Norwegian School of Economics
Bergen, Spring 2017

Technology Adoption in Norway: Organizational Assimilation of Big Data

Truc Nguyen

Truls Engebretsen Petersen

Supervisor: Trond Vegard Johannessen

Master Thesis in Business Analysis and Performance Management & Strategy and Management

NORWEGIAN SCHOOL OF ECONOMICS

Executive Summary

As data permeates and drives the digital evolution, the role of Big Data becomes increasingly essential. Big Data is making its presence known in almost every industry, and has the potential to not only transform the business world, but society at large. Given that companies in Norway are still in the early stages of making use of Big Data, studying factors affecting adoption of Big Data technology in Norway is critical and timely.

Grounded in the Diffusion of Innovation (DOI) theory, Technology Acceptance Model (TAM), and Technology-Organization-Environment (TOE) framework, an integrative model is developed for studying factors affecting adoption of Big Data technology in three aggregated stages of assimilation; initiation, adoption-decision, and implementation. The model specifies three technological characteristics (relative advantage, complexity, and security), three intraorganizational factors (organizational size, top management support, and IT expertise), and three interorganizational factors (competitive pressure, external support, and privacy) as determinants of assimilation.

The proposed model is tested using survey data collected from 336 executives in medium to large companies in Norway. Employing a multinomial logistic regression, this study finds that six predictor variables (relative advantage, complexity, security, top management support, IT expertise, and competitive pressure) are significant and can distinguish non-adopters and adopters in the assimilation stages. Of the six factors identified in the model, three (security, top management support, and competitive pressure) are found to play a vital role in all stages of Big Data assimilation, while two factors (complexity and IT expertise) are critical to the implementation and routinization of Big Data technology.

The results indicate that the model is suited for studying organizational adoption of Big Data technology. Moreover, given the scarcity of research into determinants of adoption in the Big Data literature, the research model offers a suitable point of departure for future studies on Big Data adoption. Finally, the findings have important implications for practitioners and researchers.

Acknowledgements

This thesis was written by two students as part of the Master of Science in Economics and Business Administration at the Norwegian School of Economics. We cooperated with DXC Technology, a technology company providing global IT services, helping clients harness the power of innovation to thrive on change. The thesis was written over the course of one semester, and accounts for 30 credits within the Business Analysis and Performance Management & Strategy and Management master programs.

We are grateful for the opportunity to work closely with two professionals at DXC Technology, Charlotte Irgens and Cato Saele, on a subject that is relevant for their industry and of interest to them. We would especially thank them for providing us with relevant reports on the topic of study. We enjoyed the process, and hope that our findings will be of practical importance to their company, as well as to other companies.

We would also take this opportunity to thank our supervisor, Trond Vegard Johannessen, for his valuable feedback. Further, we would like to thank Aiko Yamashita, Associate Professor at Oslo and Akershus University College of Applied Sciences, for her suggestions and insightful comments on Big Data in the early phases of this thesis. Lastly, we would thank representatives at Sparebank 1 Forsikring, Dagfinn Røed and Hanne Rygg Sorteberg, for meeting us on short notice to discuss the topic of Big Data.

Contents

List of Tables

List of Figures

1. Introduction

1.1 Topic and Research Questions

To say that there is strength in numbers has never been more appropriate, as the data revolution has brought about a new era: One in which a menagerie of digital devices is generating a flood of data sweeping through academia, business, government, and all parts of society, with the data itself being seen as a new type of asset. And as businesses come to discover the value of data and seek to harness its potential, we observe a growing interest in the notion of Big Data, which promises increased innovation, productivity, and future economic growth, from which not only businesses but society at large could benefit (Bollier, 2010). Big Data is expected to facilitate and catalyse change in almost every industry, and has the potential to make unprecedented changes to the way we live, work, and think (Mayer-Schönberger & Cukier, 2013). The applications and power of Big Data are still emerging, and while this paper studies the adoption of Big Data technology, the concept entails so much more than technological change: Big Data represents a transformation of how future enterprises will be managed.

Like most emerging trends, there is a lot of confusion surrounding Big Data. The term has become ubiquitous both in academic and business literature, with vague and inconsistent definitions hampering development of the discipline (Stuart & Barker, 2013). To achieve clarification on the essential characteristics of Big Data, Mauro et al. (2016) proposed the following definition: “Big Data is the Information asset characterized by such a high Volume, Velocity and Variety to require specific Technology and Analytical Methods for its transformation into Value” (Mauro, et al., 2016). This definition highlights the multifaceted nature of Big Data, and identifies the four essential components of the concept: information, technology, methods, and impacts. To narrow down the scope of the present study, this thesis focuses on organizational adoption of the second component of Big Data, namely *technology*. In a field where the distinction between adopters and non-adopters is obscure, studying firms’ acquisition and use of such technology provides a logical starting point for exploring adoption of Big Data.

With research by Gartner (2016) suggesting that more than three-quarters of companies are investing or planning to invest in Big Data, understanding the factors that influence organizational adoption of Big Data technology is crucial and timely. Yet, reviews of more than 200 journal articles and conference proceedings on Big Data show that little research has been done on the factors affecting adoption (Salleh & Janczewski, 2016; Rahman, 2016; Chen, et al., 2016). Of the research that exists on the subject (e.g. Agrawal, 2015; Nam, et al., 2015; Sun, et al., 2016), few have specifically studied the technological component. Moreover, despite strong growth in the European technology market for managing, analysing, and accessing Big Data, Norway is predicted to have among the lowest growth rates in Western Europe (Carnelley & Schwenk, 2016a), making research into which factors that are likely to affect adoption of Big Data technology by Norwegian companies important. The *research objective* of this thesis is therefore to develop a model for studying adoption of Big Data technology in Norway, specifically to be tested on medium to large businesses.

Innovation adoption research, which primarily concerns the adoption of information systems (IS) and information technology (IT), has produced a variety of competing and complementary models for studying adoption. Rogers' (1983) Diffusion of Innovations (DOI) and Davis' (1986) Technology Acceptance Model (TAM) represent two of the most influential theoretical contributions to the innovation adoption literature, and have been extensively used by researchers to study adoption of a rich variety of technological innovations (Hameed, et al., 2012a). Both DOI and TAM share the same premise that potential adopters evaluate an innovation based on their perception of its characteristics, and postulate that innovations with more favourable characteristics are more likely to be adopted. However, although the perspective offered by DOI and TAM may contribute to our understanding of the adoption of Big Data technology, it is based on models originally developed for studying the adoption of technological innovations by individuals making autonomous choices, whereas the acquisition and deployment of Big Data technology is an organizational-level decision influenced by intra – and interorganizational factors. Specifically, the application of DOI and TAM at an organizational level has received substantial criticism due to excluding the possibility of influences from organizational and environmental factors (Chau & Tam, 1997; Hameed, et al., 2012a). Accordingly, attempts to integrate key insights from DOI and TAM with the Technology-Organization-Environment (TOE) framework (Tornatzky & Fleischer, 1990), a contextual framework covering salient organizational and environmental factors, are made in this thesis.

Although the TOE framework, an organizational level technology adoption framework, remains prominent and widely utilized in research on organizational adoption, there is seemingly no universal set of factors believed to affect adoption. The absence of a single, unified theory that permits researchers to predict the extent to which an organization will adopt a given technological innovation suggests researchers must develop their own tailored models (Fichman, 1999). For this reason, a literature review has been conducted to identify factors potentially relevant to the context of Big Data adoption. Additionally, besides the more broadly generalizable factors extracted from extant literature, extensions were also found relevant to cover important aspects distinctive to Big Data. Salleh and Janczewski (2016) found that despite being prevalent “themes” in the literature, security and privacy issues of Big Data have yet to be addressed empirically in adoption research. This thesis will therefore further the work of Salleh and Janczewski by attempting to achieve clarification on how privacy and security issues may be affecting adoption of Big Data technology.

Based on our review of adoption and Big Data literature, an integrative research model for the study of organizational adoption of Big Data technology is developed, where hypothesized relationships between factors grouped within the technological, organizational, and environmental context are based on the DOI, TAM, and TOE framework. Accordingly, the following *research question* is presented for this study:

RQ_1: Which technological, organizational, and environmental factors affect adoption of Big Data technology?

Additionally, as the adoption of technology by organizations is considered a stage-based process rather than a binary event, this thesis studies adoption in terms of assimilation; the multi-stage, sequential process by which an organization becomes aware, acquires, deploys, and routinizes new technology (Meyer & Goes, 1988). By studying the process of organizational adoption, known as assimilation, it is possible to reveal how determinants have differential effects at different stages of adoption. This leads to the second *research question* of the study:

RQ_2: To what extent do the technological, organizational, and environmental factors have differential effects at the different stages of adoption?

1.2 Thesis Outline

Chapter 2 introduces Big Data and offers a working definition for the present study. This is followed by an introduction to three of the pressing challenges of Big Data, before discussing the current state of Big Data adoption in Norway.

Chapter 3 introduces the theoretical foundation for this study by presenting a background on the innovation adoption literature, followed by a discussion of diffusion research and the stages of IT innovation adoption in organizations. Next, to develop this paper's research model, the Diffusion of Innovation (DOI), Technology Acceptance Model (TAM), and Technology-Organization-Environment (TOE) framework are presented. The chapter proceeds to discuss the application of these models for studying organizational adoption and presents a tentative research model, which is developed further by discussing and stating hypotheses for relevant constructs from DOI, TAM, and TOE research. Lastly, a discussion of the conceptualization of Big Data technology adoption is presented, before proposing the final research model.

Chapter 4 details the methodology; the process by which the hypotheses derived from the research model were empirically tested and research questions were answered. The choice of approach to pursue this thesis' research objectives are discussed as follows: A presentation of the research design is given, followed by an overview of the sampling and data collection, and finally, a discussion of non-response bias.

In Chapter 5, the data analysis, referring to the inspecting, cleansing, transforming, and modelling of data, is presented. The goal of the data analysis is to obtain sufficient statistical information to answer the research questions of the study. This chapter presents the preliminary analysis and descriptive statistics, followed by two multivariate analysis techniques; factor analysis and multinomial logistic regression.

Chapter 6 presents the results and discusses each of the factors identified in this study in relation to the technological, organizational, and environmental context in which they were presented in the proposed research model. Theoretical and managerial implications are presented, and finally, an evaluation of the study's limitations and potential directions for future research are offered.

2. Big Data

Over the past decade, we have witnessed the unfolding of the Internet of Things, advancements in machine learning, and technological breakthroughs in areas including robotics, artificial intelligence, virtual reality, autonomous vehicles, facial recognition, medical diagnostics, and fraud detection (Pareek, 2015). *Big Data* has emerged as the new frontier of these IT-enabled innovations and opportunities presented by the megatrend referred to as the digital information revolution. As the activities of institutions and businesses are digitized, new sources of data and technology are propelling our society into a new era: one in which an unprecedented richness of data exists on virtually any topic of interest. The potential advantages of utilizing this data have been broadly recognized (Brynjolfsson, et al., 2011), and the exponential creation of data by new data generating sources has gained attention by business, government, and academia through efforts to harness and analyse Big Data (Goes, 2014). Whereas the public, academic, and scientific sectors see Big Data as an opportunity to improve our understanding of society and the world, businesses are eyeing the opportunity to gain technology-based competitive advantages.

Like most emerging trends, there is a lot of confusion surrounding Big Data, and a common terminology is still evolving. According to Mauro et al. (2016), the degree of popularity of the Big Data phenomenon has not been accompanied by a rational development of an acceptable vocabulary. The term has become ubiquitous both in academic and business literature, with vague and inconsistent definitions hampering development of the discipline (Stuart & Barker, 2013). Thus, the purpose of the following chapter is first to introduce the concept of Big Data and present a working definition based on Mauro et al. (2016)'s review of more than 1,400 conference papers and journal articles on the topic of Big Data. This should clarify the role of the present study in relation to existing Big Data literature. Furthermore, as Big Data is an emerging field, a brief introduction to some of the current challenges of Big Data are presented, followed by a discussion of Big Data adoption in Norway.

2.1 Defining Big Data

Information: The "3 Vs"

The first attempt at defining the Big Data phenomenon was by Doug Laney from the META Group (now Gartner) in 2001 (Ylijoki & Porras, 2016). Without mentioning the term explicitly, Laney (2001) introduced the "3 Vs", underpinning the increase in data volume, velocity, and variety. *Volume* refers to the quantity of data that is generated at an exponential rate, with data sets ranging from terabytes to zettabytes in size. *Velocity* relates to the increased speed at which data is available and requires near real-time processing to maximize the value of data. *Variety* refers the multiplicity of data types generated from a range of sources, including social networks, mobile phones, traffic cameras, and various sensor (Hashem, et al., 2015). As such, Big Data generally refers to data sets characterized by the "3 Vs".

However, data is simply raw symbols with no significance beyond its existence, while *information* is data that has been processed and attributed substantive meaning. Hence, later studies have pointed out that these data characteristics are insufficient to explain the multifaceted nature of Big Data (Jain, et al., 2016). Several authors have therefore extended the "3 Vs" by adding features such as *veracity* (Ularu, et al., 2012; Miele & Shockley, 2013), *value* (Gantz & Reinsel, 2011; Fan & Bifet, 2012; Dijcks, 2013), *variability* (Fan & Bifet, 2012), and *visualization* (Chen, et al., 2012), making up a total of "7 Vs". Consequently, Big Data has become a volatile term which has led to different interpretations (Ylijoki & Porras, 2016).

Technology: A Prerequisite for Using Big Data

Specific technological needs come hand in hand with the utilization of Big Data, as dealing with data sets characterized by high volume, velocity, and variety, require computational power and storage that the average information technology system is unable to provide (Mauro, et al., 2016). Technology refers to hardware (e.g., storage and servers) and software (e.g., applications) that enable the accessing, managing, and analysing of Big Data. Several technologies have emerged to deal with Big Data, including Hadoop, MapReduce, CouchDB, Cassandra, Pig, Hive, MongoDB, and AsterData (PwC, 2015). Although these technologies are not exclusively used for Big Data, their application on datasets that fit the characteristics of Big Data classifies them as Big Data technologies.

According to Microsoft (2013), Big Data involves the application of "serious computing power to seriously massive and often highly complex sets of [data]". In dealing with large data sets beyond the ability of traditional systems, popular technologies include Hadoop, as it enables the distributed processing of data across multiple, remotely located commodity machines (or *nodes*) (Shvachko, et al., 2010). Rather than relying on expensive high-end hardware, Hadoop brings scalable parallel computing to commodity hardware, which makes the utilization of Big Data affordable (Ularu, et al., 2012). Furthermore, the technological requirements go beyond dealing with the volume of the Big Data; to include issues arising from larger and faster transmissions of data, as well as the constraints on data storage caused by the capacity of storage devices. Thus, while Big Data is not confined to the realm of technology, the issues of storing, processing, and analysing Big Data are critical technological challenges that suggest Big Data technologies are a necessary prerequisite for using Big Data.

Methods: Business Intelligence and Analytics

The *value* component has become a core concept of Big Data, as data provides no value by itself. The requirements needed to make proper use of Big Data are often referred to as Business intelligence and Business Analytics. The umbrella term *Business Intelligence (BI)* became popular in the 1990s and refers to "a broad category of applications, technologies, and processes for gathering, storing, accessing, and analyzing data to help business users make better decisions" (Watson, 2009, p. 491). In general, the term is applied in connection with the use of data that are stored in traditional databases and/or warehouses (Johannessen, 2017). However, the era of Big Data has become an of enabler of analytics. As new kinds of data emerged in the mid-2000s, traditional BI tools were no longer sufficient to harness the potential of data with high volume, velocity, and variety (Davenport, 2013). Consequently, *Business Analytics (BA)* was introduced to represent the key analytical component of BI, divided into three "phases": descriptive, predictive, and prescriptive. The first phase, *descriptive analytics*, is commonly referred to as the traditional BI tools that help organizations understand what happened in the past. This type of analytics uses historical data and identifies patterns to improve decision-making[1]. The second phase, *predictive analytics*, seeks to determine the best solution or outcome among various choices and uses statistical models to evaluate what could happen. The third phase, *prescriptive analytics*, not only focus on what will happen and when it happens, but also why it will happen. Prescriptive analytics recommends decision alternatives for taking advantage of opportunities or mitigate risks by using optimization, simulation, graph analysis, heuristics, and machine learning to name a few (Raj, 2014). According to Rijmenam (2013), these three types of analytics should co-exist; none exceeds the other, but are complementary in obtaining a complete overview of an organization.

[1] Decision-making is defined as a process of choosing one or more possible alternatives as course of action for attaining one or more goals (Al-Tarawneh, 2012).

Though BI and BA are at times treated separately, some take the stance that they interchangeable, while others argue they are distinct but connected tools (Gnatovich, 2006). Chen et al. (2012) use *Business Intelligence and Analytics (BIA)* as a unified term referring to "the techniques, technologies, systems, practices, methodologies, and applications that analyse critical business data to help an enterprise better understand its business and market and make timely business decisions" (p. 1166). Accordingly, BIA can be regarded as the practices needed to derive value from Big Data. The emergence of Big Data thus represents the latest chapter in BIA (Gartner, 2013; Wixom, et al., 2014).

Impacts: The Value Component

Big Data is expected to have a strong impact on almost every industry, with the potential to dramatically transform our society (Bollier, 2010; Mauro, et al., 2016). As the applications and power of Big Data are still emerging, discussing the impacts of Big Data unequivocally goes beyond the scope of this paper. However, Big Data is already forcing companies to reconsider their organization and business processes due to the availability of data that can be transformed into information to underpin a competitive advantage in data-driven markets (McAfee & Brynjolfsson, 2012). A substantial appeal of Big Data is that it can fundamentally change our understanding of decision-making, with wide implications for the way business compete and operate (EMC, 2013; Schrage, 2016). McAfee and Brynjolfsson (2012) argue that as the tools and philosophies of Big Data spread, our "long-standing ideas about the value of experience, the nature of expertise, and the practices of management" will change. A study from MIT concluded that companies engaged in data-driven decision-making were, on average, 5% more productive and 6% more profitable than their competitors (Brynjolfsson, et al., 2011). Furthermore, Tambe (2014) examined the extent to which early adopters of Big Data technology would have distinct advantages over their competitors. The study demonstrated that firms' investments in such technology, for the period 2006 to 2011, were associated with 3% faster productivity growth. This performance gap is predicted to continue growing as more relevant data are generated (EY, 2014). Similarly, the European Commission (2016) predicts that the use of Big Data by the top 100 EU manufacturers could lead to savings worth €425 billion. For the year 2020, employing BIA on Big Data could bring the EU economic growth by an additional 1.9%, equivalent to a GDP increase of €206 billion. Regardless, these impacts are just the tip of the iceberg, as the power and applications of Big Data are still emerging.

2.1.1 A Consensual Definition

Figure 1: The four main components of Big Data (Source: Own drawing based on Mauro et al., 2016)

Evidently, Big Data is a multifaceted concept. A recent review of Big Data literature by Mauro et al. (2016) identified four common "themes", i.e., prevalent concepts representing the four main components of Big Data; information, technology, methods, and impacts. *Information* refers to the data-related aspects of Big Data and is commonly associated with the "3 Vs"; volume, velocity, and variety. *Technology* relates to the technological needs for processing data and is a prerequisite for making use of Big Data. *Methods* are the techniques that can be applied in BIA to get meaningful and actionable information. Lastly, *impacts* refer to the influence Big Data has on business, government, and society, and is associated with value creation. Figure 1 illustrates these four main themes in the Big Data literature. Based on this classification of Big Data, Mauro et al. (2016) proposed the following consensual definition comprising all four components:

> "Big Data is the Information asset characterized by such a High Volume, Velocity and Variety to require specific Technology and Analytical Methods for its transformation into Value" (Mauro, et al., 2016, p. 129)

This definition is compatible with the usage of terms such as *Big Data Technology*, the focus of the present study, and is therefore considered the working definition of Big Data.

2.2 Big Data Challenges

While Big Data is predicted to have a deep transformational impact on all aspects of society, deriving valuable information from Big Data is no ordinary task. To further complicate the matter, the value of Big Data is often case-dependent, where some data are highly valued by one company but worthless to others (Ylijoki & Porras, 2016). With no clear path to value, even companies that believe in the transformative power of Big Data are left questioning how to achieve it. Thus, the real challenge is discovering value in the data (Jin, et al., 2015; Court, 2015). As appealing as the concept of Big Data may be, companies lacking skills and resources will have a difficult time managing the messy data that are available to them. More so, data scientists spend up to 80% of their time trying to make sense of data rather than generating new business insight (Jin, et al., 2015). Without the necessary data handling resources and a system supporting the use of data, making decisions in a timely manner might be unattainable, and potentially limit the effectiveness of companies (EMC, 2013). Further, when employing Big Data, on-premise solutions may involve significant operational risks and expensive infrastructure. The ongoing maintenance of these system can discourage organizations from using it (Khanna, 2016). Other challenges include the need to ensure that the right infrastructure and appropriate data governance are in place (Manyika, et al., 2011).

Thus, companies face considerable challenges in leveraging the value of Big Data. The aim of following section is not to present a comprehensive review of these challenges, but rather a brief overview of three pressing concerns that are thought to hinder the adoption of Big Data; namely the lack of data skills, privacy and security matters.

Data Skills Gap

Data scientists are high in demand. The skill set that these highly qualified professionals must possess is “a solid foundation in math, statistics, probability, and computer science” (Davenport & Patil, 2012). More so, data scientists must carry out the task of uncovering hidden patterns, identifying trends, or discovering valuable information from raw data – this is essential for any organization that intend to derive business advantage from Big Data (Manyika, et al., 2011). Not only are data scientists required to master Big Data technologies, computer languages, and techniques, but they also need to possess the necessary business acumen to create models that can be applied to genuine business problems. Consequently, qualified candidates are scarce, which makes recruitment problematic and expensive (Carnelley & Schwenk, 2016c).

According to Manyika et al. (2011), there is a need for 1.5 million managers and analysts in the United States alone. These professionals have the capacity to understand the business aspect of Big Data and contribute to deriving valuable new information. Similarly, 140 000 to 190 000 data scientist positions in the United States will remain unfilled by 2018. In fact, data scientists represented less than 2% of the global population of data workers in 2015. According to the International Data Corporation (IDC), the demand for data scientists will continue to grow significantly, representing a compounded annual growth rate of 14.3% between 2015-2020, which is much faster than the demand for data workers (Carnelley & Schwenk, 2016c). On LinkedIn, statistical analysis and data mining were the second most desirable skills to have in 2016 (Fisher, 2016). This shortage of talents necessary to make the most of Big Data is a pressing challenge and one that companies and policy makers must address (Manyika, et al., 2011).

However, the data skills gap is being tackled from several different angles. For instance, universities have established degree programs to seed the next generation of analytically literate employees. More so, graduates that possess relevant qualifications not only in mathematics, statistics, and computer science, but also social science and economics, are recruited by businesses and government with the intention to bring together multidisciplinary teams of data scientists. Furthermore, IDC believes that a part of the skills shortage can be addressed by utilizing Self-Service Business Intelligence (SSBI) (Carnelley & Schwenk, 2016c). *SSBI* is defined as “the facilities within the BI environment that enable BI users to become more self-reliant and less dependent on the IT organization. These facilities focus on four main objectives: easy access to source data for reporting and analysis, easier and improved support for data analysis features, faster deployment options such as appliances and cloud computing, and simpler, customizable, and collaborative end-user interfaces” (Imhoff & White, 2011, p. 5). To the extent that SSBI is viable, companies can utilize current employees in the organization that are less specialist and technical; to mitigate the problem of skill shortages by democratizing data access and analytics usage (Court, 2015). Nevertheless, data science will never be entirely self-service, but the goal is to blend skills from across the organization. Together with the right self-service tool, effective training, collaborative working practices, and a robust approach to data governance, companies will be able to harness more of the opportunities within their data, as well as maximising the agility and effectiveness of their workforce (Carnelley & Schwenk, 2016c).

Privacy

The use of Big Data raises concerns as it challenges key privacy principles. Privacy is the "right of individuals or cooperative users to maintain confidentiality and control over their information when it is disclosed to another party" (Porambage, et al., 2016, p. 37). As increasing amounts of data are collected about individuals, it is becoming more important than ever to safeguard fundamental principles of privacy (Datatilsynet, 2013). In particular, the massive retention and analysis of demographic, socioeconomic, behavioural, financial, and other transactional data by companies, authorities, and other large organizations present privacy issues in almost every field where Big Data is utilized (Ali, et al., 2016; Jain, et al., 2016). From a privacy perspective, the challenge is to ensure individuals have sufficient control over their own data to prevent misuse and abuse by data owners (i.e., companies that make use of Big Data) and data brokers (i.e., companies that collect data for the purpose analysing and licensing/selling information to other organizations). From the perspective of data owners and data brokers, the challenge is ensuring compliance with data regulations, while simultaneously preserving data utility (i.e., the value of their data). Though the compliance with privacy regulations ensure that consumers' personal information are being appropriately managed, the protection of personal data has become increasingly challenging as data are multiplied and shared ever more widely (European Commission, 2016).

As ever more data are available, the costs of storage are plunging, and the desire to retain data is increasing, companies, governments, and other large organizations are building massive collections of data (i.e., Big Data sets). With this comes the increasing privacy and security concerns related to storage, access, and usage of these data. Furthermore, the risk of data breaches[2] is increasing. Relatively recent incidents in the United States, such as Edward Snowden's NSA leaks and the massive security breach at US retail chain Target, where more than 40 million records containing names, addresses, and credit card information were exposed, highlight the consequences for both individuals and data owners (Macaskill & Dance, 2013; Bloomberg, 2014; Picchi, 2014). High profile data breaches, such as the attack on Target, have pushed consumers to withdraw from businesses that do not sufficiently protect personal data. According to the European Commission (2016), trust has become the key determinant of the adoption of Big Data technology in Europe. Consumers are increasingly concerned about privacy, and loss of trust translates into lost opportunities and revenues for

---

[2] A data breach is a security incident where individuals gain unauthorized access to sensitive, protected or confidential data.

businesses. However, companies’ compliance with privacy legislation goes a long way in protecting consumers, and the recently introduced General Data Protection Regulation (GDPR) was designed to protect and empower all EU citizen’s data privacy by reshaping the way organizations across the region approach data privacy. The new, unified privacy law for EU will replace the current legal framework by May 2018 (Lord, 2017), and is a prominent example of a new wave of universal privacy regulations that is forcing businesses to rethink how they collect, manage, and govern access to personal data.

The *Netflix Prize*, which began in 2006, raised many concerns surrounding privacy. This was an open competition for the best collaborative filtering algorithm that would predict how much someone would enjoy a movie based on their movie preferences. The winner team surpassed Netflix’s own algorithm for predicting ratings by 10.06% (Chen, et al., 2012). For this competition, Netflix provided datasets that were constructed to preserve the privacy of their userbase. However, two researchers were able to identify individual users by matching the datasets with film ratings on the Internet Movie Database (IMDb) (Narayanan & Shmatikov, 2007). This example illustrates a major challenge of Big Data from a privacy perspective, namely the risk of re-identification. Re-identification means “data that initially emerges as anonymous is identifiable again my means of various techniques” (Datatilsynet, 2013, p. 10). For the case of Big Data, individuals may be identified from data that are initially anonymized through the compilation of multiple data sets. Thus, the potential of Big Data comes with a risk; the users’ privacy is frequently at danger. With advancements in techniques and algorithms that can be used to re-identify individuals, control over personal information becomes harder to maintain.

Security

Accompanying the current digital transition is a worldwide increase in IT budgets for security. According to PAC, 70% of spending on security is on protection, 20% on detection, and 10% on the response to security threats (Lartigue, 2016). A downside of the digital evolution and the emergence of Big Data is that the digital vulnerabilities of IT users are at a greater risk. Data has become the primary target of attackers, whether for criminal activities or espionage (Lartigue, 2016). Security is the practice of “defending information and information assets through the use of technology, processes and training from: unauthorized access, disclosure, disruption, modification, inspection, recording and destruction” (Jain, et al., 2016, p. 3). While data privacy is focused on the use of governance of individual data, security on the other hand, concentrates more on protecting data from malicious attacks and the misuse of stolen data.

The digital transformation of companies, governments, and other large organizations is generating huge volumes of data, which are captured and stored at extensive data centres. These data centres, known as modern data warehouses, data lakes, or data reservoirs, often comprise companies' most valuable information assets (Lartigue, 2016), and needs to be protected from three main groups of security threats; technical faults, internal threats, and external threats. *Technical faults* refer to abnormalities or defects of a component, equipment, or system that may lead to failure. In the event of technical failures, it is critical to have adequate data protection. However, backing up Big Data environments are subject to financial constraints, and even in the best protected data centres, data can be physically damaged and lost (Lartigue, 2016). An incident in Belgium, where Google's data centre was struck by lightning, damaging several hard disks and leading to permanent loss of data[3], illustrates this (Greenberg, 2015). While *internal threats* come from the unauthorized access to data within a company or organization, *external threats* come from remote unauthorized access (e.g., hackers). By introducing data governance that limits the employees' access to the information system, internal threats can be minimized. On the other hand, external threats are becoming increasingly numerous and harder to overcome.

As very few organizations are likely to build a Big Data environment in-house, many companies use a third-party cloud solution for their Big Data deployments. In fact, cloud computing is one of the technologies that has been a precursor and facilitator to the emergence of Big Data, and the concepts are inextricably linked; it has enabled companies to store large amounts of data, and been especially useful for smaller organizations that do not have sufficient storage capacity (Hashem, et al., 2015). More so, cloud computing allows organizations to consolidate data from all sources and do it at a Big Data scale (Khanna, 2016). However, using a third-party cloud solution has its security threats. At present, there is not a single supplier providing a standard and robust security solution for Big Data environments, which can be problematic when Big Data brings together a large amount of data, including sensitive information that must both be protected from intrusions and hidden from most users of the system (Lartigue, 2016). As the consequence of potential data breaches might be

[3] The accident caused 0.000001% of Google's data being permanently lost.

disastrous for companies, decision-makers need to evaluate the risk of using a third-party cloud solution for their Big Data deployments.

2.3 Big Data Adoption in Norway

The sudden drop in oil prices in 2014 has contributed to the ongoing transformation of the Norwegian economy from being oil-driven towards a more diverse industrial landscape, where data is becoming the new oil powering the future of the information economy. This transformation has been a priority on the Norwegian Government's agenda for a long time, as well as the digitalization and modernization of the public sector (Regjeringen, 2016b). In 2016, the Norwegian Government had two digital agendas specifically for Big Data; first, to consider strategies on the use of Big Data in the public sector, and second, to monitor the technology developments of Big Data. The latter will make it easier for companies to exploit and understand Big Data technology (Regjeringen, 2016a).

Nevertheless, according to PAC (2016b), there are two main factors holding back investments in Big Data technology. First, data scientists are both scarce and costly to hire. Particularly in Norway, with a limited population of 5 million people, the challenge is finding qualified professionals. Educating new data scientists and bringing in foreign talent might be costly options, albeit necessary ones. Fortunately, the number of applicants for IT studies at the undergraduate level increased by 31% in the period between 2016-2017 (Gjerde, 2017). This is a welcoming trend that could relieve some of the skill shortages in Norway. Ironically, due to admission limits for these programs, only a portion of these applicants will be admitted, and Norwegian institutions have yet to figure out how to satisfy this increasing demand. The second factor holding back investments is the fact that Big Data comes with high overhead costs, which means that only the largest firms are able to justify the investments. However, technology and applications for Big Data are becoming easier to use and costs are rapidly declining, which is easing this problem (PAC, 2016b).

