science & Technology (IS&T) at UNO.

Through our literature review we gained insight in what challenges Big Data implementations convey. We observed that there exist literature explaining Big Data challenges in a broad sense, not specific to emergency management. Sivarajah et al. (2016) presents a conceptual classification of Big Data challenges, dividing the challenges into data, process, and management challenges. Data challenges concerns the characteristics of the data itself, process challenges refers to how the data is processed and analysed, from capturing the data to interpreting and presenting the end results, while management challenges covers how the data is accessed, handled and governed (Sivarajah et al., 2016, p. 269-274). We used the conceptual classification of Big Data challenges as a basis for our study, and created a revised set of challenges in the context of the specific field of emergency management. Our findings concluded that all but one of the existing challenges were kept, while three new challenges emerged while adding the context of emergency management.

We consider this study as relevant to the study direction of Information Systems, since the study not only focuses on the technical aspects of the Big Data technology, but also the interaction between the technology and the people that are applying it. One of the ambitions of the study is that it is relevant to the research facility CIEM at the University of Agder. The research center conducts research on how new developments within information and communication technology offers new opportunities for more effective emergency and crisis management through the integration of information from a variety of data sources (CIEM, n.d.). Their work can be associated with the research field of crisis informatics, which is the research field that this study is contributing to.

1.1 Disposition

This part of the introduction chapter will introduce the remaining chapters of the study, and acts as a more detailed overview as to how each chapter is described compared to the table of contents.

Chapter 2: Theoretical Background
Presents existing research and literature on relevant fields to our study. The chapter is divided into concepts (2.1) and literature review (2.2). The following concepts is elaborated: *Big Data, Emergency* and *Emergency management,* while the literature review (2.2) presents how the literature were conducted and its results.

Chapter 3: Research Approach
Explains the methodology used in the study, where the research perspective is presented first (3.1). Thereafter the research strategy (3.2) and research design (3.3) is presented. The research design includes information on selection of informants (3.3.1), how the data collection was conducted (3.3.2) and how the data analysis was carried out (3.3.3). Lastly, the criteria for evaluating the data quality is presented (3.4).

Chapter 4: Research Context
Presents the context of which the study is conducted. This includes the organisations that the informants are associated with. The organisations is divided into emergency management organisations (4.1) and research facilities (4.2).

Chapter 5: Results
Presents the results from the data collection, ordered in the different classifications of challenges originating in Sivarajah et al. (2016). These are divided into data challenges (5.1), process challenges (5.2) and management challenges (5.3). In addition to the existing challenges, the new challenges that emerged under the data collection is presented (5.4).

Chapter 6: Discussion
Reviews the results against the literature, to drive them towards a conclusion. The discussion is ordered into two parts, first the theoretical contributions (6.1) which is organised by the classifications of challenges; data challenges (6.1.1), process challenges (6.1.2), management challenges (6.1.3) as well as new challenges emerged (6.1.4). Recommendations on how practitioners can overcome the mentioned challenges is then presented (6.2).

Chapter 7: Conclusion and Implications
Answers our research questions, thereafter discussing implications for the research, before suggestions for future research is provided.

2.0 Theoretical Background

The existing theory, research and literature is presented in this chapter. This creates the foundation on which the study is based upon and is referred to throughout the report. The chapter is divided into two parts: theoretical concepts describing terms and notions used extensively in the study, then the presentation of the literature that has been conducted prior to the study.

2.1 Concepts

This part covers the concepts, notions, terms used throughout the thesis, and will define them for a greater understanding of the study conducted. The existing research on the different concepts within our problem domain is presented and described, which includes: *Big Data, emergency* and *emergency management*.

Big Data

Big Data is a process that facilitates the decision making, through analysis of large amount of data of different types, from a variety of sources, to produce a stream of knowledge (Fertier, Barthe-Delanoë, Montarnal, Truptil & Benaben, 2016, p. 4; Power, 2014). Among research and academic communities, public organisations and the business sector the Big Data term has been known for a good while: in 2001 Gartner defined Big Data as *"*high volume, high velocity and high variety*"* (Laney, 2001; Meier, 2015). These three characteristics are known as the 3V's (Sivarajah et al., 2016, p. 269). Volume refers to the sheer amount of data that is collected (Laney, 2001, p. 1), e.g. large data-sets consisting of terabytes, petabytes, zettabytes of data - or even more (Sivarajah et al., 2016, p. 269). Velocity means the high speed at which the data is produced, where the high speed often is a result of real-time data, such as monitoring of cyber security. The variety refers to data heterogeneity (Gudivada, Jothilakshmi & Rao, 2015, p. 1). From a historical perspective, mass-scale computing has been conducted in a long time. The recent 3V's definition is due to recent exponential increases in telecommunication bandwidth that connects a network of centralised and decentralised data storage systems, which are processed thanks to digital computational capacities (Hilbert, 2015, p. 136).

There are other publications characterizing Big Data with additional terms, such as the 4V's (volume, velocity, variety, and variability) of data (Liao, Yin, Huang & Sheng, 2014; Sivarajah et al., 2016, p. 269) and 6V's (volume, velocity, variety, veracity, variability and value) of data (Gandomi & Haider, 2015; Sivarajah et al., 2016, p. 269).

Organisations are investing heavily in Big Data technologies to utilise its many possibilities. Considering that Big Data is associated with other technologies, like machine learning and artificial intelligence, it is certainly of interest by many. Big Data technologies requires a cross-collaborative team with IT representation to evaluate the opportunities they present and what they require to be deployed. Altimeter Digital Transformation Survey show that Big Data is the fourth prioritised technology investment planned for 2019 (Solis, 2018, p. 24).

Emergency

An emergency is a situation that may cause harm to the population or damage to property. It is a natural or man-made event (Al-Dahash, Thayaparan and Kulatunga, 2016, p. 1193). The term is used different and sometimes interchangeably: *"*Emergency is sometimes used interchangeably with the term disaster, as, for example, in the context of biological and technological hazards or health emergencies" (UNISDR, 2017). The United Nations defines a disaster as the following:

> A serious disruption of the functioning of a community or a society at any scale due to hazardous events interacting with conditions of exposure, vulnerability and capacity, leading to one or more of the following: human, material, economic and environmental losses and impacts. (UNISDR, 2017)

The term disaster is often described differently by the various scholars because of the system by which they were explained and based on their causes and consequences. Emergency and disaster are also terms used interchangeably with the term crisis (Al-Dahash et al., 2016, p. 1191-1192). A crisis is defined as "a disruption that physically affects a system as a whole and threatens its basic assumptions, its subjective sense of self, and its existential core" (Pauchant & Mitroff, 1992, p. 15).

The literature presents some differences between the terms of disaster, crisis and emergency. "Frequently, these terms are used interchangeably, but they actually could mean three different things" (Al-Dahash et al., 2016, p. 1198). To better illustrate the differences Al-Dahash et al. (2016) created a diagram based on their systematic review of literature, as seen in Figure 1.

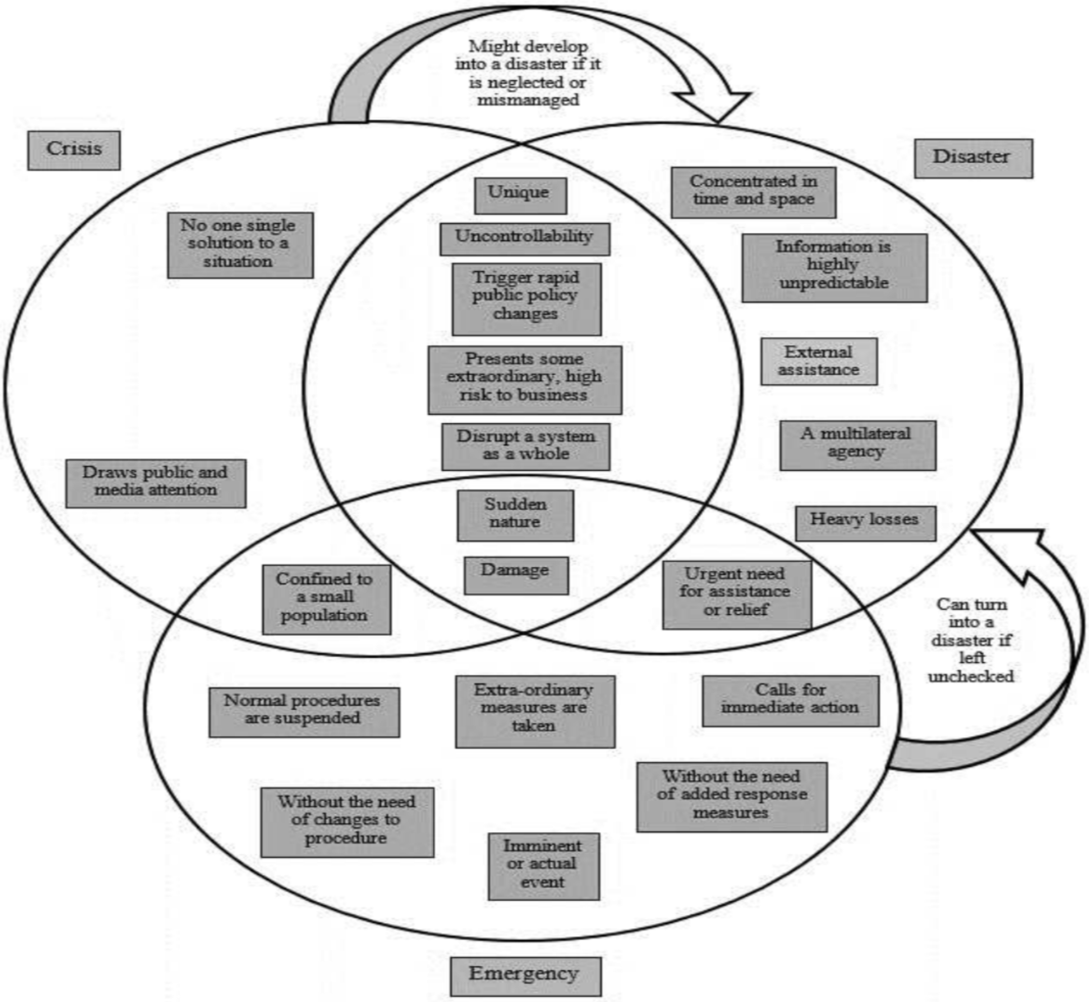

Figure 1: Overlapping of the terms disaster, crisis and emergency (Al-Dahash et al., 2016, p. 1197)

During this study the terms crisis, disaster and emergency will be used interchangeably.

There are four main stages in a disaster: preparation, response, recovery and prevention/mitigation. This is referred to as the disaster management cycle and describes the process by which individuals, communities and organisations prepare for, respond to and recover from extreme events (Wood, Boruff & Smith, 2013, p. 149). This study will focus on the challenges emergency management organisations are facing when implementing Big Data technology based tools into their operations during the response stage of the disaster.

Emergency management
Disaster management, crisis management and emergency management are often used interchangeably and largely overlap in their definition. United Nations Office for Disaster Risk Reduction (UNISDR) define disaster management as *"The organisation, planning and application of measures preparing for, responding to and recovering from disasters",* and also states that *"Emergency management is also used, sometimes interchangeably, with the term disaster management"* (UNISDR, 2017). Emergency management is described as a systematic process with primary aim to reduce the negative consequences and effect of disasters, hence safeguarding people and social infrastructure (Arslan, Roxin, Cruz & Ginhac, 2018, p. 2; Norris et al., 2015, p. 2). Natural disasters, terrorism, epidemics and migrant crises are all instances that require precise management and coordination of emergency personnel and emergency operators. These are situations that threaten important values and/or weakens one's ability to perform important functions (Eriksen, 2011, p. 13).

Emergency management includes several activities as the definition of disaster management states: *"the integration of all activities required to build, sustain and improve the capabilities to prepare for, respond to, recover from, or mitigate against a disaster"* (Arslan et al., 2018, p. 2; Norris et al., 2015, p. 2). The activities of prevention and preparedness focus on risk management, while response and recovery focus on crisis management. These activities are not independent and sequential as the phases of response and recovery are initiated instantaneously, whereas recovery operations can go on for months. The ultimate success of emergency management activities is influenced by the data collected during the preparedness and prevention phases, and effective emergency management can be achieved by reducing the uncertainty in the information about possible states of affected objects, external influences, as well as the interdependence within the system (Arslan et al., 2018, p. 2; Velev & Zlateva, 2011, p. 1).

A growing need for preparedness for emergency response has emerged over many years, the risk of natural disasters has increased due to progressively changing weather patterns in the recent decade. Incidents in the early 20th century already highlighted limitations to response systems, which makes effective emergency response challenges for the emergency management. Modeling, simulation and visualisation techniques can help address many of the challenges brought forth by the need for emergency response preparedness (Jain & McLeal 2003, p. 1).

During our endeavour researching the Big Data Challenges we will mainly focus on emergency- and crisis management, but there will definitely be aspects of disaster management as well. We will not differentiate between the terms for the challenges we may discover in this study. This is because the informants, both researchers and practitioners, will have had experiences across the terms and could use them interchangeably. Some may have experienced a crisis developed into a disaster because of mismanagement, and other may have experienced an emergency left unchecked. It is argued that an emergency is a situation which may be an impending crisis (Al-Dahash et al., 2016, p. 1198). Considering their interchangeability, but with some different meaning, we do not see their differentiation a critical point of this study, but rather a limitation.

2.2 Literature Review

The literature review was undertaken to provide insight to what Big Data challenges that have been identified in the litterature, and to what can be of relevance to the field of emergency management. This led us to a gap in the literature which proves a need for research on Big Data challenges in emergency management. Results from the literature review will describe the challenges uncovered with a general sense that can be applied to the context. The results from the literature review is discussed up against the results from the data collection in chapter 6.0 Discussion. This chapter explains how the literature review was conducted. The results from the literature review is presented in 2.2.1 Data Challenges, 2.2.2 Process Challenges and 2.2.3 Management Challenges. The publications identified through the literature review is based mainly on findings from the Scopus and Google Scholar database. The library database at University of Agder, Oria, was also used. Relevant publications and literature have been identified by going through the reference list in relevant books and publications. In addition, publications has been identified through expert advises, where professors and associates were able to present us with relevant articles and publications.

The approach of the literature review is based on the model presented by Okoli and Schabram (2010), including the following stages:

- *Planning:* Includes the purpose of the literature review as well as protocol and training. The purpose of the review is to gain insight into existing research and publications on the field. Training on literature reviews were conducted through the course IS-420-1: Current Topics and Research Areas in Information Systems at University of Agder, Kristiansand, Norway.
- *Selection:* Includes searching the literature and the practical screen. For the practical screen of the selection of literature, the important criteria were that the content should include specific Big Data challenges and/or the emergency management context.
- *Extraction:* Includes the quality appraisal and data extraction. To ensure the quality appraisal when extracting data from the literature it was important that the literature discussed challenges of relevance in the context of emergency management. Extraction of the data has been done through a concept matrix. In the concept matrix the challenges presented in Sivarajah et al. (2016) is included together with the context of emergency management.
- *Execution*: Includes analysis and findings, as well as writing the literature review. The findings is based on the challenges Sivarajah et al. (2016).

Table 1 presents the criteria the search was based upon. For the search certain keywords were used together with the subject areas. This generated 340 results, by which we thereafter used our inclusion and exclusion criteria to sort out relevant papers. We sorted out the articles which did not have Big Data challenges and the emergency management context, thereafter excluded the articles which only covered emergency management challenges and those which were too technical. We did conclude include some articles without the emergency management context because of their content of relevance.

Search Library	Scopus, Google Scholar, Oria
Keywords	Big Data, Big Data Technology, Emergency Management, Crisis Management, Disaster Management, Crisis Response, Challenges, Issues
Subject Areas	Computer science and information systems, information systems social science, emergency management, crisis response, crisis informatics
Total Search	340
Selected and Reviewed	30
Language	English
Inclusion	Big Data challenges, emergency management context
Exclusion	Emergency management challenges, very technical articles

Table 1: Literature Review Process

Many case studies reveals how there are huge beneficial value in Big Data if it is used correctly, but through the evolution of Big Data there have been uncovered a wide set of challenges related to the technology. There are research challenges abound, ranging from heterogeneity of data, inconsistency and incompleteness, timeliness, privacy, visualisation, and collaboration, to the tools ecosystem around Big Data (Jagadish et al., 2014, p. 86). Sivarajah et al. (2016, p. 265), presents a systematic literature review that summarises the challenges of Big Data in a conceptual classification based on data life cycle:

- *Data Challenges:* Which relate to the characteristics of the data itself.
- *Process Challenges:* Which related to a series of "how" techniques, e.g. how to capture data, how to integrate data, how to transform data, how to select the right model for analysis and how to provide results.
- *Management Challenges:* Which relates to e.g. privacy, security, governance and ethical aspects.

The challenges based on the data life cycle is presented in figure 2 below.

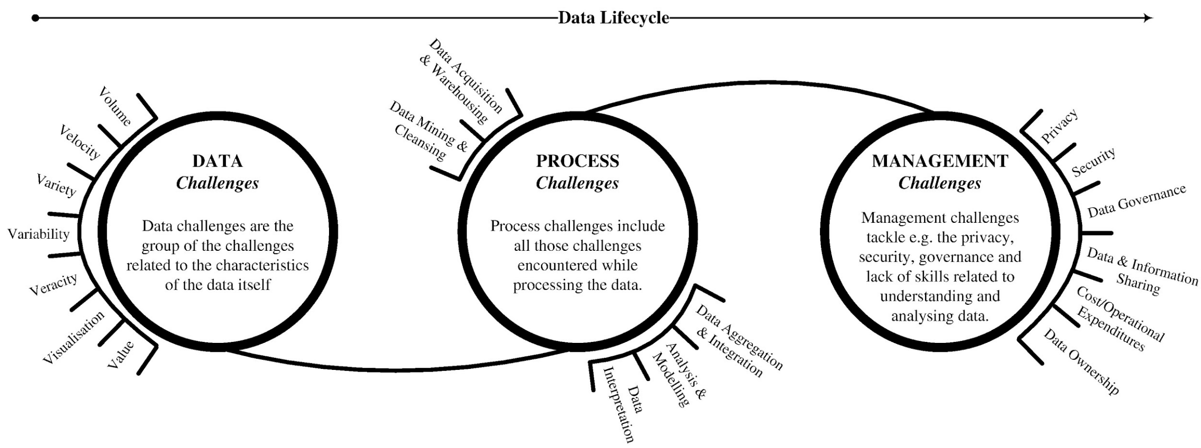

Figure 2: Conceptual classification of BD challenges (Sivarajah et al. 2016, p. 265)

Misconceptions in literature reviews often include a large focus on authors and publishers rather than the concepts in the articles (Webster & Watson, 2002). In order to avoid such misconceptions and focus towards the concepts in the literature, a concept matrix is presented. The publications are presented in the rows while the concepts, in this case the challenges with the use of Big Data, are presented in the columns. The literature review results is based on the conceptual classification of Big Data challenges model, and has divided the challenges into each data life cycle stage. To present an overview of the study's context of emergency management, a separate column for the context is provided.