Industry Characteristics

Norway has the smallest IT service and software market in the Nordic Region. Notably, the manufacturing market, with Norwegian giants such as Statoil, Hydro, Orkla, and Yara, have a disproportionately large influence on the overall software market (Hallberg & Ahorlu, 2016; PAC, 2016b). Nonetheless, BI software is the second fastest growing market in Norway after SaaS[4]. There are three combined factors driving this trend. First, a growing volume of data is being generated. Second, the increasing number of techniques for the analysis of this data. Third, the development of more user-friendly tools that are used to derive insights, such as data visualization and real-time analytics (PAC, 2016a). Large companies in Norway, especially within the oil and gas sector, are increasingly characterized to be ready for Big Data solutions. According to PAC (2016b), the decline in oil prices will rush investments in Big Data solutions because of the need to spend money where the biggest return lies. Sectors such as banking, retail, and the public sector, also carry huge potential. As businesses seek to remain competitive in an increasingly data-driven marketplace, and the available data sources continue to grow, Big Data technology can be a source of competitive advantage (PAC, 2016c). However, as the GDPR takes effect in 2018, businesses must recognize and comply with the new data protection regulations when considering the use of Big Data. Like most Big Data adopters around the world, Norwegian businesses face challenges surrounding privacy and security when undertaking Big Data initiatives. One of the segments that will fare best in Norway is therefore the evolving cyber security landscape, which is driving investments in security software to keep pace with these challenges (PAC, 2016b).

Furthermore, Norway generally scores high on international rankings in information and communications technologies (ICT), such as digital skills and infrastructure. For instance, Norway scores well above the EU average on all five dimensions[5] mentioned in EU's *Digital economy and society index*. Compared to 28 EU countries, Norway is overall ranked as number two on this performance index after Denmark (Regjeringen, 2016a). However, in terms of Big Data technology, there is a different story to tell. For the period between 2014-2019, IDC predicts Norway to have a compounded annual growth rate of 20.9% in Big Data technology and services, which is less than the Western European average of 22.7%. Notably, for the year 2018, Norway is predicted to have the lowest growth rate in Western Europe of 18.8% (Carnelley & Schwenk, 2016a). IDC's forecast for the year 2020, indicates that Norway also has one of the lowest shares in the BA software market in Western Europe. Sweden, for instance, has an estimated share of 4.8%, compared to Norway's 1.8%. Greece, Ireland, and Portugal, are the only countries predicted to have lower shares than Norway (Carnelley &

---

[4] SaaS (Software as a service) is a "software distribution model in which a third-party provider hosts applications and makes them available to customers over the Internet. SaaS is one of three main categories of cloud computing, alongside infrastructure as a service (IaaS) and platform as a service (PaaS)" (Rouse & Casey, 2016).

[5] Connectivity, human capital, use of internet, integration of digital technology, and digital public services.

Schwenk, 2016b). While Norway is quite advanced with regard to ICT, these numbers suggest that the adoption rate of Big Data technology in Norway is slower than its fellow European countries. Thus, many businesses in Norway are likely to be in the early stages of Big Data adoption.

3. Theoretical Framework

This chapter introduces the theoretical foundation for the present study by presenting a background on the innovation adoption literature, followed by a discussion of diffusion research and the stages of information technology (IT) innovation adoption in organizations. Next, to develop this paper's research model, the Diffusion of Innovation (DOI), Technology Acceptance Model (TAM), and Technology-Organization-Environment (TOE) framework will be presented. The chapter then proceeds to discuss the application of these models for studying organizational adoption and presents a tentative research model, which is developed further by discussing and stating hypotheses for relevant constructs from DOI, TAM, and TOE research. Lastly, a discussion of the conceptualization of Big Data technology adoption is presented, before proposing the final research model.

3.1 Background on Innovation Adoption

An innovation can be defined as any idea, product, program, or technology that is new to the adopting unit (Premkumar & Roberts, 1999). Innovation has been extensively studied and has a long history as a multi-disciplinary field, with research conducted in disciplines such as economics, management, education, sociology, organizational studies, information technology, and many others (Rogers, 1983). Despite diversity across these disciplines, they are unified by their concern with three basic research questions, one of which this paper seeks to contribute:

"What determines the propensity of an organization to adopt a particular innovation"
(Fichman, 1999)

A significant amount of research has been conducted to better understand factors influencing the adoption of innovations. Innovation adoption research has produced a variety of competing and complementary models, each suggesting different sets of determinants of adoption. While theories on innovation adoption were originally developed to examine the adoption by individuals making autonomous choices (Davis, 1986; Fichman, 1992), recent research have extended innovation theory to include more complicated adoption scenarios, such as by organizations (Rogers, 1983; Kwon & Zmud, 1987; Tornatzky & Fleischer, 1990). At both the individual and organizational level of analysis, research can be divided into two main approaches; process research and **antecedent factor research** (King, 1990). Innovation process research examines the sequence of events that constitute the process of innovation adoption, and is generally more qualitative by nature. Innovation antecedent factor research[6], on the other hand, focuses on identifying and examining the determinants of innovation adoption (Hameed, et al., 2012a). This thesis is consistent with the latter research approach, as the purpose is to identify and examine determinants of the adoption of Big Data technology.

While no single, unified theory of innovation adoption exists, innovation adoption research has produced a variety of competing and complementary models and frameworks (Fichman, 1999; Hameed, et al., 2012a). Rogers' (1983) Diffusion of Innovations (DOI), Davis' (1986) Technology Acceptance Model (TAM), and Tornatzky and Fleischer's (1990) Technology-Organization-Environment (TOE) framework are among the most influential and commonly used theoretical perspectives on IT innovation adoption (Hameed, et al., 2012a). These have been extensively used by researchers to study adoption of a rich variety of innovations, including organizational adoption of Big Data (Nam, et al., 2015; Agrawal, 2015; Sun, et al., 2016). By defining Big Data technology as an IT innovation, DOI, TAM, and the TOE framework become relevant for Big Data adoption.

Diffusion Research

Diffusion research examines how innovations spread, and can be traced back to the observations of the French scholar Gabriel Tarde, described as an "intellectual far ahead of his time in thinking about diffusion" (Rogers, 1983, p. 40). In his book, *The Laws of Imitation*, Tarde (1903) originated several key diffusion concepts, including what we today refer to as the S-curve of diffusion. While not calling the concepts by their present-day names, he did recognize that the rate of adoption of an innovation had "a slow advance in the beginning, followed by a rapid and uniformly accelerated progress, followed again by a progress that continues to slacken until it finally stops" (p. 127). The early concepts of diffusion studied by Tarde, as well as by British, German, and Austrian diffusionists, laid the foundation for several decades of diffusion research in the social sciences (Rogers, 1983; Stacks & Salwen, 2009).

By reviewing a substantial number of diffusion studies, Everett Rogers, a professor in rural sociology, observed that the diffusion process displayed patterns and regularities, even across conditions, innovations, and cultures (Stacks & Salwen, 2009). In his book, *Diffusion of Innovations*, Rogers synthesized these findings into a theory of the adoption of innovations

[6] Antecedent factor research (King, 1990) is sometimes referred to as variance research (Hameed, et al., 2012a) and adopter research (Fichman, 1999).

among individuals and organizations (Rogers, 1983). Rogers' seminal work in the diffusion of innovations is the second most cited publication in the social sciences (Green, 2016).

Stages of Innovation Adoption

The organizational adoption of an innovation is not a binary event but rather a stage-based process that unfolds over time (Fichman, 1992, p. 197). Studies on organizational innovation adoption therefore target distinct stages on the *adoption continuum;* the stages used to describe the adoption process. As such, ambiguity in the conceptualization of the adoption construct can lead to issues with misinterpretation and misunderstandings of both the research model and results (McKinnie, 2016). This section will therefore define adoption and review innovation diffusion literature, which will serve as the basis for the development of this paper's conceptualization of the Big Data adoption construct (Ch. 3.3.5).

Information systems (IS) adoption research is grounded in the theoretical framework of diffusion of innovations (Rogers, 1995). From a technological diffusion perspective, adoption describes the organizational effort directed toward diffusing an IT innovation throughout the firm (Cooper & Zmud, 1990). According to Rogers (1995), the adoption of an innovation starts with the firm's initial awareness, knowledge, and evaluation of the innovation. These initial stages include "both identifying and prioritizing needs and problems on one hand, and searching the organization's environment to locate innovations of potential usefulness to meet the organization's problems" (Rogers, 1995, p. 391). Together, the initial stages constitute initiation, defined as "all of the information gathering, conceptualization, and planning for adoption of an innovation, leading up to the decision to adopt" (Rogers, 1983, p. 364). Following the decisions to adopt comes restructuring or re-invention of the innovation to fit the organizational needs, clarification of the role and purpose of the innovation, and routinization of the innovation by incorporating it into the regular activities of the firm. Together, these latter stages constitute implementation, which Rogers (1983, p. 364) defined as "all of the events, actions, and decisions involved in putting an innovation into use". Figure 2 illustrates a simplified adoption process.

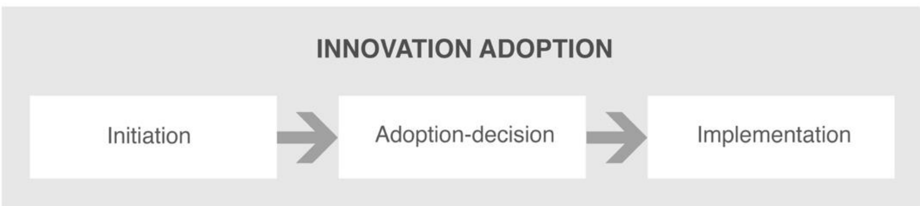


Figure 2: A simplified adoption process (Source: Own drawing based on Rogers, 1983)

Rogers' (1995) description of the adoption process implies that after an innovation is adopted, it needs to be accepted, adapted, routinized, and institutionalized into the firm. The idea that adoption and routinization are two distinct stages is consistent with what assimilation theory calls an *assimilation gap*[7], i.e., innovation usage tends to lag behind adoption (Zhu, et al., 2006a). In other words, "widespread acquisition of an innovation need not be followed by widespread deployment and use by acquiring firms" (Fichman & Kemerer, 1999, p. 256).

Rogers (1995) describe a sequential process by which adoption of innovations by organizations, starting with awareness or knowledge of the innovation, results in the introduction and use of a product, process, or practice that is new to the adopting organization (Hameed, et al., 2012a). This idea that adoption of IT innovation by organizations is a multi-stage, sequential process has been widely recognized. Based on Lewin's (1947) change model[8], Zmud and colleagues (Kwon & Zmud, 1987; Cooper & Zmud, 1990) proposed a model of IT adoption as a six-stage process starting from initiation and progressing through adoption, adaptation, acceptance, routinization, and infusion. The stages in this model of organizational IT adoption are presented in Table 1.

Table 1: Adoption stage model (Cooper & Zmud, 1990, p. 124-125)

Stage	Definition
Initiation	A match is found between the innovation and its application in the organization
Adoption	A decision is reached to invest resources to accommodate the implementation effort
Adaption	The IT application is available for use in the organization
Acceptance	The IT application is employed in organizational work
Routinization	Usage of the IT application is encouraged as a normal activity; The IT application is no longer perceived as something out of the ordinary
Infusion	The IT application is used within the organization to its fullest potential

Building on the above, more recent research explain IT adoption as a process moving through awareness, interest, evaluation, commitment, limited deployment, partial deployment,

[7] The term assimilation gap was introduced by Fichman and Kemerer (1999) to explain why information technology may be widely acquired, but sometimes only sparsely deployed among the acquiring firms. Assimilation gap was defined as "the difference between the pattern of cumulative acquisitions and cumulative deployments of an innovation across a population of potential adopters" (Fichman & Kemerer, 1999, p. 258).

[8] Kurt Lewin's (1947) well-known change model proposed organizational change as a three-step procedure: Unfreezing, changing, and refreezing. Initiation is associated with unfreezing; adoption and adaption are associated with changing; and acceptance, routinization and infusion are associated with refreezing.

and general deployment (Fichman, 2001; Rai, et al., 2009; McKinnie, 2016). These studies employ up to a seven-stage process of IT adoption to identify organizations' current stage in the adoption process.

Our literature review also highlights that studies in accordance with innovation antecedent factor research, i.e., researchers that primarily study the determinants of innovation adoption rather than innovation adoption as a process of change, take two distinct approaches. The first approach refers to studies that operationalize adoption as a dichotomy; whether the organization is an adopter or not (Iacovou, et al., 1995; Thong, 1999). These studies adapt a simplified conceptualization of adoption to reflect the complete, multi-stage adoption process that firms face. In a study of electronic data interchange (EDI) adoption, Iacovou et al. (1995) defined EDI adopters as those that possessed the capability to transact via EDI, and non-adopters as those that did not possess this capability. Similarly, studying IS adoption, Thong (1999) used a dichotomous measure, defining adopters as organizations using at least one software application. The second approach refers to studies that use a multi-item scale to operationalize the entire process of adoption. Both McKinnie (2016) and Rai et al. (2009) used items consistent with a seven-stage adoption process. Zhu et al. (2006a), on the other hand, developed multiple items for each of the three adoption stages presented by Rogers (1995). Using a multi-item scale enables researchers to reveal how determinants of adoption have differential effects at the different stages of adoption.

The literature makes compelling arguments for both simplifying the conceptualization of adoption, as well as using a multi-item scale to reflect the complete adoption process. Whereas IT adoption may simply be seen as an adoption-decision, the conceptualization is often extended to include pre-adoption and post-adoption stages (Hameed, et al., 2012a).

3.2 Theoretical Models of IT Innovation Adoption

3.2.1 Rogers' Diffusion of Innovations (DOI)

The diffusion of innovation model was developed by Rogers in 1962 to explain how innovations spread over time through a social system between actors (Rogers, 1983), where an actor may be any societal entity, including individuals, groups, or organizations (Wejnert, 2002). As actors in a social system communicate and influence each other, their probability of adopting an innovation is affected. According to Rogers (1983), innovations can be adopted or rejected by individual members of a system, or by the entire social system through a collective or an authority decision.

Rogers (1983) used the term adoption for when the decision to accept and use an innovation had to be made. At the individual level, which has been the dominant focus of traditional diffusion research, the adoption-decision is optional. By **optional decisions**, Rogers states that the choices to adopt or reject an innovation are made by the individual independent of the decisions of other members in a system. Therefore, the distinctive aspect of optional innovation decision is that the individual is the unit of decision-making, rather than the social system. However, when considering an organization as the system in which the innovation decision occurs, decisions can be made on behalf of the entire social system through a collective or an authority decision (Rogers, 1983). **Collective decisions** are "choices to adopt or reject an innovation that are made by consensus among the members of a system", whereas **authority decisions** refer to "choices to adopt or reject an innovation that are made by a relatively few individuals in a system who possess power, status, or technical expertise" (Rogers, 1983, pp. 29-30). Consequently, from an organizational level of analysis, the social system is the unit of decision-making, rather than the individual. The decision to adopt or reject innovations within a formal organization will usually fall within the collective or authority decision category, as the decision is generally made by top management (Rogers, 1983; Premkumar & Roberts, 1999).

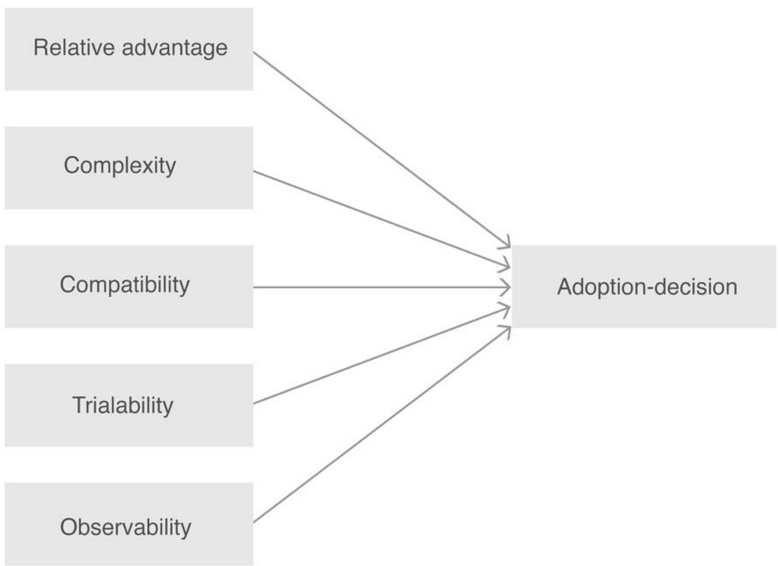


Figure 3: The Diffusion of Innovations Model (Source: Own drawing based on Rogers, 1983)

Understanding what drives the diffusion of a technological innovation requires an understanding of organizations' adoption behavior. Rogers (1983) provides a classification of five generalized attributes of innovations, and postulates that it is ultimately the adopter's perception of these attributes that affect the rate of adoption of the innovation; relative advantage, complexity, compatibility, trialability, and observability. A model of these attributes, herein defined as the Diffusion of Innovations (DOI), is presented in Figure 3.

Relative advantage is "the degree to which an innovation is perceived as being better than the idea it supersedes" (Rogers, 1983, p. 213). The nature of the innovation and the context within which it is adopted largely determine what type of relative advantage is important to the adopters. **Complexity** refers to "the degree to which an innovation is perceived as relatively difficult to understand and use" (Rogers, 1983, p. 230). While any innovation can be classified on a simplicity-complexity continuum, the perception of complexity may vary between adopters. **Compatibility** is "the degree to which an innovation is perceived as consistent with the existing values, past experiences, and needs of potential adopters" (Rogers, 1983, p. 223). The general notion is that an innovation that is more compatible entails less uncertainty for the adopter. **Trialability** concerns "the degree to which an innovation may be experimented with on a limited basis" (Rogers, 1983, p. 231). New ideas that can be tested out will generally be adopted more rapidly. Lastly, **observability** refers to "the degree to which the results of an innovation are visible to others" (Rogers, 1983, p. 232). Concerning technology, the software component of a technological innovation is typically less apparent to observations.

In regard to technological ideas, which comprise the majority of innovations studied in diffusion research, Rogers' five attributes of innovations are typically assumed to be direct antecedents of innovation adoption-decisions (Rogers, 1983; Arts, et al., 2011; Puklavec, et al., 2014). Besides complexity, which is believed to have a negative influence on adoption, each of the attributes are thought to positively affect adoption-decisions.

3.2.2 The Technology Acceptance Model (TAM)

Building on the principles of the Theory of Reasoned Action (TRA) (Fishbein & Ajzen, 1975), one of the most fundamental and influential theories of human behaviour (Venkatesh, et al., 2003), TAM has been widely used to predict the acceptance and use of IT in organizations (Davis, 1986; 1989; 1993; Davis, et al., 1989). Whereas TRA[9], drawn from social psychology, is designed to predict and explain a wide range of human behaviour across contexts, TAM is an adaptation of the TRA specifically tailored to the IS contexts (Venkatesh, et al., 2003). As such, the purpose of TAM is to offer an explanation for the determinants of acceptance and use of IT (Davis, et al., 1989). Comparable to how traditional diffusion research focuses on the individual's innovation adoption-decision, the majority of studies employing TAM have targeted the individual users' acceptance and use of technology.

In the IS field, TAM is generally considered the most influential theory for describing an individual's acceptance and use of technological innovations (Lee, et al., 2003), and has been widely studied and empirically supported by a substantial number of IS researchers (Legris, et al., 2003; Ma & Liu, 2004; King & He, 2006). The TAM, presented in Figure 4, consists of five central elements: perceived usefulness, perceived ease of use, attitude towards use, intention to use, and actual use.

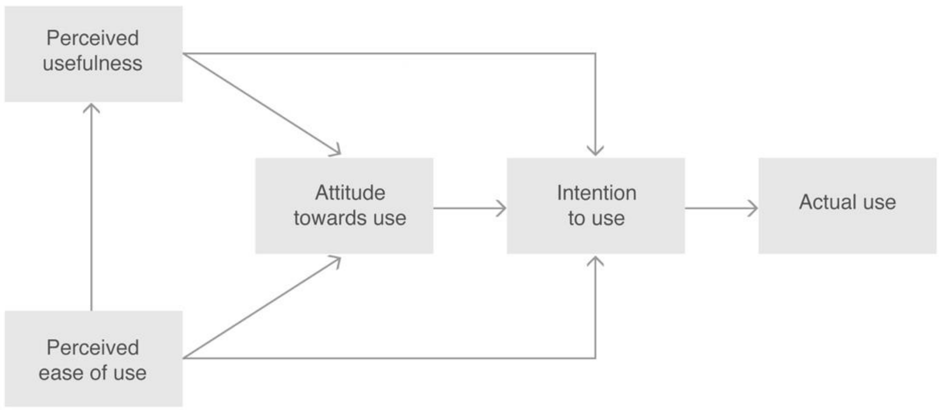


Figure 4: The Technology Acceptance Model (Source: Own drawing based on Davis et al., 1989)

According to TAM, a user's perception of the technology's usefulness and ease of use jointly determine the user's attitude towards use. Davis (1989) defines **perceived usefulness** as "the degree to which a person believes that using a particular system would enhance his or her job performance" (p. 320), and **perceived ease of use** as "the degree to which a person believes that using a particular system would be free of effort" (p. 320). A person's attitude towards use, in turn, determine intention to use. Building upon the fundamental presumption

[9] A detailed discussion of TRA is not performed as it falls outside the scope and purpose of this paper. However, some reference is made to emphasize the role of the TRA as a reference paradigm within which TAM was originally developed (Davis, 1986).

of TRA[10], TAM assumes that a person's intention to use ultimately translate directly into actual use. This implies that actual use should be predictable from measures of intention to use, and that any other factors that influence actual usage do so indirectly through intention to use (Davis, et al., 1989).

In the original model (Davis, et al., 1989), usefulness is seen as a direct antecedent of intention to use, although partly mediated by attitude. Ease of use, on the other hand, was thought to be fully mediated by perceived usefulness and attitude. However, subsequent studies employing TAM have found that ease of use has a direct influence on intention to use (Davis, et al., 1989; Venkatesh & Davis, 1996; Venkatesh, et al., 2003).

3.2.3 The Technology-Organization-Environment (TOE) Framework

A higher unit of analysis of IT adoption is that of the organization and the environment in which it operates. Extending the analysis of individual adoption to organizations requires a considerably more expansive framework that captures both the individual and organizational determinants of innovation adoption (Hameed, et al., 2012a). Wejnert (2002) proposed a framework where determinants of innovation adoption were grouped into three major components; characteristics of innovations, characteristics of innovators, and environmental context. Similarly, Li et al. (2011) described a framework that classified decision factors into three dimensions; decision entity factors, decision object factors, and context factors. However, the most recognized attempt at identifying and categorizing determinants of IT adoption in organizations is presented by Tornatzky and Fleischer (1990) in their book; *The Process of Technological Innovation*. They propose a framework of how the determinants of IT adoption can be grouped into three contextual elements that influence the adoption and implementation of technological innovations in organizations.

The TOE framework, presented in Figure 5, recognizes that adoption of technological innovations is influenced by a range of factors in the context of the technology, organization, and external environment (Tornatzky & Fleischer, 1990). Being an organizational-level framework, TOE explains that these three elements of a firm's context stimulate and influence the technology innovation adoption-decision.

[10] An individual's intention to use is understood as "the immediate causal determinant of the user's overt performance of that behaviour" (Davis, 1986).

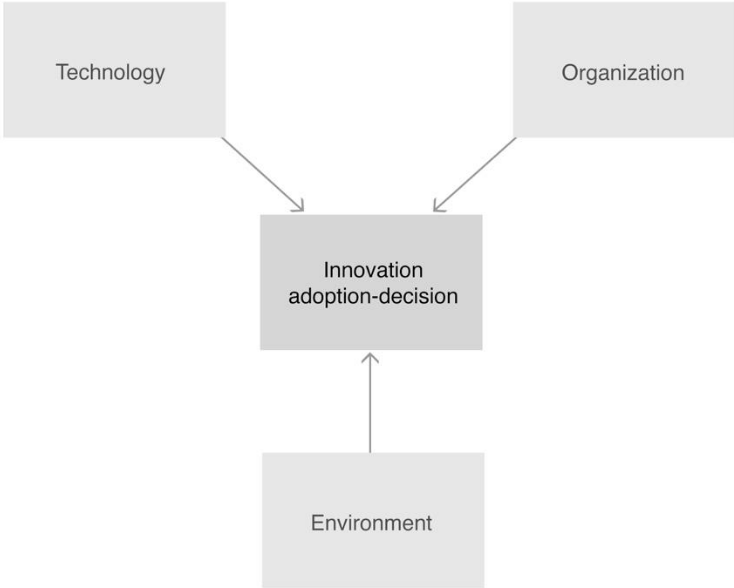


Figure 5: The Technology-Organization-Environment framework (Source: Own drawing based on Tornatzky & Fleischer, 1990)

The **technological context** represents both the internal and external technologies available to the firm, which include existing technologies inside the firm, as well as the pool of available technologies in the market (Tornatzky & Fleischer, 1990). Researchers have primarily taken two different approaches to studying the technological context. The first approach focuses on the specifics of the innovation being studied. In describing the technological context for a given adoption-decision, most studies focus solely on the innovation characteristics; that is, the features or characteristics of the technological innovation itself (Hameed, et al., 2012a; Hoti, 2015). While the characteristics studied by researchers are numerous, our literature review highlights the most significant characteristics of technological innovations that influence adoption (see Appendix C.1). Drawing heavily on diffusion research and DOI (Rogers, 1983), the literature has consistently found relative advantage[11], complexity[12], and compatibility as key factors influencing adoption (Davis, 1989; Igbaria, et al., 1997; Thong, 1999; Premkumar & Roberts, 1999; Grandon & Pearson, 2004; Ifinedo, 2011; Boonsiritomachai, 2014; Hung, et al., 2016). Some researchers have also extended the number of innovation characteristics studied to include observability and

[11] Relative advantage is frequently referred to as expected or perceived benefit (Iacovou, et al., 1995; Chwelos, et al., 2001; Nam, et al., 2015).

[12] Complexity and perceived ease of use are seen as pair-wise opposites, relating to how innovations are perceived as relatively difficult to understand and use.

trialability (Ramdani & Kawalek, 2009; Alrousan, 2014; Boonsiritomachai, 2014), and cost (Premkumar & Roberts, 1999; Hameed, et al., 2012b), albeit with varying success. The second approach to describe the technological context relates to researchers that look beyond the characteristics of the innovation. These researchers include a measure of the organization's readiness for the technology being studied, contending that adoption is contingent on complementary factors such as technological readiness, infrastructure, and maturity (Ling, 2001; Malladi & Krishnan, 2013).

The **organizational context** represents the intraorganizational environment, and describes the characteristics and resources of organizations that facilitate or constrain the adoption of technological innovations (Tornatzky & Fleischer, 1990). The common approach to define the organizational context is in terms of the descriptive measures that the IT literature has identified as determinants of organizational adoption. Our literature review highlighted the most significant determinants (see Appendix C.2), where organizational size has been the most frequently examined factor in the study of organizational innovation adoption (Hameed, et al., 2012b). Researchers have consistently found a significant relationship between size and IT adoption (Premkumar & Roberts, 1999; Zhu, et al., 2004; Buonanno, et al., 2005). Secondly, top management support and leaders' attitude towards change are also found as top predictors (Premkumar & Roberts, 1999; Al-Isma'ili, et al., 2016; Hung, et al., 2016). Thirdly, IS competence and IT expertise are seen as major factors in the adoption of new technologies (Li, et al., 2011; Hameed, et al., 2012b; Nam, et al., 2015). Furthermore, Iacovou et al. (1995) are among several researchers to test a multidimensional factor of organizational readiness; referring to the degree to which an organization has the awareness, resources, commitment, and governance to adopt IT. Some studies have also expanded the organizational context to include organizations' level of formalization and centralization, resource availability and slack, the existence of a product/project champion, as well as organizational culture (Hameed, et al., 2012b; Puklavec, et al., 2014).

Lastly, the **environmental context** represents the external, or interorganizational, environment in which the organization conducts its business (Tornatzky & Fleischer, 1990). The literature often defines this arena in which organizations conduct business in terms of the external pressure, external support, and regulatory environment that they are subjected to (Appendix C.3). While some studies examine competitive pressure (Premkumar & Roberts, 1999; Boonsiritomachai, 2014), industry and market pressure (Al-Isma'ili, et al., 2016), and partner pressure (Chwelos, et al., 2001) as separate measures, others investigate a combined factor under the umbrella term external pressure (Grandon & Pearson, 2004). Iacovou et al.

(1995) defined external pressure as any influence from a firm's competitors, industry, or trading partners, on the organization's adoption-decision. Secondly, the support infrastructure for technology is also understood to impact adoption. Puklavec et al. (2014) considered providers of innovations as a unique group of partners that offer external support. Here, external support refers to the "availability of support for implementing and using an innovation" (Puklavec, et al., 2014, p. 189). Premkumar and Roberts (1999) argue that organizations are more willing to risk adopting new technologies when there is adequate support for the technology. Lastly, the third factor pertaining to the environmental context relates to the regulatory environment. Ifinedo (2011) studied whether government support, the assistance provided by the authority, encouraged adoption of IT innovations. Using the same factor, others have tested how the adequacy of institutional frameworks and business laws governing the use of innovations affect adoption (Zhu, et al., 2004; Nam, et al., 2015).

In most of the studies employing the TOE framework, researchers have treated factors of the technological, organizational, and environmental context as direct antecedents of adoption and implementation. Although researchers seemingly agree with Tornatzky and Fleischer (1990) that the three TOE contexts influence adoption, there is seemingly no universal set of factors for each technology or context that is being studied.

3.3 Model and Hypothesis Development

As presented, theories from innovation adoption literature offer different, yet complementary models for adoption of IT in organizations. DOI and TAM have a solid theoretical foundation and consistent empirical support (Premkumar & Roberts, 1999; Venkatesh, et al., 2007). Nonetheless, despite being extensively used in the study of innovation adoption, Hameed et al. (2012a) found that DOI and TAM were utilized mainly for individual level of analysis. In particular, DOI has received substantial criticism in its applications at an organizational level (Chau & Tam, 1997). Having originally been developed to study innovation adoption by individuals, DOI and TAM share a common limitation: They exclude the possibility of influences from intraorganizational and interorganizational factors (Lee & Cheung, 2004). Consequently, while DOI and TAM are generally applicable to IT adoption by autonomous individuals, modifications and extensions are needed to account for external influences in more complicated adoption scenarios, such as by organizations.

A common approach to study IT adoption at the organizational level is by integrating DOI and TAM with a contextual framework that covers the organizational and environmental antecedent factors (Hameed, et al., 2012a). This paper applies a multiple perspective approach through the inclusion of the TOE framework, intended to compensate for the limitations of DOI and TAM by including organizational and environmental factors believed to affect adoption. Hence, this paper develops a theoretical model for adoption of Big Data technology by combining DOI and TAM with the TOE framework. This makes it possible to bring the analysis of Big Data adoption by firms in Norway to an organizational level. The remainder of this chapter is dedicated to discussing the concepts and constructs from the three leading perspectives of IT innovation adoption in order to establish this paper's research model. The tentative research model is illustrated in Figure 6.

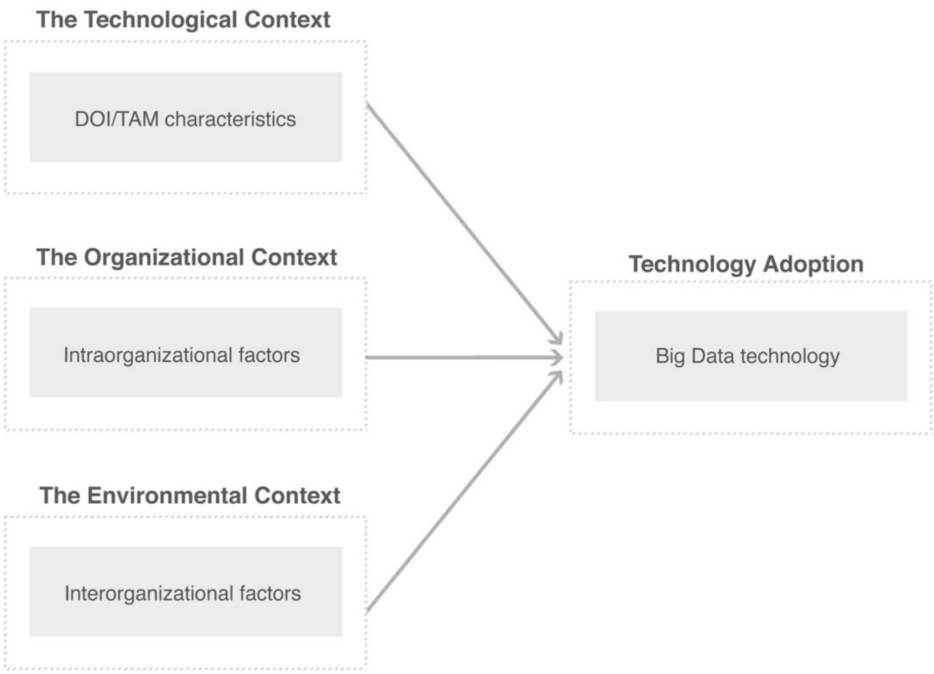


Figure 6: Tentative research model

3.3.1 The Technological Context

A review of IT adoption literature indicates that characteristics of innovations are the main focus of most IT adoption studies (Hameed, et al., 2012a; Puklavec, et al., 2014). Both DOI and TAM share the same premise that potential adopters evaluate an innovation based on their perception of its characteristics, and postulate that innovations with more favourable characteristics are more likely to be adopted (Rogers, 1983; Davis, 1986). Value-oriented characteristics such as perceived usefulness (Davis, 1989) and relative advantage (Rogers, 1983), effort-oriented characteristics such as perceived ease of use (Davis, 1989) and complexity (Rogers, 1983), in addition to compatibility (Rogers, 1983) have repeatedly been found as major factors influencing adoption of innovations (Li, et al., 2011). According to Tornatzky and Klein (1982), who conducted a meta-analysis of 75 innovation adoption studies to identify salient innovation characteristics, only relative advantage, complexity, and compatibility were recommended as consistently correlated with the adoption of an innovation. This study places these characteristics of the innovation – as described by DOI and TAM – within the technological context.

Relative Advantage and Perceived Usefulness

Both Rogers' relative advantage (Rogers, 1983) and Davis' usefulness (1989) are the most consistently relevant constructs from DOI and TAM, and have found strong support in IT adoption studies as one of the top predictors of adoption (Plouffe, et al., 2001; Venkatesh, et al., 2003). Though both concepts appear relevant for this study, researchers have acknowledged their similarities (Davis, et al., 1989; Moore & Benbasat, 1991; Plouffe, et al., 2001; Kulviwat, et al., 2007). While usefulness and relative advantage are conceptually distinct (Kulviwat, et al., 2007, p. 1065), they are commonly operationalized using nearly identical multi-item scales (Davis, 1989; Plouffe, et al., 2001; Ifinedo, 2011). Additionally, when comparing usefulness and relative advantage, Kulviwat et al. (2007) found a strong relationship between the two, suggesting that the constructs cover very similar concepts[13]. Thus, as the two concepts are considered highly empirically and conceptually related, this thesis includes only one construct.

Whereas usefulness "reflects the belief that an innovation helps perform a function" (Kulviwat, et al., 2007, p. 1065), relative advantage is the "degree to which an innovation is perceived to be *better* than its precursor" (Kulviwat, et al., 2007, p. 1065). As Moore and Benbasat (1991) argue, an innovation is usually developed with a particular purpose in mind, and must be perceived to fulfil this purpose *better than* its precursors if it is to be adopted. The concept of relative advantage thus has an intuitive appeal, as it is a highly generalizable concept. Moreover, although Davis (1989) does not explicitly use the term "relative", his definition of usefulness is in relative terms; the value of an innovation lies in its ability to perform a function *better* than the alternative. For these reasons, given the intuitive appeal of Rogers' (1983) construct and its common occurrence in IT adoption literature, the term relative advantage is used in this thesis.

---

[13] Kulviwat et al. (2007, p. 1072) found a correlation of 0.96 between usefulness and relative advantage.

Innovation adoption in an organizational context is almost exclusively associated with utilitarian information systems (Heijden, 2004). A utilitarian system, as opposed to a hedonic system, aims to provide instrumental value to the adopter. Instrumentality implies there are objectives external to the interaction between the adopter and the innovation, such as improving effectiveness, efficiency, and productivity. Accordingly, the definition of relative advantage in a utilitarian perspective draws attention to benefits external to the innovation adopter interaction, namely improving performance (Heijden, 2004).

The relative advantage of an innovation is postulated in DOI to positively influence adoption (Rogers, 1983). A rational adoption-decision by an organization, according to Premkumar and Roberts (1999), naturally involves evaluating the performance impacts of the new technology. The impact of relative advantage on adoption-decisions have repeatedly been shown in innovation adoption literature (e.g., Tornatzky & Klein, 1982; Ifinedo, 2011). Thus, it is hypothesized that the greater the perceived relative advantage of Big Data technology, the more likely it is to be adopted:

H_1: Perceived relative advantage is positively associated with adoption of Big Data technology

Complexity and Perceived Ease of Use

Effort-oriented constructs such as complexity and perceived ease of use have been widely studied and found to be factors influencing IT adoption in studies based on DOI and TAM (Venkatesh & Davis, 1996; Arts, et al., 2011). However, prior research has noted substantial similarities between the construct definitions and their operationalization (Venkatesh, et al., 2003). Indeed, perceived ease of use in TAM and complexity in DOI are so alike that they have been seen as pair-wise opposites (Moore & Benbasat, 1991; Plouffe, et al., 2001; Puklavec, et al., 2014). Consistent with Rogers (1983), both constructs can be understood to represent opposing ends on a simplicity-complexity continuum on which an innovation can be classified.

Complexity in the IT adoption literature bears a negative association in the sense that it represents hurdles to be overcome by potential adopters (Venkatesh, et al., 2003). The greater the perceived complexity of an IT innovation, the higher the cost of adoption and subsequent behavioural change becomes, inhibiting potential adopters from following through with adoption (Arts, et al., 2011). Hence, as complexity more intuitively represents the idea of an inhibiting factor, this thesis covers this construct using the DOI nomenclature.

The complexity of an innovation is determined by the degree to which it is perceived as difficult to understand and use, and is postulated to be negatively related to innovation adoption and implementation (Rogers, 1983). Researchers have repeatedly found that complexity functions as a barrier to innovation adoption (e.g., Tornatzky & Klein, 1982; Grandon & Pearson, 2004). Thus, it is hypothesized that the greater the perceived complexity of Big Data technology, the less likely it is to be adopted:

H_2: Perceived complexity is negatively associated with adoption of Big Data technology

Perceived Compatibility

The similarities between Rogers' (1983) relative advantage and complexity, and Davis' (1989) perceived usefulness and ease of use are clear. However, the DOI includes three additional characteristics of innovations that are not found in TAM. With regard to compatibility, DOI presents it as a measure of the degree to which it is perceived as being consistent with the existing values, past experiences, and needs of the potential adopter (Rogers, 1983; Premkumar & Roberts, 1999). This conceptualization of compatibility has garnered considerable critique, as the inclusion of *needs* in the original definition is considered a source of confounding with relative advantage (Moore & Benbasat, 1991). Thus, studies have reported difficulties in distinguishing between the two constructs (Tornatzky & Klein, 1982; Moore & Benbasat, 1991; Plouffe, et al., 2001; Arts, et al., 2011). This has led some researchers to treat compatibility and relative advantage as a single construct (Taylor & Todd, 1995; Li, et al., 2011). Li et al. (2011), for instance, argued that incompatibility would be reflected in the adopters' perception of whether adopting the innovation would be advantageous to the firm, and consequently encapsulated compatibility within relative advantage. Similarly, when compatibility and relative advantage loaded together in Taylor and Todd's (1995) statistical analysis, they opted to combine the two. Moreover, as the constructs have been found to be perceived as identical by respondents (Moore & Benbasat, 1991), researchers have been called to reconceptualise compatibility in innovation adoption studies (Karahanna, et al., 2006).

This thesis regard compatibility as a multidimensional construct entailing normative or cognitive compatibility (i.e., what adopters feel or think about an innovation) and operational or practical compatibility (i.e., compatibility with what adopters do) (Tornatzky & Klein, 1982), with emphasis on the latter. While the former refers to compatibility with values or norms, operational compatibility represents congruence with the existing practices or processes of the adopter (Tornatzky & Klein, 1982). This second interpretation closely resembles interpretations of compatibility in more recent research, which regard compatibility in terms of the degree to which an innovation is perceived by the firm as being consistent with existing methods for executing their mission (Ifinedo, 2011; Puklavec, et al., 2014).

Research on compatibility as an antecedent of IT adoption has produced varying results, and the inclusion of the construct separately from relative advantage has been questioned. However, due to the clear conceptual difference between the two, others argue that the concepts should be included separately (Moore & Benbasat, 1991). Moreover, in their meta-analysis on innovation characteristics, Tornatzky and Klein (1982) found compatibility to be among the few consistent antecedents of innovation adoption. Similarly, Plouffe et al. (2001) found that removing the compatibility construct from PCI[14] – an extended version of DOI – negatively impacted their prediction of adoption. Hence, compatibility from DOI is included in this thesis.

As an attribute of innovations, compatibility has been postulated in DOI to have a positive influence on adoption (Rogers, 1983). Empirical support to corroborate this relationship has previously been found (e.g., Tornatzky & Klein, 1982; Thong, 1999; Arts, et al., 2011). The acquisition and use of Big Data technology can bring in significant changes to the practices and processes of businesses, and resistance to change is expected (Premkumar & Roberts, 1999). With regard to the normative aspect of compatibility, it is believed that a business' values and beliefs must be compatible with any changes following the adoption of an innovation. Thus, it is hypothesized that the greater the perceived compatibility of Big Data technology, the more likely it is to be adopted:

H_3: Perceived compatibility is positively associated with adoption of Big Data technology

Perceived Observability and Trialability

Lastly, the DOI proposes the inclusion of the observability and trialability constructs. While the theoretical arguments for the inclusion of these constructs are sound (e.g., observability and trialability reduces uncertainty for the adopter) (Rogers, 1983), Tornatzky

[14] The Perceived Characteristics of Innovations (PCI) model utilized in some research (e.g., Plouffe, et al., 2001), is an alternative name for the DOI model used within this research.

and Klein (1982) found the two constructs not to be consistently related to adoption. Furthermore, observability and trialability are rarely utilized in IT adoption research at the organizational level (e.g., Puklavec, et al., 2014). The literature offers two potential explanations for this. First, Rogers (1983) suggests that technology innovations have two components; a hardware aspect and a software aspect. Software dominant innovations, according to Rogers (1983, p. 232), have less observability. Accordingly, Big Data technology, as a software dominant innovation, is less apparent to observations. Secondly, with regards to trialability, Moore and Benbasat (1991) found that it had significantly less weight than other constructs in an organizational context. For these reasons, in accordance with the recommendations of Tornatzky and Klein (1982) to eliminate redundant innovation characteristics, this thesis does not investigate these constructs further.

3.3.2 The Organizational Context

The organizational context refers to the group of intraorganizational factors that are believed to influence adoption. A review of innovation adoption literature has been performed, in an effort to avoid the inclusion of superfluous constructs.

Organizational Size

There has long been an interest in the effect of the size of a firm on various aspects of business activity (Daniel & Grimshaw, 2002), particularly in the study of organizational innovation adoption (see Damanpour (1992) and Hameed et al., (2012b) for meta-analyses). A multiplicity of opinions exists as to the role that organizational size plays in the process of innovation adoption (Zhu, et al., 2004). On one hand, the size of an organization is believed to affect a number of organizational aspects, including resource availability, decision-making, and organizational structure (Rogers, 1995; Hameed, et al., 2012b). Smaller firms are believed to suffer from resource poverty[15] (Welsh & White, 1981), which may constrain experimentation, acquisition, and implementation of new innovations (Premkumar & Roberts, 1999). Research also indicate that small firms are uniquely characterized by financial constraints, lack of professional IT expertise, and short-term management perspectives (Welsh & White, 1981; Al-Isma'ili, et al., 2016), all of which frequently occur as factors in IT adoption

[15] Resource poverty was introduced by Welsh and White (1981) to describe that the size of small businesses creates a special condition – resource poverty – which distinguishes small firms from their larger counterparts.

research (e.g., Hameed, et al. 2012b). By contrast, larger firms possess greater resource availability and slack, enabling the mobilization of adequate resources for adoption (Tornatzky & Fleischer, 1990; Zhu, et al., 2004). On the other hand, larger firms are less agile and may suffer from inertia, affecting adoption of new innovations (Hannan & Freeman, 1984). Smaller firms, by contrast, are believed to require less communication, less coordination, and less influence to gather support for adoption (Zhu, et al., 2004).

Despite the frequent occurrence of organizational size as a factor in the study of organizational innovation adoption, empirical support for the effect of size has been inconclusive[16] (Hameed, et al., 2012b). Nevertheless, most research postulates a positive relationship between organizational size and IT innovation adoption (e.g., Premkumar & Roberts, 1999), including research on Business Intelligence adoption (Malladi & Krishnan, 2013; Hung, et al., 2016). Given the lack of consistency in research on the effect of size on IT adoption, this is clearly an area in which this thesis might improve our understanding. Thus, it is hypothesized that the greater the size of the organization, the more likely Big Data technology is to be adopted:

H_4: Organizational size is positively associated with adoption of Big Data technology

Top Management Support

Top management support has been a recurring critical factor in IS adoption research (Thong, et al., 1996), and is believed to play a crucial role in all stages of innovation adoption (Hameed, et al., 2012b). Literature suggests that top management support, defined herein as the degree to which top management understands the importance of Big Data technology and the extent to which it is involved in related initiatives (Park, et al., 2015), is an essential criterion for organizational innovation adoption for two primary reasons. Firstly, studies have found that top management support is critical for creating a conducive environment for innovation adoption (Premkumar & Roberts, 1999). By virtue of their leadership role, top management are responsible for creating a supportive organizational climate that facilitates receptivity (Thong, et al., 1996). Secondly, top management possess the authority to provide and mobilize sufficient organizational resources for motivating, acquiring, and implementing innovations (Premkumar & Roberts, 1999). The vital role of top management in allocating

[16] The meta-analysis performed by Hameed et al. (2012b) found that organizational size only had a weak significance to IT adoption.

resources for adoption and enabling associated activities has repeatedly been emphasized by researchers (e.g., Hung, et al., 2016; Al-Isma'ili, et al., 2016). Additionally, a review of Big Data literature and our preliminary interview with Business Intelligence professionals[17], highlighted a third reason for the significance of top management support: As the benefits of Big Data to an adopting organization are highly contextual, the *business case* for adoption is often ambiguous. Given the importance of the aforementioned construct of relative advantage, it is unlikely that adoption will take place before the strategic value of Big Data technology is recognized by top management. Hence, a third reason as to why top management support plays a crucial role in Big Data adoption is proposed.

Findings from previous research suggest that top management support is positively related to the adoption of new technologies in small and large organizations across a range of IS innovations (Thong, et al., 1996; Premkumar & Roberts, 1999; Hameed, et al., 2012b). Thus, it is hypothesized that the greater the top management support for Big Data technology, the more likely it is to be adopted:

H_{5}: Top management support is positively associated with adoption of Big Data technology

IT Expertise

It has been suggested that a highly skilled, knowledgeable, and experienced workforce is a key factor affecting the adoption of IT and innovations (Ettlie, 1990; Lucchetti & Sterlacchini, 2004). IT expertise, the experience of IT employees in terms of skill and knowledge (Hameed, et al., 2012b), has been widely studied in adoption literature under highly similar terms such as IS competence (Nam, et al., 2015), IS knowledge (Thong, 1999), IT competence (Zhu, et al., 2006b), employee skill (Meyer & Goes, 1988), and technology readiness (Ifinedo, 2011). A relatively recent meta-analysis of the relationship between organizational characteristics and IT innovation adoption in organizations identified IT expertise as one of the major factors facilitating adoption (Hameed, et al., 2012b). Hence, empirical support for the positive influence of IT expertise on IT adoption across a range of innovations has been found, including in the context of Big Data adoption (Nam, et al., 2015).

[17] Norwegian interview transcripts are available per request.

Inadequate levels of IT expertise have been shown to be a barrier to IT adoption (Hameed, et al., 2012b). In particular, smaller businesses have historically been found lacking in specialized IT knowledge and technical skills (Thong, 1999). Premkumar and Roberts (1999) suggest that such firms may be unaware of new technologies or may not want to risk adoption, possibly due to an inability to integrate the innovation in a way that resolves work-related problems (Yeh, et al., 2014). It is also proposed that firms may be tempted to postpone adoption until they have sufficient internal expertise (Thong, 1999). Conversely, firms that possess higher levels of IT expertise are more likely to accept innovations as they have been found to have a better understanding of the potential benefits arising from adoption (Ifinedo, 2011). Accordingly, our literature review shows that research postulates a positive relationship between IT expertise and organizational technology adoption. Thus, it is hypothesized that the greater the internal IT expertise, the more likely Big Data technology is to be adopted:

H_6: IT expertise is positively associated with adoption of Big Data technology

Organizational Resources

The final organizational attribute included in this thesis is resources, which refers both to organizational resource availability (Boonsiritomachai, 2014) and resource slack (Li, et al., 2011). IS literature suggests that the intention of organizations to adopt new technology is affected by the availability of financial, technological, and human resources (Hameed, et al., 2012b). Financial resources refer to the availability of capital for investment in technology innovations, for implementation of subsequent changes, and coverage of ongoing expenses during usage (Iacovou, et al., 1995). With regard to Big Data, Nam et al. (2015) found that the availability of financial resources had a significant influence on adoption by Korean firms. Technological resources refer to the level of IT sophistication, in terms of IT usage and management, as well as the IT infrastructure installed in the organization (Hameed, et al., 2012b). Iacovou et al. (1995) state that highly sophisticated organizations (in terms of IT) are "less likely to feel intimidated by new technology, possess a superior corporate view of data as an integral part of overall information management, and have access to the required technological resources" (p. 469). Lastly, human resources refer to the existing IT knowledge within an organization. The influence of human resources on adoption has previously been explored in the discussion of IT expertise.

Managers, and by extension, organizations, are believed to be more supportive of the adoption of new technology when financial, technological, and human resources are available

(Chong, et al., 2009). However, resource availability may not be sufficient for an organization to adopt Big Data technology. Li et al. (2011) postulate that organizations need resource slack. Resource slack refers to those resources an organization has acquired which are not committed to an existing business operation, and subsequently can be used in a discretionary manner (Dimick & Murray, 1978). Previous research has suggested that slack positively impacts willingness to adopt, as it enables organizations to act more boldly (Bourgeois, 1981; Singh, 1986). Similarly, slack may encourage risk-taking as excess resources allow the organization to absorb costs associated with potential failure (Singh, 1986; Li, et al., 2011).

Although the resource construct has repeatedly been studied in IS adoption research, data to support the relationship between resources and organizational adoption has been inconclusive (Hameed, et al., 2012b). Yet, the majority of IS literature argue for a positive relationship between organizational resources and innovation adoption (e.g., Boonsiritomachai, 2014; Nam, et al., 2015). Given the lack of empirical support for the effect of resources on IT adoption, this is another area in which this thesis might improve our understanding. Thus, it is hypothesized that the greater the organizational resources, the more likely Big Data technology is to be adopted:

H_7: Organizational resources is positively associated with adoption of Big Data technology

3.3.3 The Environmental Context

The environmental context refers to the group of interorganizational factors that are believed to influence adoption. A review of innovation adoption literature has been performed, to ensure the inclusion of the most relevant and recurring constructs.

Competitive Pressure

Economists have long since recognized the strategic significance of IT. Porter and Millar (1985) analysed the strategic rationale underlying the relationship between competitive pressure and IT innovations (Zhu, et al., 2004). They suggested that adoption of IT by businesses alter the competitive environment in three crucial ways. Firstly, IT innovations have the potential to change industry structure, and in doing so, altering the rules of competition (Porter & Millar, 1985). This may lead to environmental uncertainty and more intense competition, which are believed to increase both the need for and the rate of innovation adoption (Thong, 1999). Secondly, IT innovations have a strong effect on competitive advantage in terms of both cost and differentiation, giving companies new ways to outperform their rivals (Porter & Millar, 1985). Competitive pressure may demand innovation adoption in order to maintain a market position established on the basis of a competitive advantage whose sustainability is threatened by novel IT. Such a view is consistent with the resource-based view (Peteraf, 1993) and the activity-based view (Porter, 1996), both well-known perspectives on firm-level competitive advantage in the field of strategy and management. Thirdly, IT innovations contribute to the emergence of entirely new value offerings and businesses, which may intensify competition (Porter & Millar, 1985). Chau and Tam (1997) found that more intense competition is associated with higher IT use and innovation adoption. Moreover, their findings suggest that organizations tend to have a reactive rather than proactive attitude towards adopting IT; that is, satisfaction with the current state leads to lower incentives to adopt. It follows that more intense competition may provide incentives for IT adoption.

For these reasons, it is believed that competitive pressure accelerates innovation adoption as firms seek to leverage new IT, not only to survive, but also to outperform competitors. Previous research on organizational IT adoption has recognized competitive pressure as an important antecedent of adoption (Premkumar & Ramamurthy, 1995; Premkumar & Roberts, 1999; Iacovou, et al., 1995; Grandon & Pearson, 2004). Furthermore, all of the reviewed studies postulated a positive relationship between competitive pressure and organizational IT adoption. Thus, it is hypothesized that the greater the competitive pressure, the more likely Big Data technology is to be adopted:

H_8: Competitive pressure is positively associated with adoption of Big Data technology

External Support

Research on IT adoption by organizations shows that external support, defined as the “availability of support for implementing and using an information system” (Premkumar & Roberts, 1999, p. 472), is a key factor in the IT adoption process (Puklavec, et al., 2014). While the availability of external support is hypothesized to increase a firm’s willingness to try novel technologies (Premkumar & Roberts, 1999), it is also believed to serve as a means by which organizations may compensate for a lack of internal IT expertise (Thong, 1999). Thong et al. (1996) found that in the absence of internal IT expertise, firms tend to seek the support of consultants and vendors. In other words, earlier research has seen external support as a source of external IT expertise (Thong, et al., 1996; Caldeira, 1998; Thong, 1999). However, our literature review suggests that the majority of adoption studies utilizing external support as a research variable has done so in the context of small and medium-sized enterprises (e.g., Caldeira, 1998). Less attention appears to have been given to the construct in adoption studies of large organizations. Thus, this is another area in which this thesis might improve our understanding.

Findings from previous studies suggest a positive relationship between external support and the adoption of new IT (Premkumar & Roberts, 1999; Al-Isma'ili, et al., 2016; Hung, et al., 2016). It is therefore hypothesized that the greater the availability of external support, the more likely Big Data technology is to be adopted:

H_9: External support is positively associated with adoption of Big Data technology

3.3.4 Model Extensions

The literature on innovation adoption has been described as fragmentary, contradictory, and even beyond interpretation (Meyer & Goes, 1988). The absence of a unified theory that permits researchers to predict the extent to which an organization will employ a given innovation, has produced a body of research that is arguably less than the sum of its parts (Meyer & Goes, 1988, p. 897). The result is a research literature offering a great number of competing and complementary theories and models of adoption. Fichman (1999) contends that the absence of a general theory of innovation suggests researchers should develop models tailored to the specific innovation. Even so, some research variables and relationships are more broadly generalizable. Therefore, the research model proposed in this thesis is primarily based upon more generalizable variables and relationships. However, extensions are believed necessary to cover potentially important aspects that are distinctive to adoption of Big Data technology.

Chapter 2.2 highlighted security and privacy issues as distinctive barriers to organizational adoption of Big Data. Multiple Big Data surveys have also found that security and privacy issues are among the hindering factors of adoption (e.g., Heudecker & Lisa, 2014; Filkins, 2015). In their literature review on Big Data’s security and privacy related concerns, Salleh and Janczewski (2016) discussed how these concerns may affect Big Data adoption by organizations. This thesis will attempt to supplement this discussion by integrating security and privacy related issues of Big Data into the proposed research model and subsequently test hypotheses for their relation to adoption of Big Data technology. Consequently, this section will discuss the relevance of security and privacy issues for organizational adoption within the technological and environmental context of the TOE framework.

Privacy Regulations in the Environmental Context

The emergence of Big Data raises important privacy issues, some of which are discussed in Chapter 2.2. With falling costs of storage and increasing demand for retention of demographic, behavioural, financial, and other transactional data for analytic purposes, firms are faced with the challenge of upholding privacy. From a legal perspective, firms are pressured to ensure compliance with data protection acts, rules and regulations, while simultaneously preserving data utility, that is, the value of their data. Salleh and Janczewski (2016) postulate that one environment-related factor of organizational adoption of Big Data is the issue of privacy and its associated rules and regulations. With the introduction of EU's General Data Protection Regulation, set to replace the current Data Protection Act in Norway in 2018, firms will face an even greater obligation to protect personal identifiable information (Datatilsynet, 2015). In particular, firms working with Big Data are under pressure to deliver on legal expectations without compromising their business goal (Salleh & Janczewski, 2016).

To the best of our knowledge, few researchers have empirically investigated the relation between privacy and adoption of Big Data technology. Considering the attention privacy has been given in Big Data publications, it appears worthy of further investigation. Thus, this thesis includes Big Data *privacy* as an environmental factor. According to Zhu et al. (2004), the regulatory environment, a factor closely resembling the privacy construct proposed herein, has previously been recognized as a critical factor in diffusion research based on the TOE framework. Furthermore, a recent study on Big Data adoption by Nam et al. (2015) found support for a relationship between privacy regulations[18] and Big Data adoption. As such, it is hypothesized that the greater the perceived difficulty of complying with privacy related regulations, the less likely Big Data technology is to be adopted:

H_{10}: Privacy regulations are negatively associated with adoption of Big Data technology

Security Issues in the Technological Context

The technology-related security issues associated with Big Data's unique characteristics are frequently found in Big Data publications (e.g., Alshboul, et al., 2015, Salleh & Janczewski, 2016). According to Salleh and Janczewski (2016), the volume, velocity,

---

[18] Nam et al. (2015, p. 4796) used the term *government pressure* to study the effect of "privacy issues regarding collecting and managing personal data" for adoption of Big Data.

and variety of Big Data presents security concerns that necessitate the employment of newer and improved security solutions and mechanisms to ensure confidentiality, integrity, and availability of data. Their discussion of security can be summarized as follows: Firstly, the enormous volume of data collected and created in a typical Big Data environment invites new security issues that must be met with security technologies and solutions that are able to scale with the size of data sets and distributed nature of Big Data. Secondly, the rate at which data is generated and the speed of how it should be analysed and acted upon is believed to amplify security issues commonly found in traditional data environments. Lastly, the collection of structured, semi-structured, and unstructured data offers new challenges in terms of providing restrictions for access and security policies that fit each source of data (Salleh & Janczewski, 2016).

Based on the above, Salleh and Janczewski (2016) argue that adopters of Big Data will face distinctive security issues. Furthermore, they postulate that these issues will influence an organization's adoption of Big Data through perceived complexity and compatibility. This thesis intends to test this proposition by integrating Big Data security, as a multidimensional construct, in the research model. Thus, the following conceptualization of security is proposed: (1) the technological challenges posed by Big Data reflects the *complexity* in providing effective security, and (2) the level of preparedness of organizations in embracing the security challenges that comes with Big Data can be attributed to the organization's *compatibility* of their current security mechanisms with those required by Big Data (Salleh & Janczewski, 2016). It is therefore hypothesized that the greater the perceived security issues of Big Data technology, the less likely it is to be adopted:

H_{11}*: Security issues are negatively associated with adoption of Big Data technology*

3.3.5 Dependent Variable: Assimilation of Big Data Technology

Categorizing firms as adopters and non-adopters of Big Data is a daunting task, especially considering the ambiguity surrounding the term. In particular, due to the diverse and contradictory definitions in both academic and business literature, the basis for evaluating the Big Data capabilities of a firm can be somewhat difficult to justify. Fortunately, Mauro et al. (2016) proposed four main components that characterize the Big Data concept: information, technology, methods, and impacts (see Ch. 2.1.1). The **technology** component of Big Data, the focus of the present study, refers to hardware (e.g., storage and servers) and software (e.g., applications) that enable the accessing, managing, and analysing of data sets characterized by the three V's (volume, velocity, and variety); which is what we tend to associate with Big Data. As such, technology is deemed a prerequisite for using Big Data, thus offering an intuitive way of thinking about the Big Data capabilities of a firm. Logically, firms that have acquired Big Data technology possess the capability to use Big Data, whereas firms that have yet to acquire this technology do not possess this capability. Henceforth, firms are categorized as adopters and non-adopters based on their possession of Big Data technology.

The earlier discussion of the stages of innovation adoption shows that technology adoption by organizations is a stage-based process progressing in a sequential manner. The *ideal* organizational innovation adoption study, according to Tornatzky and Klein (1982), should fully account for this process, which has come to be known as **assimilation** (Meyer & Goes, 1988). Whereas diffusion is the process by which a technology spreads across a population of organizations, assimilation refers to the adoption process within organizations; stretching from initial awareness, to acquisition and deployment of the technology (Meyer & Goes, 1988). Hence, drawing upon the innovation diffusion literature (Rogers, 1983; Fichman, 1999), Big Data adoption is defined in terms of assimilation; the sequence of stages from (1) a firm's initial awareness and evaluation of Big Data technology, (2) to the formal allocation of resources for its acquisition and deployment, and finally (3) to its incorporation of the technology into the regular activities of the firm.

While some studies depict assimilation as a six-stage process (Cooper & Zmud, 1990; Fichman, 2001), others use a seven-stage model (Rai, et al., 2009; McKinnie, 2016). This thesis proposes an aggregated, three-stage model of Big Data adoption, where **initiation**, **adoption-decision**, and **implementation** represent three stages of assimilation. The present study will therefore use the term *adoption* generically in the context of organizations, and is inclusive of the three-stage assimilation process. This is consistent with Rogers' (1983) description of the adoption process, which has been adapted by a number of innovation adoption researchers (e.g., Zhu, et al., 2006; Nam, et al., 2015). Thus, this thesis studies the adoption of Big Data technology, modelled as a three-stage process defined as assimilation, to identify determinants of adoption and to reveal whether they have differential effects at the different stages.

3.3.6 Final Research Model

Based on the literature, theoretical arguments, and empirical support presented in this chapter, a research model is proposed for the study of organizational adoption of Big Data technology. The integrative model combines three key constructs from two of the most commonly applied theories in IS adoption research – DOI and TAM – with the organizational level technology adoption framework, known as the TOE framework, and specific extensions made relevant by research on security and privacy issues distinctive to Big Data.

The result is a research model that enables the investigation of factors in the context of technology, organization, and environment that affect the adoption of Big Data technology. The model includes 11 factors within the three contexts of the TOE framework: relative advantage, complexity, compatibility, organizational size, top management support, IT expertise, organizational resources, competitive pressure, external support, privacy, and security. Each factor is hypothesized to influence the adoption of Big Data technology, and have potentially differential effects at the different stages of adoption – initiation, adoption-decision, and implementation – collectively referred to as assimilation. The final research model is presented in Figure 7. In agreement with the TOE framework, constructs are presented according to their corresponding context.