	Big Data Challenges in Emergecy Management																			
		Data Challenges							Process Challenges					Management Challenges						Context
	Articles	Volume	Velocity	Variety	Variability	Veracity	Visualisation	Value	Data Acquisition & Warehousing	Data Mining & Cleansing	Data Aggregation & Integration	Analysis & Modelling	Data Interpretation	Privacy	Security	Data Governance	Data & Information Sharing	Cost/Operational Expenditures	Data Ownership	Emergency management
Abdullah et al. (2017)	1	x		x			x		x	x	x	x	x				x		x	x
Arslan et al. (2018)	2	x		x			x	x	x		x			x				x		x
Boulos et al. (2011)	3	x	x	x	x	x	x	x						x	x			x		x
Fertier et al. (2016)	4	x	x	x	x	x											x			x
Hilbert (2013)	5			x	x	x		x		x		x		x	x	x		x		
Hilbert (2015)	6			x	x									x	x	x	x	x		
Meier (2015)	7	x	x	x	x	x	x	x	x	x		x		x	x	x	x	x		x
Qadir et al. (2016)	8	x	x	x	x	x		x						x						x
Rahman et al. (2017)	9	x			x			x	x	x				x		x	x			x
Vayena et al. (2015)	10			x	x	x		x						x		x	x			
Watson et. al (2017)	11	x	x	x	x	x		x						x	x	x				x
Zicari (2014)	12	x	x	x	x	x	x	x	x	x		x	x	x	x	x				
Chen et al. (2012)	13	x	x	x	x		x			x		x	x	x					x	
Sivarajah et al. (2016)	14	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
Guevara et al. (2015)	15					x		x		x	x	x	x							x
Athanasia & Stavros (2015)	16			x		x		x		x		x								x
Khan et al. (2012)	17													x	x	x	x			
Barnaghi et al. (2013)	18	x	x	x	x	x		x	x		x	x	x			x				x
Gudivada et al. (2015)	19	x		x	x	x		x	x	x	x	x	x	x	x	x	x			
Saleh et al. (2013)	20	x	x						x	x		x		x	x					
Chen et al. (2013)	21	x	x	x	x			x	x	x	x	x	x							
Bertot & Choi (2013)	22					x			x					x	x	x	x			
Zuboff (2015)	23	x	x			x		x	x			x	x	x	x					
Jin et al. (2015)	24	x	x	x	x	x		x		x		x	x							
Ofli et al. (2016)	25	x	x	x		x		x										x		x
Whipkey & Verity (2015)	26	x	x	x	x	x		x		x		x	x	x	x	x		x		x
Duffield (2013)	27											x								x
Swaminathan (2018)	28					x		x					x							x
Lu et al. (2012)	29																			x
Jagadish et al. (2014)	30	x	x	x	x	x			x	x	x	x	x	x			x	x	x	

Table 2: Concept Matrix on Big Data challenges in Emergency Management by Ø. S. Fongaard & T. F. Nestaas, 2019.

2.2.1 Data Challenges

Sivarajah et al. (2016, p. 269) defines data challenges as *"the group of the challenges related to the characteristics of the data itself"*.

Volume

The large scale and big volume of data, which can be terabytes, petabytes, zettabytes of data, is a challenge on its own. The heterogeneity, ubiquity and dynamic nature of the different data generation resources and devices, and the enormous scale of the data itself, make determining, retrieving, processing, integrating and inferring the physical world data, (e.g. environmental data, business data, medical data, surveillance data) a challenging task (Barnaghi, Sheth & Henson, 2013; Sivarajah et al. 2016, p. 269).

Velocity
"The challenge of velocity comes with the need to handle the speed with which new data is created or existing data is updated". (Chen et al., 2013, p 158). This issue is often seen in modern times where mobile phones and other digital devices produce large amounts of data, floods of information is generated, and not from a specific point. These are mainly datasets that are generated through large complex networks including data generated by the proliferation of digital devices, which are positioned ubiquitously resulting in driving the need for real-time analytics and evidence-based planning (Sivarajah et al., 2016, p. 273). The technology for streaming data has been researched for years to find ways to handle the high velocity of information, but the streaming systems are still limited because of the increasing volume of today's sensors placed on a high number of vast devices (Chen et al., 2013, p 158).

Variety
The different types of data can often define variety (e.g. multiple data formats with structured and/or unstructured text, images, multimedia content, audio, video, sensor data or noise). Data challenges related to the variety (i.e. diverse and dissimilar forms) of data are also deemed a challenge (Sivarajah et al. 2016, p. 269). Variety is defined as one of the challenges because of the diversified data types, together with timely response requirements (velocity), and uncertainties in the data (veracity) (Jin, Wah, Cheng & Wang, 2015, p. 1). Technical challenges can include multiple data streams of varying data types using sensors in urban sensing (Boulos et al., 2011, p. 5).

Variability
Data, in which the meaning constantly change over time is variability. Out of the seven V's of Big Data it is an essential feature, but is often confused with variety (Sivarajah et al., 2016, p. 273). Variability can be an extremely important factor for application that use real-time and current data where the data will not stay static. Data may also change form as it is processed, applying new meanings. Applications for detecting fraud, such as some government sponsored healthcare programs requires modeling and processing of dynamically changing data (Gudivada et al., 2015, p. 3).

Veracity
Veracity can be referred to as the imprecision, inaccuracy and inconsistency of large datasets that may give and/or support a wrong perceived meaning of the data (Sivarajah et al., 2016, p. 269). Should it be possible to overcome the hard task of making sense of Big Data, the second major problem is the veracity of it (Meier, 2015, p. 33). The data quality and imprecision of the whole situation can be a crucial factor in decision making. Validity of information for trustworthiness is one issue (Athanasia & Stavros, 2015, p. 6). Challenges that come with veracity are then many since the trustworthiness of the information is a decisive factor. How well understood the sampling biases are, if there are doubts, imprecision and inaccuracy of measures, fabrications of data, misstatements or untruths and misplaced evidence in the data are some of the challenges concerning veracity (Zicari, 2014, p. 109).

Visualisation
More recently Big Data and Big Data Analytics have been used to describe the datasets and analytical techniques in applications that are so large and complex that they require advanced and unique data storage, management, analysis, and visualisation technologies.
Producing all this large information from terabytes to exabytes retrieved from sensors to social media data to readable, and representing the key information and knowledge extracted

effectively through using different visual formats makes visualisation a challenge for Big Data technology (Chen, Chiang & Storey, 2012, p. 1166).

Value
Big Data does not produce any value, the value comes only from what we infer from it, and why we need Big Data Analytics to extract it (Zicari, 2014, p. 105). Value is what knowledge and edge from vast amounts of structured and unstructured data that one can extract. Big Data researchers considers the value as an essential feature, as somewhere within that data (Sivarajah et al., 2016, p. 273).

2.2.2 Process Challenges

Sivarajah et al. (2016, p. 273) defines process challenges as "the group of challenges encountered while processing and analysing the data that is from capturing the data to interpreting and presenting the end results".

Data Acquisition & Warehousing
The aim of Data Acquisition is to aggregate information in a digital form for further warehousing and analysis since the data may not be alike because it originates from different sources (Abdullah, Ibrahim & Zulkifli, 2017, p. 409). Acquiring, storing and loading of Big Data is a huge job. Much of the data can be filtered out "on-line" but since the raw data is often too voluminous the option of storing it at the warehouse might not be possible, and with high velocity the challenge increases (Jagadish et al., 2014, p. 89). Data management and warehousing is considered the foundation of Business Intelligence & Analytics 1.0 (Chen et al., 2012, p. 1166).

Data Mining & Cleansing
Data mining and cleansing is the challenge that relates to extracting and cleaning data from a collected pool of large scale unstructured data (Sivarajah et al., 2016, p. 273). Cleansing of data is the process for identifying incomplete and unreasonable data and are to undergo the cleansing process. The datasets will then be filtered for categorization using methods of extraction so that the semantics and correlation may be obtained (Abdullah et al., 2017, p. 409).

Data Aggregation & Integration
The aggregation is the role of creating knowledge from raw data that is mined and cleansed from large unstructured data (Barnaghi et al., 2013, p. 7; Sivarajah et al., 2017, p. 274). Big Data's value primarily comes through integrating massive heterogeneous datasets (Gudivada, 2015, p. 3). Because of the vast variety of data sources present in times of a disaster, a need for the challenge of integration and aggregation of data and to make effective visualisations from it, hereby value, is created (Arslan et al., 2018, p. 1)

Analysis & Modelling
After the data acquisition, warehousing, mining, cleansing, aggregation and integration, the next phase is the data analysis and modelling. Since Big Data is often noisy, unreliable, heterogeneous and dynamic by nature, older methods for analysis and modelling with focus on solving intricacy of relationships between schema-enabled data might not fit well to the analysis and modelling (Shah, Rabhi & Ray, 2015; Sivarajah et al., 2016, p. 274). There is a need to use data analysis and modelling to anticipate what might happen in the future, not just knowing what is happening in the current (Chen et al., 2013, p. 161).

Data Interpretation
This step entails visualising the data in a way that makes the data understandable for the users (Sivarajah et al., 2016, p .274). Data interpretation challenges refers to challenges relating to understanding the output or visualising and sharing the results of the analysis and modelling (Zicari, 2014, p. 110).

2.2.3 Management Challenges

Sivarajah et al. (2016) defines management challenges as "a group of challenges encountered, for example while accessing, managing and governing the data".

Privacy
Concerns about privacy in Big Data is one of the most common concerns out there (Hilbert, 2013, p. 27). Any data on any person inevitably raise privacy issues (Nature Editorial, 2007, p. 637). There are a lot of data warehouses that contain sensitive and personal data, which results in legal, ethical and privacy concerns with accessing such data. There are challenges relating to ensuring such data is used correctly, tracking how it is used, transformed and managing its lifecycle (Zicari, 2014, p.111). In particular, location based information being used in Big Data Analytics processes poses clear privacy concerns (Sivarajah et al., 2016, p. 274.)

Security
The challenges in Big Data regarding security are not that different from the challenges facing systems using traditional data. Malware is a consistent threat, combined with a lack of adequate security controls to ensure the information is resilient, analysing logs, network flows and intrusion detection creates data security challenges on its own. (Sivarajah et al., 2016, p. 274).

Data Governance
Besides data growth, data centre infrastructure and the ability to provide scalability, data governance for describing what data is warehoused, analysed and accessed is termed as one of the biggest challenges IT managers face (Sivarajah et al., 2016, p. 274). Organisations perceive data governance as a way to warranting data quality, improving and leveraging information as well as maintaining its value as a key organisational asset (Otto, 2011; Sivarajah et al., 2016, p. 274)

Data & Information Sharing
There are Big Data challenges related to sharing of data between different departments and organisations, including overlap amongst applications, duplication in stored information and confusion around the responsibilities of each business unit and application (Tekiner & Keane, 2013; Abdullah et al., 2017, p. 407). There are also challenges related to the sheer amount of data that would be needed to be shared, and the processing power that is demanded to complete such data transfers (Jagadish et al., 2014, p. 92).

Cost/Operational Expenditures
The constant increase in data of all forms has resulted in a need of sophisticated Big Data processing in data centers. These data centers are often spread across several countries to embed resilience and spread risk, which in combination with intensive processing operations results in high storage and data processing costs (Sivarajah et al., 2016, p. 275). If you would

need to gather information based on nodes and sensors there are challenges related to the cost of the nodes (Boulos et al, 2011, p. 5). Organisations need to carefully weigh the benefits of the Big Data technology up towards it costs (Whipkey & Verity, 2015, p. 21).

Data Ownership

Besides privacy, the data ownership issues are one of the main unsolved social issues regarding Big Data technologies. There are challenges regarding who owns the data on social media, e.g. Is it the user or Twitter that owns the "*tweets*"? Along with the ownership of data arises the issue of controlling and its accuracy (Sivarajah et al., 2016, p. 275).

3.0 Research Approach

Through this chapter the research approach, strategy and design is presented. The criteria for the evaluation of the research quality are then further presented. This chapter has the purpose of informing and giving the reader insight into which approach has been used to answer our research question.

3.1 Research Perspective

This research has been conducted with a qualitative research approach to answer the research question. This means that the main objective with the research has been to gain a high level of understanding, instead of the opportunity to generalise the results from a bigger set of informants such as a quantitative research approach. Qualitative research approaches aim at understanding the humans and their underlying social and cultural contexts, where the goal is to understand the problem from the side of the participant (Myers, 1997). The purpose of this research was to gain insight into the challenges individuals in emergency management organisations are experiencing with Big Data technology, and their interpretations of this. Through qualitative interviews we have gained insight into the individuals personal experience, knowledge, opinions or similar information of interest (Gripsrud, Olsson & Silkoset, 2015, p. 90).

3.2 Research Strategy

The research strategy is a step-by-step plan of action that gives direction to the thoughts and efforts, which enabled us to conduct research systematically and on schedule to produce quality results and detailed reporting (Dinnen, 2014). The research strategy applied in this study can be described as a semi-delphi study. The delphi study or delphi method is a prognosis technique that is suitable to gather information from a number of experts, which together can predict the future prospects on a certain area. The delphi research strategy is interactive, where several rounds of interviews and analyses are made before presenting the results (Sander, 2017). This research is not identified as a delphi study, but rather a semi-delphi study because it is restricted to one round of interviews with the informants, due to time limits. A semi-delphi study in this case means that we identified a number of emergency management experts on the field of Big Data, and conducted an in-depth interview with each of them to gather information about the research question.

There are several advantages with this research strategy; the research is based on statements from experts on the relevant field, all the informant opinions and stances are anonymous and because of the individual interviews there are no dominant personalities affecting other informants (Sander, 2017). The use of semi-delphi study as a research strategy is justified because of the way it allows experts on the relevant field to comment on the challenges of using Big Data technologies in crisis management. We then compared the findings from the interviews to the existing model presenting Big Data challenges in Sivarajah et al. (2016). Through this research strategy we were able to add new, remove irrelevant, and verify existing challenges into the existing model.

3.3 Research Design

A research design can be defined as a blueprint for your research, dealing with at least four problems: what question to study, what data are relevant, what data to collect and how to analyse the results (Frankfort-Nachmias & Nachmias, 1992, p. 77-78). Chapter 3.3 is based on this definition. The research question has already been presented in the introduction: “Big Data Challenges in Emergency Management”, following is the presentation on how informants were selected, how data was collected and how the data analysis was conducted.

3.3.1 Selection of Informants

Due to the relation between the research and the CIEM, we already had a established network of contacts, with a relevant connection to the field of our study. Criteria for the selection of informants for our study included experience and/or knowledge on Big Data technology used in crisis situations. The selection criteria for the informants were developed in cooperation with advisors from the University of Agder.

The selection of informants in the research included people with hands-on experience on using Big Data technology in crisis situations. These informants consisted of people from the emergency management organisations Standby Task Force (SBTF), ACAPS, Direct Relief, Médecins Sans Frontières, the Emergency Medical Communication Centre (EMCC) Southern Norway, Red Cross and Kristiansand municipality. The selection also included researchers with experience on the use of Big Data technology in crisis situations from their own research. The selection of these informants consisted of people from the research facilities CIEM, Western Norway Research Institute, Qatar Computing Research Institute (QCRI) and the College of Information Science at UNO. The selection was extended to include both people with hands-on experience and researchers on the field to get a wider perspective on the research question from different point of views. All the informants are completely anonymised, according to NSD’s guidelines. Therefore, each informant and their statements is identified by an ID number together with the date of the statement, e.g. *Informant ID2, 18.03.2019.*

The selection method used in this research is a combination of references from experts and advisors on the University of Agder, self-selection and the snowball method. The initial method that was used were references from experts and advisors from our own university, which gave us some informants to start with. The snowball method was used with these informants, which resulted in them referencing us to other relevant informants. The snowball method, also known as snowball sampling or chain-referral sampling, is a non-probability (non-random) sampling method used when the characteristics to be possessed by samples are rare and difficult to find. The method is based on referrals from initial subjects to generate additional subjects (Dudovskiy, 2018). We used the snowball method because the availability of informants was sparse and geographically spread around the world. In addition it was necessary at times to obtain a referral to achieve contact with the right informants. A lot of the technology-based emergency management experts and practitioners are occupied with their work, and they were hard to get hold of without a referral from other contacts. Through the whole process, self-selection was also used as a selection method, both on the organisational and individual level. We used self-selection extensively with the CIEM organisation, where we examined what publications each researchers had published, and chose them based on the relevance of their earlier research.

The contact with each informant were done through e-mail. The informants were briefly introduced to our research on mail before the interviews, and were presented a general interview guide to get a overview of the categories of questions we would present on the interview. This interview guide was modified to the informants expert area in each interview.

We required each informant to sign a declaration of consent document in accordance with Norwegian data protection guidelines (Norwegian Center for Research Data, 2019). As part of the snowball method, we asked each informant at the end of the interview if they had ideas for other relevant informants for our study, and asked for their contact information as well as a referral. Further on we contacted these informants with information about the project as well as an inquiry for participation in the project as an informant.

Even though we did not present the specific requirements or specifications for the participating informants, some of the contacted informants felt that they did not have the necessary competence or did not have enough time to allocate for the interview. In this case we often experienced that we were referred to other, more suitable informants in the same organisation. These informants were often in a better position to answer our questions than the initial informants that were contacted.

3.3.2 Data Collection

The method used for data collection has been a hybrid between deductive and inductive. We have used an abductive methodology, since practical scientific research cannot be based either on pure deduction or induction (Svennevig, 2001, p. 1). *"*Central to any scientific process is the inferential step from some initial puzzling fact to some theoretical hypothesis which can explain it. This inferential process is called abduction by the pragmatist philosopher Charles S. Peirce*"* (Svennevig, 2001, p. 1). By using an abductive methodology for data collection we were able to revise on previous findings from Sivarajah et al. (2016), as well as be open to changes and additions from our data collection. Abduction is a process of gaining new knowledge to supplement what we already have (Svennevig, 2001, p. 2).

The data collection was conducted through interviews with informants that possessed relevant knowledge or experience from Big Data technology in crisis situations and emergency management. We conducted most of the interviews as individual interviews. Individual interviews were suitable for our research because it enabled each informant to speak freely and without boundaries such as hierarchical structures or social constructs. However, there were some of the informants that insisted on conducting the interview in pair, due to time limitations in their personal schedules.

The interviews in this study has been conducted as semi-structured interviews. Semi-structured and unstructured interviews are more suitable than structured interviews for explorative research that aims to discover new information (Oates, 2006, p. 188). We have chosen a semi-structure approach to the interviews because the primary purpose of our study was to discover new challenges emerging related to Big Data and emergency management, but at the same time checking if the already identified Big Data challenges were present in emergency management as well. By performing semi-structured interviews we were able to have a list of predefined themes that covered the identified challenges, as well as letting the interviewee being able to speak freely, and introduce issues of their own that they saw as relevant to our topic. Figure 3, originating from Jacobsen (2005), explains the different types of interviews, ranging from completely closed, to completely open interviews. We have conducted the second most open type of interview, a interview with an *"*interview guide with topics, in a set order and open answers only*"* (Jacobsen, 2005, p. 145).

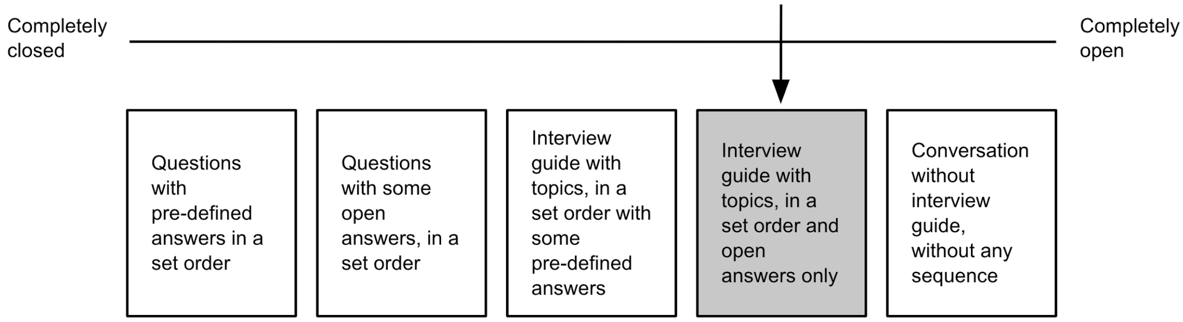

Figure 3: Levels of structuring of an interview (originally from Jacobsen, 2005, p. 145)

When conducting our interviews, we used a general interview guide as a basis for our interviews to gain a certain level of structure. It is important that qualitative interviews are not arranged too structured, but at the same time it needs some structure to not be too unstructured (Jacobsen, 2005, p. 144-147). The interview guide was based on acquiring information on already identified challenges in the literature, as well as identifying potential new challenges with the use of Big Data technology in emergency management. The choice of using a interview guide is reasoned with the assumption that it would be suitable to link the information collected from the interviews to the existing research identified in the literature review.

A total of 15 interviews were conducted with a total of 16 informants (one of the interviews was conducted as a pair interview). Before the interview the informants received information about the interview and were informed about the interview length of between 30-60 minutes. The information included the purpose of the study, details about us as students, their rights and the interview guide. By providing the interview guide, we were able to let the interviewees prepare themselves for the questions that we were going to ask them.