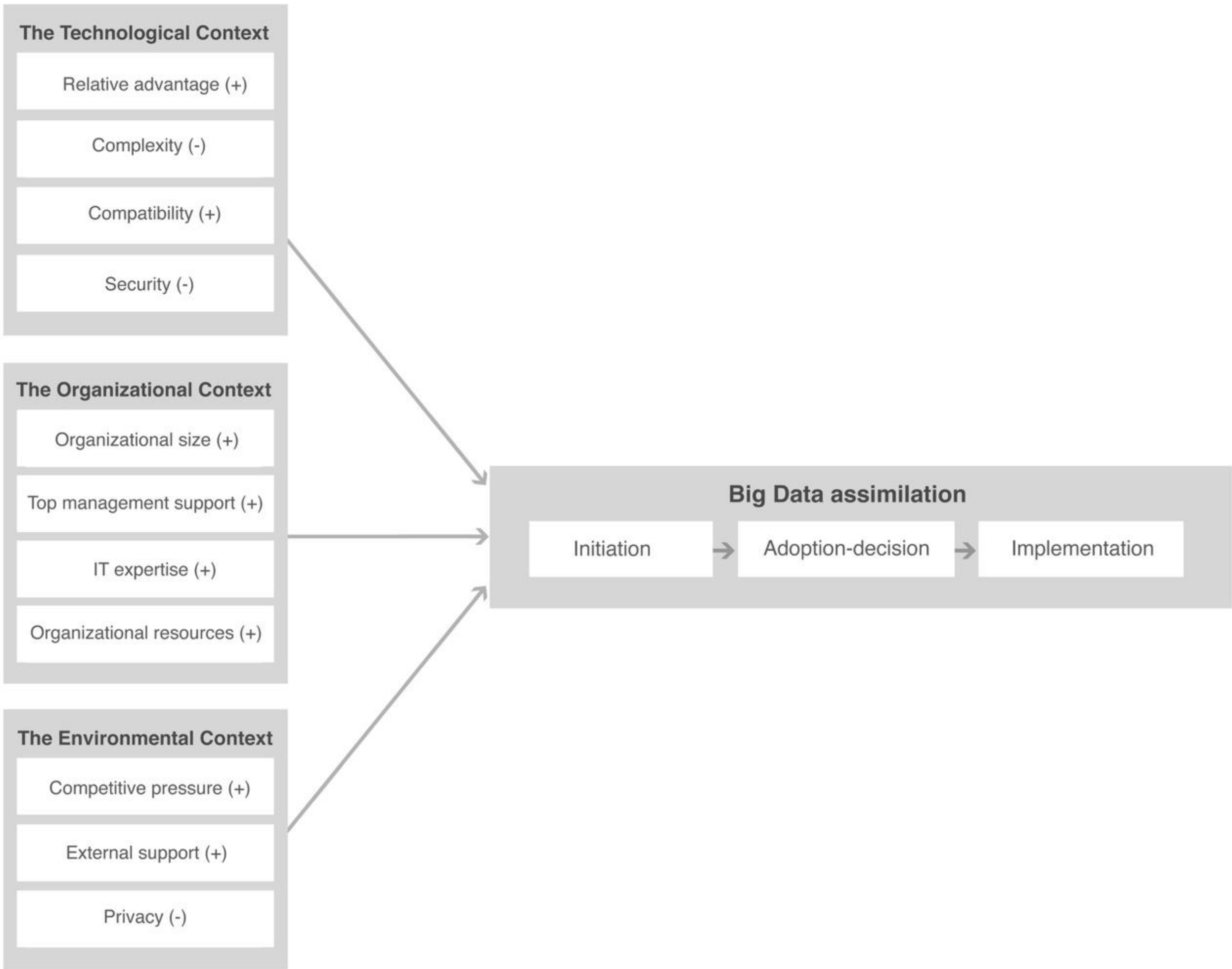


Figure 7: Research model

4. Methodology

The previous chapter developed a research model to investigate the determinants of organizational adoption of Big Data technology. This section details the methodology; the process by which the hypotheses derived from the research model were empirically tested and research questions answered. Methodology concerns the rationale for the procedures and techniques used to collect, analyse and interpret data, and is essential to any research as it allows the reader to critically evaluate the study's validity and reliability (Saunders, et al., 2012). While there is no ideal research methodology, steps have been taken to ensure the choice of methods is consistent with accepted practices in the chosen field of study. The choice of approach to pursue the thesis' research objectives is discussed as follows: A presentation of the research design is given (Ch. 4.1), followed by an overview of the sampling and data collection (Ch. 4.2), and finally, a discussion of non-response bias (4.3).

4.1 Research Design

The nature of the research questions dictates the design, which is the overall plan for answering the research questions. Exploratory, descriptive, or explanatory research designs are usually chosen depending on the purpose of the study and how much research currently exists on the topic (Saunders, et al., 2012).

The objective of this thesis is to identify and examine determinants of Big Data technology adoption by organizations. The thesis starts addressing the research questions by performing a review of literature related to technology and innovation adoption. Based on the presented literature, a research model was developed which allows the investigation of technological, organizational, and environmental factors that affect the organizational adoption of Big Data technology. Several hypotheses are then derived from the model and subsequently tested by collecting and analysing data. Research findings that are consistent with the hypotheses serve to validate our proposed relationships, whereas findings that are inconsistent with the hypotheses leads to rejection. This sequential research approach is consistent with *deductive reasoning* (Blaikie, 2010), also known as the *top-down* approach.

As technology adoption is extensively covered by existing research, the present study, for the most part, builds upon existing literature and pre-defined concepts. Hence, the study is primarily of a descriptive and conclusive nature, as it aims to describe and explain relationships between variables based on theoretically grounded expectations about how and why the variables should be related.

However, our review of adoption literature also identified significant research gaps. More specifically, few studies have explored adoption from an organizational unit of analysis, let alone in the context of Big Data. Additionally, security and privacy were found to be novel constructs yet to be explored in empirical research on Big Data adoption. Thus, as discussed in Chapter 4.1.3, entirely new measurement items were developed for these constructs. For these reasons, the thesis is also considered to be of an exploratory nature, as it seeks to validate existing theory – with new extensions – in a novel context.

Furthermore, this study is quantitative by nature, as the data collection generated numerical data which was subsequently subjected to the statistical techniques as discussed in Chapter 5. Since hypothesis testing in the social sciences requires statistical analysis, a quantitative approach was deemed the most appropriate. Quantitative research is typically associated with the deductive approach employed in this thesis (Saunders, et al., 2012). By contrast, a qualitative research design could have been used to provide a richer and more detailed insight into the organizational adoption of Big Data. However, for the specific purpose of testing pre-defined hypotheses regarding a wide range of theoretical constructs, the highly detailed and controlled procedures associated with quantitative research was preferred.

4.1.1 Research Strategy

Researchers can choose between a variety of research strategies, including experiments, surveys, archival research, case studies, ethnography, action research, and grounded theory (Saunders, et al., 2012). In research on technology adoption, surveys and case studies have been the dominant strategies at the individual and organizational level (Galliers & Land, 1987; Choudrie & Dwivedi, 2005).

The survey strategy was chosen for this thesis due to convenience and for parsimony reasons. As students at the Norwegian School of Economics, we were freely able to use Qualtrics, an internet-based survey software, for the design, distribution, and collection of survey data. This made it possible to reach a large and geographically disperse group of organizations, while at the same time collecting data about each individual respondent in an effective and inexpensive manner. Additionally, a self-administered survey has the benefit of allowing the respondents to answer anonymously and at their own convenience, and is therefore perceived to be less likely to contaminate or distort the respondent's answers (MacKenzie & Podsakoff, 2012; Saunders, et al., 2012).

Another reason for choosing the survey strategy is that it provides a better basis for generalizing, allow for replicability, and permit some degree of statistical power (Bouchard, 1993, p. 367). As noted by Tornatzky and Klein (1982), case studies are often insufficient to provide a basis for generalizing about adoption and innovation processes, whereas surveys are "methodologically adequate as they permit replicability and some degree cross-study comparability" (p. 29).

Furthermore, survey research is typically recommended when you have clearly defined independent and dependent variables and a specific model of the expected relationship which can be tested against observations of the phenomenon (Pinsonneault & Kraemer, 1993). By opting for the survey strategy, this thesis was capable of studying adoption of Big Data technology by organizations (1) in its natural setting, (2) in the present, and (3) without controlling or affecting the independent and dependent variables.

Lastly, a cross-sectional survey was naturally chosen for this study, given the time constraints inherent to writing a master thesis. In contrast to longitudinal studies, with the capacity to track change and development over time, a cross-sectional study provides data from a population at a single point in time (Saunders, et al., 2012). Such data is appropriate for describing the relationship between research variables and make predictions.

4.1.2 Questionnaire Design

A number of factors were considered in designing the survey to ensure the reliability of the responses, which are susceptible to biases and errors (MacKenzie & Podsakoff, 2012; Saunders, et al., 2012; Neuman, 2014; Planing, 2014).

A primary concern for this research was to motivate respondents to participate and respond accurately. Online surveys are generally plagued by low response rates, which is problematic when (1) attempting to generalizing findings, and (2) the statistical analysis employed requires a minimum sample size (see Ch. 5.3.3). Thus, a cover letter aimed at motivating participation was developed to accompany the survey invitation (see Appendix B). Addressing the respondents by their first and last name, the cover letter explained why participation was requested and how the responses would be used. Empirical evidence has demonstrated that respondents are more likely to participate and exert cognitive effort to answer questions accurately when motivation is high and the topic is perceived to be relevant

(MacKenzie & Podsakoff, 2012; Neuman, 2014). Therefore, in addition to stating the importance of participation for this master thesis, attempts were made to explain how participation could be useful for their organization. Furthermore, to solicit participation and encourage honest responses, respondents were given the opportunity to receive a summary of the research findings. Offering incentives for participation, such as rewards or feedback, can be a potential remedy for common method bias (MacKenzie & Podsakoff, 2012) and non-response bias (Simmons & Wilmot, 2004).

Regarding participant error, another concern was the respondents' perceptions of Big Data. Given inconsistent definitions of the term in both academic and business literature, respondents may possess vastly different understandings of the subject. In addition, the newness of Big Data and the lack of familiarity with the term may imply respondents have yet to understand and form opinions on the subject. The lack of experience in thinking about the survey subject may impair the respondents' abilities to answer questions, which could consequently lead to systematic errors (MacKenzie & Podsakoff, 2012). Thus, the definition of Big Data technology as employed in this thesis was presented in the beginning of the survey to ensure that the respondents and researchers shared a mutual understanding of the subject of the survey (see Appendix A).

Once committed to participate in the survey, several factors may cause response bias by decreasing the respondent's ability to answer accurately. Issues with the complexity of questions, item ambiguity, and double-barreled questions are common sources of method bias in survey research (MacKenzie & Podsakoff, 2012; Planing, 2014; Neuman, 2014). To avoid these issues, the conventional approach is to leverage existing scales that have already been validated in extant literature. Where applicable, measures used in this study are based on previously used and/or validated scales, with slight modifications to fit the research setting (see Ch. 4.1.3). These items are close-ended questions worded in a positive manner. To address the issue of common scale attributes (i.e., similarly worded questions), some studies shift between negatively and positively worded statements (MacKenzie & Podsakoff, 2012). However, to prevent confusion with reverse-worded items, and to avoid altering the content validity or conceptual meaning of scales, all adapted items were worded positively.

Designing the survey structure also involves a number of trade-offs with regard to potential sources of bias. Survey structure refers to the grouping of items, the order in which they appear, the number of items per page, and survey length. Regarding the grouping of items, intermixing of items from different constructs was avoided. While the practice of intermixing items has been recommended as a remedy for common method bias (Kline, et al., 2000), it may also increase the inter-construct correlations at the same time as decreasing the intra-construct correlations (Podsakoff, et al., 2003). As this thesis studies the relationship between conceptually distinct constructs, inter-construct correlation error was regarded as a greater concern than intra-construct correlation error. Hence, items of the same construct were placed successively. Due to the number of constructs in the study, questions were spread equally across four pages as to avoid survey fatigue common to longer surveys. A compromise was also sought between the number of items per construct and the total survey length, as long surveys are typically attributed lower response rates than short surveys (Saunders, et al., 2012). Thus, to limit response time, the number of items per construct was restricted to 3-5 items.

The items included in the survey were originally developed in English while the survey took place in Norway where the official language is Norwegian. Thus, as the majority of the respondents were native Norwegians, the survey was also translated from English to Norwegian using the *direct translation technique* suggested by Usunier (1998), cited in Saunders et al. (2012, p. 442). The English and Norwegian versions of the survey were developed concurrently over several iterations, with feedback and adjustments from two professionals from academia, including our supervisor, two associate professionals from the IT industry, and one manufacturer from the sample group. The *pre-testing* provided feedback which led to several adjustments. In particular, many phrases and wordings were changed to be easier to comprehend. The feedback also provided an opportunity to check the face validity[19] of the survey, and thus indications of which phrases were poorly translated and needed revision.

Furthermore, two interviews were conducted with three professionals working respectively with data warehousing, business intelligence, and Big Data to elaborate on the definition of Big Data[20]. These interviews highlighted that asking respondents to categorize themselves as adopters or non-adopters of Big Data based on a dichotomous question (i.e., “yes” or “no”) was problematic for two important reasons. First, the ambiguous nature of Big Data may imply that respondents are unaware of what constitutes adoption – specifically in the context of this research – and are therefore unable to accurately categorize themselves. Thus, the operationalization of the dependent variable, namely Big Data adoption, was

[19] Face validity refers to whether the items “appear logically to reflect accurately what it is intended to measure” (Saunders, et al., 2012, p. 671).

[20] Norwegian interview transcripts are available per request.

changed from a dichotomous variable to a 7-point scale representing the complete, multi-stage adoption process (also referred to as assimilation). Second, clear definitions of Big Data and Big Data technology were needed to provide the respondents with a basis for answering questions accurately. As such, two steps were taken to ensure the respondents and researchers shared a common understanding of the terminology in the survey. Firstly, a more extensive definition was provided in the introductory text. Secondly, a distinction was made between Big Data and Big Data technology, to enable respondents to answer questions more accurately. The survey can be seen in its entirety in Appendix A.

4.1.3 Operationalization of Constructs

With the exception of organizational size, each of the constructs developed in Chapter 3.3 represent latent, unobservable variables that cannot be measured directly. It is necessary to operationalize these constructs into theoretically meaningful and measurable variables (also called measurement items). To ensure consistency with previous research, the conventional approach is to leverage existing scales.

Most of the constructs employed in this study have measurement items that are proposed and/or validated in IS and IT adoption research. However, to ensure these items were appropriate for the context of this study, some adjustments were needed. Specifically, three measures for **relative advantage** were taken from Davis (1989) and Moore and Benbasat (1991), but reworded to fit the organizational unit of analysis. These items captured the utilitarian nature of the construct, as they measured the perceived instrumental value of adopting Big Data technology in terms of productivity, effectiveness, and performance. Additionally, one item was adopted from Chaveesuk (2010) to reflect the technology's role as strategic decision-aid. Concerning **compatibility**, three items representing the multidimensionality of the construct (cognitive, operational, and system compatibility) were adapted from Chaveesuk (2010). A fourth item was developed to cover data compatibility, which was believed to be another important dimension of the construct ("this technology is compatible with data captured at my company"). The measurement items for complexity, top management support, IT expertise, organizational resources, competitive pressure, and external support were also adapted from past research, but only with minor adjustments to fit the context of this study.

Table 2: Final measurement items

Construct	Item		Adapted from
Relative advantage	RA1	This technology improves my company's performance	Davis (1989) Moore and Benbasat (1991) Chaveesuk (2010)
	RA2	This technology improves my company's productivity	
	RA3	This technology improves the effectiveness of my company's operations	
	RA4	This technology provides my company with valuable information for decision making	
Complexity (R)	CX1	My company finds it easy to get this technology to do what we want it to do	Davis (1989)
	CX2	My company's interaction with this technology is clear and understandable	
	CX3	My company finds this technology easy to use	
	CX4	It is easy for my company to become skillful at using this technology	
Compatibility	CM1*	This technology is compatible with the data captured at my company	Chaveesuk (2010)
	CM2	This technology fits well with my company's existing operating practices	
	CM3	This technology is compatible with my company's IT infrastructure	
	CM4	Using this technology is consistent with my company's values and beliefs	
Security (R)	SE1*	The skills required to ensure data security when using this technology are easy for my company	
	SE2*	It is easy for my company to integrate security policies for this technology	
	SE3*	My company has adequate tools and mechanisms to provide effective data-protection when using this technology	
	SE4*	My company has security capabilities to adopt this technology	
	SE5*	My company has security policies that suits the different types of data in the company when using this technology	
Organizational size		What is your company size, by employees? Less than 50/ 50-100 / 101-150 / 151-250 / 251-400 / More than 400	
Top management support	MS1	Top management believe that investment and expenditure in this technology is worthwhile	Yeh et al. (2014) Premkumar and Roberts (1999)
	MS2	Top management believe that this technology has potential strategic value	
	MS3	Top management support is important to provide the resources for my company to adopt this technology	
IT expertise	IE1	Our IT employees have equal or better technical knowledge than our competitors	Ravichandran and Lertwongsatien (2005)
	IE2	Our IT employees have the ability to quickly learn and apply new information technologies	
	IE3	Our IT employees have the skills and knowledge to manage IT projects in the current business environment	
Organizational resources	OR1	My company has the technological resources to adopt this technology	Boonsiritomachai (2014)
	OR2	My company has the financial resources to adopt this technology	
	OR3	My company has no difficulties in finding all the necessary resources (e.g. funding, people, time) to adopt this technology	
Competitive pressure	CP1	We believe we would lose our customers to our competitors if we did not adopt this technology	Qian et al. (2016) Premkumar and Roberts (1999)
	CP2	We feel it is a strategic necessity to use this technology to compete in the marketplace	
	CP3	We believe that our competitors get many advantages through adopting this technology	
	CP4	Many of our competitors are going to adopt this technology in the near future	
External support	ES1	There are businesses in the community which provide support for use of this technology	Premkumar and Roberts (1999)
	ES2	There are agencies in the community who provide training on this technology	
	ES3	Technology agencies actively market this technology by providing incentives for adoption	
Privacy	PR1*	My company's use of this technology is limited by data protection acts, rules and regulations in Norway	
	PR2*	My company finds it challenging to protect data privacy when adopting this technology	
	PR3*	My company finds it difficult to comply with privacy related regulation when using this technology	
	PR4*	My company finds it difficult to meet legal expectations concerning the use of Big Data without compromising our business goals	

*Items developed for this study by the authors, (R) – Reverse coded items

The remaining constructs developed specifically for this research were operationalized by novel items. **Security** was developed to test the propositions of Salleh and Janczewski (2016). They argued that adopters of Big Data would face distinctive security issues in terms of complexity and compatibility. Thus, two sets of items were developed to capture the multidimensionality of the security construct. First, inspired by the wording in Davis' (1989) measurement items, two items were developed to reflect the complexity in providing effective security (e.g., it is easy for my company to integrate security policies for this technology). Second, three items were developed to reflect the compatibility issues of current security mechanisms with those required by Big Data (e.g., my company has security capabilities to adopt this technology). Regarding **privacy**, four items inspired by Salleh and Janczewski (2016) were developed to reflect the perceived difficulty of complying with privacy related regulation (e.g., my company finds it difficult to comply with privacy related regulation when using this technology). The final measurement items can be seen in Table 2.

All of the items were presented as statements accompanied by an ordinal scale, where the numbers reflect how strongly the respondent agrees or disagrees with the statement (Saunders, et al., 2012). Specifically, a 7-point Likert scale ranging from "Strongly disagree" (1) to "Strongly agree" (7) was used to measure the respondents' perception of all constructs. An odd number of response categories was chosen as it allows for neutral responses.

Finally, for **Big Data adoption,** a seven item Guttmann scale was developed to operationalize the aggregated, three-stage model of Big Data assimilation discussed in Chapter 3.3.5. Each of the seven items correspond to a distinct assimilation stage: (1) non-awareness, (2) awareness, (3) interest, (4) evaluation/trial, (5) commitment, (6) limited deployment, and (7) general deployment. The scale is similar to the one that Fichman and Kemerer (1997) used to assess adoption of software process innovations, the scale Rai et al. (2009) used to measure assimilation of electronic procurement innovations, and the scale that McKinnie (2016) used to operationalize the adoption of cloud computing. Respondents were asked to identify their current stage in regard to the adoption of Big Data technology. Thus, organizations were classified according to the highest stage achieved on the Guttmann scale as of the time the survey was taken. The measurement items for Big Data adoption can be seen in Table 3.

Table 3: Measurements items for Big Data assimilation

Stage	Criteria to Enter stage	Item	Stage in research model
1. Non-awareness	The organization is unaware of Big Data technology	My company is not familiar with Big Data technology	**Initiation**
2. Awareness	The organization is aware of Big Data technology	My company is familiar with Big Data technology and/or has considered using it	
3. Interest	The organization is committed to actively learn more about Big Data technology	My company is planning to use Big Data technology within the next 24 months	
4. Evaluation/trial	The organization has initiated evaluation or trial of Big Data technology	My company has launched pilot projects or initiatives for evaluating and/or trialling Big Data technology	**Adoption-decision**
5. Commitment	The organization has committed to use Big Data technology in a significant way	The acquisition of specific Big Data technologies are planned, in progress, implemented or cancelled	
6. Limited deployment	The organization has Big Data technology but a program of limited use	My company has Big Data technology, but we have yet to establish a program of regular use	
7. General deployment	The organization has Big Data technology and a program of regular use	My company has Big Data technology, and we have established a program of regular use	**Implementation**

4.2 Sampling and Data Collection

The target population for this study has been defined as the total population of publicly registered medium to large enterprises in Norway. Medium to large enterprises is herein defined as businesses with revenues greater than €10 million (NOK 85 million) and whose number of employees exceed 50 (EU, 2003). A complete list of the entire population – 3 172 enterprises – was obtained from Proff.no[21] on February 10th, 2017.

Eligible respondents from each enterprise were considered individuals best qualified to speak about the enterprise’s adoption activities. As the organizational adoption of Big Data technology is presumed to be an **authority decision** made by top management, the chosen representative from each enterprise (i.e., the respondent) was either the CEO, CIO, CTO or of a similar executive position. Unfortunately, the list from Proff.no did not contain contact information (i.e., e-mail addresses) pertaining to representatives from each company. Consequently, contact information for eligible respondents had to be collected by the authors; ergo, approximately 160 hours were spent reviewing each individual company’s publicly

---

[21] Proff.no is an official distributor of enterprise information from Brønnøysund Register Centre, the Norwegian government agency responsible for the management of numerous public registers in Norway, including the enterprise register (Brreg, 2017).

available information. Ultimately, the e-mails of 2 625 eligible respondents were registered, representing 82,8% of the entire population of the study[22].

An invitation to participate in the survey was sent by e-mail to the 2 625 eligible respondents, representing equally many unique enterprises. The online survey was administered in the period between 29th of March and 26th of April 2017. In total, 507 (19,3%) respondents opened the survey, of which 403 (79,5%) responses were recorded as fully complete. Of these 403 participants, 107 showed interest in receiving a summary of the research findings; illustrating the relevancy of the subject of the present study. A summary of the data collection and response rate is presented in Table 4. Two follow-up reminders were sent out during the data collection period to motivate participation. Evidently, the stimuli had a considerable effect on participation.

Table 4: Data collection and response rate

	Sent out	Completed surveys	Response rate
Initial survey invitation	2 625	159	6.1%
Reminder 1	2 439	157	6.4%
Reminder 2	2 255	87	3.9%
Total		403	15.35%

Furthermore, the sample size in empirical research should be sufficiently large as to be representative of the given population in the study. Researchers suggest that the larger the sample size the lower the likely error in generalizing to the population (e.g., Krejcie & Morgan, 1970; Saunders et al., 2012). Different methods have been suggested for determining the sample size of a given population. Krejcie and Morgan's (1970) formula for determining sample size has been widely used by researchers, particularly in survey research (c.f., Bartlett, et al., 2001). According to Krejcie and Morgan's criterion, the sample size for a population of 3 172 enterprises, given a confidence level of 95%, a margin of error of 5%, and a response distribution of 50%, is required to be at least 343. Moreover, the relationship between sample size and total population displays diminishing returns; as the population increases the sample size increases at a diminishing rate and remains constant at 384. Thus, the initial sample size

[22] An overview of the sample characteristics for all respondents is presented in Table 5 (Ch. 5.2).

of 403 complete responses appear more than adequate and should lend itself to generalizations about the population.

4.3 Non-response Bias

According to Dillman (2001), cited in Lindner (2001), there are four main sources of error in survey research: sampling error, coverage error, measurement error, and non-response error. As any of these errors increase in a survey research, “the results and recommendations of the study become increasingly suspect and decreasingly valuable as evidence of the characteristics in the target population” (Lindner, et al., 2001, p. 43). Whereas the procedures for handling the first three sources of error are discussed in Chapter 4.1.2 and 4.2, we have yet to discuss the handling of non-response error. This is especially important given the low response rate attributed to online survey research (Saunders, et al., 2012; Neuman, 2014).

Non-response error occurs when respondents included in the sample fail to participate or provide usable responses and are different from those who do with regard to the characteristics of interest in the study (Lindner, et al., 2001). To ensure the external validity of survey findings, researchers must consider whether the results of the survey would be the same if a 100% response rate had been achieved. Miller and Smith (1983) state that non-response error is of a concern even for studies with response rates as high as 90%, as the “data gathered from self-selected respondents may not represent the characteristics of the entire sample” (p. 45). Similarly, Lindner et al. (2001) write that non-response error can threaten the external validity of a study if less than 100% response rate is achieved.

Non-response bias can be categorized into two types (c.f. Yan & Curtin, 2010). The first is item non-response, which occurs in the absence of answers to specific questions in the survey. In designing the survey on Qualtrics, all responses were forced as to avoid this type of error. The second type of bias is unit non-response, which refers to the complete absence of participation in the survey from a respondent. Unit non-response is typically unavoidable in survey research. Given the response rate of 15.35% (Table 4), unit non-response error is a potential threat to the external validity of this study. Fortunately, researchers have developed a number of procedures to test for non-response bias to ensure that the external validity of the research findings are not threatened. The statistical procedure used to test for non-response bias is detailed in Chapter 5.1.2.

5. Data Analysis

In the following chapter, the data analysis, referring to the inspecting, cleansing, transforming, and modelling of data, will be presented. The goal of the data analysis is to obtain sufficient statistical information to answer the research questions of the study.

As this thesis is interested in the relationships between multiple independent variables (measurement items) and one dependent variable (assimilation of Big Data technology), multivariate analysis[23] is used. Although various techniques exist, the appropriate multivariate technique depends on the characteristics of the dependent and independent variables (Hair, et al., 2010). Since the dependent variable is categorical and the independent variables are both categorical (e.g., organizational size) and continuous (e.g., complexity), multinomial logistic regression is deemed appropriate. Multinomial logistic regression allows for the prediction of a categorical outcome, such as stage of adoption, from a set of categorical and continuous *predictor variables*. In particular, using a logistic regression, it is possible to evaluate the probability of a company being in a specific stage of adoption, given their pattern of responses to the thirty-seven measurement items. However, these items are thought to measure 10 underlying constructs and can subsequently be condensed into a smaller set of factors that lend themselves more easily for use as predictor variables in a logistic regression. Thus, a factor analysis will be performed to reduce the dimensionality of the data, before delving into the multinomial logistic regression.

Accordingly, this chapter presents the preliminary analysis (Ch. 5.1) and descriptive statistics (Ch. 5.2), followed by two multivariate analysis techniques; factor analysis (Ch. 5.3) and multinomial logistic regression (Ch. 5.4).

5.1 Preliminary Analysis

Employing an online survey raises concerns with data quality. Thus, a data screening process and assessment of biases were conducted to ensure the data was applicable, reliable, and valid for further statistical analysis, and is presented in the following sections.

---

[23] Multivariate analysis refers to "all statistical techniques that simultaneously analyse multiple measurements on individuals or objects under investigation" (Hair, et al., 2009, p. 4).

5.1.1 Data Screening

To ensure data quality, a data screening process was performed to remove participants who were not sufficiently motivated to provide accurate responses or did not belong to the target population. While missing data were not an issue in this study, as only complete survey responses (N=403) were taken for further screening, outliers were assessed and removed in order to avoid significant influences on descriptive values and inferential statistics computed from the data (Stevens, 1984). Though outliers do not generally exist in Likert scale data, as extreme values range from 1 to 7 (Gaskin, 2017a), extremely inconsistent responses or invariant answering patterns can be considered outliers. According to Liu and Zumbo (2007), there are three possible sources of outliers; the first refers to errors that occur during the data collection and preparation phase. This particular source of outliers was eliminated as the dataset was automatically exported directly from Qualtrics to SPSS, minimizing the threat of human error in the transition process toward the data analysis. The two remaining sources of outliers are categorized as; unpredictable measurement-related errors from participants, and recruitment of participants that do not belong to the target population (Liu & Zumbo, 2007), which will be discussed in the following.

According to DeSimone et al. (2015), there are three common data screening techniques: direct, statistical, and archival. The *archival screening* method, useful for detecting the second source of outliers, involve the examination of the respondents' pattern response behaviour over the course of the survey. The archival technique is intended to screen respondents who; respond inconsistently across similar items, respond inconsistently across dissimilar items, respond too quickly, and respond the same way to all items (DeSimone, et al., 2015, p. 172). Meade and Bartholomew (2012) suggested using a long string of the same response option being selected repeatedly to indicate lack of effort or careless responding. Screening responses with 6 to 14 invariant responses in a row has been recommended by researches (DeSimone, et al., 2015). A response pattern approach, also called *longstring approach* or *invariant responding* (Huang, et al., 2011), was therefore applied in the data screening. However, as most of the items in the survey were scored in the same direction, which implies that a longer string of invariant responses is more likely, the cutoff point chosen to represent invariant responses in a row was determined by the number of questions on two consecutive survey pages. More specifically, respondents with identical answers to all items on a single page, for two consecutive pages (N=11), were deleted[24]. To further minimize any random, careless, or inconsistent responding, the *response time approach* was employed (Huang, et al., 2011). Using this approach, the time that each respondent spent taking the survey was identified and respondents that spent less than 210 seconds[25] (N=22) on completing the survey were eliminated.

Finally, the third source of outliers refers to the recruitment of participants that do not belong to the target popution, which this study has defined as medium to large businesses in Norway. Control questions were introduced in the survey to identify respondents that did not fit this criterion; namely the size of the respondent's firm measured by annual revenue and number of employees. Hence, companies with less than NOK 85 million in annual revenue and less than 50 employees (N=23) fell outside the sample definition and were removed. Moreover, as this study had defined eligible respondents as executive management, respondents who did not hold an executive position (N=11) were deleted. Thus, the final sample consisted of 336 companies.

Looking back at Chapter 4.2, the recommended sample size for the target population was 343. This gives the sample after data screening a margin of error of 5.06%[26], which should be adequate for generalization about the population.

5.1.2 Assessing Sampling and Method Biases

Several procedural remedies were implemented in the questionnaire design (Ch. 4.1.2) to prevent sampling and method bias. However, no research design can account for all sources of biases. Therefore, statistical tools were employed to assess the presence of biases in the sample.

Non-response Bias

Non-response bias is a challenge facing studies using surveys as a method of data collection. When there are significant differences between those who responded to the survey

---

[24] E.g. one page could consist of all items rated "Strongly agree", while the successive page could consist of all items rated "Neutral".

[25] It takes on average 300 seconds (5 minutes) to complete the survey, spending less than 210 seconds would be considered improbable if the respondent carefully read through all questions.

[26] For a sample size of 336 from a population of 3172, given a confidence level of 95% and a response distribution of 50%, the margin of error is 5.06% (Krejcie & Morgan, 1970).

and those who did not, the collected data has the potential of not representing the target population (Draugalis & Plaza, 2009). Atif and Richards (2012) illustrate this error in the following equation:

Respondents' characteristics = population characteristics +/- non-response bias

The error is the difference between the survey estimate and the actual population value; the equation indicates that if there were no non-response bias, then the sample (respondents' characteristics) would be representative of the population.

To determine the representativeness of the sample, an assessment of non-response bias is recommended when response rates are low. A low response rate indicates that those who responded have a greater chance of being self-selected (Lewis, et al., 2013). As a significant proportion of the sample (84.65%) failed to participate in the study, it was important to investigate whether respondents differ from non-respondents, which could lead to biases in the dataset and affect the validity of the survey results (Atif & Richards, 2012).

A common approach to test for non-response bias is by employing the extrapolation methods suggested by Armstrong and Overton (1977), which are based on the assumption that subjects who respond less readily have characteristics more like non-respondents. "'Less readily' has been defined as answering later, or as requiring more prodding to answer" (Armstrong & Overton, 1977, p. 2). The most common type of extrapolation method is *wave analysis*. "Wave" refers to the response generated by a stimulus, e.g., a follow-up reminder. Individuals who respond in later waves are assumed to have responded because of the increased stimulus and are expected to be similar to non-respondents. Performing a comparison of respondents and non-respondents is generally an accepted procedure and widely used method in quantitative research to identify non-response bias. If no differences are found between early and late respondents, an assumption can be made that non-response bias is unlikely to affect the sample results (Lindner, et al., 2001).

To test for non-response bias, early and late respondents were identified based on the time of their recorded response. The first 69 respondents were considered early respondents because their responses were recorded on the first day of sending out survey invitations. The last 69 respondents were considered late respondents due to the efforts exerted to obtain them (two reminder emails were sent out before they decided to take the survey). As the dataset consisted of both nominal (demographic information) and ordinal (Likert scale items) data, two tests were applied in order to account for the data's distinct characteristics.

A Pearson Chi-squared test was employed to determine whether any differences existed between early respondents and late respondents by comparing them to the demographical information obtained in the study (Appendix D.1). However, one of the conditions to use a Chi-squared test was not met for two variables (industry and years in business) because expected count for some cells were less than 5; Fisher's Exact test was therefore applied on these variables (McDonald, 2014). The results from the tests revealed that none of the nominal variables were significant ($p > 0.05$), which indicate no difference between early and late respondents with respect to their executive position, industry, annual revenue, number of employees, and years in business. Additionally, as Likert scale items have distinct characteristics: discrete instead of continuous values, tied numbers, and restricted range (Winter & Dodou, 2010, p. 1), it was appropriate to employ a non-parametric procedure that accounted for the ordered nature of Likert scales. Thus, the Mann-Whitney U test was used on the remaining items to determine whether responses by early respondents differ significantly from late respondents (Appendix D.2). The results reveal that; of all the responses on the 37 items, only 3 items were found significant ($p < 0.05$). Although this could suggest non-response bias, it may be attributed to randomness due to the number of variables tested. Moreover, since the majority of items were non-significant, this paper concludes that no discrepancies between early and late respondents were found. Hence, non-response bias is unlikely to affect the sample.

Characteristics of Respondents

As the organizational adoption of Big Data technology is assumed to be an authority decision made by executive management – both IT managers and non-IT managers were included in this study. Various managerial positions have been the subjects for studies of organizational technology adoption (Gibbs & Kraemer, 2004; Zhu & Kraemer, 2005; Zhu, et al., 2006a; Soares-Aguiar & Palma-dos-Reis, 2008; Boonsiritomachai, 2014; Nam, et al., 2015). However, one concern that arises is that IT managers and non-IT managers might have different perceptions on the constructs that are hypothesized to influence the adoption of Big Data technology. In a study of technology acceptance, Zhu et al. (2006a) performed a Kolmogorw-Smirnov (K-S) test to compare sample distributions of IS managers and non-IS managers. They computed composite constructs before employing the K-S test and found that two out of ten constructs in their study were significant. However, the significant variables represented objective characteristics of the firm (firm size and global scope), and answers to them were less likely to be influenced by subjective opinions. Hence, they concluded that positions of the respondents did not not cause significant biases.

However, since the present research employes different constructs than Zhu et al. (2006a), their conclusion cannot be generalized to this study. Thus, to investigate this concern, a similar test was conducted on the composite constructs (see Ch. 5.2) to determine if there were any differences between non-IT managers (CEO/President/VP/Managing director/Other C-level executive) and IT managers (CIO/IT director/Technology director) (see Appendix D.3). A non-significant K-S test ($p > 0.05$) suggests that the sample distributions of the two independent groups do not differ statistically. Out of the ten constructs, only *top management support* yielded significant ($p < 0.001$), suggesting that IT managers and non-IT managers have different perceptions of this construct in relation to Big Data technology. On average, non-IT managers scored higher (mean = 5.73) on top management support than IT managers (mean = 5.14), indicating that non-IT managers place a greater importance on top management support when adopting Big Data technology, while IT managers place a relatively lower importance on this construct. Overall, as nine other constructs were not significant, this paper will conclude that the differences in positions do not cause significant biases. However, for further analysis, one may be cautious about drawing any conclusion regarding top management support.

Common Method Bias

The influence of common methods bias has become a major concern in survey-based research (Podsakoff, et al., 2003). Bagozzi and Yi (1991) defined common method bias as the "variance that is attributable to the measurement method rather than to the construct of interest" (p. 426). A widely used technique for determining the presence of common method bias is Harman's single-factor test. By using this test, all items are loaded onto one single factor in an exploratory factor analysis. The basic assumption is that common method bias is unlikely to affect the data if the total variance for the single factor is less than 50% (Podsakoff, et al., 2003; Eichhorn, 2014). Appendix D.4 shows the result of this test for all items in this study and reveals that the variance is less than the indicated threshold. However, Harman's single-factor test has been criticized for its limitations and researchers have more recently recommended using a confirmatory factor analysis to address this bias. Consequently, by using this approach, there is still a chance that common method bias is present even with the low variance value (Podsakoff, et al., 2003; Gaskin, 2017b). Nonetheless, as the single factor in the exploratory factor analysis only accounted for 29.97% of the variance, this paper will not take the analysis further, but conclude that common method bias is unlikely to be a major concern.

5.2 Descriptive Statistics

The descriptive statistics are used to present the basic features and summary of the data in study. In the following, statistics of the sample demographics and composite constructs are presented.

Sample Demographics

The characteristics of the complete sample (N=336) are presented in Table 5, categorized by the assimilation stages. The table illustrates that the majority of respondets were CEO/President/VP/Managing directors (61.9%), while the remaining respondets were either CIO/IT director/CTOs (22.6%) or possessed other executive positions (15.5%). Moreover, while almost a third of the companies (N=101) were currently in the adoption-decision stage of Big Data assimilation, only a small proporition of the sample were placed in the implementation stage (N=44). The majority of companies, however, were still in the initiation stage (N=191).

The companies represent various industries, specifically; the manfacturing industry (19.3%), and construction, agriculture & materials (12.5%) correspond to almost a third of the sample. However, these industries are less represented in the implementation phase (< 10%) than they are in the initiation phase (> 40%), suggesting that few companies within these industries have adopted Big Data technology. Conversely, a number of companies within banking & insurance (15.9%), retail & wholesale (11.4%), information technology (13.6%), and entertainment, media & tourism (13.6%) are in the implementation stage, and have thus come further with Big Data technology than other industries. While the manufacturing industry (14.9%) and energy & utiltities (12.9%) were most frequent counted in the adoption-decision phase, not as many companies from these industries have moved into the implementation stage. Hence, these numbers suggest that industries are represented differently along the assimilation stages of Big Data technology.

Table 5: Sample characteristics

	Initiation (N=191) %	Adoption-decision (N=101) %	Implementation (N=44) %	Full sample (N=336) %
Role				
CEO/President/VP/Managing director	60.2	61.4	70.5	61.9
CIO/IT director/Technology director	26.2	17.8	18.2	22.6
Other C-level executive	13.6	20.8	11.4	15.5
Industry				
Banking and insurance	4.2	6.9	15.9	6.5
Manufacturing	24.6	14.9	6.8	19.3
Construction, agriculture and materials	18.8	5.0	2.3	12.5
Telecommunications	0.5	2.0	2.3	1.2
Transport, logistics and post	6.8	6.9	2.3	6.3
Energy and utilities	6.3	12.9	6.8	8.3
Retail and wholesale	9.4	6.9	11.4	8.9
Services	6.3	9.9	6.8	7.4
Public sector and healthcare	6.3	5.9	4.5	6.0
Information technology	3.1	7.9	13.6	6.0
Entertainment, media and tourism	0.5	8.9	13.6	4.8
Education and scientific research	5.2	3.0	6.8	4.8
Other	7.9	8.9	6.8	8.0
Annual revenue				
85-150 million NOK	16.8	12.9	13.6	15.2
150-300 million NOK	24.1	16.8	9.1	19.9
300-500 million NOK	17.8	17.8	13.6	17.3
500-1000 million NOK	18.8	16.8	9.1	17.0
More than 1000 million NOK	22.5	35.6	54.5	30.7
Number of employees				
50-100	37.7	21.8	27.3	31.5
101-150	17.3	16.8	13.6	16.7
151-250	15.2	18.8	11.4	15.8
251-400	11.0	8.9	6.8	9.8
More than 400	18.8	33.7	40.9	26.2
Years in business				
Less than 5 years	5.2	1.0	2.3	3.6
5-10 years	4.2	3.0	4.5	3.9
11-20 years	15.2	14.9	18.2	15.5
21-30 years	14.7	16.8	15.9	15.5
Longer than 30 years	60.7	64.4	59.1	61.6

Total percentages may not equal 100 due to rounding of numbers

Additionally, almost a third of the companies (30.7%) earn more than NOK 1 000 million in annual revenues. These companies are spread unevenly across the assimilation stages – with more of the larger companies in the implementation phase (54.4%). The same trend is observed with regard to employees; companies with more than 400 employees (26.2%) are more prevelant in the implementation phase (40.9%) than they are in the initiation phase (18.8%). This could suggest that organizational size has an effect on the adoption of Big Data technology.

Lastly, while more than half of the companies (61.6%) are older than 30 years, almost the entire sample (> 90%) have been in business for at least 10 years. Only a small proportion of the sample (< 8%) have been in business for less than 10 years.

Composite Constructs

Table 6: Mean of composite constructs

	Assimilation stages					
	Initiation		Adoption-decision		Implementation	
	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
Technological context						
Relative advantage	5.20	1.03	5.70	0.70	6.10	0.84
*Complexity	4.14	1.04	4.05	1.03	3.33	1.03
Compatibility	4.45	1.04	5.24	0.84	5.94	0.76
*Security	3.76	1.10	3.43	1.01	2.67	1.01
Organizational context						
IT expertise	4.81	1.11	5.03	1.18	5.77	1.01
Top management support	5.16	1.03	6.00	0.70	6.58	0.51
Organizational resources	4.23	1.17	5.02	1.00	5.38	1.06
Environmental context						
External support	4.78	0.94	4.98	0.95	5.27	0.92
Competitive pressure	4.15	1.22	5.28	0.90	5.80	0.91
Privacy	3.81	0.95	3.92	1.17	3.82	1.31

* Reversed coded construct

Composite variables were computed for each construct by averaging their corresponding items, presented in Table 6. The table summarize the average of all responses for each hypothesized construct; from 1 (strongly disagree) to 7 (strongly agree). For instance, companies within the initiation stage perceive lower relative advantage (5.20) of Big Data technology than companies in adoption-decision (5.70) and implementation (6.10) stages.

Accordingly, Table 6 provides an indication of the relationship between the proposed constructs and the assimilation of Big Data technology

Within the technological, organizational, and environmental context, the mean for each construct increases for every stage in the assimilation process, suggesting companies within the initiation stage score lower on a specific construct than companies in the adoption-decision stage, and companies within the adoption-decision stage score lower on a specific construct than companies within the implementation stage. This implies that companies have different perceptions of the specific constructs when belonging to different assimilation stages. The exception is privacy, as the mean seem not to vary (3.81 – 3.92 – 3.82). Additionally, because the complexity and security items have been reversed coded to coincide with the hypotheses (i.e., the constructs are negatively associated with the adoption of Big Data technology), the mean for these constructs are decreasing for every stage of assimilation; suggesting companies within the initiation stage perceive higher complexity and security (more difficult) than companies within the adoption-decision and implementation stages.

5.3 Factor Analysis

Before testing the hypotheses developed in Chapter 3.3, it is necessary to assess whether the items included in the survey (Ch. 4.1.3) truly measure the underlying constructs they are intended to. This section will therefore describe the statistical technique – factor analysis – employed to investigate the relations between the measurement items and the constructs they are believed to represent.

Factor analysis provides a good starting point for multivariate analysis, as it provides insight into the inter-relationship among variables and the underlying structure of the data. For the purpose of this research, factor analysis gives the opportunity to transform a large set of variables (i.e., measurement instruments) into a smaller number of new variables (i.e., constructs), called factors. By computing factor scores, these factors can in turn be incorporated into the subsequent analysis as *predictor variables*.

Thus, this section begins by making a distinction between observed and latent variables, followed by a stepwise discussion of the factor analysis performed in this research. The purpose of the analysis is to reduce the number of variables into a set of factors with corresponding values that each represent a construct in the research model.

Observed and Latent Variables

Social science research often deals with two types of variables; those that are measured directly (observable variables)[27] and those that cannot be measured directly but rather inferred from observable variables (latent variables) (Field, 2009; Planing, 2014). In the present research, the constructs that comprise the research model are considered latent variables, whereas measurement items used to operationalize these constructs are treated as observable variables.

Chapter 4.1.3 described the operationalization of latent variables in this research. A model of the relationship between measurement items and the construct they are intended to represent is called a measurement model (Planing, 2014). The rationale behind the measurement model is that the combined answers to multiple observable items better represent the complexity of a construct than any single item could alone. Consequently, a measurement model offers greater richness in measurement, captures the nuances of a construct, and enables the researcher to assess how reliably the constructs have been measured (Easterby-Smith et al., 2008, cited in Planing, 2014).

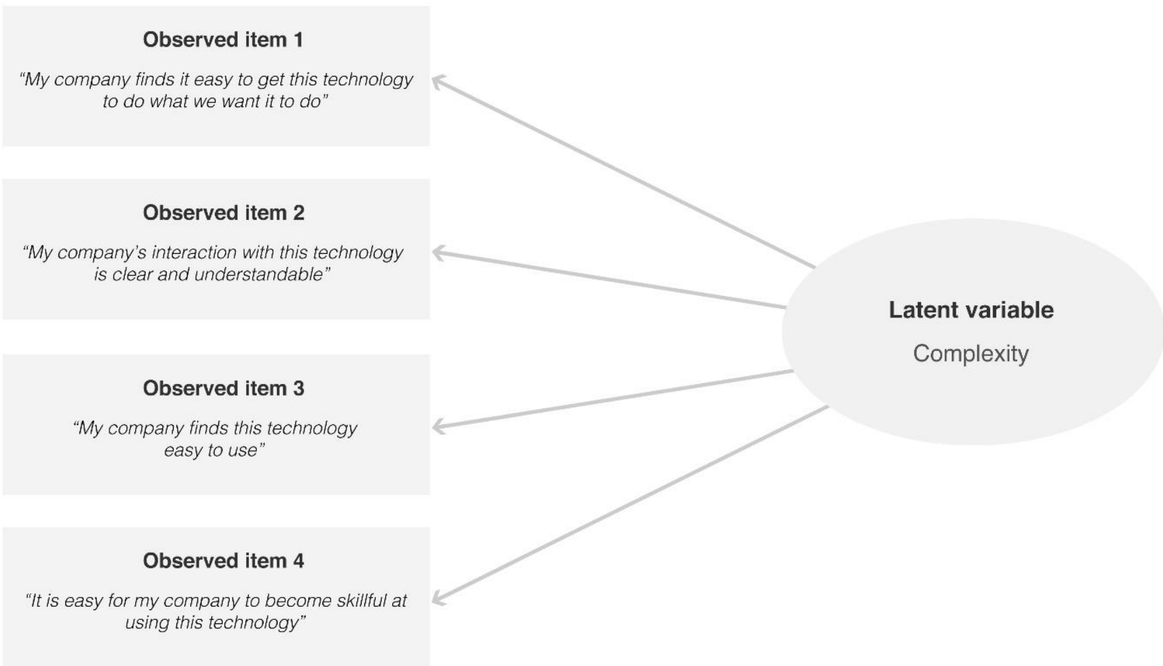


Figure 8: Measurement model for the latent variable complexity (Source: Own drawing based on Planing, 2014)

[27] Organizational size is an observable variable and is consequently used directly as a predictor variable in the multinomial logistic regression (Ch. 5.4).

An example of the measurement model for **complexity** is presented in Figure 8. Here, complexity is thought to reflect the observed items 1, 2, 3, and 4. The stronger the influence of the latent variable on the observed variable, later referred to as factor loading, the higher the correlation between the observed variables (Field, 2009; Planing, 2014).

5.3.2 Exploratory Factor Analysis (EFA)

The statistical method used to examine relationships between observed and latent variables is called factor analysis (Field, 2009). Not to be confused with principal component analysis[28] (PCA), factor analysis represents a variety of statistical techniques operating on "the notion that measurable and observable variables can be reduced to fewer latent variables, called factors" (Yong & Pearce, 2013, p. 80). These factors, in turn, can be used to represent underlying constructs (e.g., complexity) (Costello & Osborne, 2005). There are two methods of factor analysis; exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Whether an EFA or CFA is appropriate for a given study largely depends on the purpose of the study and the "level of knowledge of the underlying factors" (Planing, 2014, p. 201). In general, CFA is appropriate when a hypothesized structure explaining the relationship between variables have been established *a priori*; that is, when the number of factors *and* their relation to the observed variables are specified prior to data collection (Baglin, 2014). CFA also typically assumes each observed variable only load on one factor. Conversely, EFA permits each observed variable to load on any identified factor, and is thus appropriate if, *a priori*, a structure for the variables' relationship cannot be assumed with confidence (Matsunaga, 2010; Planing, 2014).

Aligned with the exploratory nature of this study and given the novelty of the research context, the observed variables were analysed using the EFA approach. Although the literature review in Chapter 3 provided theoretical and empirical support for the constructs included in this research, it is necessary to explore their operationalization in the context of organizational adoption of Big Data for three crucial reasons. First, measurement items were not only

[28] There is considerable debate between the use of factor analysis and PCA as approaches to locating underlying dimensions of a dataset, where the general notion seems to be that whenever PCA is used it should not be described as a factor analysis, and one should not impute substantive meaning to the resulting components (c.f., Costello & Osborne, 2005; Tabachnick & Fidell, 2007; Field, 2009; Baglin, 2014). Although PCA seeks to reduce the number of variables into a smaller set of variables (called components), it does so without regard to any underlying structure caused by latent variables. On the other hand, factor analysis is concerned with identifying the underlying factor structure that explains the relationship between the observed variables. Since this research is interested in latent variables, factor analysis was chosen as it is more theoretically aligned to the goals of the research (Matsunaga, 2010).

extrapolated from studies of entirely different adoption scenarios than Big Data technology, but also changed to fit the organizational level of analysis. Second, constructs developed specifically for this research were operationalized by entirely new, unproven measurement items (privacy and security). Third, all measurement items were translated from their original language (English) to Norwegian. For these reasons, there are likely to be a number of problematic measurement items. Since EFA does not apply a priori theory about which items belong to which constructs, it is considerably better at identifying poor measurement items. Accordingly, EFA was chosen for this research.

5.3.3 Assumptions

Before performing an EFA, the literature suggests evaluating the sample size, normality of data, and factorability. Although there are few strict assumptions, satisfying minimum requirements may greatly enhance EFA solutions (Tabachnick & Fidell, 2007).

Sample Size

While strict rules for sample size for EFA have waned (Costello & Osborne, 2005), a general rule of thumb is to have 5-10 participants per variable up to a total of 300, beyond which test parameters tend to stabilize regardless of participant to variable ratio (Tabachnick & Fidell, 2007; Field, 2009; Yong & Pearce, 2013). Because factor analysis is based on the correlation matrix of the variables involved, and correlation coefficients tend to be less reliable for smaller samples, a larger sample is always preferred. Tabachnkick and Fidell (2007, p. 613) cite Comrey and Lee's (1992) guide regarding sample size: 50 is very poor, 100 is poor, 200 is fair, 300 is good, 500 is very good, and 1 000 as excellent. This research has a sample size of 336 and a participant to variable ratio of 9:1, which according to extant literature is satisfactory.

Normality

Normality may enhance the solution of a factor analysis and it is therefore beneficial if variables are normally distributed[29] (Tabachnick & Fidell, 2007). The normality of single variables can be assessed both visually and statistically. Statistically, normality may be

[29] When statistical inference is used to determine the number of factors, multivariate normality is assumed (Tabachnick & Fidell, 2007). Although univariate normal distribution is no guarantee for multivariate normal distribution, it is considered a sufficient indicator of normality (Hair, et al., 2010).

assessed by performing the Shapiro-Wilk test[30], and by calculating the skewness and kurtosis[31] values (Appendix E.1). Visually, frequency histograms and normal probability plots may reveal deviations from normality.

Inspection of the skewness and kurtosis values indicate that only one variable, MS3 ("top management support is important to provide the resources for my company to use this technology"), has a value exceeding an acceptable threshold (±1.96) (Hair, et al., 2010). While this may indicate normality, deviations were observed for several variables upon visual inspection of the frequency histogram and normal probability plots[32]. Moreover, the Shapiro-Wilk test was significant for all variables ($p < 0.05$), suggesting that one should refrain from assuming normal distribution. However, as most variables are responses to positively worded questions measured along a Likert scale, deviations from normality are expected due to the ordinal nature of the scale. Fortunately, while normality is recommended, it is not required for EFA – nor does is it guaranteed to degrade the solution (Tabachnick & Fidell, 2007; Hair, et al., 2010). Nonetheless, these findings will have implications for the choice of factor extraction method, as discussed later (Ch. 5.3.4).

Factorability

The literature suggests using the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) and Barlett's test of sphericity to test for factorability; that is, to assess the appropriateness of the dataset for factor analysis (c.f., Tabachnick & Fidell, 2007; Hair, et al., 2009). According to Kaiser (1974, cited in Planning, 2014), KMO ranges from 0 to 1 and can be interpreted as follows: 0.8 or above as meritorious, 0.7 or above as middling, 0.6 or above as mediocre, 0.5 or above as miserable, and below 0.5 as unacceptable. Barlett's test of sphericity, on the other hand, is a significance test whose associated significance level must be less than 0.05 for factor analysis to be considered appropriate. For the present research, KMO of 0.894 and Barlett's test of sphericity ($p < 0.01$) indicate that the initial data is suitable for factor analysis (see Appendix E.2).

[30] The Shapiro-Wilk test compares the shape of a sample distribution to the shape of a normal curve.

[31] Skewness is a measure of symmetry of the distribution and kurtosis is a measure of the peakedness of the distribution (Tabachnick & Fidell, 2007).

[32] Frequency histograms and normal probability plots were generated for each of the 37 measurement items, but not included in the appendix of this thesis due to the sheer number of figures.

5.3.4 Components of EFA

Before presenting the EFA, we would like to bring to the reader's attention that EFA is a complex, multi-step process involving several choices regarding factor extraction method, rotation, retention, and factor scores. Without delving into great detail, this section documents the major choices taken in this research, before presenting the final EFA solution.

Extraction Method

In regard to factor extraction methods in EFA, the literature suggests using principal axis factoring (PAF) and maximum likelihood (ML) (Costello & Osborne, 2005; Winter & Dodou, 2012; Baglin, 2014). Although numerous other extraction methods exist, including unweighted least squares, generalized least squares, alpha factoring, and image factoring, PAF and ML remain the most popular and tend to give the best results (Winter & Dodou, 2012). Costello and Osborne (2005) and Baglin (2014) argue that the choice between ML and PAF should be based on the nature of the underlying distribution of the data. For this reason, the distribution of data should be examined prior to choosing extraction method (see assessment of normality in Ch. 5.3.3). Fabrigar et al. (1999) state that when the assumption of normality is violated, as in the case of this research, PAF is the recommended extraction method.

Furthermore, a comparison of ML and PAF in EFA by Winter and Dodou (2012), found that PAF generally outperforms ML. They also found that PAF and ML display opposite tendencies regarding over-extraction (specifying too many factors) and under-extraction (specifying too few factors); PAF was better in over-extraction and ML in under-extraction. Applying some of the stricter factor retention criteria to the factor analysis suggests this research extracts too many factors (see later discussion of factor retention). Hence, PAF was chosen as the factor extraction method for EFA in this research.

Rotation Method

As the result of a factor analysis is generally uninterpretable without a rotation[33], researchers must decide upon a factor rotation to improve interpretability and the utility of the factor solution (Tabachnick & Fidell, 2007; Field, 2009). In general, a decision is required between orthogonal rotation and oblique rotation: orthogonal rotation produces factors that

[33] Without rotation, variables would load highly on the most important factor and have small loadings on all other factors. Rotation helps interpretation by "rotating" factor axes so that variables are loaded highly on only one factor (c.f., Field, 2009).

are uncorrelated, while oblique rotation allows factors to correlate. Researchers have historically favoured orthogonal rotation due to its (mathematical) simplicity and because the factors are more easily interpretable, whereas modern research seem to favour oblique rotation[34] (Osborne, 2015). Hence, the orthogonal rotation is unlikely to be considered "best practice". However, as the oblique rotation gave unintelligible and erratic results when employed on the data in this study, an orthogonal rotation was chosen. More specifically, the equamax rotation was chosen as the orthogonal rotation method for this research. Mulaik (1972, cited in Tabachnick and Fidell, 2007), reported that equamax may behave erratically unless researchers specify the number of factors. Fortunately, this research employs a priori criterion for factor retention, which is discussed next.

Factor Retention

A key question when performing a factor analysis is how many factors to retain (i.e., how many factors to keep for further analysis). When deciding upon the number of factors to retain, a conceptual foundation of how many factors ought to be found should be combined with empirical evidence regarding how many factors can be reasonably supported (Hair, et al., 2010). Since there is no exact basis for determining the number of factors to retain, the literature suggests relying on multiple criteria when deciding on the appropriate number of factors (Fabrigar, et al., 1999; Tabachnick & Fidell, 2007; Hair, et al., 2010; Field, 2009).

The simplest criterion for factor retention is known as the Kaiser's criterion (Kaiser, cited in Field, 2009). Kaiser suggested retaining all factors with eigenvalues greater than 1, while ignoring those with smaller eigenvalues. This criterion, also known as latent root criterion (Hair, et al., 2010), is based on the idea that eigenvalues represent the amount of variation explained by its associated factor and that eigenvalues greater than 1 represent a considerable amount of variation (Field, 2009). Jolliffe (1972), however, argued that Kaiser's criterion was too strict, and proposed retaining all factors with eigenvalues greater than 0.7.

Another common criterion is known as the scree test. Cartell (1966) suggested plotting a graph of each eigenvalue against its associated factor to identify the optimum number of factors to be extracted. The graph (i.e., scree plot) suggested by Cartell has a distinct shape; there is initially a steep downward curve that gradually becomes horizontal. Cartell (1966)

---

[34] EFA was initially calculated by hand or with computers having severely limited computing power. As a simpler mathematical technique, orthogonal rotation was logically preferred. However, with modern computers there are few barriers to performing oblique rotation (c.f., Osborne, 2015).

argues that the point at which this curve first begins to straighten out indicates the cut-off point for selecting factors.

A third criterion is based on achieving a specified cumulative percentage of total variance by the retained factors. According to Hair et al. (2010), "the purpose of this criterion is to ensure practical significance for the derived factors by ensuring that they explain at least a specified amount of variance" (p. 107). While research in the social sciences can be satisfied with accounting for less variance than in the natural sciences, Hair et al. (2010) suggests factors should account for at least 60% of the variance.

The last criterion relevant for the present research is the *a priori* criterion (Hair, et al., 2010). When applying this criterion, a pre-defined number of factors can be extracted based on the expectations of the researchers. Though less empirically justifiable, the criterion is useful for testing a theory, model, or hypothesis (Hair, et al., 2010). Thus, in the proceeding factor analysis, the a priori criterion will be applied. Moreover, rather than discarding the aforementioned criteria, they will be used in conjunction with the a priori criterion to support the choice of a fixed number of factors.

Factor Scores

Following the verification of a factor structure, factor scores[35] can be computed and used as variables in the subsequent multivariate analysis. Conceptually, factor scores are composite variables representing how much an individual respondent would score on a factor (DiStefano, et al., 2009; Hair, et al., 2010). Consequently, lower values on the items measuring a specific construct (factor) will result in a lower factor score, while higher values on the items measuring a construct will result in a higher factor score.

The literature distinguishes between two types of factor score computation methods; non-refined and refined (c.f., DiStefano, et al., 2009). Non-refined methods are relatively simple procedures, such as summing loadings on a factor, while refined methods are considered more sophisticated as they attempt to retain the relationship between factors. Thus, consistent with our choice of an orthogonal rotation, and to maintain orthogonality of factor scores (i.e., uncorrelated factor scores), a refined method was chosen. More specifically,

[35] Factor scores indicate a respondent's relative placement on a factor (i.e., construct).

regression factor scores[36] were obtain for the final factor solution from SPSS to be used as predictor variables in the subsequent multinomial logistic regression (Ch. 5.4).

5.3.5 Final EFA Solution

The following section describes the process by which problematic items were removed from further analysis and a simple factor structure was achieved in a final EFA. An assessment of the initial reliability (i.e., internal consistency) of the measurement items, is followed by a discussion of the identification and removal of problematic measurement items through respecifying the EFA. Lastly, the final EFA is presented with a simple factor structure.

Initial Reliability Assessment

Cronbach's alpha (α), the most common measure of scale reliability, was used to assess the reliability of the scales for each of the constructs in this research[37]. Cronbach's alpha value ranges from 0 to 1, where values closer to 1 indicate a greater degree of reliability of the scale (Field, 2009). Though there seems to be some disagreement regarding the acceptable cut-off point, an alpha above 0.7 is generally considered acceptable (Kline, cited in Field, 2009).

Internal consistency of each factor and its corresponding items are presented in Appendix E.4, which shows acceptable values between 0.7 and 0.8 for all factors. However, Field (2009) recommends dropping items that lead to substantial improvement in overall reliability. As can be seen in Appendix E.4, dropping three items (MS3, ES3, and RE1) would cause a considerable increase in overall reliability. Moreover, the corrected item-total correlations of the same items were considerably lower than their counterparts. Thus, moving into the EFA, careful attention was given to these items.

Respecifying the EFA

An exploratory factor analysis, with principal axis factor (PAF) extraction method, was initially conducted on all 37 items with orthogonal rotation (equamax). Given the a priori criterion of 10 factors, the number of factors were fixed for the analysis. The cumulative

[36] The advantage of the regression method is that it maximizes validity by providing the highest correlation between a factor score and the corresponding factor (DiStefano, et al., 2009).

[37] Reliability means that a measure consistently reflects the construct it is intended to measure (Field, 2009).

variance explained is presented in Appendix E.3, while the initial communalities[38] and factor loadings (rotated pattern matrix) are presented in Appendix E.5.

To reach a simple factor structure, several iterations of the EFA was conducted to identify problematic measurement items by evaluating each item's communality, factor loading, and cross-loadings. Based on a literature review to identify best practices in EFA, a set of criteria was established for the removal of poor items. These criteria are presented in Table 7.

Table 7: Criteria for removal of problematic measurement items

Pre-established criteria for item deletion	Threshold	References
Communalities	Lower than 0.5	Worthington and Whittaker (2006), Hair et al. (2010), Field (2009)
Factor loading	Lower than 0.4	Worthington and Whittaker (2006), Tabachnkick and Fidell (2007), Hair et al. (2010), Field (2009), Yong and Sean (2013)
Cross-loading	Greater than 0.4 or a difference of less than 0.2 from the item's highest factor loading	Worthington and Whittaker (2006), Tabachnick and Fidell (2007)
Factor to item ratio	Delete factors with less than two items, unless: - The two items are highly correlated (r > 0.7) and - The two items are fairly uncorrelated with other items	Worthington and Whittaker (2006), Tabachnick and Fidell (2007), Yong and Sean (2013)

The respecification of the EFA was done through the removal of items that failed to satisfy the criteria in Table 7. Of the 37 measurement items in the initial EFA, 12 were dropped over successive respecifications. Three items were dropped due to low communalities (< 0.5), three items were dropped due low factor loadings (< 0.4), two items were dropped on the basis of the factor to item ratio criterion, while the remaining four items were dropped due issues with convergent validity (AVE < 0.5 or CR < 0.7) (AVE and CR are discussed further in Chapter 5.3.6). In total, all items associated with compatibility (CM1-CM4) and organizational resources (OR1-OR3) were dropped, while one item was dropped from security (SE5), competitive pressure (CP3), privacy (PR1), top management support (MS3), and external support (ES3). The final model was respecified by deriving a new factor solution without these variables.

[38] A variable's communality represents the variance accounted for by the factor solution (Hair, et al., 2010).