There were some of the informants that reserved themselves for a maximum interview length of around 30 minutes, while others accepted the time needed to get through all the questions extensively. Most of the interviews were done through video conference tools such as Skype or Skype for Business, because of the difference in geographical locations. Meanwhile there were some of the interviews that we were able to conduct as face-to-face interviews due to the fact that the informants were based nearby geographically. We preferred face-to-face interviews if it was possible, due to the extensive ability to read the informants interpretation through other factors such as body language and facial expressions. This was not always possible on video conferences due to mediocre internet connections. Table 3 presents an overview over the conducted interviews.

Organisation	Duration (hours:minutes)	Type of Interview
Standby Task Force	1:14	Individual, Video Conference
ACAPS	1:02	Individual, Video Conference
Direct Relief	0:59	Individual, Video Conference
Centre for Integrated Emergency Management	0:31	Individual, Video Conference
Western Norway Research Institute	0:53	Pair, Video Conference
Qatar Computing Research Institute	0:39	Individual, Video Conference
Centre for Integrated Emergency Management	0:52	Individual, Face-to-face
Centre for Integrated Emergency Management	0:48	Individual, Face-to-face
College of Information Science & Technology at UNO	0:40	Individual, Video Conference
EMCC Kristiansand	0:36	Individual, Face-to-face
Médecins Sans Frontières Spain	0:39	Individual, Video Conference
Centre for Integrated Emergency Management	1:12	Individual, Video Conference
Red Cross - British Red Cross	0:46	Individual, Video Conference
Kristiansand Municipality	0:22	Individual, Video Conference
Centre for Integrated Emergency Management	0:46	Individual, Video Conference

Table 3: Overview of interviews

3.3.3 Data Analysis

To analyse the data we conducted a thematic analysis. Thematic analysis is a qualitative analysis, that were used because of its flexibility and its ability to identify, analyse, and report patterns within data (Braun & Clarke, 2006, p. 4-6). The data analysis had its basis in the following phases:

Phase 1: Familiarising yourself with your data
In the first phase we familiarised ourselves with the data. This involved transcribing the interviews and repetitive reading. At this stage we obtained an overall and general sense of the information, and reflected on the meaning behind it. Each informant was unique and had a different perspective on their ideas. Their experience and hands-on examples gave us insight and credibility to the challenges which they raised. For a better understanding of the data by us as researchers, the role of transcribing was evenly dealt, for every other interview the roles changed. In addition, the one that did not transcribe were handed the task of repetitive reading and taking notes to familiarise himself with the data.

Phase 2: Generating initial codes
In this phase we started to produce initial codes from the data. The transcription was sorted and organised into meaningful data groups. These groups included challenges presented by Sivarajah et al. (2016), as well as unfamiliar and new emerging challenges which arose

during the coding. This included themes that were not directly related, but could have interesting aspects at a later stage.

Phase 3: Searching for themes
Searching for themes refocused the analysis at a broader level of themes, rather than codes. We sorted the different codes, and then saw them in correlation with a Big Data challenge in the context. Many parts of the code correlated to several challenges, therefore it was important to sort them in an organised manner to map the themes and challenges. The interviews were coded into four main themes for the challenges of Big Data in emergency management: Three classifications as described by Sivarajah et al. (2016) and one for new challenges emerged. The new challenges emerged were then complemented to the existing model.

Data challenges were identified as challenges in relation to the data form, standards, characteristics, speed, trustworthiness and similar.

Process challenges were concerned with the challenges in connection with the processing and analysis of data. This could include acquisition, storage, aggregation, cleansing and mining of data. The process challenges are viewed differently from each informants experience and context, making it essential to understand the environment in which they experienced the challenges.

Management challenges were determined by the managerial challenges that were related to the information and data. Identification of these included the understanding of how each of the informant's organisation managed its data, or if they even did so themselves. Challenges in connection with management of data were often related to the security, privacy, sharing, governance and ownership of the information they used.

New challenges emerged were identified by factors which could not be characterised in direct relation with the earlier classification of challenges by Sivarajah et al. (2016), and would not fit its description, or had no relation at all. Examples on such challenges were lack of knowledge, time-pressure, resource scarcity and data access.

Phase 4: Reviewing themes
After the identification of the main themes of challenges, they were refined. During this phase the task of dropping, dividing or keeping the different challenges was finalised. Most of the challenges we identified were kept, though one new emerged challenge was dropped, as well one divided into two challenges. The challenges this affected were mostly classified as *new challenges emerged*.

Phase 5: Defining and naming themes
During this phase we combined the identified and refined challenges. The Big Data challenges were then set together and organised coherent and consistently to produce the chapter of results. The phase of combining the results for presentation refined the challenges to create the best possible understanding and definition of the themes and challenges.

Phase 6: Producing the report
The last phase involved the write-up of the report. We had at this point provided evidence from 16 informants, a mix of both practitioners and researchers in the relevant field. For the presentation of the results we assembled the results for evidence, supplemented by examples

from the informants to provide a deeper understanding. The veracity of the study is dependent upon the validity of the qualitative study and the strategies of choice, as well as the acknowledgement of the limitations of the study. The understanding of the study could also bring new perspectives to the research, questions about Big Data's role in emergency management which have not been mentioned before.

3.4 Evaluation of Data Quality

The standard criteria for judging the data quality of research is validity and reliability (Munkvold, 1998, p. 19). Validity is usually divided into internal validity and external validity. Internal validity entails that the correct data has been obtained, and applies to which extent the causality in the interviews is intended. External validity concerns if our findings may be transferred to other contexts and situations, it therefore revolves around generalization of the findings. Reliability refers to which degree you can trust that your results are reliable (Gripsrud et al., 2015, p. 49-51). Table 4: "Parallel criteria to evaluate interpretive studies", presented by Guba & Lincoln (1989), is used as data quality criterias because it is appropriate to interpretive studies.

Parallel criteria	**Goal**	**Tactic**
Credibility	Establishing the match between the constructed realities of respondents (or stakeholders) and those realities as represented by the evaluator and attributed to various stakeholders	Field work and longitudinal observation Discussion of data and results with external peers and informants (member checks)
Transferability	Presenting a sufficient detailed account of the findings as to enable judgment by the reader of how these findings can be transferred to other contexts	Thick description
Dependability	Ensuring that methodological changes and the interpretive process are documented so that the reader can follow the process and the choices made by the researcher	Making the research process explicit
Confirmability	Ensuring that the data, interpretations and results are grounded in the context and not just a result of the researcher's imagination	Making data available Describing the logic used for moving from data to the final results

Table 4: Parallel criteria to evaluate interpretive studies (Guba & Lincoln, cited in Munkvold, 1998, p. 20).

Credibility was achieved through verification of statements by the informants used in the results before final delivery and publication. Each informant was sent the statements presented in the results originating from them, accompanied with the context, for them to verify. Through the verification the informants had the possibility to comment or explain the interpreted results if it did not reflect its true meaning. There were made several corrections to the results after the verification. E.g., Informant ID5 correct his statement saying that 1-3 % of Tweets were posted with a geolocation, instead of his initial statement saying that only 0,5 % of Tweets were posted with a geolocation. Informant ID22 corrected an important sentence, where we had misunderstood the informant and written the meaning in positive format, when it was meant as a negative format, saying: *"On the other side, it has to be assured that those who are eligible to access the data are not intending to use it for bad purposes"*. The verification of the informant's statements has given the study credibility and validity. 14 of the total 16 informants were able to provide verification and eventual modifications and corrections to their statements.

Transferability was achieved by presenting the results in a detailed manner so that the reader might find reason and ground to transfer them to other contexts. The revised model were open for further revisions for other contexts, and many of the findings from the different organisations participating were traceable for better transferability.

Dependability was achieved through a orderly described research process, describing the choices in the process, the interpretations and assumptions that has been conducted. This is seen through our research approach to ensure the possibility for other researchers to achieve the same methodological approach.

Confirmability was achieved by making as much data available as possible without identifying the informant. The organisations of the informants is presented, and will ensure that the data, our interpretations and results are grounded in the context of emergency management and not anyones fantasy.

4.0 Research Context

As mentioned in 3.3.1 Selection of Informants the criteria for the selection of informants for our study included experience and/or knowledge from/on Big Data technology used in crisis situations. The selection of informants in the research includes people with hands-on experience from Big Data technology in crisis situations as well as researchers with experience on Big Data technology in crisis situations from their own research. Therefore, research context is divided into two main types of organisations: emergency management organisations and research facilities. This chapter presents the emergency management organisations and the research facilities where the study has been conducted, as well as an overview of the sources to our data collection.

4.1 Emergency Management Organisations

Emergency management organisations are the organisations conducting the on-field crisis response activities such as search, rescue and providing fundamental needs such as food, water and shelter. Following is the description of the different emergency management organisations we have conducted the data collection for this research in. The informants for the interviews in our study represented their organisations anonymously as people with hands-on experience from the use of Big Data technology in crisis situations.

Standby Task Force (SBTF)

Standby Task Force, abbreviated SBTF, is a non-profit organisation incorporated in 2014 in Delaware, US, and are a global network of trained and experienced digital volunteers working together online. They were set up in 2010, and started with engagements in disasters at such a scale that the affected country asked for international help. They have been active in many natural disasters since then, where typical disasters range from weather caused disasters such as typhoons, orcanos, floods to earthquakes. They are also supporting humanitarian agencies with election monitoring and other projects (SBTF, n.d.). The organisation has always consisted of approximately five to eleven people, living worldwide. They are divided into the core team, the coordinators and the directors.

They work with systemizing information from social media in relation to crisis situations, in cases where the organisations that are directly involved in on-site operations does not have the time or resources to work with social media data. SBTF keep their operations to nature disasters and other neutral crises, and do not interfere with hostile environments such as conflicts and wars. Their mission is to provide volunteer online digital responses to humanitarian crisis, local emergencies and issues of local or global concern (HDX, n.d.).

ACAPS

ACAPS is a non-profit organisation that provides independent humanitarian analysis to help crisis responders to better understand how to address the worlds disaster. It was established in 2009 as a joint project between the Norwegian Refugee Council and Save the Children. The organisation consists of a team of 28 professionals based in Geneva and in the field, reinforced by a pool of experts. ACAPS is managed and hosted by NORCAP (NRC, 2018).

ACAPS is involved in four main activities: humanitarian analysis, tailored support, capacity building and advocacy. Their analysis supports the humanitarian community by providing up-to-date information on more than 40 key crisis around the globe - they cover every crisis as and when it happens. They are offering tailored support such as analytical products, scenario building workshops and field assessments on request, such as the Briefing Notes for the START Fund (START Network, n.d.). They consistently try to strengthen the humanitarian sector as a whole through capacity building, and have gained a strong reputation for its technical know-how and the quality of its sources. They also campaign for humanitarian organisations to use more information direct from the field to help them make better-informed decisions (ACAPS, n.d.).

Direct Relief
Direct Relief is a humanitarian aid organisation that operates in all the United States as well as more than 80 countries around the world. Its mission is to improve the health and lives of people affected by poverty or emergencies without regard to politics, religion or the ability to pay for their services. It was established as early as in 1948 as a California-based non-profit corporation under the name William Zimdin Foundation. It changed its name to Direct Relief Foundation in 1957 and based its services on providing postwar assistance to enable people affected by World War Two to help themselves. In 1962 it became licensed as a wholesale pharmacy to secure prescription medicines for use abroad. Today Direct Relief provides appropriate and specifically requested medical resources to community-based institutions and organisations both throughout the world and across the United States (Direct Relief, n.d.).

Direct Relief seeks to break the vicious cycle consisting of sickness and poorness: sick people who do not receive care cannot work, they get poor or stay poor, and people who are poor are at higher risk of getting sick. They provide better access to health services through their medical assistance programs which equip health professionals working in resource-poor communities to meet the challenges of diagnosing, treating, and caring for people (GuideStar, n.d.).

Médecins Sans Frontières (MSF) Spain
Médecins Sans Frontières, abbreviated MSF, also known as Doctors without Borders, is a international, independent medical humanitarian organisation. It was founded in 1971 in France by a group doctors and journalists after the war and famine in Biafra, Nigeria. The organisation included 300 volunteers when it was founded, and is now a global movement, including over 42 000 people with staff from over 150 countries. MSF provides assistance to populations in distress, to victims of natural or man-made disasters and to victims of armed conflict (MSF, n.d.).

Emergency Medical Communication Centre (EMCC) Southern Norway
Emergency Medical Communication Centre is a call center which answers the medical emergency number 113 and does surveillance of the ambulance transport. EMCC is manned by medical personnel, most commonly nurses and paramedics (Ludvigsen & Nylenna, 2019). The EMCC department in Southern Norway receives emergency calls after the Norwegian index for medical assistance, and coordinates ambulance missions in Aust- and Vest-Agder as well as the municipalities of Nissedal and Fyresdal in Telemark county and Lund municipality in Rogaland county. In 2016 the EMCC received around 132 700 conversations by phone in and out of the central, of which nearly 48 000 were emergency calls (NAKOS, 2019).

Red Cross - British Red Cross
British Red Cross provides support in emergencies, refugee support, independent living services, and first aid and humanitarian education. They operate both in their own right and as part of the International Red Cross and Red Crescent Movement, the world's largest

humanitarian network, which has more than 17 million volunteers across 190 countries. They refuse to ignore people in crisis and their vision is of a world where everyone gets the help they need in a crisis. More than 21,500 volunteers and 4,100 staff at the British Red Cross work together to help individuals and communities prepare for, cope with and recover from emergencies. This includes everything from disasters and conflicts to individual injuries and personal challenges (British Red Cross, 2019; Disasters Emergency Committee, 2018).

Kristiansand Municipality
All Norwegian municipalities has a fundamental responsibility to both take care of the public safety and security. This means that all Norwegian municipalities must fulfill the requirement of the municipal emergency preparedness obligation and civil protection law, with associated regulations. Preventive social security is to be safeguarded by the Planning and Building Act. The municipalities are also responsible for coordinating the municipality's work on public safety and emergency preparedness (Daatland, 2019).

Work on public safety and preparedness in the Kristiansand municipality exists at an overall level and in all sectors. The emergency preparedness work in the municipality is managed by an emergency response- and security unit. The unit has the overall responsibility for ensuring that crises are handled correctly. The emergency response unit shall also ensure that all sectors have updated ROS analyses, contingency plans and alert lists. The sector itself is responsible for carrying out necessary analyses and plans. The emergency response unit is also responsible for the municipality's comprehensive ROS analysis, coordinating emergency preparedness and exercises across sectors, and with external partners throughout the region. In addition, the unit is the coordinator of the municipality's crisis management and the work committee for crisis management (Daatland, 2019).

4.2 Research Facilities

The selection of informants in this research includes researchers with experience on the use of Big Data technology in crisis situations from their own research. Therefore, we have used the context of research facilities and contacted relevant informants in different research facilities which aims at different research on Big Data technologies in the emergency management field. The informants for the interviews in our study represented these organisations anonymously as researchers with experience on the use of Big Data technology in crisis situations.

Centre for Integrated Emergency Management (CIEM)
CIEM is a multidisciplinary research center at the University of Agder, in Norway. It was established in 2011, and currently consists of 24 researchers in addition to the management and the alumni. This team includes some of the world's leading experts in sensor technology and wireless communications, artificial intelligence, decision support, analysis of social media, and cooperation and coordination. CIEM operates as a top research priority at the University (CIEM, 2019).

The center conducts research that includes how new developments within the information and communication technology (ICT) offers new opportunities for more effective emergency and crisis management through the integration of information from a variety of data sources. The Center for Integrated Emergency Management is dealing with issues at both global and national levels, and is working with governmental and nongovernmental emergency operators, research institutions and industry internationally (CIEM, 2019).

Western Norway Research Institute
The Western Norway Research Institute is an international assignment-based research institute located in Sogndal in Western Norway. The research institute is organised as a foundation and was established in 1985. The institute consists of a total of 30 employees, and has competence in the academic fields of social studies, natural sciences, technology and the humanities (Western Norway Research Institute, n.d.).

The research institute is coordinating the Big Data and Emergency Management project, abbreviated to the BDEM project. The project aims through cooperation with universities all around the world to establish a long term partnership where excellent education is to be embedded in excellent research in Big Data and emergency management. The project started in 2017 and is due to last for 36 months (Big Data and Emergency Management, 2019).

Qatar Computing Research Institute
The Qatar Computing Research Institute, abbreviated QCRI, is a national research institute in Doha, Qatar. It operates under the Hamad bin Khalifa University. It was established in 2019 by Qatar Foundation for Education, Science and Community Development - a private, non-profit organisation that is supporting Qatar's transformation from carbon economy to knowledge economy. The purpose of the institute is to support Qatar Foundation's mission by helping to build Qatar's innovation and technology capacity. It focuses on tackling large-scale computing challenges that address national priorities for growth and development (QCRI, n.d.).

Researchers at the Qatar Computing Research Institute has developed a free and open source software that automatically collects and classifies tweets that are posted during humanitarian crises, which is called Artificial Intelligence for Disaster Response, abbreviated AIDR. The open source software is available for humanitarian or crisis response organisations that are familiar with using Twitter as a data source. It is aiming at solving problems relating to the large scale of data produced via social media during crisis situations, which makes it difficult for humans to manage them on their own, as well as the problems relating data being too rich and complex for machines to successfully process them. AIDR is combining the "best of both worlds" by combining human and machine intelligence (Jikimlucas, 2016).

College of Information Science & Technology (IS&T) at UNO
The College of Information Science & Technology, also referred to as IS&T, is part of University of Nebraska Omaha. It was established in 1996 to address the growing global need for knowledgeable professionals in the area of information technology, and aims to be a premier college with excellence in education, research, and service in the disciplines necessary to meet the needs of their students and the communities they serve. They offer five undergraduate bachelors programs in: Bioinformatics, Computer Science, Cybersecurity, Information Technology Innovation and Management Information Systems. (University of Nebraska Omaha, n.d.).

5.0 Results

This chapter presents the results of and findings in the data collection of the study. The results is based on interviews with a total of 16 informants relevant to our study. The selection of informants included seven people with hands-on experience from the use of Big Data technology in crisis situations, and nine people from research facilities.

The results related to the challenges introduced by Sivarajah et al. (2016) is presented first, divided into data challenges, process challenges and management challenges. Then the new emerged challenges identified in the data collection are presented. At each challenge there is provided information on how many informants of the total of 16 that identified each challenge with Big Data technology in emergency management. To get a full overview of the challenges identified by each informant, refer to 9.2 Appendix B: Overview of identified challenges.

5.1 Data Challenges

Data challenges is referred to as the group of challenges that are related to the characteristics of the data itself (Sivarajah et al., 2016, p. 265).

Volume
Identified as a challenge by a total of 7 informants.

When asked what kinds of challenges Informant ID12 were facing when collecting data from their sources the response was *"I think the main challenge is that there is too much of it"* (Informant ID12, 15.03.2019). A similar response was received when Informant ID24 were asked about their biggest challenges when handling their data, and referred to it as "information overload" (Informant ID24, 05.04.2019).

Huge amount of data can be overwhelming for emergency and disaster organisations. During a serious crisis the municipality gets overloads of information as well. Using their ordinary "document-center" becomes a challenge itself because of the large volume of information (Informant ID29, 03.05.2019). Informant ID30 further elaborated on the data they need to collect in connection with beneficiaries they are helping, but because of the sheer volume of that data it is very hard to collect it quickly (Informant ID30, 12.04.2019).

Velocity
Identified as a challenge by a total of 2 informants.

An informant tells how different V's of Big Data has challenges and how they drive velocity.

> If you think of the different V's of Big Data - volume, velocity, veracity etc. So each of the V's has different challenges. You can say that because of the rapid increase in the Big Data, that is produced by different producers and different people. For example a case of social media, during emergencies, lot of tweets are sent by different people from different places, or lot of videos are also posted on social media, so there is a lot of real-time data coming in. (Informant ID4, 19.03.2019)

Challenges related to data velocity are linked with other challenges such as data variety and high volume. Receiving high volume data containing a high degree of variety in real-time streams pose both computational and data processing related challenges, particularly when the results has to be produced in real-time (Informant ID14, 22.05.2019). Informant ID14 provided an example of how the challenge of velocity causes scalability challenges:

> Let's suppose you have a dashboard that visualise different types of processed data in real-time by querying and retrieving data from storage (e.g., Postgres). The same storage is being updated with a high velocity and high volume data stream. Such high speed insertion and retrieval pose several scalability challenges, ultimately the system will not be able to properly entertain your users with the right data at the right time. So this scalability is one issue, which can not only be resolved by extending your infrastructure, but also you need to come up with robust indexing techniques, robust ways to insert your data to the data storage, and robust ways to select your data from the large data storage. (Informant ID14, 22.05.2019)

Variety
Identified as a challenge by a total of 11 informants.

There is a wide variety of data sources being used to gather information in emergency management. Being skilled on all the varieties of data sources that are available to you, ranging from satellite imagery, medical records and material data flows, to population sensors and social media requires a lot of resources. All the varieties of data sources has different approaches, and managing these widely and varying information sources is a challenge (Informant ID12, 15.03.2019).

The specific data is varying on the type of disaster - floods, earthquake, fires or typhoons each have their own characteristics, making the data varying (Informant ID5, 19.03.2019). E.g. a house damaged by an earthquake does not look the same on pictures as a house damaged by a typhoon or a flash flood (Informant ID3, 08.03.2019). *"I think the most critical challenge is to deal with this vast variety of data"* (Informant ID22, 03.05.2019).

> The other challenge is variety, right? Social media data is highly sparse. It contains so many different types of small mundane information, that sometimes our machine do not have any idea how to process this. Event or sub-event detection techniques can be employed to solve and filter out such noise. (Informant ID14, 22.05.2019)

Different formats of data is creating challenges for emergency management organisations. Not all formats are compatible with each other (Informant ID22, 03.05.2019). People use different languages, different identifiers for geographical locations, different databases, different types of coordinates, all making challenges for Big Data driven tools (Informant ID2, 18.03.2019). There are variations in systems like address systems, which spreads widely around the world: every country has a different address system (Informant ID12, 15.03.2019). Variety challenges are also relating to different data structures, where some data is structured, some is semi-structured or completely unstructured. Inserting all of that data into a structure format is very challenging (Informant ID4, 19.03.2019).

Variability
Identified as a challenge by a total of 3 informants.

There are great limitations to the level of analysis that can be carried out by Big Data technologies. This is because the information landscape is not suitable for Big Data, there is a lot of unknowns, and it is very dynamic. One of the challenges with data variability is the "small N big P problem", which can be described by a high number of variables within each data unit (Informant ID9, 14.03.2019). Further elaboration of "the small N big P problem":

> If you have an excel spreadsheet, typically when you build an algorithm what you have is maybe 10 variables, and 10.000 observations. Based on the passenger seeing that dataset you will begin to "tease" out some relationships. If you're smart you can build an algorithm that can help you somehow for the data. And what he is saying is for this you have an small N so you only have 15 villages or 15 units to analyse, but then you have a big P - maybe 200 variables, you have a big variability within that data. Small N big P. So you don't have a household survey of 15.000 households, you have directs observation and focus group discussion from 15 villages. (Informant ID9, 14.03.2019)

Variability of data is challenging for machine learning. E.g. Twitter messages of a catastrophe are not a uniform type of data, it is very chaotic. People write in different languages, use different words, misspell words, and it is a really hard for machine learning to figure this out in a short matter of time (Informant ID3, 08.03.2019). *"I think that those are the big lumps of it's information space, the uncertainty of the analytical framework, it's the changing nature of things, and it's the pressure to make decisions quickly. Those are some of the key issues"* (Informant ID9, 14.03.2019).

Veracity
Identified as a challenge by a total of 13 informants.

Big Data faces a lot of challenges in emergency management and one that is given much attention in the data collection is veracity. It is given much attention in the context of social media and news in the sense of the doubts of its trustworthiness, reliability, validity and accuracy.

> One of the big challenges which is related to the veracity of the data, not all data sources or twitter users can be trusted.. so determining the credibility of those people who pose during a disaster is very very challenging. (Informant ID14, 22.05.2019)

The topic of rumours or fake news is another research direction in this area, because it is hard to identify if a message is a rumour or a real message, particularly in a crisis. When analysing Big Data from social media such as Twitter, you are facing the challenges on how to identify people who just want to mess with those lines of communication (Informant ID21, 12.04.2019). *"What I know from my colleagues is that in many cases the challenges of using Big Data is confirmation of the source"* (Informant ID19, 08.04.2019). The biggest challenge is that there is many publishers of information. The media gets the hold of released information and blows it up in a way that the public gets their opinion of what is happening. While those who are formally going to handle the situation, like the police, cannot stand by that information as it might not be true (Informant ID29, 03.05.2019). If the data is from an external source it will be more timely but probably less reliable (Informant ID20, 02.04.2019). A fundamental issue is how you use information coming through social media. How can you analyse if you do not know if the information that you have collected is trustworthy or not? There are big problems validating the accuracy of the information (Informant ID23, 04.04.2019). "Quality is also something like the reliability in terms of does someone have an agenda behind posting that, is it probably fake news, does it seem to be coming from a human but doesn't really come from a human" (Informant ID20, 02.04.2019).

If you collect a batch of Twitter messages and clean them for retweets, you will in some cases find messages containing what you need, but risks it being false information. E.g. can a famous person obtain the same name as the town being hit by the disaster. In this case you will see a vast amount of messages about this famous person in the data collection for the disaster. These messages would have to be cleansed out, and is something that would have been problematic for artificial intelligence (Informant ID3, 08.03.2019).

The literature defines information quality in terms of the consistency, the accuracy, the objectivity, and/or the completeness of the information. These are the four main attributes of the information quality (Informant ID22, 03.05.2019). The challenges of these attributes is presented by informants in the data collection *"*challenge of understanding the data because sometimes we get the information but missing the context*"* (Informant ID21, 12.04.2019). Others define information quality different, but keeps the aspect of veracity: "I think there are two aspects to that, the first is the data quality in the narrow sense, so asking question like timeliness accuracy, availability" (Informant ID20, 02.04.2019).

> So they have a document there saying that this road is closed, and then organisations like IFRC has another document for the same date and they will be saying that the road is open. So a lot of information is there but a lot of things are not consistent and a lot of information is not usable. (Informant ID22, 03.05.2019)

Further issues entails a lack of information about how the data was collected, which standards people adhere to, or if it is representative (Informant ID2, 18.03.2019). "Usually it's just the quality of the data coming in. It is not of a high enough quality even if you do a nice template, people like to work in excel and mobile data collection. Still it's a lot of time checking over the details" (Informant ID30, 12.04.2019).

Visualisation
Identified as a challenge by a total of 4 informants.

The challenges with Big Data could be related to the data, to techniques that is used, and to the presentation - the visualisation or the way we present the results to the end users (Informant ID14, 26.03.2019). Informant ID21 elaborates further on this:

> Usability for me is usable visualisation, because you don't need to have a fancy chart that doesn't say anything. That is also perhaps important as a challenge when it comes to a crisis, I think, because they don't want to have any complex things to see during so many information already. Having Big Data and you don't have a way to make this information efficiently presented, then it is hard for them to use it. The information is perhaps useful according to our perspective, but practitioners don't have a way to use it because it is not visualised and summarised in a good way. You can grasp quickly for example. (Informant ID21, 12.04.2019)

Making information look understandable is an issue itself. The hard part is to analyse and sort out the information, find out which is reliable, and then try to synthesize the information. Then you have to make the information look understandable for the decision makers. This was experienced during the 2015's earthquake in Nepal. The most difficult part is to synthesize the information into something understandable for decision makers. This process is referred to as "from data to information to knowledge" (Informant ID22, 03.05.2019). One informant experienced a movement away from visualisation dashboards, because they have seen such complicated ones that a excel spreadsheet might be more suitable to visualise data than the dashboards (Informant ID30, 12.04.2019).

> So we're thinking a lot about how to simplify the visualisations for people who don't have as high data literacy, how can we turn it into a printed report where there is low bandwidth, where you can just print it off and have it in the field, but still be data driven. (Informant ID30, 12.04.2019)

Value

Identified as a challenge by a total of 9 informants.

There are several examples on how an emergency management organisation being able to collect a lot of data, but being unable to use it because of its lack of relevance. If the data do not provide a relevance to what you are looking for, it is lacking value, as Informant ID2 elaborates:

> Now you have modern communication and information technologies, that allow you to also remotely collect a lot of data, but that means all of a sudden you have a platter of data ranging from drones to satellite imagery to social media data to all kinds of other information and footage. And, but the same time you don't know anymore what the point and purpose was for collecting that information, so you have a lot, and that means you created a haystack of information, and if you really, you would start looking for a piece of information it's really looking for a needle in a haystack, or can be! (Informant ID2, 18.03.2019)

A similar data value challenge occurs with Twitter data, since it is only usable if it is "geotagged", meaning the location is pinned along with the Tweet. Only 1-3 % of Tweets posted are pinned with a geolocation, meaning that 97-99 % of the Tweets posted is unusable for analysis that use location as outputs. Meanwhile other analysis techniques such as discovering street names in the Tweets can be used as an alternative (Informant ID5, 19.03.2019). Informant ID4 spoke about the same problem: "The major problem here is location, if the location is not properly given in the tweet or in social media, then it is difficult to help people. That is one of the challenges" (Informant ID4, 19.03.2019).

The relevance of the information will always change depending on who is going to use it for decision making. You can have data with very high quality and reliability at hand, being unable to use it because it is not relevant to certain stakeholders. It is a challenge that the information might be good at addressing a certain thing, but it is not specified for who the information is good for (Informant ID21, 12.04.2019). Informant ID20 elaborates this part of the challenge as well: "The data quality in the narrow sense, so asking question like timeliness, accuracy, availability. These typical attributes that you can put to data. Really asking do we have the right data at the right time, with sufficient level of detail" (Informant ID20, 02.04.2019). Informant ID9 summed up the challenge of data value:

> There is real crap data out there, and there is real gold. Unless you know what you are dealing with, you can with visualisation programs today you can do amazing analysis

> and beautiful products that's absolute garbage, because your data is garbage. (Informant ID9, 14.03.2019)

5.2 Process Challenges

Process challenges is referred to as the challenges encountered while processing the data (Sivarajah et al., 2016, p. 265).

Data Acquisition & Warehousing
Identified as a challenge by a total of 3 informants.

The challenge of storage can be experienced differently. Informant ID21 have experienced challenges like missing data lines that is lost, and data getting corrupted during the process of collecting the data. This results in a lack of data for storage:

> Yeah because the storage challenge, at that time I think Nepal earthquake was in 2015, and that time perhaps I just, if you want to do it quick you use normal laptop first, and then if you say, what I did I create the storage, the SQL database, and I just set up at my laptop, my storage is as big as my laptop can accommodate, so that is perhaps the challenge. (Informant ID21, 12.04.2019)

In regard to how you store your data, there is a challenge posing a trade off between storing internally or externally, as presented by Informant ID20:

> And you could argue that for example if you store your data in-house, you have a better guarantee for the data being available, but at the same time you might have a lack of timeliness, because you need to feed it into your own systems. If you get it from an external source you get it more timely but probably less reliable. (Informant ID20, 02.04.2019)

Data Mining & Cleansing
Identified as a challenge by a total of 7 informants.

Data quality is a big challenge and an important research area. Data quality in health and emergency management is very important, and there are big challenges removing all the noise (e.g., irrelevant or routine data in social media data) from the data (Informant ID4, 19.03.2019). Data needs to be clean to make common operational datasets work, and the cleansing is essential to enable joint analysis, which is always a challenge. Data cleansing for metadata is also a massive problem (Informant ID9, 14.03.2019). The task of mining and cleansing data affects the timeliness. Users of the collected data is experiencing a cycle of having to go back, clean and ask the person who collected the information about further details to assure the data quality. This has to be done before the analysis stage (Informant ID30, 12.04.2019).

> Yeah if we could get rid of that, 80% sort of cleaning it means we could focus a lot more on analysis because we capacity wise always sort of struggled. You can think of some sort of great analysis you can do, but you may not have time. (Informant ID30, 12.04.2019)

Data Aggregation & Integration
Identified as a challenge by a total of 9 informants.

Aggregation and integration of data to make better use of it, both for internal and external use is an issue. High insertion and selection rates are creating aggregation and integration challenges, were it is not able to properly entertain all the users with the right data at the right time. One of the challenges with data aggregation and integration is the coding. An example of this is how to differentiate the message belonging to a specific information need, how to do this quickly, and how to sort the information into acceptable categories (Informant ID21, 12.04.2019). There are challenges arising for international emergency management organisations because of difficulties with data stored in infrastructure. This information is rarely categorised and not all of it is usable (Informant ID22, 03.05.2019). "Another thing is hard to tell what formats data is in, not all data is structured. Some data is semi-structured or completely unstructured. Putting all of that kind of data in to a structure format is very challenging" (Informant ID4, 19.03.2019). Integration challenges includes duplication of data, which can cause strange matchings if you e.g. are looking to trace people with a specific ID and/or name (Informant ID30, 12.04.2019).

Incorporating data is an issue that creates both challenges for past events, because of historical paper data, as well as for real-time data and future data. Taking advantage of the possibilities integration of the critical and relevant data gives can be difficult. Incorporation of parameters such as weather data and field level data during a disaster is a challenge (Informant ID26, 04.04.2019).

> Think about the tunnels in Kristiansand, when you have all those tunnels, I was just reading an article that there was a water leakage in the tunnel, that is a problem because it can cause cracks, landslides, and ultimately failure. But some tunnels have sensors that measure groundwater seepage. The real challenge is that during or prior to a crisis incident, we do not have the capabilities to combine this information with other data about the infrastructure to provide a proactive ability to make actionable decisions. (Informant ID26, 04.04.2019)

Informant ID20 informs that predictive models are used some places, but these are rather straightforward models. They are not integrated, and not leveraging a high level of AI or Big Data technology with integration of different types of data (Informant ID20, 02.04.2019).

Analysis & Modelling
Identified as a challenge by a total of 7 informants.

Analysis and modelling of data is a big challenge in emergency management, reasoned by the quality of the data (Informant ID4, 19.04.2019). There is a tradeoff between privacy and the analysis. A good analysis requires a large amount of data, and when this data is collected with personal information (e.g. username and addresses) from sources like social media there might be a danger of violating privacy regulations (Informant ID5, 19.03.2019). Making a readable analysis can pose challenges, Informant ID22 shared that the hard part was to analyse and sort out the information, finding which data is reliable and which one is not. Then to try to synthesize the information and the data. Finally the challenge of making the data look good occured, as well as transforming it into something that is understandable for the decision makers (Informant ID22, 03.05.2019).

A challenge is to develop techniques for analysis and modelling which can summarise the crisis situation. Instead of presenting raw tweets to the decision makers, a set of information summarising the situation for the last 5-10 hours would be preferred. A such analysis and modelling process would be a step forward on how to present the decision makers data, where they will get more benefit out of it (Informant ID14, 26.03.2019).

> I would say compared to other sectors, our analysis is reasonably rudimentary still, a lot of the analysis is just like overlapping two spatial layers and working on the intersection, there is not too much modelling and prediction. One, because the data is not timely enough which is why we started to look at these lightweight workflows in parallel, and an example of that is: I did some machine learning modelling of migrations going through the Balkans route, and it can predict for some countries, number of arrivals every day actually, maybe two days ahead, two within 20%, and that meant people on the ground would be resource more accurately what they would need for the amount of people coming through because they fluctuated by I think up to magnitude to about ten, like ten times from say 300 to 3000, but what we were finding was the data was so slow, so sometimes the data would be two days old and if you are predicting two days ahead you are just predicting what they are seeing right now in the field. (Informant ID30, 12.04.2019)

Data Interpretation
Identified as a challenge by a total of 7 informants.

It is important to understand the information in a correct context. A message does not necessary mean what it explicitly says, the context is sometimes needed to get a full understanding - which can pose as a challenge (Informant ID21, 12.04.2019). Other views on the challenge of data interpretation entails the use of methodology behind conclusions as told by Informant ID12. “Because the minute someone sees a key value on your results they just tune right out. That I think is a real challenge” (Informant ID12, 15.03.2019).

> I can give you a lot of refugee information, then either it can answer the questions like “how do we help the people in need there?” or it can answer the question “how do we build bigger walls in the balkans, so that nobody gets here?.” And that can twist the same information about refugee streams in two different ways, and relates to the processing. (Informant ID2, 18.03.2019)

The question of data quality, if you have enough data, and if it organised correctly brings Informant ID20 to reference research concerning interpretation of information:

> If you put a camera some point where people typically are fighting, the fighting in the city won’t go down they will just fight at a different place. Or I think there even was some findings particularly from the UK that people are fighting in front of the camera because if they are victorious they will be able to see it. (Informant ID20, 02.04.2019)

How you interpret the data, and from what perspective and from what standpoint may affect how you perceive the information. Through the interviews we were able to get insight into this in the context of crisis management. To anonymise informants we will refer to the organisation as Organisation X.

> If you are Organisation X you will want to shape your data so that it provides the output that this is a food-crisis. For example let's say you are Organisation X and you have an interest in doing a big food distribution program, you can then go out and do a food security analysis which shows you that 67% of the household only have two meals a day or one meal a day. Obviously that is a massive need for food, that's great. But then at the same time maybe there is plenty of food on the market, so you need to do some price-monitoring of the market so that you understand whether that there is a crisis for food or availability, no sorry access or availability! Is it because people can't afford the food or is it because the food is not there? Now that may change your intervention, so instead of just trucking and food, you may want to do a cash distribution, but that may change Organisation X's role, so they may have an interest in the trucking side of things. (Informant ID9, 14.03.2019)

5.3 Management Challenges

Management challenges is referred to as the challenges that tackle the privacy, security, governance and lack of skills related to understanding and analysing the data (Sivarajah et al., 2016, p. 265).

Privacy
Identified as a challenge by a total of 9 informants.

There will always be a trade-off between the privacy and the analysis. If you are supposed to conduct a thorough analysis, you need more data which in many cases poses dangers of privacy breaches, in particular for the social media data that is easily related to private information (Informant ID5, 19.03.2019). There is a need to be really mindful of the privacy concerns that surrounds the activity of utilizing data sources that involves social media data - transactional activity of people using the platforms. It is a ethical question of how you are engaging sources that could contain potential sensitive information about populations that you do not want to reveal (Informant ID12, 15.03.2019). The technology of today's digital devices enhances privacy concerns in emergency management:

> Even if you use all of anonymization technology, if you look at the geographical profiles if you have geolocation on all the time, it is typically, I would say with 95% accuracy possible to get back to the single person, simply from things like every morning you start from a certain point outside, then you stop at another certain stop, then you know this must be the home address and this must be the employer, and typically you only have one-to-one match of home addresses and employer. That's then already enough to know who someone is. (Informant ID20, 02.04.2019)

The human factor also plays a role in privacy challenges, since humans make mistakes. People are sharing things they were not supposed to share, and when they are working under pressure during crises they can get carried away when they find useful information, not thinking about the privacy paradox behind it (Informant ID3, 08.03.2019). Another human factor that creates privacy challenges is the interpretation of the regulations. The privacy laws, in this case specifically the Norwegian privacy laws, is to a certain level interpretable, meaning one person could interpret the law differently than the next. This could lead to people breaching privacy regulations, or lead to organisations holding back information due to a too strict understanding of the laws and regulations (Informant ID29, 03.05.2019). International organisations are facing challenges with legislation differentiating in different countries, as informant ID4 elaborates: “Every country have their own way of looking at this, and they handle it in completely different ways” (Informant ID4, 19.03.2019).

Security
Identified as a challenge by a total of 7 informants.