Final Pattern Structure

Table 8: Final pattern matrix, communalities, and Cronbach's alpha

	CX	RA	SE	IE	CP	PR	MS	ES	
Cronbach's alpha	.880	.899	.869	.866	.865	.866	.933	.893	**Communalities**
CX1	.711								.607
CX2	.613								.624
CX3	.830								.762
CX4	.782								.729
RA1		.681							.721
RA2		.803							.784
RA3		.770							.779
RA4		.616							.558
SE1			.699						.639
SE2			.891						.923
SE3			.669						.560
SE4			.597						.560
IE1				.799					.686
IE2				.861					.810
IE3				.733					.602
CP1					.731				.708
CP2					.834				.922
CP4					.567				.531
PR2						.804			.654
PR3						.953			.923
PR4						.731			.549
MS1							.863		.934
MS2							.781		.829
ES1								.882	.837
ES2								.867	.788

Complexity (CX), Relative advantage (RA), Security (SE), IT expertise (IE), Competitive pressure (CP), Privacy (PR), Top management support (MS), External support (ES)

Extraction Method: Principal Axis Factoring. Rotation Method: Equamax with Kaiser Normalization.[a]

a. Rotation converged in 8 iterations.

Factors loadings < 0.4 are suppressed

The final EFA was rerun with the remaining 25 items, with KMO of 0.855 and a significant Bartlett's test of sphericity ($p < 0.01$), indicating sufficient sampling adequacy and correlation between items (Appendix F.1). As the items measuring compatibility and organizational resources were entirely dropped, the a priori criterion for factor retention was adjusted accordingly (down to 8). The final analysis produced a simple factor structure (Table 8) in which each item loaded highly onto one and only one factor, while satisfying the item criteria in Table 7. The items that loaded highly on the same factors suggested that factor 1 represents complexity (CX), factor 2 relative advantage (RA), factor 3 security (SE), factor 4 IT expertise (IE), factor 5 competitive pressure (CP), factor 6 privacy (PR), factor 7 top management support (MS), and factor 8 external support (ES). Regarding the factor retention criteria, the solution can be supported by Jolleffie's (1972) criterion (i.e., keep factors with eigenvalues > 0.7) and the variance criterion (Appendix F.2), and Cartell's (1966) scree test (Appendix F.3). In sum, the final EFA produced as simple structure of eight factors explaining 72% of the cumulative variance.

However, it should also be noted that extracting factors with less than three variables is somewhat unconventional, and is only considered reliable when the variables are highly correlated with each other ($r > 0.7$), and fairly uncorrelated with other variables (Worthington & Whittaker, 2006; Tabachnick & Fidell, 2007; Yong & Pearce, 2013). As can be seen in Appendix F.4, this is the case for both items measuring factor 7 (top management support) and 8 (external support). Organizational resources, however, were dropped entirely as the items, OR2 and OR3, did not meet this criteria after OR1 was dropped due to low factor loading. Compatibility, on the other hand, despite having two items satisfying all item retention criteria, had to be removed due to issues with internal reliability and construct validity (Appendix F.5).

5.3.6 Assessing Reliability and Validity

Once interpretability is adequate, and factor and item retention is justified, the last step is to verify the factor structure by establishing reliability, convergent validity, and discriminant validity.

Concerning reliability, the degree of consistency between multiple measurements of a construct (Hair, et al., 2010), constructs were tested in terms of Cronbach's alpha (α) and composite reliability (CR), also known as latent variable reliability. Both measures exceeded their recommended lower limit, $\alpha > 0.7$ and $CR > 0.7$, and are presented in Table 8 and 9 respectively (Field, 2009; Hair, et al., 2010). Convergent validity, which concerns the degree to which measures of the same construct are correlated (Hair, et al., 2010), was tested in terms of the average variance extracted (AVE). The AVE for each construct, as presented in Table 9, exceeds the recommended value; $AVE > 0.5$ (Hair, et al., 2010). Lastly, discriminant validity, which refers to the degree to which constructs differ from one another, is established when all constructs share more variance with its own items than with other constructs (Hair, et al., 2010). This can be tested by checking whether the square root of the AVE is larger than the correlation between constructs (Agarwal & Karahanna, 2000). Table 9 shows that all AVE square roots (along the diagonal) are greater than the inter-construct correlations. Thus, reliability, convergent validity, and discriminant validity have been established.

Table 9: Validity assessments

	Convergent validity		Discriminant validity							
	CR	**AVE**	CX	RA	SE	IE	CP	PR	ES	MS
CX	.826	.545	**.738**							
RA	.811	.520	.029	**.721**						
SE	.810	.522	.022	.012	**.722**					
IE	.841	.639	.024	-.013	.036	**.799**				
CP	.758	.517	-.010	.078	.001	.000	**.719**			
PR	.872	.696	-.004	-.001	-.003	.002	.000	**.834**		
MS	.807	.677	.017	.027	.007	.027	.049	-.002	**.823**	
ES	.866	.764	.013	.012	.016	.014	.021	-.005	.001	**.874**

Complexity (CX), Relative advantage (RA), Security (SE), IT expertise (IE), Competitive pressure (CP), Privacy (PR), Top management support (MS), External support (ES).

Numbers along the diagonal (bolded) are the square root of each constructs AVE, while off-diagonal numbers are inter-construct correlations. For discriminant validity, diagonal numbers should be greater than the off-diagonal numbers.

5.4 Multinominal Logistic Regression

Concerning the choice of logistic regression, a viable alternative to multinomial logistic regression (MLR) would be the ordinal logistic regression. As the assimilation stages of Big Data technology have a naturally increasing order – with companies starting at the initiation stage, entering the adoption-decision stage, and then, lastly, reaching the implementation stage of Big Data technology – it may seem unreasonable to neglect ordinality (as in the case of MLR). However, a simplifying assumption is made when applying ordinal logistic regression, which is the *proportional odds*[39] assumption. This assumption implies that the predictor variables have the same effect on the odds for each cumulative category of the dependent variable[40] (Hair, et al., 2009; Hosmer, et al., 2013). Thus, in regard to research question two (Ch 1.1), ordinal logistic regression would not permit an examination of the differential effects that the factors may have at the different stages of assimilation. More

[39] The proportional odds assumption is also known as the *parallel lines* assumption.

[40] For instance, the odds of being in the initiation phase *versus* adoption-decision phase increase/decrease with a factor of X for one unit increase in the predictor variable, controlling for all other variables in the model. Also, the odds of being in the initiation *and* adoption-decision phase *versus* implementation phase increase/decrease by the same factor for one unit increase in the predictor variable, controlling for all other variables in the model. The categories in the dependent variable are ranked, and one must compare lower categories to the categories ranked above them.

specifically, it is possible that the effects of the predictor variables on the dependent variable are not uniform, but different throughout the assimlation stages. By assuming proportional odds, this notion would be disregarded. For this reason, a **multinomial logistic regression** was chosen. However, to illustrate that neglecting ordinality does not produce a sub-optimal model, the output from an ordinal logistic regression can be seen in Appendix G. Without going into detail, the proportional odds assumption was met (test of parallel lines: $p > 0.05$) and the ordinal logistic regression yielded almost similar results as the MLR (Appendix J), albeit with a weaker model fit.

Thus, to test the research model and the corresponding hypotheses, the factor scores from the previous section (Ch. 5.3) will be used as predictor variables in a MLR. This analysis makes it possible to investigate if, and to what degree, the predictor variables can predict the odds of different outcomes of the dependent variable; that is, different stages of assimilation. In the following section, the assumptions and model specifications of the MLR model will be presented, and model fit will be assessed.

5.4.1 Assumptions

In order to perform a MLR, several assumptions have to be met, including; independence of errors, absence of outliers, absence multicollinearity, and linearity of the logits (Tabachnick & Fidell, 2007; Field, 2009; Hair, et al., 2010).

Firstly, the *independence of errors* assumption states that responses of different cases are independent of each other, so that each response comes from a different, unrelated case. This implies that the same respondent should not be measured at different points in time (Field, 2009; Tabachnick & Fidell, 2007). As the questionnaire design ensured that respondents could only respond to the survey once, and the data collection happened in one specific period, this assumption was met. In regard to *outliers*, a data screening procedure was performed in Chapter 5.1.1 to remove inconsistent responses and invariant answering patterns, and is thus not considered a problem for this analysis.

Moreover, *multicollinearity* is present when there are strong correlations between two or more predictor variables, which imposes a threat to the validity of the multivariate analysis (Field, 2009). As discussed in Chapter 5.3.4, an orthogonal rotation and regression factor score method were employed, which minimized correlations between factors. Nonetheless, to prove the absence of multicollinearity, the tolerance and the variance inflation factor (VIF) was used as indicators. Generally, a tolerance of less than 0.2 and a VIF greater than 10 suggest a serious multicollinearity problem (Field, 2009; Tabachnick & Fidell, 2007). Hence, absence of multicollinearity can be assumed with tolerance and VIF levels around 1.0 (see Appendix H.1).

Finally, to perform a logistic regression there must be a *linear relationship* between the continuous predictors and the logit of the outcome variable. The Box-Tidwell test was therefore employed to detect departure from linearity (Hosmer, et al., 2013). This method adds the interaction terms of the form $x\ln(x)$, with x representing each of the continuous predictor variables, to the regression model. Significant coefficients are evidence of a nonlinear relationship between the predictor variables and the logit (Hair, et al., 2010). Linearity of the predictor variables was thus evaluated by examining the significance of this interaction, of which none were found significant ($p > 0.05$) (see Appendix H.2).

Hence, it is reasonable to conclude that all assumptions are met for MLR.

5.4.2 Model Specifications

In MLR, the outcome variable is categorical, and the predictor variables are either continuous or categorical. Thus, the model treated the factor “organizational size” as categorical; a dummy variable consisting of medium-sized companies (1) and large companies (0). The remaining eight predictor variables were treated as continuous. Significance levels of 1% and 5% were used (corresponding to p-values of .01 and 0.05). Moreover, as this study’s research model has a solid theoretical foundation, the forced entry method was most appropriate to employ. This method places all predictors into the regression model in one block, and parameter estimates are calculated for each block (Field, 2009). A control variable for industry effects was not included as it did not change the results. Appendix I show the case processing summary for the MLR.

5.4.3 Assessing Model Fit

The 2-log likelihood test, pseudo R^2, predictive accuracy, and likelihood ratio tests were used to assess model fit[41] (Hair, et al., 2010; Tabachnick & Fidell, 2007).

[41] The Pearson Chi-square and deviance statistics are common estimates for goodness-of-fit. However, these test statistics are sensitive to sample size and can be inflated by low expected frequencies, which happens when the model includes continuous covariates (causing many empty cells) (Field, 2009). For this reason, the authors have chosen to not present these test statistics as they are not reliable for the regression model.

2-log Likelihood Test

To assess the model fit, a *2-log-likelihood (-2LL)* test was performed by comparing the intercept-only model, which serves as a good baseline, with the final model (includes the intercept and all predictors). The minimum value for -2LL is 0, which corresponds to a perfect fit. This implies that the lower the -2LL value, the better the fit of the model[42]. A well-fitting model is significant at $p < 0.05$, implying that the full model is significantly different from the one with the intercept-only (Hair, et al., 2010). Appendix I.1 shows that the test was significant ($p < 0.01$), demonstrating an improvement over the intercept-only model, and that the predictors are related to the outcome.

Pseudo R-Square

Several *pseudo R-square* measures have been developed to represent overall fit, and Appendix I.2 presents two of these; Cox and Snell (0.417) and Nagelkerke (0.490). These statistics are most frequently reported for logistics regression and operate in the same manner, with higher values indicating greater model fit (Tabachnick & Fidell, 2007). The pseudo R^2 indicates how useful the factors are in predicting the outcome variable and reflects the amount of variation accounted for by the logistic model, with 1 indicating a perfect model fit. However, the Cox and Snell R^2 is limited in that it cannot reach the maximum value of 1, and is more difficult to interpret (Field, 2009). Hence, the Nagelkerke R^2, a modification of the Cox and Snell with a range of 0 to 1, is generally preferred. The value of 49% indicates that the model is useful in predicting Big Data assimilation.

Predictive Accuracy

Table 10: Classification

	Predicted			
Observed	Initiation	Adoption-decision	Implementation	Percent Correct
Initiation	163	25	3	85.3%
Adoption-decision	38	52	11	51.5%
Implementation	4	18	22	50.0%
Overall Percentage	61.0%	28.3%	10.7%	70.5%

[42] Note that the ordinal logistic regression has a higher -2LL (Appendix G) than MLR (Appendix I.1), indicating that the MLR model has better fit.

One method to determine the *predictive accuracy* of the model is to classify correctly the outcome category in cases for whom the outcome is known (Tabachnick & Fidell, 2007). The classification in Table 10 shows that the model can predict correctly 85.3% of the cases for initiation, 51.5% for adoption-decision, and 50% for implementation. Overall, the model correctly classified 70.5% of all cases.

Likelihood Ratio Tests

The results of the *likelihood ratio tests* in Table 11, can be used to ascertain the significance of predictors in the model (Field, 2009). Three predictor variables are significant at p < 0.01 level (security, competitive pressure, and top management support), and two predictor variables are significant at p < 0.05 level (complexity and IT expertise). Relative advantage is close to being significant (p = 0.075). The likelihood ratio test is an overall statistic that tells which factors significantly predict the outcome category, however, they do not tell specifically what the effect is (Tabachnick & Fidell, 2007).

Table 11: Likelihood ratio tests

	Model Fitting Criteria	**Likelihood Ratio Tests**		
Effect	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	456.421[a]	.000	0	.
Complexity	463.422	7.001	2	.030
Relative advantage	461.601	5.180	2	.075
Security	475.694	19.273	2	.000
IT expertise	468.108	11.687	2	.003
Competitive pressure	528.451	72.030	2	.000
Privacy	457.516	1.095	2	.578
Top management support	522.494	66.073	2	.000
External support	456.811	.390	2	.823
Size	457.991	1.571	2	.456

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

6. Results and Discussion

Grounded in the innovation adoption literature, this thesis has incorporated three theoretical perspectives – DOI, TAM and the TOE framework – to build an integrative research model for studying adoption of Big Data technology at the firm level in Norway. Consistent with factor research in the adoption literature (c.f., King, 1990; Fichman, 1999; Hameed, et al., 2012a), the primary purpose of this study was to identify salient factors affecting the adoption of Big Data technology (RQ_1). The secondary purpose was to reveal whether these factors varied across three aggregated stages of organizational adoption; initiation, adoption-decision, and implementation, collectively referred to as assimilation (RQ_2).

In the multinomial logistic regression (MLR), the relationship between the factors proposed in the research model and the assimilation stages of Big Data technology were assessed. The research model included 11 factors that were hypothesized to affect the assimilation of Big Data technology (Ch. 3.3). However, compatibility and organizational resources were omitted from the analysis, as the measurement items used to operationalize these factors failed to satisfy the retention criteria in the EFA (Ch. 5.3.5). Hence, only nine factors were included in the MLR.

The purpose of this chapter is to use the results of the MLR to answer the research questions stated in Chapter 1.1, and to discuss the implications of these findings. Thus, this chapter first interprets the MLR results to identify which technological, organizational, and environmental factors affect the assimilation of Big Data technology by companies in Norway (Ch. 6.1). It then proceeds to discuss the findings and hypotheses results (Ch. 6.2), before assessing the managerial (Ch. 6.3) and theoretical (6.4) implications of the study. Finally, an evaluation of the study's limitations and potential directions for future research are offered (Ch. 6.5).

6.1 Interpretation of the Multinomial Logistic Regression

There are two main uses of a MLR. First, to predict group membership, by calculating the odds for a company to belong in one of the assimilation stages; initiation, adoption-decision, or implementation. Second, to provide knowledge of the relationship and the strength among variables (Field, 2009). The results of the MLR are presented in Table 12: See Appendix J for a complete model output.

Table 12: Multinomial logistic regression - parameter estimates

	Model 1 *Initiation vs. Adoption-decision[a]*			Model 2 *Adoption-decision vs. Implementation[b]*			Model 3 *Initiation vs. Implementation[a]*		
	B	Exp(B)	Sig.	B	Exp(B)	Sig.	B	Exp(B)	Sig.
Intercept	-.521		.037	-2.217		.000	-2.737		.000
Complexity	.161	1.175	.309	-.559	.572	**.011***	-.398	.672	.091
Relative advantage	.216	1.242	.227	.393	1.481	.131	.609	1.839	**.028***
Security	-.333	.717	**.032***	-.749	.473	**.004***	-1.082	.339	**.000****
IT expertise	.166	1.181	.294	.655	1.925	**.007***	.821	2.273	**.001***
Competitive pressure	1.239	3.453	**.000****	.723	2.060	**.034***	1.962	7.112	**.000****
Privacy	-.055	.947	.720	-.161	.852	.397	-.215	.806	.304
Top management support	1.087	2.965	**.000****	1.083	2.954	**.005***	2.170	8.759	**.000****
External support	-.028	.972	.855	-.118	.889	.598	-.146	.864	.534
[Size=Employees 50-250]	-.379	.684	.210	.112	1.118	.785	-.268	.765	.545
[Size=Employees > 250]	0[c]	.	.	0[c]	.	.	0[c]	.	.

a. The reference category is: Initiation
b. The reference category is: Adoption-decision
c. This parameter is set to zero because it is redundant.
* Significance at < 0.05 level, ** Significance at < 0.01 level

The MLR consists of three models, each comparing companies from two assimilation stages. The first two models compare consecutive stages: Model 1 compares companies within the initiation stage with companies within the adoption-decision stage, while Model 2 compares companies within the adoption-decision stage with companies within the implementation stage. Model 3 differs in that it compares non-consecutive stages; that is, companies within the initiation stage are compared with companies within the implementation stage. For each comparison, the reference category is set to be the group representing the lower level of assimilation. Hence, in this way, each model presents the odds for a company to be in the higher level of assimilation when controlling for the predictor variables.

Statistical Significance

To interpret the MLR results, the Wald statistic (refer to Appendix J for value) is used to assess the statistical significance of each predictor variable in explaining the outcome variable, and indicates whether the β-coefficient (B) for a predictor is significantly different from zero. If so, then the predictor variable is believed to make a significant contribution to the prediction of the outcome (Tabachnick & Fidell, 2007; Hair, et al., 2010). Accordingly, Table 12 shows that six of the nine predictor variables were significant and could distinguish between companies in the assimilation stages ($p < 0.05$). These factors were; **complexity, relative advantage, security, IT expertise, competitive pressure,** and **top management**

support. Noteworthy, complexity and relative advantage were only significant in Model 2 and 3 respectively, while IT expertise was significant in two models (Model 2 and 3), suggesting that companies perceive factors differently in relation to Big Data technology when belonging to different assimilation stages. The non-significant factors; organizational size[43], external support, and privacy did not affect the assimilation of Big Data technology, and will thus not be interpreted at this point.

Direction of Association

The sign of the β-coefficients in Table 12 indicates the direction of the association, which reflects the changes in the outcome variable associated with changes in the predictor variable. A positive β-coefficient means that an increase in the predictor variable is associated with an increase in odds to end up in the later stages of assimilation, and vice versa for a negative relationship (Hair, et al., 2010). As seen in Table 12, complexity and security are negatively associated with Big Data assimilation, while the remaining four significant predictor variables have a positive directional association, consistent with the hypotheses developed in Chapter 3.3.

We observe in **Model 1** (Table 12) that competitive pressure and top management support are positively associated with the adoption-decision of Big Data technology when comparing consecutive assimilation stages (i.e., initiation versus adoption-decision). This suggests that competitive pressure and top management support drive companies' decision to adopt the technology, as an increase in these factors is associated with an increase in odds for a company to be in the adoption-decision stage. The opposite is found with regard to security, which is negatively associated with the adoption-decision. This suggests that security inhibits companies' decision to adopt the technology, as an increase in this factor is associated with a decrease in odds for a company to be in the adoption-decision stage.

In **Model 2** (Table 12), we find that top management support, IT expertise, and competitive pressure are positively associated with the implementation of Big Data technology when comparing consecutive assimilation stages (i.e., adoption-decision versus implementation). This suggests that top management support, IT expertise, and competitive pressure facilitate companies' implementation of the technology, as an increase in these factors

[43] With regard to organizational size in Table 12, large companies are set to be the reference category: With a change in company size from large companies (0) to medium-sized companies (1), the odds for a medium-sized company to be in the latter stage of assimilation is presented.

is associated with an increase in odds for a company to be in the implementation stage. The opposite is found with regard to complexity and security, which are negatively associated with implementation. This suggests that complexity and security inhibits companies' implementation of the technology, as an increase in these factors are associated with a decrease in odds for a company to be in the implementation stage.

Finally, we observe in **Model 3** (Table 12) that relative advantage, top management support, IT expertise, and competitive pressure are positively associated with Big Data implementation when comparing non-consecutive assimilation stages (i.e., initiation versus implementation). This suggests that these factors facilitate companies' implementation of the technology, as an increase in these factors is associated with an increase in odds for a company to be in the implementation stage. The opposite is found with regard to security, which is negatively associated with implementation. This suggests that security inhibits companies' implementation of the technology, as an increase in this factor is associated with a decrease in odds for a company to be in the implementation stage. However, as this model compares non-consecutive assimilation stages, the significant factors do not necessarily offer a meaningful interpretation as it is assumed that companies must undergo the entire assimilation process (i.e., unable to move directly from the initiation stage to the implementation stage). Nonetheless, Model 3 offers a broader perspective on some of the significant factors, such as relative advantage and complexity, which will be discussed further in Chapter 6.2.

Odds Ratios

Table 13: Odds ratio comparison

	Model 1 *Initiation vs. Adoption-decision*[a]		Model 2 *Adoption-decision vs. Implementation*[b]		Model 3 *Initiation vs. Implementation*[a]	
	Odds ratio[c]	Sig.	**Odds ratio**[c]	Sig.	**Odds ratio**[c]	Sig.
Top management support	2.965	.000**	2.954	.005*	8.759	.000**
Competitive pressure	3.453	.000**	2.060	.034*	7.112	.000**
Security	.717	.032*	.473	.004*	.339	.000**
IT expertise			1.925	.007*	2.273	.001*
Relative advantage					1.839	.028*
Complexity			.572	.011*		

a. The reference category is: Initiation
b. The reference category is: Adoption-decision
c. Odds ratio = Exp(B)
* Significance at < 0.05 level, ** Significance at < 0.01 level

To better understand the differential effects that the factors have at each stage of assimilation, the odds ratio will be examined, which is the exponent of the β-coefficients, representing Exp(B) in the model output (Table 12) (Hair, et al., 2010). Table 13 shows an excerpt of Table 12, with odds ratios for the significant factors. The odds ratio is an indicator of the change in odds resulting from a unit change in the predictor variable. For instance, when examining all three models, we observe that a unit increase in competitive pressure will correspond to an increase in odds of being in the latter stages of assimilation with a factor of 3.453, 2.060, and 7.112 for Model 1, 2, and 3 respectively. More specifically, from Model 1 (Table 13), we can say that for a unit increase in competitive pressure, the odds of a company being in the adoption-decision stage are 3.453 times higher than the company being in the initiation stage. Thus, the odds ratio signifies the relative influence a predictor variable has on the outcome variable.

As the odds ratios are the exponents of the β-coefficients, and the exponent of 0 (no effect) is 1, an odds ratio of 1 corresponds to a relationship with no direction. Thus, odds ratios greater than 1 reflect a positive relationship and odds ratios less than 1 represent negative relationships (Hair, et al., 2010). Specifically, the further away the odds ratio is from 1, in any direction, the stronger the influence of the predictor variable. It follows that for a significant factor to have an odds ratio larger than 1 (positive β-coefficient), the companies in the later stages of assimilation must have scored relatively higher on that particular factor than companies in the reference stage (i.e., the lower stage of assimilation of the two stages being compared). For a significant factor with an odds ratio less than 1 (negative β-coefficient), the companies in the later stages of assimilation must have scored relatively lower on that particular factor than companies in the reference stage. A non-significant factor indicates that the perception of that particular factor is not statistically different between companies in two assimilation stages.

6.2 Discussion of Research Findings

For the purpose of testing the research model and corresponding hypotheses, a MLR was utilized to predict the assimilation of Big Data technology for 336 companies in Norway using nine theoretically derived constructs[44] as predictor variables. A test of the full model

[44] Two constructs – compatibility and organizational resources – were omitted during the preceding data analysis.

against a constant only model was statistically significant, indicating that, collectively, the predictor variables reliably distinguished between companies within different assimilation stages. In particular, the research model showed good fit, with Nagelkerke R^2 of 0.49, and a predictive accuracy of 70.5%. The results indicate that the research model is suited for studying organizational adoption of Big Data technology. Moreover, given the scarcity of research into determinants of adoption in the Big Data literature (Salleh & Janczewski, 2016; Rahman, 2016; Chen, et al., 2016), the research model offers a suitable point of departure for future studies on Big Data adoption. Table 14 summarizes the findings regarding the hypotheses.

Table 14: Hypothesis testing

Hypothesis	Model 1 *Initiation vs. Adoption-decision*	Model 2 *Adoption-decision vs. Implementation*	Model 3 *Initiation vs. Implementation*
Technological context			
H1: Relative advantage	Not supported	Not supported	**Supported***
H2: Complexity	Not supported	**Supported***	Not supported
H11: Security	**Supported***	**Supported***	**Supported****
Organizational context			
H4: Organizational size	Not supported	Not supported	Not supported
H5: Top management support	**Supported****	**Supported****	**Supported****
H6: IT expertise	Not supported	**Supported***	**Supported***
Environmental context			
H8: Competitive pressure	**Supported****	**Supported***	**Supported****
H9: External support	Not supported	Not supported	Not supported
H10: Regulatory environment	Not supported	Not supported	Not supported

H3 and H7 were not tested

*Significant level at 0.05, **Significant level at 0.01

The following section discusses each factor in relation to the technological, organizational, and environmental context in which they were presented in the research model.

6.2.1 The Technological Context

As Rogers (1983) argued, adoption of innovations is related to the characteristics of the innovation as perceived by potential adopters. This study posits that security, complexity, and relative advantage will influence the adoption of Big Data technology by firms in Norway. Figure 9 summarizes the findings for the factors within the technological context as presented in Table 12.

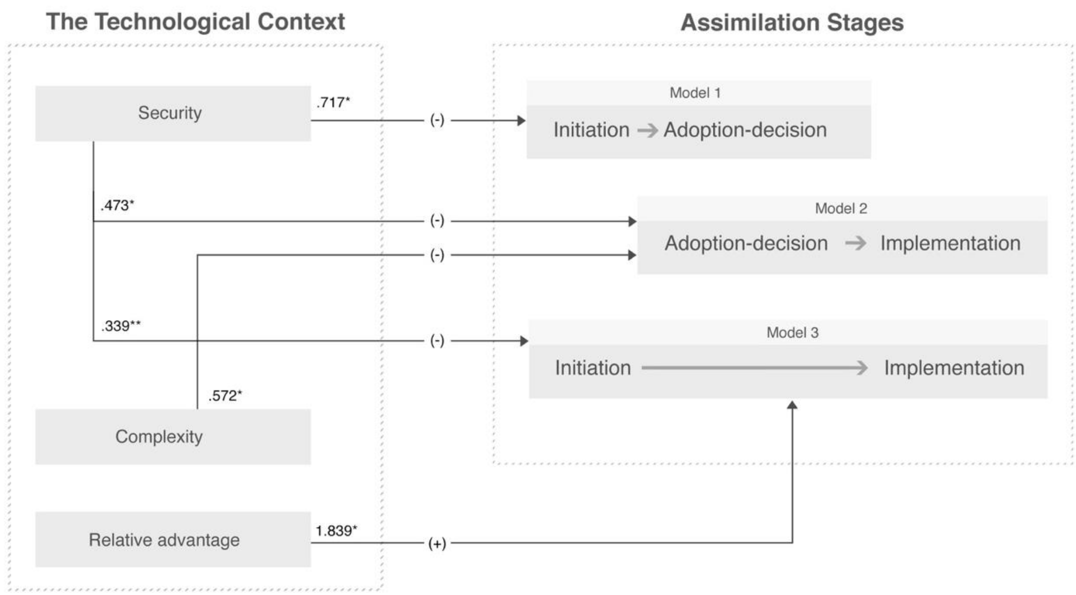


Figure 9: Findings within the technological context

Security

Concerning security, this thesis is one of the first studies to empirically investigate the influence this factor has on the adoption of Big Data technology. The results indicate that there is a significant negative relationship between security and adoption of Big Data technology, which corroborates the hypothesis presented by Salleh and Jancewski (2016). Security is also found as the only significant factor within the technological context to discriminate between firms in all stages of assimilation; Figure 9 illustrates that security is significant in all three models, with odds ratios less than 1. This implies that companies in the initiation stage perceive security as more challenging than companies in the adoption-decision stage (Model 1), and companies in the adoption-decision stage perceive security as more challenging than companies in the implementation stage (Model 2). These findings suggest that the security issues associated with Big Data's unique characteristics are hindering the assimilation of Big Data technology by firms in Norway. Moreover, the odds ratio in Model 1 (0.717) is higher than the odds ratio in Model 2 (0.473), which indicates that security is a stronger inhibitor of implementation than it is for the decision to adopt Big Data technology[45]. This suggests that security concerns are more prevalent for companies in the adoption-decision stage than for companies in the initiation stage. However, given the novelty of the security construct in research on adoption of Big Data, finding external empirical support for these results prove difficult. Nonetheless, the findings (i.e., security is hindering adoption) corroborate current promotional literature on Big Data (e.g., Intel, 2012; CSC, 2013; IDG, 2016).

Complexity

Support was found for a negative relationship between complexity and adoption of Big Data technology, albeit only when comparing the latter stages of assimilation (i.e., adoption-decision versus implementation). Figure 9 illustrates that complexity is only significant in Model 2, with an odds ratio less than 1. This implies that companies within the adoption-decision stage perceive Big Data technology as more difficult to understand and use than companies within the implementation stage. Thus, complexity is found to inhibit the implementation of Big Data technology. Contrary to expectations, however, complexity was not found significant when comparing non-adopters (initiation) with adopters (adoption-decision and implementation). A possible explanation for this finding may be that prior to adoption, at the initiation stage, companies underestimate the role of complexity and are overconfident about the usability of the technology (Arts, et al., 2011). Consequently, when firms eventually acquire the technology, as in the adoption-decision stage, they come to perceive it as considerably more complex than anticipated, which in turn inhibits implementation. Following this reasoning, it is possible to deduce that companies' initial misconception regarding the complexity of Big Data technology may lead to an assimilation gap; where widespread acquisition of the technology is not followed by widespread implementation (Fichman & Kemerer, 1999).

Relative Advantage

Based on the Diffusion of Innovation theory (Rogers, 1983), relative advantage was expected to be the most influential determinant of organizational adoption of Big Data technology. A large body of research has consistently found support for a relationship between

[45] The further the odds ratio is away from 1, in any direction, the greater the influence of the predictor variable.

relative advantage and technology adoption, both at the individual and organizational level (e.g., Rogers, 1983; Iacovou, et al., 1995; Igbaria, et al., 1997; Thong, 1999; Grandon & Pearson, 2004; Quaddus & Hofmeyer, 2007; Chwelos, et al., 2011; Boonsiritomachai, 2014, Al-Isma'ili, 2016). This research, however, found no support for a significant relationship between relative advantage and adoption of Big Data technology for any of the consecutive stages of assimilation. These findings suggest that relative advantage may not be as influential in the context of Big Data as one would expect given the prevalence of the construct in adoption literature. Interestingly, Nam et al. (2015) and Agrawal (2015) also failed to find support for relative advantage when studying adoption of Big Data, which supports the proposition that relative advantage may be redundant in studies of Big Data adoption. As to why relative advantage is insignificant, we offer two plausible explanations consistent with Nam et al. (2015). First, due to the hype surrounding Big Data, the benefits of Big Data technology may be well communicated and widely known by both adopters and non-adopters, making it difficult to distinguish between the two groups. Second, expectations may have fallen short for adopters; particularly for companies that have yet to fully implement the technology. Alternatively, adopters have yet to harness the full potential of the technology. Notwithstanding the above, support for a positive relationship between relative advantage and adoption of Big Data technology was found when comparing non-consecutive assimilation stages, illustrated in Figure 9 (Model 3), with an odds ratio greater than 1. This implies that companies in the initiation stage perceive less benefits from using Big Data technology than companies in the implementation stage, suggesting that companies successfully implementing Big Data technology accrue benefits not apparent to companies in the earliest stages of assimilation.