Emergency management organisations are experiencing challenges related to security. There are normally data breaches each year in big organisations (Informant ID22, 03.05.2019). The organisations often find themselves working in hostile environments, in areas of war for example. If you store information that is sensitive, it can harm people if it ends up in the wrong hands:

> So if you say that in this province in Syria this protection-problem, that’s one thing, if you say this village with 300 people living in it have massive protection problems, you are creating a danger for those guys. (Informant ID9, 14.03.2019)

> So if you are WFP and you have distribution lists with names, that data is highly sensitive and a vulnerability. (Informant ID9, 14.03.2019)

Data breaches can reveal the locations of humanitarian practitioners in the area, and cause harm to their lives in conflict zones. An important part of the security challenge is the tradeoff between availability and security. It is important to not have a too high level of confidentiality in the data storage, because the people who need the data will struggle to get access to it. On the other side, it has to be assured that those who are eligible to access the data are not intending to use it for bad purposes (Informant ID22, 03.05.2019).

Data Governance
Identified as a challenge by a total of 4 informants.

Informant ID12 names information management as one of the biggest challenges for emergency management organisations. Information management is referred to as the ability to effectively manage a large number of varied data sets, making sure that information management is appropriately done (Informant ID12, 15.03.2019). Informant ID12 further elaborated on the challenges with information management.

> I guess I would say there is a bunch of things involved in information management, making sure that information systems are organised precisely and that they are clean, that you are producing clean data on a regular basis, you know that metadata has been controlled, those kind of details are important. (Informant ID12, 15.03.2019)

Structure and management of the data is important when you are under pressure. In a disaster you have unclear responsibilities and unclear structures. The decision maker must have a framework for management - making this a challenge (Informant ID2, 18.03.2019).

> Because you have so many actors, because everything happens on a time-pressure, because the decision problems that you need the information for are not clear, there’s a lot of, or there is a lack of structure in the data-collection and it’s a bit, you have some sort of randomness and that is owing to, the ill defined nature of many problems. (Informant ID2, 18.03.2019)

Data & Information Sharing
Identified as a challenge by a total of 10 informants.

There are a lot of humanitarian organisations conducting data collection during disasters, but due to interoperability problems there is little data sharing between the organisations (Informant ID2, 18.03.2019). We were able to gain insight in the data and information sharing challenges that occured between two humanitarian organisations that operated in the earthquake in Nepal in 2015. To anonymise the informants, we will refer to the organisations as organisation X and organisation Y. Organisation X had an advanced information grid that they use in all crisis around the world, while organisation Y had a less advanced way of storing their information. Due to their different way of storing information internally in the organisation, these two organisations were not able to share information effortless between each other (Informant ID3, 08.03.2019). Each emergency management organisation often have information that could be used by other organisations in their own information silos, but the information cannot be shared between the organisations because of the different information silos from one organisation to the other. Breaking these silos is one of the biggest challenges when seeking to build resilience and better prevention of crises and catastrophes (Informant ID23, 04.04.2019). Informant ID22 elaborates this problem further: “Each emergency response organisation has their own information management system and they hardly collaborate with each other and finding out the common points between these platforms are a bit difficult” (Informant ID22, 03.05.2019).

There are challenges related to data and information sharing for infrastructure owned or operated by private organisations. The main issue is that sharing data about bridges for example can impact competitive advantage. Informant ID26 elaborated on this point:

> In our bridge infrastructure health monitoring project, we saw that there is a lot of hesitancy by private companies (owners/maintainers, builders and designers) to provide data or allow us to collect data using IoT for continuous health monitoring research. To address this concern, we have developed data sharing agreements with them that addresses data privacy/confidentiality concerns while being used for academic research. (Informant ID26, 04.04.2019)

Trust is another factor hindering data and information sharing. If the organisations does not know each other, they will hold back information. If they know each other well, they will have less problems sharing information between them (Informant ID29, 03.05.2019).

Cost/Operational Expenditures
Identified as a challenge by a total of 2 informants.

Informant ID20 elaborated on the challenges concerning cost and operational expenditures: “From a practical point of view you always need to bring in cost, and this is kind of the counterbalance to everything. So more you want to have of any aspect the more it will cost, and there for natural reasons there are limitations to where this can lead” (Informant ID20, 02.04.2019). The UN complains about this, when you work at the amount of money they earn through fundraising as a relative to what it would take for them to make analysis for crisis response, they are far away from meeting their funding targets (Informant ID12, 15.03.2019).

Data Ownership
Identified as a challenge by a total of 1 informant

Informant ID2 were the only interview subject that mentioned data ownership challenges: "And then of course there are questions of ownership, who owns data, where can you store it, how long can store it?" (Informant ID2, 18.03.2019).

5.4 New Challenges Emerged

In this part of the chapter we will present the challenges presented by our data collection that were not identified as Big Data challenges in the context of emergency management by Sivarajah et al. (2016).

Lack of Knowledge
Identified as a challenge by a total of 6 informants.

The data collection gave results showing there is a lack of knowledge on how to use and utilise Big Data technology in crisis management. This lack of knowledge appears in different stages of the technology. There is a certain lack of knowledge on how to use the technology, as analogical, classical ways of doing work under stressful situations seems more favourable (Informant ID20, 02.04.2019). There is a lot of hype around the term Big Data, but the reality is that very few people in the emergency management field knows how to analyse, how to sort, and how to make good outputs from Big Data, the skills are not there (Informant ID22, 03.05.2019).

There is a lack of knowledge among decision makers on how to read the outputs of Big Data analysis: "I would say that among decision makers, statistical literacy is pretty low. You can't just hand someone the results of your analysis, and tell them 'here is the key value'. They just won't understand it" (Informant ID12, 15.03.2019). The technology platforms is complicated for practitioners: "When they want to look into specific things for these databases they can hardly find out where to look into, how to collect all those information in a usable format, and then how they can use it for further analysis" (Informant ID22, 03.05.2019). This creates a need for the analysts to make outputs that cannot be misunderstood, outputs that is put in a very clear and direct way (Informant ID12, 15.03.2019).

Data Access
Identified as a challenge by a total of 11 informants.

The paradigm for emergency management has often been that there is too little information (Informant ID2, 18.03.2019). Reasons for this can be that disasters like typhoons, hurricanes or earthquakes knocks out the communications network, meaning there is no transactional or user data to collect (Informant ID3, 08.03.2019; Informant ID4, 19.03.2019). Informant ID14 labeled data scarcity as the number one challenge that his emergency management organisation were facing: "During a disaster the data scarcity is one of the challenges. Having less data means less decision making. Having more data means more informed decision making" (Informant ID14, 26.03.2019). Informant ID30 elaborates more on how collecting data is a challenge:

> It's a real struggle to get data from the field in a disaster response, so thinking about Mozambique at the moment, we collected about where organisations were working,

> where the affected people are, so we have needs assessments going on. So flying planes over trying to identify where the affected communities are, but it's that sort of sudden thing where we're working in a situation were it instantly changes over night, it's not always slow unsets. (Informant ID30, 12.04.2019)

Using Big Data technology that are basing its data collection on social media data can also face challenges relating to data access. There are different usage of different social media platforms in different countries, and there are people that are not using social media. In those cases you will have to rely on the emergency service numbers to get access to data about the victims of the disaster (Informant ID4, 19.03.2019).

Resource Scarcity
Identified as a challenge by a total of 5 informants.

There are a lot of challenges relating to the scarcity of resources in the organisations that work around a crisis. Informant ID12 (15.03.2019) named resources the biggest challenge his/hers organisation is facing. Informant ID26 (04.04.2019) informed that they do not have the human capital to do the data collection that they desire. The sheer volume of data that is produced by internet connected devices has increased dramatically the last years, and the emergency management organisations are not able to increase their human resource capacity in step with the information flow coming their way (Informant ID12, 15.03.2019).

> I think a lot of it has to with that we don't have enough people to do the work. And that is true in a lot of other non-profit humanitarian work - the need and the complexity of the problems dramatically outstrips the investment in those problems, and to some degree that is why these organisations exist, it is because we are dealing with all these extraordinalities of a complex world with many problems. But on the analyst side we don't have close to enough people to actually do the analysis, so we end up not doing a whole lot of things that actually should be done, but there isn't enough time a day to do it - there isn't enough people that actually carry out all of the analytical work. Even in much larger organisations this is the case. That kind of persistent under staffing is something which I think ripples into everything. You are designing things in advance as you know you will never have enough people to do this. I still think it is important to do but it will never be out of it. So you are always kind of behind in that process. (Informant ID12, 15.03.2019)

The resource scarcity creates challenges beyond the organisation itself, it creates challenges for other organisations that are trying to use the information they have collected. The resource scarcity makes it hard for organisations to clean and sort out the relevant data with good metadata. They do what they need for themselves, but do not have time to clean the irrelevant data - which might be relevant to other organisations, since they are uploading it to online sharing platforms (Informant ID3, 08.03.2019).

Time Pressure
Identified as a challenge by a total of 7 informants.

Due to the fact that crises and catastrophes often has a risk of costing lives, time pressure is a natural challenge that exist with Big Data technology in crisis management, that would not necessarily be a challenge in other Big Data usage areas. Informant ID20 elaborates more on the nature of time-pressure in emergency management:

> In crisis management you often need to very quickly make decisions, because the situation will otherwise worsen, particular if you have free resources at hand, you need to do something. So if you for example have 100 helpers at hand and there is flooding in the city, you need to do something, and you can't wait for you have all information with sufficient quality. (Informant ID20, 02.04.2019)

The Big Data data collection tools has to be set up quickly, because if a lot of time us used to setting up a collection system you will lose critical information (Informant ID21, 12.04.2019). In some organisations, data collection tools is only used in drills because they do not have time to set it up in a real crisis (Informant ID29, 03.05.2019). Due to the nature of crises or catastrophes, decision makers are under severe pressure to make decisions quickly (Informant ID9, 14.03.2019). Informant ID12 elaborates more on this point:

> You need to produce information within a time frame that is relevant to someone to actually use that information and do something with it. So it's always very practical, and that practicality produces a kind of "speed" imperative as a constant challenge that increases when you get more information you have to handle. (Informant ID12, 15.03.2019)

Informant ID19 tells more about how emergency management workers are struggling with time pressure:

> The challenges that they have that I know is time - they don't have time. They are working under very difficult conditions, and are working probably 15 hours a day. And time is always a challenge for this. They do it, but it is always a challenge to find the time for everything I would say. (Informant ID19, 08.04.2019)

Every aspect of the Big Data process from collection, through cleansing to visualisation takes time, in particular cleansing of unsorted data. The aspect of time pressure hinders a thorough and extensive analysis of crisis data (Informant ID30, 12.04.2019).

Data Neutrality
Identified as a challenge by a total of 3 informants.

The challenge of data neutrality is referring to that there is no such thing as objective information. The data or information is always tainted, and it is always somebody's data collection and interpretation behind it (Informant ID2, 18.03.2019). Someone can always have an agenda behind the information, or being able to shape the output to suit the decision point (Informant ID20, 02.04.2019; Informant ID9, 14.03.2019). Moreover, data neutrality challenges often arise in conflict situations, because of the data sensitivity. Emergency management organisations need to principled, meaning it has to adhere to certain standards of neutrality and impartiality. E.g. if an artificial intelligence system is programmed so that it is structurally overlooking or favouring a specific minority, that will be a problem. If the information you run the algorithms on is not neutral, it will have no value (Informant ID2, 18.03.2019).

Digital Divide
Identified as a challenge by a total of 3 informants.

Digital divide refers to the divide between the people that has a digital footprint, and the ones that do not. When moving disaster response solutions over to digital platforms, you will be facing the same challenges as you would in any other digital platform with a humanitarian responsibility, what do you do about the people who do not have a digital footprint in an AI world? (Informant ID2, 18.03.2019). Many people are not using social media or the digital communications at all. Informant ID4 experienced this when testing Big Data technology in crisis response in Norway:

> My experience with emergency handling in Norway, especially the 2014 floods in this area, we experienced that there where very few tweets. First of all there are a lot of old people staying in more remote regions, and they don't use Twitter and social media. In that case we have to rely on the emergency numbers - phone calls. The emergency services have special staff for handling these kind of calls. These calls are a way of getting the emergency information. That is why it is different to make everything automatic. That is the problem, many people are not using social media. (Informant ID4, 19.03.2019)

Europeans are used to be operating in data-rich environments, but this is not the case all around the world: "if you had to do Big Data analysis in the Central African Republic, you'd struggle. It is not a data-rich environment, you simply don't have very good data" (Informant ID9, 14.03.2019). Some people might be poorer than others, and therefore not be producing the same amount of data (Informant ID9, 14.03.2019).

Sensemaking
Identified as a challenge by a total of 7 informants.

In the context of a crisis the issue of making sense of an ambiguous situation is something that often has to be done with a lot of uncertainty and under time pressure. "So making these decision under uncertainty is a real challenge in data quality, because if you don't have the data or the better the data you have the better you can make decisions." (Informant ID20, 02.04.2019). It is difficult to process data if you do not know the purpose of which it is being processed for. A lot of the data that is processed, is processed with a decision maker in mind, which often is not present in the field. The decision maker uses it for the purpose of advocacy, rather than for operational decision making. There is a lot of differences in the contexts of strategic response and operational decision making (Informant ID2, 18.03.2019).

Situational understanding and situation recognition is one of the biggest challenges in this field. Because when you are providing emergency services during the crisis the first thing you have to do is to get an understanding of the situation (Informant ID4, 19.03.2019). The challenge is to extract and learn about the issues that the crisis response are facing in the field (Informant ID14, 26.03.2019). Sensemaking can also represent the data signals received, how to make sense out of a situation from the data.

> Maybe half the cell towers are knocked out in the south, and you're not getting all the data, maybe electricity is down, maybe people in the center is poorer and don't have as many phones so you don't get as many signals, so you actually don't know whether there is signal or noise, is it data noise or is it a signal? Unless you have an analytical

> framework which helps you make sense out of it, this is also why cannot do a representative sample in a very kinetic situation. (Informant ID9, 14.03.2019)

Cold Start
Identified as a challenge by a total of 6 informants.

In the data collection we found that before, during and after a crisis the challenge concerning the use of Big Data for analysis is an issue. This is because an analytical algorithm or model needs training data for a specific situation, and rarely no crisis is alike (Informant ID14, 26.03.2019). Depending on the type of disaster; earthquake, flood, fire, etc., they obtain their own characteristics, and robotics is reliant on a specific situation with a specific characteristics of the disaster to work correctly (Informant ID5, 19.03.2019).
If you have trained a analytical algorithm to a certain level in catastrophes in one country on a typhoon, and then discover a similar storm somewhere else in the world, the data collection will still be in different words, other languages, totally different place names. If you use it on a different disaster like an earthquake, it will be different pictures - not blown down houses, but crumbled buildings (Informant ID3, 08.03.2019).

> The first challenge in our data processing pipeline is the cold start challenge. So that means for a new event if we have never dealt with the specific language in that region, that is happening, that it's either our system doesn't know how to process that regional data that is coming other than English language, so we have to start our system in a cold-start manner, so what we have to do; we have to provide as quickly as possible our system the supervision in the native language, so that the system gets up and start processing our data in real-time. This is one of the big challenges that hinders the process. (Informant ID14, 22.05.2019)

> The challenging part specifically for the machine learning and neural networks algorithms is that you have to train the machine for this kind of things, and we have to train algorithm for this. The training data is very very hard to find. (Informant ID22, 03.05.2019)

Phases of a crisis is also affected because a model trained in the early phase of a disaster will not be applicable in the during phase of a disaster (Informant ID14, 26.03.2019). The need for training data that is representable for the specific situation is an important point, if a model does not have the correct training data the analysis will not be valid (Informant ID9, 14.03.2019).

Culture Differences
Identified as a challenge by a total of 9 informants.

Culture differences between countries creates challenges for the use of Big Data and machine learning in emergency management. Different languages, alphabets, multiple and inconsistent placenames as well as abbreviations of location names, all create challenges for a Big Data tool to work (Informant ID3, 08.03.2019). Informant ID21 gives an example of a culture differences challenge the informant experienced during the Nepal earthquake in 2015: "Many Twitter users are using Indian characters that are uninterpretable by the machine for example" (Informant ID21, 12.04.2019).

> Let's suppose after an earthquake, depends, depending on the region, if it is happening in Nepal the types of problems that Nepalese will face would be different than the types of problems if that earthquake happens in for example in Europe right? In Italy? So because the infrastructure there is different, the living standard compared to Europe is different, and many other factors, so such factors influence the, the types of issues that the affected people have. (Informant ID14, 26.03.2019)

There are differences between countries on how you can collect data. Twitter for example is not very popular in Norway, so a data collection system basing itself on merely Twitter would meet challenges when collecting data in Norway (Informant ID20, 02.40.2019). Another challenge created by culture differences is the difference in legislation between countries. Each country have their own way of looking at e.g. privacy regulations, creating challenges when collecting personal data (Informant ID4, 19.03.2019).

Further culture difference challenges faced by humanitarian response workers are the wide variety of nationalities working together and cooperating under a crisis or disaster. The way of work is different from culture to another culture, creating cooperation challenges. These are challenges for those who are working to develop emergency management platforms. (Informant ID4, 19.03.2019). This was also experienced by Informant ID30: "It can be very difficult to coordinate across, data collection across national societies were are all working different ways" (Informant ID30, 12.04.2019).

6.0 Discussion

The chapter of discussion is reviewing the results from the data collection up against the literature identified in the literature review, to move towards a conclusion. Throughout the data collection we were at multiple occasions asked to specify what we meant by Big Data in emergency management. Initially we referred to the Big Data definition presented in chapter 2.0 Theory: a process that facilitates the decision making, through analysis of large amount of data - of different types, from a variety of sources, to produce a stream of knowledge (Fertier et al., 2016, p. 4; Power 2014), as well as the 3V's definition: "high volume, high velocity and high variety" (Laney, 2001; Meier, 2015). By using these definitions, in particular the 3V's definition, we were challenged on the use of this kind of technology in emergency management, Informant ID12 elaborated why this definition might not be suitable:

> I think the Big Data needs a little bit more specificity, just because a lot of things get labeled Big Data when they are not really, just because the term gain a certain amount of currency and people like marketing. So basically what you're doing is taking a number of those data sources, some of them are quite big, and some of which are data - but they are not like "Big Data" in the sense of they are not frequently changing. I feel like in order to constitute "Big Data" it has to be something that is a flow and time element means something. (Informant ID12, 15.03.2019)

Big Data does not entail that you need to acquire a whole warehouse of data, we want to define it as: "more data than is easily feasible to be analysed by a single person in a short amount of time" (Informant ID20, 02.04.2019).

6.1 Theoretical Contributions

Through this part of the discussion chapter we will present our theoretical contributions to the field of Big Data in emergency management. The theoretical contribution is based on the challenges presented in 5.0 Results as well as existing literature. There has been some modifications to the new challenges emerged presented in the results, and some to the existing Big Data challenges presented by Sivarajah et al. (2016), to make them applicable to the context of emergency management. Figure 4 graphically presents the modifications to the challenges.

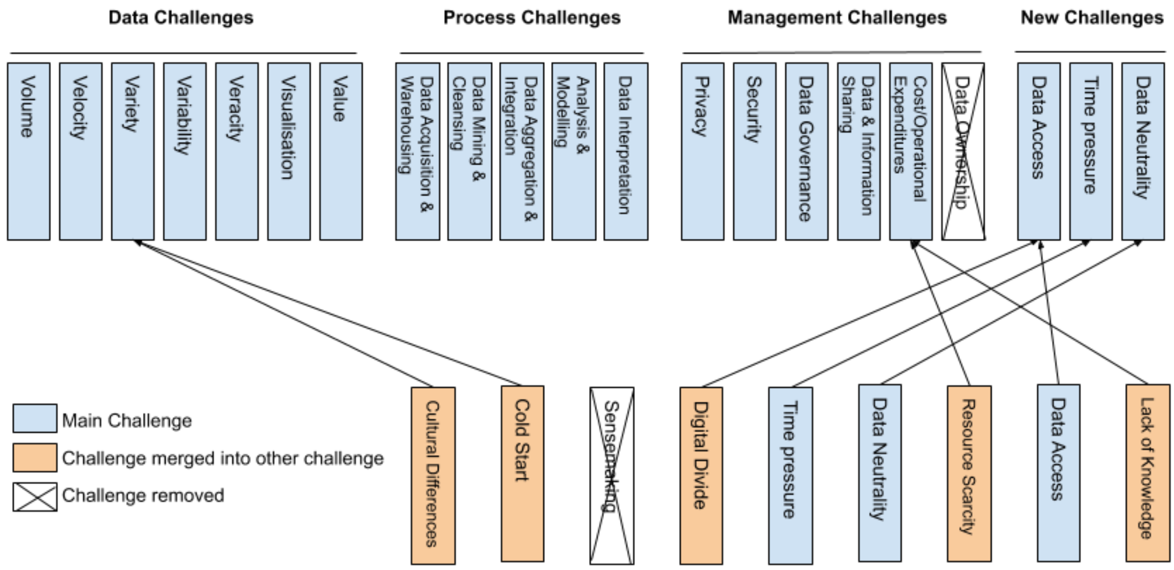

Figure 4: Graphical presentation of the modifications to the Big Data challenges by Ø. S. Fongaard & T. F. Nestaas, 2019.

Data access, time pressure and data neutrality is added to the model of Big Data challenges in emergency management, while data ownership is removed. Cultural differences and cold start is merged as an extension to the challenge of variety, digital divide is merged as an extension to data access, while resource scarcity and lack of knowledge is merged as an extension to cost/operational expenditures. Sensemaking is not considered a Big Data challenge and is removed. After these modifications we are able to add the three new challenges of data access, data neutrality and time pressure to the original classification of Big Data challenges model presented in Sivarajah et al. (2016), see figure 5.

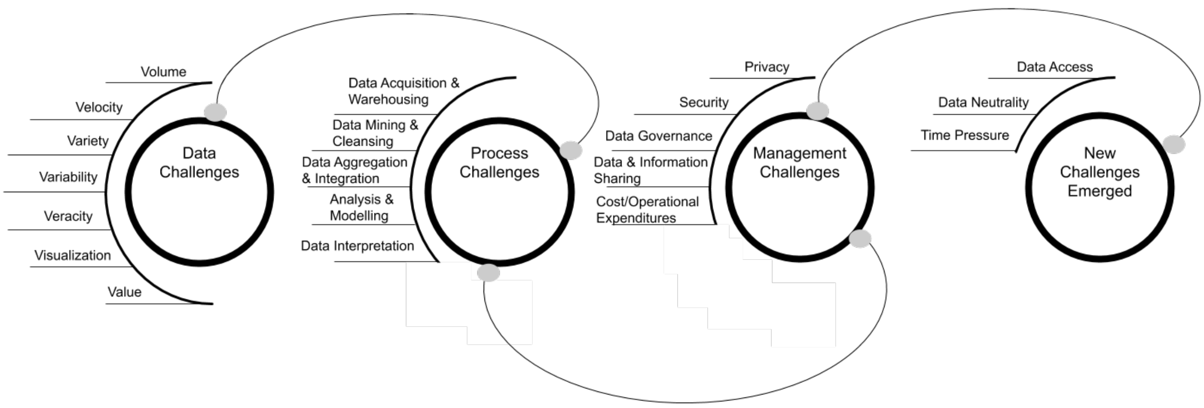

Figure 5: Graphical presentation of Big Data challenges in emergency management by Ø. S. Fongaard & T. F. Nestaas, 2019.

6.1.1 Data Challenges

The Big Data challenges presented in Sivarajah et al. (2016) includes the data challenges of volume, velocity, variety, variability, veracity, visualisation and value, also referred to as the seven V's.

The large scale and big volume of data, which can be terabytes, petabytes or zettabytes of data, is a challenge of its own. (Sivarajah et al., 2016, p. 269). This is also reflected in the results, where it was discovered that the sheer volume of the data is making it hard to collect quickly (Informant ID30, 12.04.2019). Information overload is one of the biggest challenges in emergency management (Informant ID24, 05.04.2019). On the other hand, these statements do not comprehend if it is due to the high number of terabytes of data, or the low ability to collect data - which can be caused by a whole set of other reasons, such as variety, lack of knowledge or resource scarcity. It is important to add that the challenge of data volume is only experienced in the response phase of the disaster (Informant ID24, 05.04.2019). The challenge of data volume were identified by a total of seven informants.

The data challenge of velocity appears with the need to handle the speed with which new data is created or existing data is updated (Chen et al., 2013, p. 158). Velocity creates a data challenge in emergency management because the field is reliant upon real-time data streams. The combination of data velocity and variety in real-time data streams creates challenges when trying to deliver real-time results. There is a need to develop more robust indexing techniques to insert data to the data storage to tackle the challenge of data velocity (Informant ID14, 26.03.2019). Just as the challenge of data volume, velocity is only a challenge experienced in the response phase of the disaster (Informant ID24, 05.04.2019) Data velocity were only identified as a challenge by two informants.

Different types of data, such as diverse and dissimilar forms of data creates the challenge of data variety (Sivarajah et al., 2016, p.269). Diverse and dissimilar types and forms of data is equally present in emergency management. The data sources can range from satellite imagery, medical records and material data flows, to population sensors and social media (Informant ID12, 15.03.2019). Data variety were identified by a high number of informants, a total of eleven informants identified variety as a challenge. One of the main data variety challenges in emergency management is that each disaster have their own characteristics, making data such as pictures different (Informant ID5, 19.03.2019; Informant ID3, 08.03.2019). In the results, cold start was identified as a new emerged challenge. This challenge has several similarities with the challenges related to data variety. Cold start is the challenge of deploying an analytical algorithm based on Big Data in a new setting, which it has not been feeded training data for. The training data is hard to find, because the training data has to be specific for each disaster (Informant ID22, 03.05.2019). Every difference down to the detail makes it hard for a analytical algorithm to work as well as in previous situations. Therefore the system has to be “cold started” for each disaster. In other words, this means that the variety in the data from one disaster to another, creates the cold start challenge, meaning that the cold start challenge identified in the data collection belongs as a extension to the variety challenge, not a new emerged challenge. The cold start challenge is mainly related to Big Data tools that utilises artificial intelligence and machine learning technology. Culture differences is also identified as a new emerged challenge by the informants. Culture differences entails the differences in language, alphabet, location names and surroundings, creates challenges for Big data technology in international emergency management. We see this as a second extension to data variety, since culture differences can be a synonym to culture variety. The difference in culture creates a variety challenge in the data because of different languages, alphabets and so on.

The challenge of data variability refers to how the data is constantly changing over time, because of e.g. processing (Gudivada et al., 2015, p. 3; Sivarajah et al., 2016, p. 273). The information landscape in emergency management is not suitable for Big Data, partly because of its dynamicality. Additionally, to conduct a proper Big Data analysis a high number of observations on a low number of variables is required. Crisis management entails the opposite: a low number of observations on a high number of variables - e.g. households (Informant ID9, 14.03.2019). There were a low number of informants identifying variability as a challenge with the use of Big Data technology in emergency management, but the challenge was extensively reviewed among the three informants that did identify it.

Data veracity can be defined as the imprecision, inaccuracy, inconsistency of large datasets that may give and/or support a wrong perceived meaning of the data (Sivarajah et al., 2016, p. 269). The challenge of data veracity is a known challenge among several of the informants, as a total of 13 informants identifies the Big Data challenge in emergency management. Big Data technology that bases its data collection on social media data is vulnerable for veracity challenges, as there are limitations in the sense of its trustworthiness, reliability, validity and accuracy. There are limitations on how to confirm that the sources are representative (Informant ID19, 08.04.2019), and there are big problems validating the accuracy of the information (Informant ID23, 04.04.2019).

Presenting the data in a readable manner defines visualisation. It is about representing key information and knowledge more instinctively and effectively through using different visual formats like graphs or dashboards (Sivarajah et al., 2016, p. 273). Only four informants identified visualisation as a challenge, but those who did could elaborate it well. Presenting the data on short time, with quality that gives the end user a value is a challenge during a crisis. The techniques that they use to present the data is not centered around the end user (Informant ID14, 26.03.2019). The techniques are focused on how much data that can be fed to the end user, but this might be the wrong focus. The end user often needs just a quick grasp at a simple summary or visual tool, something that can be read by anyone in a matter of seconds or minutes (Informant ID21, 12.04.2019). Operating in a crisis with Big Data that needs to be made into readable information furthers the challenges. Producing readable information from large amounts of data retrieved from sources like sensors or social media, and then presenting the key information and knowledge effectively through different visualisations is a challenge (Chen et al., 2012, p. 1166). This is referred to by Informant ID22 as “from data to information to knowledge” when making data understandable for decision makers (Informant ID22, 03.05.2019). The pressure of time and data exchange with people with low bandwidth connection, and how to present data to them creates challenges. Because of time-pressure, emergency management organisations need to develop techniques and communication to learn how to create the best visualisations for people in the field, people who need fast readable information to supplement and support their actions and choices (Informant ID30, 12.04.2019).

The knowledge and edge which we may extract from large amount of structured and unstructured data is the value of data in Big Data (Sivarajah et al., 2016, p. 273). In emergency management there is modern communication and information technologies that allow remote collection of data, giving you a vast specter of data from drones to satellite images to social media data, but you do often not know the purpose of the data anymore (Informant ID2, 18.03.2019). If the data is collected for the purpose of finding food or water it might not be to any value for those calculating amount of people who needs medicines. A challenge is that you could have very high quality data, in the sense that is reliable, but if it is not relevant or addressed to who might make use of it, it has no value (Informant ID21, 12.04.2019). There is a lot of data collected for emergency management, and willingness of sharing the data does not seem to be an issue. The problem might originate in the way data is collected, and how it is catalogued to give value for those who need it for decision support. Amazing analyses can be created, but with no value for anyone if the data is without relevance (Informant ID9, 14.03.2019). In the context of emergency management the challenge of value with Big Data seems to pressed by shortage of time. You need the right data at the right time, with enough detail to act on it, which is usually not the case during a crisis (Informant ID20, 02.04.2019). As it is identified as a challenge by a total of nine informants, it makes value one of the more acknowledged challenges. Value is also a challenge that entail different viewpoints, since the challenge has roots many places.

6.1.2 Process Challenges

The challenge of data acquisition and warehousing is related to acquiring data from diverse sources and storing it for value generation purpose. With Big Data brings complexity which can cause unforeseen problems for the Big Data engineering such as data acquisition and storage (Sivarajah et al., 2016, p. 273). Only three informants identified acquisition and warehousing as a challenge. Information from the informants showed that context related issues were experienced as well as the complexity and unforeseen problems. Problems with data getting corrupted during the collection were experienced, resulting in a lack of data. Another challenge could be the resources you have at hand, if you limited to setting up the database on your laptop you will not be able to accomodate for the vast volumes of data (Informant ID21, 12.04.2019). In the perspective of the context of emergency management, good quality data is required to act on it, preferably both timely and reliable. There seems to be a trade off on this, you can choose to store data in-house for better availability and reliability, but less timeliness because of your systems, as opposed to externally with better timeliness, but less reliable (Informant ID20, 02.04.2019).

Extracting and cleansing data from a pool of a large scale unstructured data is defined as the data mining and cleansing challenges by Sivarajah et al. (2016, p. 273). The issue of cleansing is recognised as a big challenge since you have to remove all the noise (Informant ID4, 19.03.2019). Informant ID9 supports this, particularly concerning metadata. Data needs to be clean to make common operational datasets work, and the cleansing is essential to enable joint analysis, which might prove further challenges of data and information sharing (Informant ID9, 14.03.2019). For the context of emergency management, as mentioned in connection with storage and warehousing, timeliness is important. Mining and cleansing data are a time-consuming task (Informant ID30, 12.04.2019). This challenge is mentioned by seven informants, and seems to be a big challenge for the use of Big Data in the field emergency management. It is a challenge that can be seen in connection with time pressure, having to produce clean data for processing to support decision making.

In emergency management Big Data can aggregate online activities through e.g. Twitter, but integrating them can be a big challenge. This can involve aggregating and integrating clean data, mined from large unstructured data, and maybe even from diverse types (Sivarajah et al., 2017, p. 274). These challenges can be seen by the coding of the message, how does one message differentiate itself for a specific information need or category (Informant ID21, 12.04.2019). Without aggregating and integrating the data, its use or potential might not be met. In the international emergency management organisations this is an issue. They often do not categorise their information, and much of it is unusable because of this (Informant 22, 03.05.2019). The reason can be that they do not have time for it in a crisis. Because of the variety of data sources that is present at the time of a disaster, the challenge of aggregation and integration arises. This challenge can then further affect the visualisation and value of the information gathered (Arslan et al., 2018, p. 1). Over half of the informants identified aggregation and integration as a challenge. Some mentioned the problem in a different manner, that the integration of data might cause duplication, which could be critical for data used for an emergency (Informant ID30, 12.04.2019). Duplication can further cause issues for sharing of information between organisations and departments (Abdullah et al., 2017, p. 407).

Big Data can often have the characteristics of being noisy, unreliable, heterogeneous and dynamic by nature which can make it hard to create analysis and models of it (Sivarajah et al., 2016, p. 274). In the context of emergency management, bad data quality makes it challenging to perform analysis and modelling (Informant ID4, 19.04.2019). The hard part of making a readable analysis is to analyse and sort out the information which is reliable (Informant ID22, 03.05.2019). At the same time there is a tradeoff between privacy and the analysis. The more information you obtain the better the analysis can be, but the use of personal information might hinder this process (Informant ID5, 19.03.2019). At the moment compared to other sectors emergency management organisations analysis is still reasonably rudimentary. There is not too much modelling and prediction (Informant ID30, 12.04.2019). This challenge can be in connection with the resources they have and in the context the technology has to work.

The challenges of understanding the output, visualising and sharing the results of analyses and models refers to the data interpretation. (Sivarajah et al., 2016, p. 274; Zicari, 2014, p. 110). Understanding the information at hand in a crisis could be critical, and a message does not always mean explicitly what it says. You sometimes need the full context for understanding a message (Informant ID21, 12.04.2019). How you form and process the information you send can affect what the information means when it is interpreted. The context and standpoint of an actor or organisation that provided the information may change the meaning of it (Informant ID2, 18.03.2019; Informant ID9, 14.03.2019). Information can be tainted, badly formulated and therefore interpreted badly, and in emergency management it seems to be no different.

6.1.3 Management Challenges

Data that identifies a person inevitably raise privacy issues (Nature Editorial, 2007, p. 637). In particular information containing positioning and location data poses clear privacy concerns in Big Data Analytics (Sivarajah et al., 2016, p. 274). Social media data is being used continuously in emergency management, creating a need to be really mindful about privacy regulations and concerns around engaging sources that could contain potential sensitive information (Informant ID12, 15.03.2019). Even if you are using all possible anonymization technology on geolocation data, you will be able to get back to the single person with a high accuracy (Informant ID20, 02.04.2019). At the same time it can be hard to be compliant with all privacy regulations, both since privacy legislation is differentiating in different countries (Informant ID4, 19.03.2019), but most of all because there will always be a tradeoff between a thorough analysis and privacy compliance (Informant ID5, 19.03.2019). In order to make a good analysis of a crisis situation you might have to cross privacy legislation barriers. At the same time many crisis situations involves the risk of people losing their lives, creating a new ethical discussion asking: is it okay to cross privacy legislation barriers in order to get access to the location of people in desperate need? Privacy regulations under normal situations and under crisis situations are different, under a crisis you will probably want to share your private information to crisis response workers in order for them to reach you (Informant ID21, 12.04.2019). This means that the balance of optimal privacy regulations is a whole other discussion than the optimal level of privacy regulations under normal surroundings. At the same time it is important for the crisis response to have the challenge of privacy in mind, so that only necessary breaches of privacy legislations is conducted in order to save lives in a crisis situation, therefore privacy is not removed as a management challenge.

Data security consists of threats such as malware, viruses and lack of security controls (Sivarajah et al., 2016, p. 274). Emergency management organisations are experiencing security challenges as well, as there are normally breaches each year in the big organisations (Informant ID22, 03.05.2019). It is important to split the discussion of security challenges into two main categories: Emergency management organisations working in hostile environments such as areas of war, and emergency management organisations working in non-hostile environments such as nature disasters. For those working in hostile environments, security is a big challenge if they are storing sensitive information. This is because a data breach can result in the sensitive information being used for hostile activities, e.g.: “So if you say that in this province in Syria this protection-problem, that’s one thing, if you say this village with 300 people living in it have massive protection problems, you are creating a danger for those guys” (Informant ID9, 14.03.2019). On the other side, for organisations working in non-hostile environments, it can be argued just like the privacy challenge, that it is better to focus on saving peoples lives, than to be very strict on data security. Because of the importance of data security in emergency organisations that operate in hostile environments, security is kept as a management challenge.

Organisations perceive data governance as a way to warrant data quality, improve and leverage information as well as maintain its value as a key organisational asset (Otto, 2011; Sivarajah et al., 2016, p. 274). Managing large numbers of varied data sets effectively and making sure information management is appropriately done is some of the noticeable data governance challenges in emergency management. Without a proper structure in the data collection, you will get a sort of randomness in your data (Informant ID2, 18.03.2019). Such a randomness in the data creates challenges for Big Data solutions to conduct proper analyses in emergency management.

The challenges regarding data and information sharing between different departments and organisations includes overlapping of applications, duplication in information, and confusion around the responsibilities of each business unit (Abdullah et al., 2017, p. 407; Tekiner and Keane, 2013). All these aspects of data and information sharing issues are experienced in emergency management as well. Each organisation conducting data collection has their own way of storing the information, in their own information silos. These information silos are often overlapping each other, making it hard to share data between them (Informant ID23, 04.04.2019). There is often a confusion around information and the responsibilities of each organisation because of differences in goals, resources and tasks within each organisation’s area (Informant ID24, 05.04.2019).

There were only two informants of the total of 16 that mentioned cost and operational expenditures as a challenge in emergency management specifically. Cost and operational expenditures is a consequence of intensive processing operations which results in high storage and data processing costs (Sivarajah et al., 2016, p. 275). Cost is a counterbalance to everything, if you want to have more of any aspect it will cost, and is therefore a natural limitation (Informant ID20, 02.04.2019). Among the existing challenges discovered in the data collection, some new challenges emerged as well, including lack of knowledge and resource scarcity. Lack of knowledge includes the lack of know-how to use the Big Data technology (Informant ID20, 02.04.2019). There are very few people in the emergency management field that knows how to analyse, sort, and make good outputs from Big Data, because they have not acquired the necessary skills to do so (Informant ID22, 03.05.2019). Still, none of the informants elaborating on the challenge of lack of knowledge added that the challenge originated in the challenge of cost and operational expenditures. Nevertheless, the knowledge of Big Data in the commercial sector is very well developed, years of development in front of the emergency management sector (Informant ID22, 03.05.2019). Based on this we assume the lack of knowledge challenge is caused by the aspect of cost and operational expenditures that is required to develop Big Data knowledge. Resource scarcity is another new emerged challenge discovered in the data collection. The amount of data that is produced by internet connected devices has increased significantly the last years, and the emergency management organisations have not been able to increase their human resource capacity in step with the information flow coming their way (Informant ID12, 15.03.2019). They do not have the human capital to conduct the data collection that they desire (Informant ID26, 04.04.2019). This can be caused by several reasons: the organisation can have a hard time recruiting human capital due to workplace unattractiveness, or they might not have the financial capacity to hire the necessary human resource capacity. As emergency management organisations like e.g. Médecins Sans Frontières has established itself as a global movement and obtained over 42 000 people with staff from over 150 countries (MSF, n.d.), we do not consider that the lack of human resource capacity is caused by workplace unattractiveness, but rather assume that the resource scarcity is caused by cost and operational expenditures. The new challenges emerged presented in 5.0 Results: lack of knowledge and resource scarcity is therefore merged as an extension to the challenge of cost and operational expenditures.