6.2.2 The Organizational Context

The organizational context refers to the group of intraorganizational factors influencing adoption. This study posits that top management support and IT expertise influence the adoption of Big Data technology by firms in Norway. No support, however, was found for organizational size. Figure 10 summarizes the findings for the factors within the organizational context as presented in Table 12.

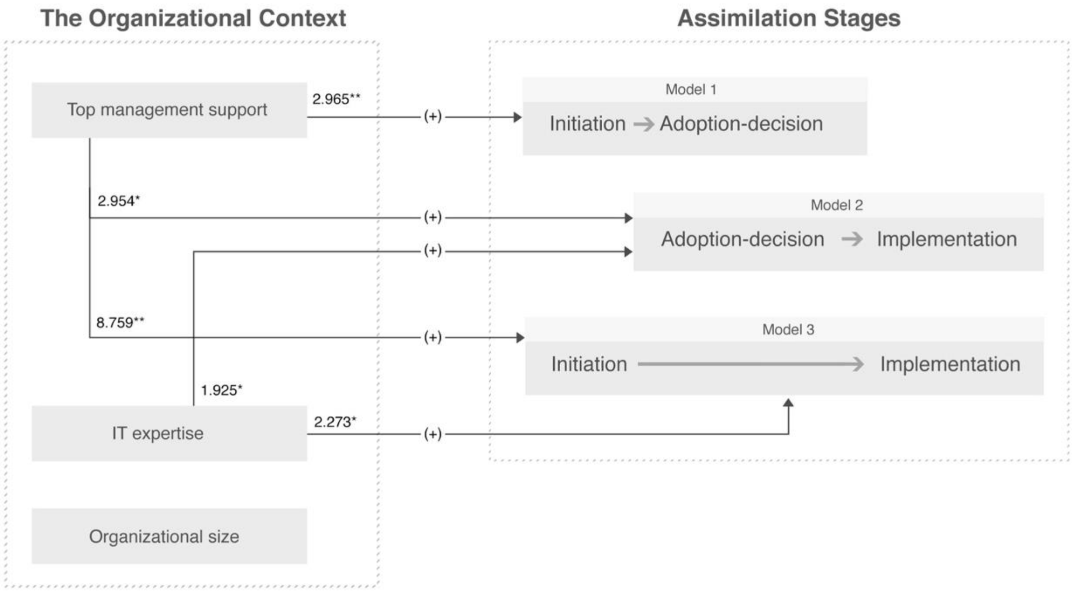


Figure 10: Findings within the organizational context

Top Management Support

Top management support has consistently been found to play a crucial role in the adoption of IT innovations (e.g., Premkumar & Roberts, 1999; Hameed, et al., 2012b; Al-Isma'ili, 2016; Hung, et al., 2016). This thesis extends these findings to the domain of Big Data, as the results indicate a significant positive relationship between top management support and adoption; Figure 9 illustrates that top management support is significant in all three models, with odds ratios greater than 1. This implies that companies in the initiation stage perceive lower top management support than companies in the adoption-decision stage (Model 1), and companies in the adoption-decision stage perceive lower top management support than companies in the implementation stage (Model 2). Moreover, the results suggest that top management support is the only significant factor within the organizational context to discriminate between firms in all stages of assimilation. This implies that top management support is a facilitator of the assimilation of Big Data technology, consistent with another recent study of Big Data adoption (Park, et al., 2015). In addition, the odds ratio in Model 1 (2.965) is nearly identical to the odds ratio in Model 2 (2.954), which indicates that top management support is equally important across every stage of assimilation of Big Data technology. This makes top management support the most consistently important driver of Big Data technology adoption.

IT Expertise

The results offer new insight into the role of employees in the context of organizational adoption of Big Data technology. Figure 10 illustrates that IT expertise is significant in Model 2 and 3, with odds ratios greater than 1. This implies that companies in the initiation stage and adoption-decision stage perceive their IT expertise to be lower than companies in the implementation stage. Surprisingly, there is no support for such a relationship between firms' perception of their IT expertise in the initiation and adoption-decision stages (Model 1). This suggests that higher levels of IT expertise facilitate implementation of Big Data technology, but does not affect the decision to adopt the technology. An alternative interpretation of these results suggests that inadequate levels of IT expertise serve as a barrier to implementation, albeit not as a barrier to adoption-decisions. This implies that firms may be postponing implementation of Big Data technology, rather than acquisition, until they have sufficient internal expertise. These results are somewhat inconsistent with previous literature's tendency to find empirical support for a positive influence of IT expertise on adoption (e.g., Thong, 1999; Hameed, et al., 2012b; Yeh, et al., 2014), which suggest that IT expertise in the context of Big Data may play a greater role in the later stages of assimilation.

Organizational Size

No support was found for organizational size, which was measured in terms of number of employees[46]. As such, organizational size does not appear to influence adoption of Big Data technology. These results are surprising, as our literature review found overwhelming empirical support for the effect of size in adoption studies (Appendix C.2). However, the reviewed studies were frequently performed on small to medium-sized enterprises (SMEs). Thus, to ascertain the role of organizational size in the assimilation of Big Data technology, more research is required as this study is restricted to medium and large businesses.

6.2.3 The Environmental Context

The environmental context refers to the group of interorganizational factors influencing adoption. This study posits that competitive pressure is a critical factor influencing the adoption of Big Data technology by firms in Norway. Privacy and external support,

[46] An alternative measure of organizational size is annual turnover. However, support was not found for organizational size for either operationalization of the construct.

however, were not found to be significant factors. Figure 11 summarizes the findings for the factors within the environmental context as presented in Table 12.

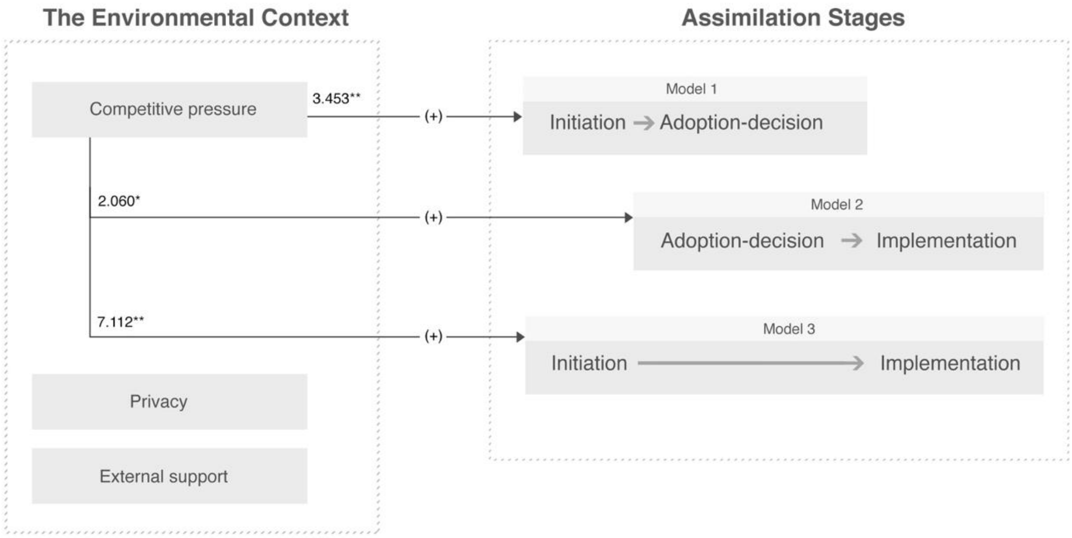


Figure 11: Findings within the environmental context

Competitive Pressure

The strategic rational underlying the relationship between competitive pressure and IT adoption put forth by Porter and Millar (1985) appears relevant to this day. The results of this study indicate a significant positive relationship between competitive pressure and adoption of Big Data technology. In particular, the findings suggest that competitive pressure is the *only* significant factor within the environmental context to discriminate between firms in all stages of assimilation, as illustrated in Figure 11, with odds ratios greater than 1. This implies that companies in the initiation stage perceive less competitive pressure than companies in the adoption-decision stage (Model 1), and companies in the adoption-decision stage perceive less competitive pressure than companies in the implementation stage (Model 2). Competitive pressure is therefore a facilitator of assimilation of Big Data technology, consistent with findings in similar adoption research (e.g., Iacovou, et al., 1999; Malladi & Krishnan, 2013; Nam, et al., 2015). Additionally, the odds ratio in Model 1 (3.453) is higher than the odds ratio in Model 2 (2.060), which indicates that competitive pressure is a stronger driver for the decision to adopt Big Data technology than it is for implementation. This implies that, due to competitive pressure, companies may feel compelled to adopt Big Data technology pre-emptively to avoid a future competitive disadvantage. In this regard, adoption of Big Data technology may be perceived a strategic necessity by firms in Norway.

Privacy

Based on the prevalence of privacy issues in Big Data literature, privacy regulations were expected to have a significant influence on assimilation of Big Data technology. However, no support was found for a relationship between privacy and assimilation. Accordingly, this study was unable to provide support for the hypothesis by Salleh and Janczewski (2016), which proposed that privacy related regulations would inhibit adoption of Big Data technology. The findings therefore suggest that privacy regulations are not hindering the assimilation of Big Data technology by firms in Norway. However, considering the imminent introduction of the General Data Protection Regulation (GDPR), set to replace the current Data Protection Act in Norway in 2018, these results are surprising. The introduction of GDRP will place a greater burden on firms in Norway, and especially on those that intend to utilize Big Data, as they will face an even greater obligation to protect personal identifiable information. Thus, while the adoption of Big Data technology is currently not hindered by privacy regulations, this is susceptible to change with the enactment of GDPR.

External Support

Considering the number of vendors and third-party agencies offering Big Data related services, external support was expected to influence the assimilation of Big Data technology. However, this study found no support for a significant relationship between external support and assimilation of Big Data technology. There are several plausible reasons for the lack of significance in this study. First, the availability of external support could be nearly the same for both adopters and non-adopters, making the factor unable to discriminate between firms in the assimilation stages. A second reason could be attributed to the size of the firms in this study. By examining earlier literature (e.g., Thong, 1999; Premkumar & Roberts, 1999; Al'Isma'ili, 2016), we find that smaller firms, in the absence of internal IT expertise, are believed to perceive external support as a source of external IT expertise. It is possible that the same assertion cannot be made with regard to larger firms, such as those in this study, with relatively better IT expertise. Alternatively, larger firms may find it more reasonable to develop internal IT expertise rather than soliciting services from vendors and third parties. However, given the number of vendors and third-party agencies offering Big Data related services, these findings are surprising and warrants more research.

6.3 Managerial Implications

The results provided in this study have several important implications for practitioners: Both management in adopting organizations and IT suppliers should benefit from understanding the factors influencing adoption of Big Data technology at various stages of assimilation.

Implications for Management

The first conclusion that can be drawn from this study and what it implies for decision-makers, is that top management support is a pivotal factor for all stages of the assimilation of Big Data technology. The role of top management has long been a subject of interest to researchers (e.g., Garrity, 1964), and studies have found that top management support is critical for creating a supportive organizational climate that facilitates receptivity towards innovation adoption (Thong, et al., 1996). Moreover, top management possess the authority to provide and mobilize sufficient organizational resources for motivating, acquiring, and implementing innovations (Premkumar & Roberts, 1999). Managers must therefore acknowledge the vital role they play in allocating resources for adoption and enabling associated activities. According to Young and Jordan (2008), this implies top managers may have to personally accept that they have the most influence on the adoption process. Thus, as top management support is expected to facilitate adoption of Big Data technology, managers must actively engage in every stage of assimilation.

Second, competitive pressure appears to accelerate the adoption of Big Data technology by firms in Norway. The influence of competitive pressure is pervasive in all stages of assimilation, which substantiates the perceived necessity of adopting Big Data technology in today's economic environment. Previous research on organizational adoption have also recognized competitive pressure as an important determinant of adoption (e.g., Premkumar & Roberts, 1999; Zhu, et al., 2004). However, although Big Data technology promises businesses benefits, competitive pressure could instigate a bandwagon effect that precipitates premature decisions to adopt. The bandwagon effect implies that firms would be more likely to use Big Data technology if others within the same industry use it. Specifically, competitive bandwagon pressure can occur because non-adopters fear below-average performance if they perceive competitors to benefit from adopting Big Data technology (Rosenkopf, 1993, p. 487). Managers should realize that successful implementation; the incorporation of the technology into regular activities, may be contingent on having adequate IT expertise to deal with complex

IT tasks, and to ensure a governed and compliant use of data. Managers must therefore carefully consider their needs and capabilities, and avoid the hype surrounding Big Data.

Third, security issues are found to hinder the adoption-decision and implementation of Big Data technology. Management contemplating adoption should be prepared for extensive pushback regarding security concerns, as security needs to be an area which receives constant attention during the entire assimilation process. Moreover, the findings imply that companies lacking compatible security tools and skills required to provide comprehensive data protection are less likely to adopt the technology. This should encourage managers to accompany commitments to Big Data technology with improvements in IT infrastructure, in addition to training and recruitment of IT personnel with security expertise. Since competencies relating to Big Data technologies and security are scarce (Lartigue, 2016), management must also be prepared to call on the skill and services of third parties. Thus, while businesses are racing to adopt Big Data technology, managers should be careful to avoid leaving security an afterthought. Fortunately, one of the segments that fares best in Norway is the cyber security landscape, which is driving investments in security software to keep pace with current security challenges (PAC, 2016b).

Fourth, perceived complexity and lack of IT expertise are highlighted as barriers to implementation of Big Data technology. Concerning complexity, our findings substantiate the notion that the greater the perceived complexity of an IT innovation, the higher the cost of behavioural change becomes following an adoption-decision, which inhibits potential adopters from following through with implementation (Arts, et al., 2011). Similarly, this study finds that poor IT expertise serves as a barrier to implementation, albeit not as a barrier to adoption-decisions. This suggests that managers should strive to cultivate a highly skilled and knowledgeable IT workforce capable of tackling complex IT task. Attention to strengthening internal IT expertise through training and recruiting may prove critical in addressing the barriers to adoption-decisions and implementation presented by security concerns and perceived complexity. However, managers should also be aware that human resources are scarce due to shortages of individuals possessing relevant IT skills (e.g., data scientists), which makes sourcing of experienced talent difficult (Carnelley & Schwenk, 2016c). This suggest the need to develop strong in-house IT expertise through training of current employees. Nonetheless, as the Big Data market expands rapidly (Carnelley & Schwenk, 2016a), the emergence of Big Data technologies not requiring highly specialized, rare technical skills or data scientist may diminish the influence of perceived complexity and lack of IT expertise on adoption in the near future.

Lastly, the perception of the relative advantage offered by Big Data technology was only found statistically different when comparing firms in the earliest stage of assimilation with those in the last stage. This suggest that the benefits of Big Data technology become apparent only after implementing the technology. Accordingly, managers of firms that are just starting their Big Data journey should be aware of the possibility that the technology's performance lags the adoption of the technology.

IT Suppliers

At present, digitalization is a common theme on the agenda of most organizations. Norwegian organizations feel a strong need, magnified by competitive pressure, to take advantage of new and emerging technologies. For businesses, Big Data technology represents both an opportunity and a challenge, and IT suppliers must create clear value propositions that address the needs of the adopters. Research on organizational adoption of technology suggests external support, the availability of support for implementing and using a technology, is a key factor influencing the adoption process. This study, however, finds no support for this claim in Norway. This suggests that IT suppliers may need to review their offerings and be prepared to adapt and evolve their product strategies to better align with the current economic environment and needs of Big Data adopters. Moreover, IT suppliers should look for opportunities to make themselves relevant to the process of assimilating Big Data technology. Specifically, this study reveals that security concerns, lack of IT expertise, and perceived complexity of Big Data technology are hindering adoption-decisions and implementation. IT suppliers and vendors should capitalize on this by providing IT competence, technical support, training, and information to take the burden off the shoulders of adopters, while simultaneously ensuring the adopters have capabilities to benefit from applying the technology in their operations. Efforts are also needed to increase awareness of the types of benefits from Big Data technology, as our findings indicate that these may not be apparent to potential adopters.

Furthermore, IT suppliers should benefit from being prepared to provide advice regarding security, as security concerns will be on the agenda of most adopters. In this regard, IT suppliers could develop a consultancy approach to counteract the security concerns of adopters. Moreover, with the introduction of the European Parliament's General Data Protection Regulation (GDPR), set to replace the current Data Protection Act in Norway on the 25th of May 2018 (Datatilsynet, 2015), this approach could be extended to tackle privacy issues and concerns. Although this study found no support to suggest that privacy issues are inhibiting the adoption of Big Data technology, this is subject to change in the foreseeable future given the imminent introduction of GDRP.

IT suppliers should focus on developing the competencies that are needed and scarce in the market. One approach to meet the many needs of potential adopters involve creating mixed teams of data scientists, experienced IT personnel, and security and data governance experts. For this to be viable, recruitment and training of new employees may be necessary. However, IT suppliers may find it difficult to deliver on all aspects of Big Data, given the scarcity of competencies relating to the technology. An alternative approach involves focusing on offerings that can cement a desirable market position. The faster IT suppliers can differentiate themselves in the marketplace, the more likely they will gain the competencies that are needed and scarce. Along the same vein, IT suppliers may find it useful to develop or join a strong business ecosystem. This would enable specializing suppliers and vendors, with diverse sets of capabilities, collectively to better address needs of the adopters of Big Data technology.

6.4 Theoretical Implications

This thesis represents an early attempt at developing a model for studying the organizational assimilation of Big Data technology. The results in this thesis provide substantive contributions to the Big Data and innovation adoption literature. Firstly, the thesis serves as one of the first theoretically informed, empirical studies of organizational assimilation of Big Data technology, and possibly the first in Norway. Secondly, the study highlights the value of integrating different theoretical perspectives within the TOE framework as applied to the assimilation of Big Data. Thirdly, the findings suggest a need for improved measurements and conceptualizations of factors in the context of Big Data adoption. Finally, the research suggests that there is value in a process-oriented approach to adoption by focusing on assimilation rather than a dichotomous (yes/no) adoption-decision. These implications are discussed in more detail in the following sections.

Organizational Assimilation of Big Data

In utilizing the TOE framework as a theoretical lens to guide this research, this thesis extends the research on Big Data as one the first theoretically informed, empirical studies on organizational assimilation of Big Data technology. In doing so, the research addresses a lack of literature providing insight into factors influencing the adoption and use of Big Data. The existing stream of literature is populated with studies describing emerging tools, technologies, and analytical techniques useful for dealing with the unique characteristics of Big Data, but is sparse with regard to adoption by firms. Although some studies have investigated the adoption of Big Data (e.g., Agrawal, 2015; Nam, et al., 2016; Park et al., 2016), the conceptualization of adoption is often poorly defined and ambiguous, as researchers struggle to judge whether firms have adopted or not. This leads to confusion and potential issues with misinterpretation and misunderstandings of both research models and results, as there is no clear distinction to inform the reader of the boundaries of the research. Moreover, the lack of a consistent and consensual definition of Big Data in the literature exacerbates this issue, as researchers and practitioners differ in understanding of the concept (Stuart & Barker, 2013). To improve upon the current literature, this thesis clearly defines and conceptualizes Big Data and adoption based on extant literature. Specifically, by limiting the scope of the study to Big Data *technology*, a clearly defined component of Big Data (Mauro, et al., 2016), and defining adoption in terms of assimilation (Meyer & Goes, 1988), a concept common in adoption research (e.g., Fichman & Kemerer, 1999; Rai, et al., 2009; McKinnie, 2016), there is little doubt about the boundary conditions of this research. Thus, the results from this study provide researchers with a fresh, holistic perspective on assimilation of Big Data technology from a well-founded theoretical perspective.

Integration of Theoretical Perspectives

This study highlights the value of integrating different theoretical perspectives within the TOE framework for studying organizational adoption. Specifically, the findings in this research suggests that factors from all three TOE contexts; technology, organization, and environment, are of primary importance to the assimilation of Big Data technology. In regard to the traditional theories most commonly utilized in adoption studies – the DOI and TAM – this research highlights the need to study multiple factors in addition to innovation characteristics. Thus, this thesis’ approach to studying adoption at the organizational level by means of integrating individual-level theories with a contextual framework covering organizational and environmental factors, seem applicable and justified in the context of Big Data adoption. Moreover, the findings herein provide support for previous research contending that the most commonly used theoretical models of technology adoption may need to be refined or extended for studying organizational-level adoption (e.g., Meyer & Goes, 1988; Fichman, 1999; Hameed, et al., 2012a). In particular, none of the constructs adapted from DOI and TAM were able to discriminate between firms in all stages of assimilation. This implies that the more broadly generalizable constructs from general technology adoption models may not be applicable to the study of assimilation of Big Data technology. Alternatively, these findings suggest the need to improve measurements and conceptualizations of factors in the context of Big Data adoption.

Improving Measurements and Conceptualization of Factors

The majority of constructs utilized in this thesis were defined in accordance with extant literature and operationalized by measurement items that were proposed and/or validated in IS and IT research. Of the 37 measurement items used to operationalize the constructs in this thesis, 12 items failed to satisfy the empirical criteria for item retention during the factor analysis. Accordingly, these items were dropped to preserve the reliability and validity of the constructs. This highlights the need for improved measurement items and conceptualization of constructs in the context of Big Data adoption. In particular, this thesis was unable to test the hypothesized influence of compatibility and organizational resources, as all of the measurement items used to operationalize these constructs failed to satisfy retention criteria in the exploratory factor analysis. Concerning compatibility, this thesis lends support to previous research calling for a reconceptualization of the construct in innovation adoption studies (e.g., Karahanna, et al., 2006). This study attempted to conceptualize compatibility as a multidimensional construct entailing cognitive compatibility (i.e., values and beliefs), operational compatibility (i.e., operating practices), system compatibility (i.e., IT infrastructure), and data compatibility, albeit unsuccessfully. A reconceptualization of compatibility in the context of Big Data adoption may consider segregating the dimensions of compatibility into different constructs with separate sets of measurement items. This could improve the substantive meaning of the compatibility construct in relation to Big Data. Similar reasoning should also apply to the conceptualization of organizational resources. A more meaningful conceptualization of the construct may entail separating technological, financial, and human resources into distinct constructs. Lastly, the traditional measurement items used to operationalize relative advantage from a utilitarian perspective (e.g., Davis, 1989; Moore & Benbasat, 1991), which relate to performance improvements (i.e., productivity and effectiveness), does not seem to reflect the strategic value that companies impute to Big Data. While relative advantage has consistently been found as one of the top predictors of adoption

(Appendix C.1), this study only found partial support[47] for the construct. This may suggest a need for additional measurement items to operationalize relative advantage in the context of Big Data adoption.

Process Orientation

The results in this thesis offer a compelling argument for a process oriented approach to studying organizational adoption of Big Data by focusing on assimilation as the dependent variable. Rather than studying adoption as a dichotomous variable (yes/no), which is common in adoption studies utilizing the TOE framework (e.g., Iacovou, et al., 1995; Thong, 1999; Premkumar & Roberts, 1999; Agrawal, 2015), this thesis extends the notion of adoption to include initiation and implementation. Specifically, when defining adoption in terms of assimilation – the sequence of stages ranging from a firm's initial awareness and evaluation of a technology, through the formal allocation of resources for its acquisition and deployment, to the incorporation of the technology into the regular activities of the firm – then the dichotomous adoption-decision variable becomes an extremely insensitive measure of adoption. That this study finds differential effects of factors on various stages of assimilation substantiates this notion. Thus, this thesis suggests that a process orientation toward organizational adoption is valuable in the context of Big Data.

6.5 Limitations and Further Research

This study offers new insight into factors affecting adoption of Big Data technology by Norwegian firms at a particularly opportune moment; despite strong growth in the European Big Data technology market, Norway exhibits growth rates lower than the Western European average (Carnelley & Schwenk, 2016a). Thus, insight into the facilitating and inhibiting factors of Big Data technology adoption should be valuable for industry players in Norway. Moreover, this study also provides impetus for future research on Big Data adoption. Future studies could help establish the results' generalizability and contribute to an improved model of organizational adoption of Big Data technology in several ways.

First, this study is based on data from a single country, and while the participants represent a wide variety of industries, the findings are not sufficient as to represent the entire

[47] Relative advantage was only found significant when comparing non-consecutive assimilation stages; firms in the initiation stage versus those in the implementation stage.

international business community. Future research could contribute to ascertain the generalizability of the findings by testing the proposed research model in, preferably multiple, other countries. Additionally, this research was limited to studying medium to large enterprises, which implies findings may not be generalizable to the wider population of businesses. According to the European Commission (2017) and NHD[48] (2012), small and medium-sized enterprises (SMEs) represent more than 99% of all European and Norwegian businesses. Future research may benefit from studying a greater variety of organizations in terms of size. This may achieve clarification on the role of organizational size as a factor in future studies of Big Data adoption.

Second, the data collection in this study focused on a single key representative (i.e., respondent) for each firm. While this approach is common in organizational research, all the responses represent the perspective of the executive management. A long-standing question in the innovation literature is "who should rate the innovation if the potential adopter is not an individual, but an organization?" (Tornatzky & Klein, 1982, p. 41). Future studies should consider obtaining responses from multiple business units, to identify which persons' responses have the best predictive utility in Big Data adoption research. Specifically, future research should gather responses from both IT managers and non-IT managers within the same firm. Although this study found no significant response bias linked to the executive positions of the respondents[49], which corroborate the findings of Zhu et al. (2006a), future research with multiple representatives for each firm might provide additional validity to the research findings.

Third, while a survey strategy was chosen for this research to encourage replicability and future cross-study comparability, an in-depth case study of Big Data adoption could provide additional insight. A case study of Norwegian firms may serve to validate the factors influencing adoption of Big Data technology found herein, and offer a more profound understanding of the role of each factor for adoption. Security in particular, as a novel construct developed specifically for this study, could benefit from further research and theoretical development. There is also an abundance of TOE factors available to researchers, and the case study approach could be useful in identifying factors relevant to Big Data adoption. Several factors not covered in this thesis may be relevant to include in future studies.

[48] The Royal Norwegian Ministry of Trade and Industry (Norwegian: Nærings – og handelsdepartementet).

[49] Of ten constructs, only *top management support* was perceived statistically different by IT and non-IT managers (Ch. 5.2).

For instance, future research may want to examine top management's perspective on the strategic value of Big Data technology, as perception of strategic value has been found to influence the approval and funding of information systems (e.g., Grandon & Pearson, 2004). Subramanian and Nosek (2001) identified three factors creating strategic value in information systems; operational support, managerial productivity, and strategic decision aid, which should be applicable to Big Data adoption. Moreover, a qualitative case study would be useful in reconceptualising some of the more generic constructs – such as compatibility, organizational resources, and relative advantage – to ensure applicability in the context of Big Data adoption. Thus, while this thesis advocated for the use of a survey strategy, the case study approach should be complementary in advancing research on adoption of Big Data.

Lastly, this study was limited by the time constraints inherent to writing a master thesis. Future studies on organizational adoption of Big Data may benefit from utilizing a longitudinal research design. A study with the capacity to track change and development over time could follow the entire Big Data adoption process through each stage of assimilation, from initial awareness through implementation and routinization. Whereas this research employed a three-stage assimilation process, longitudinal studies may find it feasible to study a less aggregated process; perhaps even the nine-step assimilation process originally conceptualized by Meyer and Goes (1988). Such an approach would advance our understanding of how Big Data technology is diffused within organizations and the factors which influence this process.

6.6 Conclusion

Big Data presents many exciting opportunities as well as formidable challenge. As investments in Big Data increase, understanding the organizational adoption of Big Data technology is crucial and timely. With the Norwegian market for Big Data technology expected to exhibit growth rates lower than the Western European average (Carnelley & Schwenk, 2016a), the main objective of this study has been to identify factors affecting the adoption of Big Data technology by companies in Norway. Grounded in the innovation adoption literature, this study identified 11 determinants of adoption under three broad contexts (technology, organization, and environment) and evaluated their influence on the three stages of organizational adoption of Big Data technology – initiation, adoption-decision, and implementation – collectively referred to as assimilation.

A survey was developed to measure the 11 determinants and data was collected from executives in 336 medium to large organizations in Norway. The preliminary analysis of the data indicated that more than half the sample (191) have yet to adopt Big Data technology, which reinforces the notion that companies in Norway are still in the early stages of adoption. A multinomial logistic regression was used to assess the relationship between the proposed determinants and the assimilation stages of Big Data technology; wherein relative advantage, top management support, IT expertise, and competitive pressure were found to facilitate adoption, while complexity and security concerns were found to inhibit adoption of Big Data technology by companies in Norway.

The findings suggest that competitive pressure will presumably drive more companies into adopting the technology, and that companies contemplating adoption must solicit the support of top management, as their active engagement plays a critical role throughout the adoption process. Furthermore, although this study finds that higher levels of IT expertise facilitate adoption, the results also demonstrate that the lack thereof may inhibit implementation of the technology. At the same time, companies face considerable challenges with adoption-decisions and implementation due to complexity and security concerns – possibly exacerbated by a lack of internal IT expertise. This suggest that companies should strive to strengthen internal IT expertise through training and recruiting in order to tackle complex IT tasks and counteract security concerns related to Big Data technology. Thus, whereas managers in adoption firms must recognize and be prepared for setbacks with implementation of the technology, IT suppliers could capitalize on this by offering services that strengthens the adopting firms' capabilities and IT expertise. The results from this study also indicate that the technology's performance may lag the adoption of the technology, so that the true benefits of adopting Big Data technology do not materialize until the technology is implemented and incorporated into the daily activities of the firm.

This study finds that the proposed research model is suited for studying organizational adoption of Big Data technology. Moreover, given the scarcity of research into determinants of adoption in Big Data literature (Salleh & Janczewski, 2016; Rahman, 2016; Chen, et al., 2016), the research model offers a suitable point of departure for future studies on Big Data adoption.

References

Agarwal, R. & Karahanna, E., 2000. Time Flies When You're Having Fun: Cognitive Absorption and Beliefs about InformationTechnology Usage. *MIS quarterly,* 24(4), pp. 665-694.

Agrawal, K. P., 2015. *Investigating the Determinants of Big Data analytics (BDA) Adoption in Asian Emerging Economies,* Patna: Chandragupt Institute of Management.

Alam, S. S. & Noor, M. K. M., 2009. ICT Adoption in Small and Medium Enterprises: an Empirical Evidence of Service Sectors in Malaysia. *International Journal of Business and Management,* 4(2), pp. 112-125.

Ali, A. et al., 2016. Big data for development: applications and techniques. *BioMed Central,* 1(2), pp. 2-14.

Al-Isma'ili, S., Li, M., Shen, J. & He, Q., 2016. Cloud computing adoption determinants: An analysis of australian SMES. *PACIS 2016 proceedings,* pp. 1-17.

Alrousan, M. K., 2014. *E-commerce Adoption by Travel Agencies in Jordan,* Cardiff: Cardiff Metropolitan University.

Alshboul, Y., Wang, Y. & Nepali, R., 2015. *Big Data LifeCycle: Threats and Security Model.* Puerto Rico, Dakota State University.

Al-Tarawneh, H. A., 2012. The Main Factors Beyond Decision Making. *Journal of Management Research,* 4(1), pp. 1-23.

Armstrong, S. J. & Overton, T. S., 1977. Estimating nonresponse bias in mail surveys. *Journal of Marketing Research,* 14(3), pp. 396-402.

Arts, J., Frambach, R. T. & Bijmolt, T., 2011. Generalizations on consumer innovation adoption: a meta-analysis on drivers of intention and behavior. *International Journal of Research in Marketing,* 28(2), pp. 134-144.