There were one informant mentioning data ownership as a challenge. The informant mentioned questions like who owns the data, where can you store it, and for how long you can store it (Informant ID2, 18.03.2019). Data ownership is referring to who the legal owner of information is, when it is collected from internet platforms like e.g. social media (Sivarajah et al., 2016, p. 275). It is noteworthy that the challenge is only introduced shortly by one of the total of 16 informants, which indicates that it might not be a very relevant challenge with the use of Big Data in emergency management. This can be explained with the same paradox that were discussed with privacy and security - in a normal situation it is important, but in the surroundings of a disaster where lives are at risk you are willing to sacrifice a bit of your rights because you want the crisis response workers to know where you are (Informant ID21, 12.04.2019). It is natural to assume that if you are willing to sacrifice your privacy rights, you are willing to give up the ownership of you data as well. It is important to add that this does only adhere to the response phase of the disaster. This means that data ownership is a Big Data challenge as presented in Sivarajah et al. (2016), but it is not regarded as a Big Data challenge in emergency management, and is therefore removed as a management challenge.

6.1.4 New Challenges Emerged

The chapter of 5.0 Results presented a total of nine new challenges emerged: lack of knowledge, data access, resource scarcity, data neutrality, time pressure, digital divide, sensemaking, cold start and culture differences. Some of these has been mentioned in the discussion already, as an extension to the existing challenges in Sivarajah et al. (2016). These include: cold start and culture differences discussed as variety challenges, and lack of knowledge and resource scarcity as cost/operational expenditures challenges. Therefore, the remaining five challenges will be discussed further in this part.

One of the most noticeable new challenges that emerged was the challenge of data access, where a total of eleven informants identified it. The challenge can be described as problems relating to being unable to collect a necessary volume of information. The challenge of data access can be viewed as an opposite to the challenge of data volume. Besides at times being overloaded by too much data (Informant ID24, 05.04.2019), the paradigm for emergency management has often been that there is too little information (Informant ID2, 18.03.2019). In nature disasters such as typhoons, hurricanes or earthquakes this can be caused by the communications network being knocked out, leaving Big Data tools dependent on data collection through e.g. social media data useless (Informant ID3, 08.03.2019; Informant ID4, 19.03.2019). The challenge of data access has direct consequences for how well the organisation can conduct its decision making. By having a sparse level of data to base itself upon, the decision makers are not able to perform informed decision making (Informant ID14, 26.03.2019). At the same time, lack of information can be used as information in a crisis. In the response of the hurricane in the Philippines in 2013, lack of information were used extensively as information describing what extent of damage each area suffered. The hurricane blew down the cell towers, which knocked out the communication network in certain areas, meaning that the crisis response organisations were able to conclude that the areas where they were not able to collect data probably were the areas that suffered the most extent of damage (Informant ID3, 08.03.2019). During the data collection, digital divide was mentioned as a challenge by a total of three informants. Digital divide refers to the divide between the people that has a digital footprint, and the ones that do not. This creates a challenge for Big Data driven tools that collect information from digital sources (Informant ID2, 18.03.2019). The challenge of digital divide is often experienced when trying to collect data from elders in the society (Informant ID4, 19.03.2019), or in less developed areas of the world (Informant ID9, 14.03.2019). The result of the digital divide is the challenge of data access - the organisations are not able to collect the necessary volume of information from certain parts of the population. Therefore, the challenge of digital divide is merged as an extension to the data access challenge. Organisations in the business sector are used to operate in data-rich environments, which is not the case for emergency management organisations, due to the nature of disasters and the areas of the world they operate in (Informant ID9, 14.03.2019). Therefore it can be said with confidence that the Big Data challenge of data access is caused by the context of emergency management, which explains why the challenge is not presented in the overall Big Data challenges in Sivarajah et al. (2016). Data access is considered as a data challenge as it describes the characteristics of the data. Much like the challenge of data volume being the size of the data pool - data access characterises the lack of access on data.

The nature of disasters and its context of risking loss of lives creates a whole other Big Data challenge: time pressure. Disasters often occur without warning, therefore a Big Data tool has to be set up quickly to avoid losing critical information, creating the challenge of time pressure (Informant ID21, 12.04.2019). In addition to losing critical information, a consequence of time pressure is that decisions are not made quickly enough, resulting in the situation worsening (Informant ID20, 02.04.2019). One informant from an emergency management organisation stated that some of their data collection tools were only used in drills and exercises, because they do not have the time to set it up during a actual crisis (Informant ID29, 03.05.2019). It can be argued that the time pressure challenge is not specific to Big Data technologies, but the emergency management field as a whole. All activities in emergency management are under time pressure in a disaster where lives are at stake, including the data collection and Big Data analysis. Nevertheless, Big Data and artificial intelligence tools are often advanced systems that is dependent upon training data for a specific situation (Informant ID22, 03.05.2019). This is described further in the cold start challenge which is discussed and placed as an extension to the data challenge of variety. It describes the characteristics of Big Data Analytics which makes it hard for it to work under the time pressure of an disaster. The challenge is not mentioned in the Big Data challenges presented in Sivarajah et al. (2016), which indicates that it is not a challenge in universal sectors conducting Big Data analyses. The context of emergency management and the nature of disasters creates the urgent need for Big Data tools to be set up quickly and produce results without delay, making time pressure a Big Data challenge specific to the field of emergency management. Time pressure is a challenge associated with the managerial part of Big Data, it is therefore a management challenge. This is because depending on how you manage your data, time pressure will be a managerial issue affecting the outcome of information for the decision maker.

The challenge of data neutrality is a challenge identified by the informants that is not covered in the existing literature of Sivarajah et al. (2016). It refers to that no information is objective, it is always tainted and there is always someones data collection behind it (Informant ID2, 18.03.2019). We will discuss the challenge of data neutrality from two different perspectives, firstly from the perspective of emergency management organisations operating and collecting information from conflict and hostile areas such as areas of war. Secondly, from the perspective of how information can be tainted during the processing stage to serve a specific purpose in the organisation. Data neutrality in the sense of organisations working in conflict and hostile areas where data is collected for a specific need, gives the data collection neutrality issues. The organisations working with such data needs to be principled, and have to adhere to certain standards of neutrality and impartiality (Informant ID2, 18.03.2019). Emergency management organisations does often keep a neutral role when working in hostile environments, therefore it can harm them if they make their decision based upon information that is not neutral. At the same time, data will never be objective, it is generated by someone or something (Informant ID2, 18.03.2019). Utilizing a standard of only using completely neutral data is therefore impossible. There has to be a balance between what the standard of impartiality and the practical ability to collect such data, which needs to be assessed by each organisation, and adapted to the environment which they are working in. The second perspective of data neutrality entails how the information is processed, and how people are able to shape the output of it to suit a certain decision point, and/or their own agenda (Informant ID20, 02,04.2019; Informant ID9, 14.03.2019). Data neutrality is the third new challenge emerged to be added in the context of emergency management to the existing model of Big Data challenges. It can be argued that it is natural for data neutrality to be excluded from the model without the context of emergency management, since other sectors utilizing Big Data technology is not working in hostile environments. At the same time, they can still face data neutrality challenges relating to the processing of data. Data neutrality is a result of how the processing of the data has been conducted, therefore it is considered a process challenge. It could be examined as a data challenge as well, but the characteristics of the data concerning neutrality falls under the challenge of veracity, and further under the challenge of value. Data neutrality will be affected by how the data is processed, how it is framed towards an agenda or goal, therefore it is natural to associate it with the process challenges.

Sensemaking is presented as a challenge by several informants. It can be defined as situational understanding and situational recognition and is one of the biggest challenges in the field of emergency management (Informant ID4, 19.03.2019). Decisions has to be made with a lot of uncertainty and under time pressure (Informant ID20, 02.04.2019). The decision maker is often not present in the field, but rather at a headquarter located somewhere else in the world, making sensemaking and situational understanding even more challenging (Informant ID2, 18.03.2019). Although sensemaking is identified as one of the biggest challenges in the field of emergency management, we do not consider it as a challenge in direct correlation to Big Data, but rather an emergency management challenge. The examples provided in the results is pointing out that it is the people working in crisis response that is having challenges with sensemaking and situational understanding, rather than the Big Data technologies. At the same time it can be argued that it is hard to gain situational understanding from Big Data analysis, but this is caused by the diversity of the other challenges presented in this research. Sensemaking is therefore removed as a challenge.

6.2 Recommendations for Practitioners

In addition to the challenges with Big Data in emergency management, we are providing a set of recommendations to help practitioners overcome some the mentioned challenges. This functions as the practical contribution of our report.

Recommendation 1: Verification of social media information
Through the data collection it has become obvious that the veracity of the data is one of the biggest challenges faced in emergency management as most of the informants stated it as a challenge. It is also something that hinders Big Data and artificial intelligence based tools with processing enough accurate training data before, during and after crises. This is referred to as the “cold start” challenge, which is an extension to the data variety challenge. To overcome the issues of veracity like trustworthiness, accuracy, fake news and other characteristics alike that affects the data quality, some techniques or solutions are needed. Our data collection shows that these issues mentioned are often experienced with the use of social media data, as there are many actors involved that may provide data and messages. Determining the credibility of those people who pose during a disaster is very challenging (Informant ID14, 26.03.2019).

> Determining if a piece of content is true or false may be extremely time consuming and often requires a great amount of background knowledge or context. In contrast, determining if a piece of content is believable is something we do every day on an intuitive basis, and hence tends to be a relatively fast operation. (Castillo, 2016, p. 6)

Meier (2015, p. 150) discusses how artificial intelligence can be used for verification of twitter data, determining if it is a rumor or not. Together with TweetCred which is using machine learning techniques to automatically score the credibility of crisis related tweets, verification with artificial intelligence should be something to look further into. Verification of social media messages is still a difficult task, but considering the time pressure to make decisions during a disaster, this is a field one should explore to increase data quality and save time.

> Content-based methods for verification are an interesting research direction, but they require multiple reports referring to the same situation, which may or may not be available at a given time. Logical inconsistencies in a message are one potential sign that the information is incorrect, but they are only one of many possible reasons in which information may be incorrect. More research is needed to understand to what extent these methods can contribute to verify claims done on social media during crises. (Castillo, 2016, p. 7)

We recommend emergency management organisations that use Big Data and artificial intelligence for their operations with data sources like Twitter, to seek solutions like TweetCred to overcome data veracity challenges and ease the pressure of time. As Castillo (2016) explains there is still need for more research on the topic, but the solutions that are present today should still be taken advantage of by the various organisations using artificial intelligence.

Recommendation 2: Alternate data sources

Data access has previously been, and still is a problem in developing countries. This can cause little to no data to extract and process for the purpose of emergency management. Some countries do however have more access to data, but during a disaster like a typhoon, the telecommunication might be knocked down causing data streams to be lost. When there are little data, there might still be information to collect. Cell towers that get knocked out during a hurricane or typhoon cannot pass through data anymore, but can give information on how big the damage area is by mapping where the towers without signal is. This practice has been used to produce a damage map in the Philippines once (Informant ID3, 08.03.2019). This can represent data, which sometimes is considered missing in a disaster. In such situations, other sources and types of data can be used to reduce the challenges relating to data access.

Recommendation 3: Data standardization

Emergency management organisations still work a lot in information silos. We experienced this during our data collection. A big challenge is also the sharing of data. Some do not share data a lot, some do not clean it well, and others process data by their own standards. To achieve cooperation in a crisis, sharing of data and information is necessary. Humanitarian Database Exchange (HDX) is a platform which many use, but the information there is not always cleansed well and organised. Looking for data that is not categorised and not marked for which purpose can be like looking for a needle in a haystack (Informant ID2, 18.03.2019). Data needs to be clean to make common operational datasets work (Informant ID9, 14.03.2019). There should be techniques on how to clean your data, simple methods of categorising to make it simple for others to look for it. Humanitarian Exchange Language (HXL) is a data standard, designed to be incredibly simple (Informant ID30, 12.04.2019). It aims to improve coordination across agencies responding in a humanitarian crisis, providing a more efficient and effective system of collecting and sharing data. HXL is led by United Nations Office for the Coordination and Humanitarian Affairs (UNOCHA). They describe it as “a simple standard for messy data” and it is facilitating the exchange and merging of data across agencies to create a more complete and accurate operational picture of a crisis (Warner & Obrecht, 2016, p. 15). It accepts that humanitarians mostly work in Excel. With this you may merge and combine data, and put it into visualisation tools that it works well with for presentation of the data (Informant ID30, 12.04.2019). HXL is a good example of how it is possible to move away from information silos, and individual standards for cleansing of data. If more organisations try to adapt to HXL as much as possible, sharing of data might be easier in a future crisis. For future adoption a 5-year speculative progress plan on growing the HXL ecosystem was presented by Johnson (2017). At this time in 2019 it would be in its second year. The first year would include encouragement of adding HXL tags to datasets for open data to increase the number of datasets in the ecosystem. The second and third year will relate to growing the ecosystem. Tools will be adapted to work with non-HDX (Humanitarian Database Exchange) data. The first tools to provide automated data interoperability through shared spatial qualities, and data collection tools output in HXL formats (Johnson, 2017).

> One of the key goals of a data standard is to help increase interoperability. By this stage with more datasets using HXL tags, information managers will have an easier time understanding datasets and using tools to merge and join different datasets. With the new geo services though this interoperability process can be sped up with the use of shared spatial qualities. (Johnson, 2017)

Year four and five will include pollinating the ecosystem and continued adaptation of tools for internal and private data environment. The plan is presented speculative and not as a formal plan (Johnson, 2017). We recommend emergency management organisations to seek the plan and adopt the use of Humanitarian Exchange Language. It is a tool that can make interoperability and shared datasets work across organisations.

Recommendation 4: Visualisation of information tailored to the field personnel
Interpretation of data visualisation by the end user can be a challenge, because of the way results are presented (Informant ID14, 26.03.2019). The end user can be someone in the field, with no need for a complex graph. Usable visualisation is good usability, you do not need a fancy chart without relevance to your situation (Informant ID21, 12.04.2019). Because of the time pressure in a crisis, emergency management organisations need to develop techniques and communication lines to learn how to create the best visualisations for their personnel in the field. We recommend that emergency management organisations are providing their personnel in the field with fast and easy information to take action and make decisions. Informant ID30 tells that their organisation are working on how to simplify the visualisations for people who do not have as high data literacy, and further how to send printed and dynamic reports over low bandwidth. This can make it possible for people in the field to print the report, and maintain a data driven crisis response (Informant ID30, 12.04.2019).

7.0 Conclusion and Implications

The aim for this study was to research which challenges emergency management organisations were facing when implementing Big Data tools to their operations, through the research question: *What challenges do emergency management organisations face when implementing Big Data technology in their operations during a crisis?* The literature review uncovered a set of Big Data challenges presented in Sivarajah et al. (2016), divided into data, process and management challenges, unrelated to the context of emergency management. Our study reviewed these challenges, and created a revised set of Big Data challenges in the context of emergency management. To do this we conducted a qualitative research where we interviewed informants relevant to the field of emergency management and technology, from two perspectives: emergency management organisations and research facilities. A total of 16 semi-structured interviews were conducted.

Through the study we verified the existing Big Data challenges presented in Sivarajah et al. (2016) except for one management challenge, data ownership. Data ownership is not applicable as a Big Data challenge in emergency management because of the people's willingness to sacrifice some rights during disasters. The willingness to sacrifice some rights can be explained by the nature of disasters where some of the population's lives might be dependent on the crisis response to reach them in a certain amount of time.

In addition we found that emergency management organisations are experiencing three further challenges when implementing Big Data tools to their operations: data access, data neutrality and time pressure. Data access is a data challenge and can be described as problems relating to being unable to collect a necessary volume of information. It can be caused by communication network being knocked out during disasters, as well as problems relating to the digital divide where some of the population is not producing any data. The challenge of data access has direct consequences for how well emergency management organisations can conduct their decision making. The challenge of time pressure is a management challenge and occurs during disasters because of the risk of losing lives. Big Data tools has to be set up quickly to avoid losing critical information, which creates a "speed" imperative. If not tackling the challenge of time pressure, emergency management organisations lose critical information, as well as the ability to make decisions quickly, which can result in the situation worsening. The process challenge of data neutrality refers to that no information is objective, it is always tainted and there is always someones data collection behind it. Emergency management organisations find themselves working in hostile areas, where the information they collect might not be neutral, favouring one side of the conflict. Data neutrality also covers the fact that information can be tainted during the processing stage to serve a specific purpose in the organisation, making it undesirable as a basis for decision making.

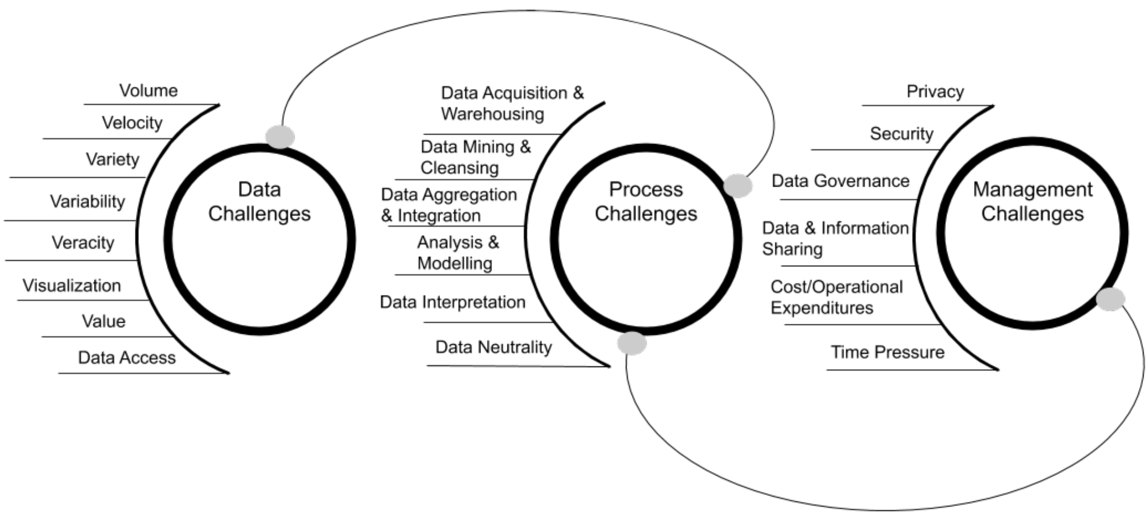

Figure 6: Conceptual classification of Big Data challenges in emergency management by Ø. S. Fongaard & T. F. Nestaas, 2019.

Through the analysis of the results and the discussion we have been able to make a revised version of the conceptual classification of Big Data challenges, originally found in Sivarajah et al. (2016). Figure 6 presents the overview of the challenges emergency management organisations are facing when implementing Big Data technology based tools in their operations during a crisis.

In addition to presenting the challenges emergency management organisations are facing with Big Data technology, we provided recommendations on how to tackle some of the challenges presented, based on the literature review and the interviews with the informants. To tackle the challenge of data veracity, we recommend organisations that are using Twitter data as their data source to look into data quality assurance tools such as TweetCred. The solution still needs more development and more research is needed on the area, but it is a possibility to explore to increase data quality. The challenge of data access can be caused by e.g. cell towers being blown down during a typhoon or hurricane, resulting in the telecommunication network being knocked out. While there is no transactional or user data to collect in a such situation, we recommend that the crisis response maps the inoperable cell towers to create a map to visualise to what extent of damage each area has suffered. To tackle the challenges relating to data and information sharing we recommend emergency management organisations to look into common data standards such as Humanitarian Exchange Language (HXL) to break down information silos and improve coordination across agencies responding in a humanitarian crisis, providing a more efficient and effective system of collecting and sharing data. In addition, we recommend emergency management organisations to develop techniques to provide fast and easy-to-read visualisations of the analysed data that is suited to their personnel in the field, to maintain a data driven crisis response.

7.1 Limitations of the study and Suggestions for Future Research

The study has been conducted as a qualitative study, with a semi-delphi study as a research strategy. We considered this as a well suited research strategy since the research is based on statements from experts on the relevant field, which are able to comment on the challenges of using Big Data technologies in crisis management. The research was limited to one round of interviews with the informants, hence the “semi” term to the delphi method. The study was limited by duration time, which resulted in a limitation in the delphi method.

For future research we would suggest an empirical evaluation of the concluding conceptual classification of Big Data challenges in emergency management. The conceptual classification, figure 6, could then be used as a framework for a case study for a crisis where Big Data technologies have been exploited.

The study has presented an overview of the challenges emergency management organisations are facing when implementing Big Data technology based tools into their operations during a crisis. There are four main stages in a disaster: preparation, response, recovery and prevention/mitigation. The study is limited to the response phase of the disaster. Further research might include the same approach to uncover challenges with Big Data technology in other stages of the crisis.

During the course of the study the terms emergency, crisis and disaster has been used interchangeably. As presented in Emergency under 2.1 Concepts these terms are referring to somewhat different situations. We would find it relevant as a future research to study Big Data challenges specific to each of the terms as a context, to uncover if there are any differences in the challenges depending on the specific context.

8.0 References

Abdullah, M. F., Ibrahim, M. & Zulkifli, H. (2017). *Big Data Analytics Framework for Natural Disaster Management in Malaysia*. Conference on Internet of Things, Big Data and Security, Malaysia.

ACAPS (n.d.). About ACAPS: What we do. Retrieved 01.04.2019 from https://www.acaps.org/about-acaps/what-we-do.html

Al-Dahash, H, Thayaparan, M and Kulatunga, U. (2016, September) *Understanding the terminologies : disaster, crisis and emergency*. Proceedings 32nd Annual ARCOM Conference, Manchester, UK.

Arslan, M., Roxin, A., Cruz, C. & Ginhac, D. (2018, January). *A Review on Applications of Big Data for Disaster Management.* The 13th International Conference on Signal Image Technology, India.

Athanasia, N. & Stavros, P. T. (2015, May).*Twitter as an instrument for crisis response: The Typhoon Haiyan case study.* Proceedings of the ISCRAM 2015 Conference, Kristiansand.

Barnaghi, P., Sheth, A. & Henson, C. (2013, December). From Data to Actionable Knowledge: Big Data Challenges in the Web of Things. *IEEE Computer Society.* Retrieved from: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6733221

Big Data and Emergency Management (2019). The BDEM Project. Retrieved from: https://www.bigdata.vestforsk.no/#intro

Boulos, M. N. K., Resch, B., Crowley, D. N., Breslin, J. G., Sohn, G., Burtner, R, ... Chuang, K. S. (2011). Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: trends, OGC standards and application examples. *International Journal of Health Geographics, 10*(67). doi.org/10.1186/1476-072X-10-67

Braun, V. & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

British Red Cross. (2019). ABOUT THE BRITISH RED CROSS. Retrieved from https://www.redcross.org.uk/about-us

Castillo, C. (2016) Veracity. C. Castillo (Red.), *Big Crisis Data: Social Media in Disasters and Time Critical Situations* (Chapter 8). Cambridge: Cambridge University Press.

Chen, H., Chiang. R. H. & Storey. V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly, 36*(2): 1165-1188.

Chen, J., Chen, Y., Du, X., Li, C., Lu, J., Zhao, S. & Zhou, X. (2013). Big data challenge: a data management perspective. *Frontiers of Computer Science, 7*(2): 156-164. DOI:10.1007/s11704-013-3903-7

CIEM (n.d.). About CIEM. Retrieved 01.04.2019 from http://ciem.uia.no/about-ciem

Daatland, S. O. G. (2019, 31. January). Kristiansand Kommune, Samfunnssikkerhet og beredskap. Retrieved from https://www.kristiansand.kommune.no/generelt/samfunnssikkerhet-og-beredskap/

Dinnen, J. (2014, March 27). Phase #2: Clearly Define Your Research Strategy. *MacKenzie Corporation.* Retrieved from: https://www.mackenziecorp.com/phase-2-clearly-define-research-strategy/

Direct Relief (n.d.). About: History. Retrieved at 01.04.2019 from https://www.directrelief.org/about/history/

Disasters Emergency Committee. (2018, 20. August). BRITISH RED CROSS. Retrieved from https://www.dec.org.uk/charity/british-red-cross

Dudovskiy, J. (2018). Snowball sampling. *Research Methodology.* Retrieved from: https://research-methodology.net/sampling-in-primary-data-collection/snowball-sampling/

Eriksen, J. (2011). Krise- og beredskapsledelse: Teamtrening. Oslo: Cappelen Damm akademisk.

Fertier, A., Barthe-Delanoë, A., Montarnal, A., Truptil, S. & Benaben, F. (2016). Adoption of big data in crisis management toward a better support in decision-making. *Proceedings of the ISCRAM 2016 Conference – Rio de Janeiro, Brazil, May 2016*

Frankfort-Nachmias, C. & Nachmias, D. (1992). *Research methods in the social sciences.* London: St. Martin's Press.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137–144.

Gripsrud, G. Olsson, U. H. & Silkoset, R. (2015). *Metode og dataanalyse*. Trondheim: Høyskoleforlaget.

Gudivada, V. N., Jothilakshmi, S. & Rao, D. (2015, April). *Data Management Issues in Big Data Applications.* The First

International Conference on Big Data, Small Data, Linked Data and Open Data (ALLDATA 2015). 16-21.

GuideStar (n.d.). Direct Relief. Retrieved 01.04.2019 from: https://www.guidestar.org/profile/95-1831116

Jikimlucas (2016, 23. October). AIDR Overview. *Github.* Retrieved from: https://github.com/qcri-social/AIDR/wiki/AIDR-Overview

HDX (n.d.). *Standby Task Force.* Retrieved 01.04.2019 from https://data.humdata.org/organization/activity/standby-task-force

Hilbert, M. (2013). Big Data for Development: From Information- to Knowledge Societies. Retrieved from SSRN: https://ssrn.com/abstract=2205145 or http://dx.doi.org/10.2139/ssrn.2205145

Hilbert, M. (2015). Big Data for Development: A Review of Promises and Challenges. *Development Policy Review. 34*(1). 135-174. https://doi.org/10.1111/dpr.12142

Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R. & Shahabi, C. (2014). Big Data and Its Technical Challenges. *Communications of the ACM, 57*(7): 86-94. https://doi.org/10.1145/2611567

Jacobsen, D., I. (2005). *Hvordan gjennomføre undersøkelser?: Innføring i samfunnsvitenskapelig metode*. (2. utg.) Kristiansand: Høyskoleforlaget.

Jain, S. & McLean, C. (2004, January). *An Integrating Framework for Modeling and Simulation for Emergency Response*. National Institute of Standards and Technology, Maryland, USA.

Jin, X., Wah, B. W., Cheng, X. & Wang, Y. (2015). Significance and Challenges of Big Data Research. *Big Data Research.* https://doi.org/10.1016/j.bdr.2015.01.006

Johnson, S. B. (2017, 2. August) Growing the Humanitarian Exchange Language (HXL) ecosystem. Retrieved from: https://medium.com/@Simon_B_Johnson/growing-the-humanitarian-exchange-language-hxl-ecosystem-32724189ef8

Laney, D. (2001). *3D Data Management: Controlling Data Volume, Velocity and Variety.* Gartner. Retrieved from: http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-ad-Variety.pdf

Liao, Z., Yin, Q., Huang, Y., & Sheng, L. (2014). Management and application of mobile big data. *International Journal of Embedded Systems, 7*(1), 63–70.

Ludvigsen, F & Nylenna, M. (2019, 4. April). AMK-sentral. I *Store medisinske leksikon*. Retrieved 2. May 2019 from https://sml.snl.no/AMK-sentral

Meier, P. (2015). *Digital Humanitarians: How Big Data Is Changing the Face of Humanitarian Response* (1. utg). Abingdon: Routledge.

MSF (n.d.) Who we are. Retrieved 25.04.2019 from https://www.msf.org/who-we-are

Munkvold, B. E. (1998). *Implementation of information technology for supporting collaboration in distributed organizations* (Doktorgradsavhandling). Norwegian University of Science and Technology, Trondheim, Norway.

Myers, M., D. (1997) Qualitative Research in Information Systems. *MIS Quarterly, 12*(2). 241-242. Retrieved from: https://www.qual.auckland.ac.nz/

NAKOS (2019). AMK Sørlandet Sykehus HF. Retrieved from: https://www.nakos.no/course/index.php?categoryid=440

Nature Editorial (2007). A matter of trust. *Nature, 449*(7163), 637–638. DOI: 10.1038/449637b

Norris, A.C.Martinez,S. Labaka,L Madanian,S. Gonzalez, J.J. & Parry, D. (2015, May). *Disaster E-Health: A New Paradigm for Collaborative Healthcare in Disasters*. In proceedings of ISCRAM 2015, Kristiansand, Norway.

Norwegian Center for Research Data (2019). Information and Consent. Retrieved from: https://nsd.no/personvernombud/en/help/information_consent/

NRC (2018, 04. June). ACAPS. Retrieved from: https://www.nrc.no/expert-deployment/what-we-do/acaps/

Oates, B. J. (2006). *Researching Information Systems and Computing.* London: SAGE Publications

Okoli, C. & Schabram, K. (2010). A Guide to Conducting a Systematic Literature Review of Information Systems Research. *Sprouts: Working Papers on Information Systems, 10*(26). https://doi.org/10.2139/ssrn.1954824

Otto, B. (2011). Organizing data governance: findings from the telecommunications industry and consequences for large service providers. *Communications of the Association for Information Systems, 29*(1), 45–66.

Pauchant, T. C. & Mitroff, I. I. (1992) *Transforming the Crisis-Prone Organization*. San-Fransisco: CA: Jossey-Bass Publishers.

Power D. J. (2014). Using 'Big Data' for analytics and decision support. *Journal of Decision Systems, 23*(2), 222-228.

QCRI (n.d.). About QCRI. Retrieved 01.04.2019 from https://www.qcri.org/about-qcri

Sander, K. (2017, August 13.). DELPHI – metoden. *Estudie.* Retrieved from: https://estudie.no/delphi-metoden/

SBTF (n.d.). about us. Retrieved 01.04.2019 from https://www.standbytaskforce.org/about-us/

Shah, T., Rabhi, F., & Ray, P. (2015). Investigating an ontology-based approach for Big Data analysis of inter-dependent medical and oral health conditions. *Cluster Computing, 18*(1), 351–367.

Sivarajah, U., Kamal, M. M., Irani, Z. & Weerakkody, V. (2016). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research, 70*(1), 263-286.

Solis, B. (2018). STATE OF DIGITAL TRANSFORMATION (2018-2019 ed.). Retrieved from: http://content.prophet.com/the-state-of-digital-transformation-2018-2019

START Network (n.d.). START Fund: Alerts Dashboard. Retrieved 01.04.2019 from: https://startnetwork.org/start-fund/alerts

Svennevig, J. (2001). Abduction as a methodological approach to the study of spoken interaction. *Norskrift*, *103*. Retrieved from: https://www.researchgate.net/publication/251398301_Abduction_as_a_methodological_approach_to_the_study_of_poken_interaction

Tekiner, F. & Keane, J. A. (2013, October). Big data framework. Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on (pp. 1494-1499). IEEE.

UNISDR. (2017, 2. February). TERMINOLOGY ON DISASTER RISK REDUCTION. Retrieved from: https://www.unisdr.org/we/inform/terminology

University of Agder (n.d.). CIEM - CENTRE FOR INTEGRATED EMERGENCY MANAGEMENT. Retrieved 01.04.2019 from: https://www.uia.no/en/research/teknologi-og-realfag/ciem-centre-for-integrated-emergency-management

University of Nebraska Omaha (n.d.). College of Information Science & Technology. Retrieved 02.05.2019 from: https://catalog.unomaha.edu/undergraduate/college-information-science-technology/

Velev, D. & Zlateva, P. (2011). An Innovative Approach for Designing an Emergency Risk Management System for Natural Disasters. *International Journal of Innovation Management and Technology*, *2*(5), 407-413. Available at: http://ijimt.org/papers/167-T00036.pdf

Warner, A. & Obrecht, A. / ALNAP. (2016). *Standardising humanitarian data for a better response: The Humanitarian eXchange Language* (HIF/ALNAP Case Study). Retrieved from: https://reliefweb.int/sites/reliefweb.int/files/resources/alnap-innovation-hxl-case-study.pdf

Webster, J. & Watson, R., T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly, 26*(2).

Western Norway Research Institute (n.d.). About the Western Norway Research Institute. Retrieved 01.04.2019 from https://vestforsk.no/en/about-western-norway-research-institute

Whipkey, K. & Verity A. (2015). Guidance for Incorporating Big Data into Humanitarian Operations. *Digital Humanitarian Network.* Retrieved from: http://digitalhumanitarians.com/sites/default/files/resource-field_media/IncorporatingBigDataintoHumanitarianOps-205.pdf

Wood, L. J., Boruff, B. J. & Smith H. M. (2013) When Disaster Strikes... How Communities Cope And Adapt; A Social Capital Perspective. *Social Capital: Theory, Measurement and Outcomes.* ISBN: 978-1-62417-822-1

Zicari, R. V. (2014). Big Data: Challenges and Opportunities. (2014) In R. (Ed.), *Big data computing* (pp. 103–128). Florida, USA: CRC Press, Taylor & Francis Group.

9.0 Appendix

Overview of the appendixes:

- 9.1 Appendix A: Interview Guide
- 9.2 Appendix B: Overview of identified challenges
- 9.3 Appendix C: Declaration of Consent for Informants

9.1 Appendix A: Interview Guide

Start question

- Do you permit us to record this interview?

1. What kind of relation do you have with the field of emergency management?
   a. Have you worked directly with emergency management in crisis situations?
      i. What crisis did you work in?
      ii. What role did your organisation play before/during/after this crisis?
      iii. What was your role during this crisis?
2. What kind of research experience do you have on Big Data in the context of emergency management?

3. Do you have experience or knowledge about how emergency management organisations do critical decision making before/under/after a crisis?
   a. What do they base their decisions making on?
      i. How do they identify and gather the information they base their decision making on?
         1. What kind of information is this?
         2. Where do they identify the information? *Ex: SMS/Twitter etc.*
         3. How do they see value in the information?
      ii. Are there any challenges associated to identifying and gathering this information?
         1. *Ex: Large volumes, variety, security, privacy, lack of competence to understand the information?*
         2. *Ex: Unreliable sources?*
         3. *Ex: Veracity or correct information? Incomplete information?*

4. Does emergency management organisations process the information in any way before they use it to support their decision making?
   a. In what way did they process the information?
      i. Any challenges related to this?
         1. *Ex: Cleansing of large data sets?*
         2. *Ex: Security, privacy, lack of competence to understand the information, considerations of the ownership of the information?*
         3. *Ex: How did they create value of the information?*
         4. *Ex: Analysis and modelling?*
   b. How do they store the information?
      i. Any challenges related to this?
         1. *Ex: Security challenges regarding storing?*
         2. *Ex: Challenges regarding large volumes of data?*
   c. Is it normal among emergency management organisations to share information used for the decision making processes, internally or externally?
      i. In that case, how do they share this information?
         1. Challenges related to this?

   a. *Ex: Security, privacy, sharing of information and ownership of data?*

5. Of all the challenges mentioned in this interview, which ones do you identify as the biggest and most important challenges?
   a. Why?
   b. How would you rate these?
   c. Do you see specific challenges harder to cope with than others?

6. Any other comments?

9.2 Appendix B: Overview of identified challenges

Overview of Identifed Challenges

Category	Interview ID	2	3	4	5	9	12	14	19	20	21	22	23	24	26	29	30
New Challenges	Cultural Differences	×	×	×			×	×		×	×			×			×
	Cold Start		×		×	×		×			×	×					
	Sensemaking	×		×		×		×		×				×		×	
	Digital Divide	×		×		×											
	Data Neutrality	×				×				×							
	Time pressure	×				×			×	×	×	×				×	
	Resource Scarcity		×				×					×		×	×		
	Data Access	×	×	×		×		×		×		×	×	×	×		×
	Lack of Knowledge		×				×			×		×				×	×
Management Challenges	Data Ownership	×															
	Cost/Operational Expenditures						×			×							
	Data & Information Sharing	×	×						×	×	×	×	×	×	×	×	
	Data Governance	×	×				×			×							
	Security		×	×		×						×		×	×		×
	Privacy	×	×	×	×	×	×			×		×				×	
Process Challenges	Data Interpretation	×				×	×			×	×			×	×		
	Analysis & Modelling			×	×		×	×			×	×					×
	Data Aggregation & Integration	×		×				×		×	×	×			×	×	×
	Data Mining & Cleansing		×	×		×	×				×	×					×
	Data Acquisition & Warehousing							×		×	×						
Data Challenges	Value	×		×	×	×	×	×		×	×	×					
	Visualisation							×			×	×					×
	Veracity	×	×	×		×		×	×	×	×	×	×		×	×	×
	Variability		×			×			×								
	Variety	×	×	×	×	×	×	×	×		×	×			×		
	Velocity			×				×									
	Volume	×		×			×	×						×		×	×

9.3 Appendix C: Declaration of Consent for Informants

Request for participation in the research project

“Challenges with the Use of Big Data technology in Emergency Management”

This is a proposition for you to participate in a research project with the purpose of exploring what challenges organizations within the emergency management domain face when they are using Big Data technology. Through this writing we will provide information about the goals for the project and what a participation will entail for you.

Purpose
We are conducting a master thesis where the purpose is to explore what challenges organizations within the emergency management domain faces when they are using Big Data technology. This is a qualitative study, where we will gather information through interviews with informants who have experience in the field described, from the organizations point of view. The gathering of information will highlight the issue: challenges with the use of Big Data technology in emergency management.

Who is responsible for the research project?
The University of Agder, Kristiansand, Norway.

Why are we asking you to participate?
You have relevant experience either from projects related to emergency management or from actual crises, where there have been collected information with the purpose for use in emergency management. The informants has been identified through the contact network of the Department of Information Systems at the University of Agder.

What does it entail for you to participate?
A participation will entail a personal interview. It will take between about 30-60min. You will be interviewed about your experiences with collecting and using information in emergency management situations.

Participation is voluntary
It is voluntary to participate in the project. If you choose to participate, you can withdraw your consent at any time without reasoning yourselves. All information about you will then be anonymized. There will be no negative consequences for you if you choose not to participate, or at a later stage decides to withdraw your consent.

Your privacy - how we store and use your information
We will only use your information for the purposes described in this writing. We treat the information confidential and in coherence with the Norwegian Privacy Law (Det Norske Personvernregelverket).

- Only the students performing the master thesis and the guiding institution will receive access and process your information

- The interview will be recorded on a recording device without any internet connection. Identifiable data will only be used in secure environments, will be coded and kept separated.
- Participants will not be mentioned og recognized in publishments. Data will be disidentified and anonymized.

What happens to my information when the project ends?
The project will be scheduled to end at 01. June 2019. Information and interview recordings will be anonymized at the end of the project, so there will be no opportunity to identify who has said what.

Your rights
As long as you can be identified in the data material, you have the right to:
- get insight into what personal information that has been registered about you,
- change personal information about you,
- delete personal information about you,
- get extradited a copy of your personal information (data portability), and to send complaints to the Privacy Protection Office (Personvernombudet) and/or Data Protection Office (Datatilsynet) in Norway about the processing of your personal information.

What gives us the right to process your personal information?
We process personal information about you based on your consent.

On behalf of the University of Agder has NSD - The Norwegian Center for Research Data AS considered that the processing of personal information in this project is in accordance with the Privacy Law in Norway.

Where can i get more information?
If you have any questions about the study, or would like to use your rights, please contact:
- The University of Agder, by Devendra Bahadur Thapa on e-Mail (devinder.thapa@uia.no), Master thesis students Thomas Farstad Nestaas on e-Mail (thomas_fn@live.no) and Øistein Syversen Fongaard on e-Mail (oistein.fongaard@gmail.com).
- NSD - The Norwegian Center for Research Data AS, on e-Mail (personvernombudet@nsd.no) or telephone: +47 55 58 21 17

Best regards

Project manager (Researcher/advisor)	Students

Declaration of Consent

“Challenges with the Use of Big Data technology in Emergency Management”

I have received and understood the information about the project *Challenges with the Use of Big Data technology in Emergency Management,* and had the opportunity to ask any further questions. I declare my consent to:

- ☐ participate in a interview
- ☐ that my interview will be recorded on a recording device
- ☐ that i can be contacted throughout the project to provide further information

I agree that my information will be processed until the project is finished, 01. June 2019

Signature Date