Atif, A. & Richards, D., 2012. *Estimating Non-Response Bias in a Web-based Survey of Technology Acceptance: A Case Study of Unit Guide Information Systems.* Greelong, 23rd Australasian Conference on Information Systems.

Baglin, J., 2014. Improving Your Exploratory Factor Analysis for Ordinal Data: A Demonstration Using FACTOR. *Practical Assessment, Research & Evaluation,* 19(5), pp. 1-15.

Bagozzi, R. & Yi, Y., 1991. Multitrait–multimethod matrices in consumer research. *Journal of Consumer Research,* 17(4), pp. 426-439.

Bartlett, J., Kotrlik, J. & Higgins, C., 2001. Organizational Research: Determining Appropriate Sample Size in Survey Research. *Information Technology, Learning, and Performance Journal*, 19(1), pp. 43-50.

Blaikie, N., 2010. *Designing Social Reserach.* 2nd ed. Cambridge: Polity Press.

Bloomberg, 2014. *Target's Data Breach: The Largest Retail Hack in U.S. History.* [Online]
Available at: https://www.bloomberg.com/graphics/infographics/target-data-breach.html
[Accessed 13 June 2017].

Bollier, D., 2010. *The Promise and Peril of Big Data,* Washington, DC: The Aspen Institute.

Boonsiritomachai, W., 2014. *Enablers affecting the adoption of Business Intelligence: A study of Thai small and medium-sized enterprises,* Melbourne: College of business: Victoria University.

Bouchard, L., 1993. *Decision Criteria in the Adoption of EDI.* Laval, Association for Information Systems, pp. 365-376.

Bourgeois, L., 1981. On the Measuremen of Organizational Slack. *Academy of Management,* 6(1), pp. 29-39.

Brreg, 2017. *The Brønnøysund Register Centre.* [Online]
Available at: https://www.brreg.no/home/
[Accessed 13 May 2017].

Brynjolfsson, E., Hitt, L. M. & Heekyung, K., 2011. *Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance?,* Massachusetts: SSRN.

Buonanno, G., Faverio, P., Pigni, F. & Tagliavini, M., 2005. Factors affecting ERP system adoption: a comparative analysis between SMEs and large companies. *Journal of Enterprise Information Management,* 18(4), pp. 384-426.

Caldeira, M., 1998. *Understanding the adoption and use of information systems/information technology in small and medium-sized manufacturing enterprises: A study in Portuguese industry,* Bedfordshire: Cranfield University.

Carnelley, P. & Schwenk, H., 2016a. *Western Europe Big Data Technology and Services Forecast, 2015-2019,* Framingham: IDC.

Carnelley, P. & Schwenk, H., 2016b. *Western Europe Business Analytics Software Forecast, 2016-2020,* Framingham: IDC.

Carnelley, P. & Schwenk, H., 2016c. *Self-Service Data Science: A Strategy for Bridging the Skills Gap,* Framingham: IDC.

Cartell, R. B., 1966. The Scree Test For The Number Of Factors. *Multivariate Behavioral Research,* 1(2), pp. 245-276.

Chau, P. & Tam, K., 1997. Factors Affecting the Adoption of Open Systems: An Exploratory Study. *MIS quarterly,* 21(1), pp. 1-24.

Chaveesuk, S., 2010. *The Determinants of the Adoption and Application of Business Intelligence: An ERP Perspective,* Victoria University: Faculty of Business and Law.

Chen, H., Chiang, R. H. L. & Storey, V. C., 2012. Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly,* 36(4), pp. 1165-1188.

Chen, Y. et al., 2016. Big data analytics and big data science: a survey. *Journal of Management Analytics,* 3(1), pp. 1-42.

Chong, A., Lin, B., Ooi, K.-B. & Raman, M., 2009. Factors affecting the adoption level of c-commerce: An empirical study. *Journal of Computer Information Systems*, 50(2), pp. 13-22.

Choudrie, J. & Dwivedi, Y., 2005. Investigating the Research Approaches for Examining Technology Adoption Issues. *Journal of Research Practice,* 1(1), pp. 1-12.

Chwelos, P., Benbasat, I. & Dexter, A. S., 2001. Research Report: Empirical Test of an EDI Adoption Model. *Information Systems Research,* 12(3), pp. 304-321.

Cooper, R. & Zmud, R., 1990. Information Technology Implementation Research: A Technology Diffusion Approach. *Management Science,* 36(2), pp. 123-139.

Costello, A. B. & Osborne, J. W., 2005. Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Practical Assessment, Research & Evaluation,* 10(7), pp. 1-9.

Court, D., 2015. *Getting big impact from big data.* [Online]
Available at: http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/getting-big-impact-from-big-data
[Accessed 25 April 2017].

CSC, 2013. *Big Data & Analytics - From Sensory Overload to Predictable Outcome,* UK: CSC Marketing & Communications Department.

Damanpour, F., 1992. Organizational Size and Innovation. *Organization Studies,* 13(3), pp. 375-402.

Daniel, E. & Grimshaw, D., 2002. An exploratory comparison of electronic commerce adoption in large and small enterprises. *Journal of Information Technology,* 17(3), pp. 133-147.

Datatilsynet, 2013. *Big Data - privacy principles under pressure,* Oslo: Datatilsynet.

Datatilsynet, 2015. *Hva betyr de nye personvernreglene for din virksomhet.* [Online]
Available at: https://www.datatilsynet.no/Regelverk/EUs-personvernforordning/hva-betyr/
[Accessed 27 April 2017].

Davenport, T. H., 2013. *Analytics 3.0.* [Online]
Available at: https://hbr.org/2013/12/analytics-30
[Accessed 4 April 2017].

Davenport, T. H. & Patil, D. J., 2012. *Data Scientist: The Sexiest Job of the 21st Century.* [Online]
Available at: https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century
[Accessed 22 April 2017].

Davis, F. D., 1986. *A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results,* New England: Massachusetts Institute of Technology.

Davis, F. D., 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS quarterly,* 13(3), pp. 319-3440.

Davis, F. D., 1993. User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. *International Journal of Man-Machine Studies,* 38(3), pp. 475-487.

Davis, F. D., Bagozzi, R. P. & Warshaw, P. R., 1989. User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management science,* 35(8), pp. 982-1003.

DeSimone, J. A., Harms, P. D. & DeSimone, A. J., 2015. Best practice recommendations for data screening. *Journal of Organizational Behavior,* 36(2), pp. 171-181.

Dijcks, J.-P., 2013. *Oracle: Big Data for the Enterprise,* Redwood Shores: Oracle Corporation.

Dimick, D. & Murray, V., 1978. Correlates of Substantive Policy Decisions in Organizations: The Case of Human Resource Management. *Academy of Management,* 21(4), pp. 611-623.

DiStefano, C., Zhu, M. & Mîndrilă, D., 2009. Understanding and Using Factor Scores: Considerations for the Applied Researcher. *Practical Assessment, Research & Evaluation,* Volume 14, pp. 1-11.

Draugalis, J. R. & Plaza, C. M., 2009. Best Practices for Survey Research Reports Revisited: Implications of Target Population, Probability Sampling, and Response Rate. *American Journal of Pharmaceutical Education,* 73(8).

EC, 2017. *What is an SME?*. [Online]
Available at: http://ec.europa.eu/growth/smes/business-friendly-environment/sme-definition_en
[Accessed 29 May 2017].

Eichhorn, B. R., 2014. *The Impact of User Involvement on Information System Projects,* Ohio: Cleveland State University.

EMC, 2013. *The evolving role of data in decision-making,* London: Economist Intelligence Unit.

Ettlie, J., 1990. What Makes a Manufacturing Firm Innovative?. *Academy of Management,* 4(4), pp. 7-20.

EU, 2003. *Commission Recommendation of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises.* [Online]
Available at: http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32003H0361
[Accessed 10 May 2017].

European Commission, 2016. *The EU Data Protection Reform and Big Data,* Brussel: European Commission.

EY, 2014. *Big Data - Changing the way businesses compete and operate,* Global: Ernst & Young Global Limited.

Fabrigar, L. R., MacCallum, R. C., Wegener, D. T. & Strahan, E. J., 1999. Evaluating the Use of Exploratory Factor Analysis in Psychological Research. *Psychological Methods,* 4(3), pp. 272-299.

Fan, W. & Bifet, A., 2012. Mining Big Data: Current Status, and Forecast to the Future. *ACM SIGKDD Explorations Newsletter,* 14(2), pp. 1-5.

Fichman, R. G., 1992. *Information Technology Diffusion: A review of Empirical Research.* Massachusetts, International Conference of Information Systems, pp. 195-206.

Fichman, R. G., 1999. The Diffusion and Assimilation of Information Technology Innovations. In: R. Zmud, ed. *Framing the Domains of IT Management: Projecting the Future Through the Past.* Cincinnati: Pinnaflex Educational Resources, Inc.

Fichman, R. G., 2001. The Role of Aggregation in the Measurement of IT-Related Organizational Innovation. *MIS Quarterly,* 25(4), pp. 427-455.

Fichman, R. & Kemerer, C., 1997. The assimilation of software process innovations: An organizational learning perspective. *Management Science,* 43(10), pp. 1345-1363.

Fichman, R. & Kemerer, C., 1999. The Illusory Diffusion of Innovation: An Examination of Assimilation Gaps. *Information Systems Research,* 10(3), pp. 255-275.

Field, A., 2009. *Discovering Statistics Using SPSS.* 3rd ed. London: SAGE Publications.

Filkins, B., 2015. *Enabling Big Data by Removing Security and Compliance Barriers,* United States: SANS Institute.

Fishbein, M. & Ajzen, I., 1975. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research.* 1st ed. Massachusetts: Addison-Wesley Publishing Company.

Fisher, C., 2016. *LinkedIn Unveils The Top Skills That Can Get You Hired In 2017, Offers Free Courses for a Week.* [Online]
Available at: https://blog.linkedin.com/2016/10/20/top-skills-2016-week-of-learning-linkedin
[Accessed 22 April 2017].

Galliers, R. & Land, F., 1987. Choosing appropriate information systems research methodologies. *Communications of the ACM,* 30(11), pp. 900-902.

Gantz, J. & Reinsel, D., 2011. *Extracting Value from Chaos,* Framingham: IDC iView.

Garrity, J. T., 1964. Top Management and Computer profits. *Harvard Business Review.*

Gartner, 2013. *Gartner's 2013 Hype Cycle for Emerging Technologies Maps Out Evolving Relationship Between Humans and Machines.* [Online]
Available at: http://www.gartner.com/newsroom/id/2575515
[Accessed 23 April 2017].

Gartner, 2016. *Gartner Survey Shows More Than 75 Percent of Companies Are Investing or Planning to Invest in Big Data in the Next Two Years.* [Online]
Available at: http://www.gartner.com/newsroom/id/3130817
[Accessed April 2017].

Gaskin, J., 2017a. *Data screening.* [Online]
Available at: http://statwiki.kolobkreations.com/index.php?title=Data_screening
[Accessed 12 May 2017].

Gaskin, J., 2017b. *Confirmatory Factor Analysis.* [Online]
Available at:
http://statwiki.kolobkreations.com/index.php?title=Confirmatory_Factor_Analysis#Harman.E2.80.99s_single_factor_test
[Accessed 13 May 2017].

Gibbs, J. L. & Kraemer, K. L., 2004. A Cross-Country Investigation of the Determinants of Scope of E-commerce Use: An Instituational Approach. *Electronic Markets,* 14(2), pp. 124-137.

Gjerde, A. S., 2017. *Stor vekst i antall søkere til it-studier.* [Online]
Available at: https://www.dn.no/talent/2017/04/20/1122/Utdannelse/stor-vekst-i-antall-sokere-til-it-studier
[Accessed 21 April 2017].

Gnatovich, R., 2006. *Business Intelligence Versus Business Analytics - What's the difference?.* [Online]
Available at: http://www.cio.com/article/2448992/business-intelligence/business-intelligence-versus-business-analytics--what-s-the-difference-.html
[Accessed 5 April 2017].

Goes, P. B., 2014. Editor's comments: Big Data and IS research. *MIS Quarterly*, 38(3), pp. 3-8.

Grandon, E. E. & Pearson, J. M., 2004. Electronic commerce adoption: an empirical study of small and medium US businesses. *Information & Management*, Volume 42, pp. 197-216.

Greenberg, A., 2015. *Google Lost Data After Lightning Hit Its Data Center in Belgium.* [Online]
Available at: http://time.com/4004192/google-data-lightning-belgium/
[Accessed 25 April 2017].

Green, E., 2016. *What are the most-cited publications in the social sciences (according to Google Scholar)?.* [Online]
Available at: http://blogs.lse.ac.uk/impactofsocialsciences/2016/05/12/what-are-the-most-cited-publications-in-the-social-sciences-according-to-google-scholar/
[Accessed 27 February 2017].

Hair, J. J. F., Black, W. C., Babin, B. J. & Anderson, R. E., 2010. *Multivariate Data Analysis.* 7th ed. United States: Pearson.

Hallberg, J. & Ahorlu, M., 2016. *Nordic Services Forecast, 2016-2020,* Framingham: IDC.

Hameed, M. A., Counsell, S. & Swift, S., 2012a. A conceptual model for the process of innovation adoption in organizations. *Journal of Engineering and Technology Management*, Volume 29, pp. 358-390.

Hameed, M. A., Counsell, S. & Swift, S., 2012b. A meta-analysis of relationships between organizational characteristics and IT innovation adoption in organizations. *Jounal of Information and Management*, 49(5), pp. 218-232.

Hannan, M. & Freeman, J., 1984. Structural Inertia and Organizational Change. *American Sociological Review*, 49(2), pp. 149-164.

Hashem, I. A. T. et al., 2015. The rise of "big data" on cloud computing: Review and open research issues. *Information Systems*, Volume 47, pp. 98-115.

Heijden, V. D., 2004. User Acceptance of Hedonic Information Systems. *MIS Quarterly,* 28(4), pp. 695-704.

Heudecker, N. & Lisa, K., 2014. *Survey Analysis: Big Data Investment Grows but Deployments Remain Scarce in 2014,* Stamford: Gartner Inc..

Hosmer, D. W., Lemeshow, S. J. & Rodney, R. X., 2013. *Applied Logistic Regression.* 3rd ed. Massachusetts: John Wiley & Sons.

Hoti, E., 2015. The technological, organizational and environmental framework of IS innovation adaption in small and medium enterprises. Evidence from research over the last 10 years. *International Journal of Business and Management,* Volume 4, pp. 1-14.

Huang, J. L. et al., 2011. Detecting and Deterring Insufficient Effort Responding to Surveys. *Journal of Business and Psychology,* 27(1), pp. 99-114.

Hung, S. et al., 2016. *Factors influencing business intelligence systems implementation success in the enterprises.* Chiayi, PACIS 2016 Proceedings.

Iacovou, C. L., Benbasat, I. & Dexter, A. S., 1995. Electronic Data Interchange and Small Organizations: Adoption and Impact of Technology. *MIS quarterly,* 19(4), pp. 465-485.

IDG, 2016. *Data & Analytics Survey,* Framingham: IDG Enterprise.

Ifinedo, P., 2011. An Empirical Analysis of Factors Influencing Internet/E-Business Technologies Adoption by SMEs in Canada. *International Journal of Information Technology and Decision Making,* 10(4), pp. 731-766.

Igbaria, M., Zinatelli, N., Cragg, P. & Cavaye, A. L., 1997. Personal Computing Acceptance Factors in Small Firms: A Structural Equation Model. *MIS quarterly,* 21(3), pp. 279-305.

Imhoff, C. & White, C., 2011. *Self-Service Business Intelligence - Empowering Users to Generate Insights,* Washington: TDWI.

Intel, 2012. *Big Data Analytics - Intel's IT Manager Survey on How Organizations are Using Big Data,* California: Intel IT Center.

Jain, P., Gyanchandani, M. & Khare, N., 2016. Big data privacy: a technological perspective and review. *Journal of Big Data,* 3(1), pp. 3-25.

Jeon, B. N., Han, K. S. & Lee, M. J., 2006. Determining factors for the adoption of e-business: the case of SMEs in Korea. *Applied Economics,* Volume 38, pp. 1905-1916.

Jin, X., Wah, B. W., Cheng, X. & Wang, Y., 2015. Significance and Challenges of Big Data Research. *Big Data Research,* 2(2), pp. 50-64.

Johannessen, T. V., 2017. *SOL17 - Business Intelligence: Introduksjon,* Bergen: Norges Handelshøyskole.

Jolliffe, I., 1972. Discarding Variables in a Principal Component Analysis. I: Artificial Data. *Journal of the Royal Statistical Society. Series C (Applied Statistics),* 21(2), pp. 160-173.

Karahanna, E., Agawal, R. & Angst, C., 2006. Reconceptualizing Compatibility Beliefs in Technology Acceptance Research. *MIS Quarterly,* 30(4), pp. 781-804.

Khanna, A., 2016. *Why Cloud Analytics is Better than Analytics.* [Online]
Available at: http://data-informed.com/why-cloud-analytics-is-better-analytics/
[Accessed 27 April 2017].

King, N., 1990. Innovation at work: The research literature. In: M. West & J. Farr, eds. *Innovation and creativity at work: Psychological and organizational strategies.* Oxford, England: John Wiley & Sons, pp. 15-59.

King, W. R. & He, J., 2006. A meta-analysis of the technology acceptance model. *Information & management,* Volume 43, pp. 740-755.

Kline, T., Sulsky, L. & Rever-Moriyama, S., 2000. Common Method Variance and Specification Errors: A Practical Approach to Detection. *Journal of Psychology,* 134(4), pp. 402-421.

Krejcie, R. & Morgan, D., 1970. Determining sample size for research activities. *Educational and Psychological Measurement,* Volume 30, pp. 607-610.

Kulviwat, S. et al., 2007. Toward a Unified Theory of Consumer Acceptance Technology. *Psychology & Marketing,* 24(12), pp. 1059-1084.

Kwon, T. & Zmud, R., 1987. Unifying the fragmented models of information systems implementation. In: *Critical Issues in Information Systems Research.* New York: John Wiley & Sons, pp. 227-251.

Laney, D., 2001. *3D Data Management: Controlling Data Volume, Velocity, and Variety,* Stamford: META Group Inc..

Lartigue, E., 2016. *Cyber Security & Big Data,* Paris: PAC SITSI.

Lee, M. & Cheung, C., 2004. Internet Retailing Adoption by Small-to-Medium Sized Enterprises (SMEs): A Multiple-Case Study. *Information Systems Frontiers,* 6(4), pp. 385-397.

Lee, Y., Kozar, K. A. & Larsen, K. R. T., 2003. The technology acceptance model: past, present, and future. *Communications of the Association for Information Systems,* 12(50), pp. 752-780.

Legris, P., Ingham, J. & Collerette, P., 2003. Why do people use information technology? A critical review of the technology acceptance model. *Information & Management,* Volume 40, pp. 191-204.

Le, H. V., Rowe, F. & Truex, D., 2012. An Empirical Study of Determinants of E-commerce Adoption in SMEs in Vietnam an Economy in Transition. *Journal of Global Information Management,* 20(3).

Lewin, K., 1947. Group Decisions and Social Change. In: T. Newcomb & E. Hartley, eds. *Readings in Social Psychology.* New York: Henry Holt and Co., pp. 340-344.

Lewis, E. F., Hardy, M. & Snaith, B., 2013. Estimating the Effect of Nonresponse Bias in a Survey of Hospital Organizations. *Eval Health Prof,* 36(3), pp. 330-351.

Lindner, J. R., Murphy, T. H. & Briers, G. E., 2001. Handling nonresponse in social science research. *Journal of Agricultural Education,* 42(4).

Ling, C. Y., 2001. *Model of Factors Influences on Electronic Commerce Adoption and Diffusion in Small- & Medium-Sized Enterprises.* Australia, School of Information Systems Curtin University of Technology.

Liu, Y. & Zumbo, B. D., 2007. The Impact of Outliers on Cronbach's Coefficient Alpha Estimate of Reliability. *Educational and Psychological Measurement,* 67(4), pp. 620-634.

Li, X., Troutt, M. D., Brandyberry, A. & Wang, T., 2011. Decisions Factors for the Adoption and Continued Use of Online Direct Sales Channels among SMEs. *Journal of the Association for Information Systems,* 12(1), pp. 1-31.

Lord, N., 2017. *What is GDPR (General Data Protection Regulation)? Understanding and Complying with GDPR Data Protection Requirements.* [Online]
Available at: https://digitalguardian.com/blog/what-gdpr-general-data-protection-regulation-understanding-and-complying-gdpr-data-protection
[Accessed 24 April 2017].

Lucchetti, R. & Sterlacchini, A., 2004. The Adoption of ICT among SMEs: Evidence from an Italian Survey. *Small Business Economics,* 23(2), pp. 151-168.

Macaskill, E. & Dance, G., 2013. *NSA files: decoded - What the revelations mean for you.* [Online]
Available at: https://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded#section/1
[Accessed 13 June 2017].

MacKenzie, S. & Podsakoff, P., 2012. Common Method Bias in Marketing: Causes, Mechanisms, and Procedural Remedies. *Journal of Retailing*, 88(4), pp. 542-555.

Malladi, S. & Krishnan, M. S., 2013. *Business Intelligence and Analytics in Organizations - An Empirical Analysis.* Milan, International conference on Information Systems.

Manyika, J. et al., 2011. *Big Data: The next frontier for innovation, competition, and productivity,* United States: McKinsey Global Institute.

Ma, Q. & Liu, L., 2004. The Technology Acceptance Model: A Meta-Analysis of Empirical Findings. *Journal of Organizational and End User Computing,* 16(1), pp. 59-72.

Matsunaga, M., 2010. How to Factor-Analyze Your Data Right: Do's, Don'ts, and How-To's.. *International Journal of Psychological Research,* 3(1), pp. 97-110.

Mauro, A. D., Greco , M. & Grimaldi, M., 2016. A formal definition of Big Data based on its essential features. *Library Review,* 65(3), pp. 122-135.

Mayer-Schönberger, V. & Cukier, K., 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think.* Boston, MA: Houghton Mifflin Harcourt.

McAfee, A. & Brynjolfsson, E., 2012. Big Data: The Management Revolution. *Harvard Business Review,* 90(10), pp. 60-68.

McDonald, J. H., 2014. *Handbook of Biological Statistics.* 3rd ed. Maryland: Sparky House Publishing.

McKinnie, M., 2016. *Cloud Computing: TOE Adoption Factors By Service Model In Manufacturing,* Atlanta: Georgia State University.

Meade, A. W. & Bartholomew, C. S., 2012. Identifying Careless Responses in Survey Data. *American Psychological Association,* 17(3), pp. 437-455.

Meyer, A. & Goes, J., 1988. Organizational Assimilation of Innovations: A Multilevel Contextual Analysis. *The Academy of Management Journal,* 31(4), pp. 897-923.

Microsoft, 2013. *The Big Bang: How the Big Data Explosion Is Changing the World.*
[Online]
Available at: https://news.microsoft.com/2013/02/11/the-big-bang-how-the-big-data-explosion-is-changing-the-world/#sHQLLiMFfHVy4ii5.97
[Accessed 13 June 2017].

Miele, S. & Shockley, R., 2013. *Analytics: The real-world use of big data,* New York: IBM Institute for Business Value.

Miller, L. & Smith, K., 1983. Handling nonresponse issues. *Journal of Extension,* pp. 45-50.

Moore, G. & Benbasat, I., 1991. Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation. *Information Systems Research,* 2(3), pp. 192-222.

Nam, D., Kang, D. & Kim, S., 2015. *Process of big data analysis adoption: Defining big data as a new IS innovation and examining factors affecting the process.* Hawaii, Hawaii International Conference on Systems Sciences.

Narayanan, A. & Shmatikov, V., 2007. *How To Break Anonymity of the Netflix Prize Dataset,* Austin: Cornell University Library.

Neuman, W., 2014. *Social Research Methods: Qualitative and Quantitative Approaches.* 7th ed. Harlow: Pearson Education Limited.

NHD, 2012. *Små bedrifter - store verdier: Regjeringens strategi for små og mellomstore bedrifter,* Oslo, Norway: Nærings- og handelsdepartementet.

Oliveira, T. & Martins, M. R. O., 2009. *Determinants of Information Technology Adoption in Portugal.* Lisboa, Proceedings of the International Conference on e-Business, pp. 264-270.

Osborne, J. W., 2015. What is Rotating in Exploratory Factor Analysis. *Practical Assessment, Research & Evaluation*, 20(2), pp. 1-7.

PAC, 2016a. *Software & IT services - Market inbrief - Norway,* Paris: PAC SITSI Research.

PAC, 2016b. *Software products - Market View - Norway,* Paris: PAC SITSI Research.

PAC, 2016c. *IT Services - Market View - Norway,* Paris: PAC SITSI Research.

Pan, M.-J. & Jang, W.-Y., 2008. Determinants of the Adoption of Enterprise Resource Planning within the Technology-Organization-Environment Framework: Taiwan's Communications Industry. *Journal of COmputer Information Systems.*

Pareek, D., 2015. *5 Reasons why the Big Data Market is Exploding.* [Online]
Available at: https://datafloq.com/read/5-Reasons-Big-Data-Market-is-Exploding/1948
[Accessed 26 April 2017].

Park, J.-H., Kim, M.-K. & Paik, J.-H., 2015. *The Factors of Technology, Organization and Environment Influencing the Adoption and Usage of Big Data in Korean Firms.* Madrid, Spain, Electronics and Telecommunications Research Institute.

Peteraf, M., 1993. The Cornerstone of Competitive Advantage: A Resource-Based View. *Strategic Management Journal*, 14(3), pp. 179-191.

Picchi, A., 2014. *Target's brand takes a massive hit amid data breach.* [Online]
Available at: http://www.cbsnews.com/news/targets-brand-takes-a-massive-hit-amid-data-breach/
[Accessed 13 June 2017].

Pinsonneault, A. & Kraemer, K., 1993. Survey Research Methodology in Management Information Systems: An assessment. *Journal of Management Information Systems,* 10(2), pp. 75-105.

Planing, P., 2014. *Innovation Acceptance.* 1st ed. Stuttgart: Springer Gabler.

Plouffe, C., Hulland, J. & Vandenbosch, M., 2001. Research Report: Richness Versus Parsimony in Modeling Technology Adoption Decisions - Understanding Merchant Adoption of a Smart Card-Based Payment System. *Information Systems Research,* 12(2), pp. 208-222.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y. & Podsakoff, N. P., 2003. Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology,* 88(5), pp. 879-903.

Porambage, P. et al., 2016. The Quest for Privacy in the Internet of Things. *IEEE Cloud Computing,* 3(2), pp. 36-45.

Porter, M., 1996. What is Strategy?. *Harvard Business Review,* Volume November-December, pp. 61-78.

Porter, M. & Millar, V., 1985. How Information Gives You Competitive Advantage. *Harvard Business Review,* 63(4), pp. 149-160.

Premkumar, G. & Ramamurthy, K., 1995. The Role of Interorganizational and Organizational Factors on the Decision Mode for Adoption of Interorganizational Systems. *Decision Sciences,* 26(3), pp. 303-336.

Premkumar, G. & Roberts, M., 1999. Adoption of new information technologies in rural small businesses. *Omega: The International Journal of Management Science,* 27(1), pp. 467-484.

Puklavec, B., Oliveira, T. & Popovič, A., 2014. Unpacking business intelligence systems adoption determinants: An exploratory study of small and medium enterprises. *Economic and business review,* 16(2), pp. 185-23.

PwC, 2015. *Hva er Big Data, og hva betyr Big Data for deg?,* Oslo: Pricewaterhouse Coopers AS.

Qian, L. Y., Baharudin, A. S. & Kanaan-Jebna, A., 2016. Factors affecting the adoption of enterprise resource planning (ERP) on cloud among small and medium enterprises (SMES) in Penang, Malysia. *Journal of Theoretical and Applied Information Technology,* June, 88(3), pp. 398-409.

Rahman, N., 2016. *Factors Affecting Big Data Technology Adoption.* Portland, Student Research Symposium.

Rai, A., Brown, P. & Tang, X., 2009. Organizational Assimilation of Electronic Procurement Innovations. *Journal of Management Information Systems,* 26(1), pp. 257-296.

Raj, P., 2014. *Handbook of Research on Cloud Infrastructures for Big Data Analytics (Advances in Data Mining and Database Management Book Series).* 1st ed. Pennsylvania: IGI Global.

Ramdani, B. & Kawalek, P., 2009. Predicting SMEs' adoption of enterprise systems. *Journal of Enterprise Information Management,* 22(1/2), pp. 10-24.

Ravichandran, T. & Lertwongsatien, C., 2005. Effect of Information Systems Resources and Capabilities on Firm Performance: A Resource-Based Perspective. *Journal of Management information Systems,* 21(4), pp. 237-276.

Regjeringen, 2016a. *Digital agenda for Norway in brief,* Oslo: Norwegian Ministry of Local Government and Modernisation.

Regjeringen, 2016b. *Nasjonalbudsjettet 2017,* Oslo: Det Kongelige Finansdepartementet.

Rijmenam, M. V., 2013. *Understanding Your Business With Descriptive, Predictive and Prescriptive Analytics [VIDEO].* [Online]
Available at: http://www.smartdatacollective.com/bigdatastartups/144436/understanding-your-business-descriptive-predictive-and-prescriptive-analytics
[Accessed 5 April 2017].

Rogers, E. M., 1983. *Diffusion of Innovation: Third edition.* 3rd ed. New York: The Free Press.

Rogers, E. M., 1995. *Diffusion of Innovations.* New York: The Free Press.

Rosenkopf, L., 1993. Institutional and Competitive Bandwagons: Using Mathematical Modeling as a Tool to Explore Innovation Diffusion. *Academy of Management Review,* 18(3), pp. 487-517.

Rouse, M. & Casey, K., 2016. *Software as a Service (SaaS).* [Online]
Available at: http://searchcloudcomputing.techtarget.com/definition/Software-as-a-Service
[Accessed 10 May 2017].

Salleh, K. & Janczewski, L., 2016. Technological, organizational and environmental security and privacy issues of big data: A literature review. *Procedia Computer Science,* Volume 100, pp. 19-28.

Saunders, M., Lewis, P. & Thornhill, A., 2012. *Research methods for business students.* 6th ed. Harlow: Pearson Education Limited.

Schrage, M., 2016. *How the Big Data Explosion Has Changed Decision Making.* [Online]
Available at: https://hbr.org/2016/08/how-the-big-data-explosion-has-changed-decision-making
[Accessed 18 April 2017].

Scupola, A., 2009. SMEs' E-commerce Adoption: Perspective from Denmark and Australia. *Journal of Enterprise Information Management,* 22(1-2), pp. 152-166.

Shvachko, K., Kuang, H., Radia, S. & Chansler, R., 2010. *The Hadoop Distributed File System.* Incline Village, Nevada, IEEE.

Simmons, E. & Wilmot, A., 2004. Incentive payments on social surveys: a literature review. *Social Survey Methodology Bulletin,* 5(1-11).

Singh, J., 1986. Performance, Slack, and Risk Taking in Organizational Decision Making. *Academy of Management,* 29(3), pp. 562-585.

Soares-Aguiar, A. & Palma-dos-Reis, A., 2008. Why Do Firms Adopt E-Procurement Systems? Using Logistic Regression to Empirically Test a Conceptual Model. *IEEE Transactions on Engineering Managment,* 55(1), pp. 120-133.

Stacks, D. W. & Salwen, M. B., 2009. *An integrated approach to communication theory and research.* 2nd ed. New York: Routledge.

Stevens, P. J., 1984. Outliers and Influential Data Points in Regression Analysis. *Psychological Bulletin,* 95(2), pp. 334-344.

Stuart, J. & Barker, A., 2013. *Undefined By Data: A Survey of Big Data Definitions,* St Andrews: School of Computer Science.

Subramanian, G. H. & Nosek, J. T., 2001. An empirical study of the measurement and instrument validation of perceived strategy value of information systems. *Journal of Computer Information Systems,* 41(3), pp. 64-69.

Sun, S., Cegielski, C. G., Jia, L. & Hall, D. J., 2016. Understanding the Factors Affecting the Organizational Adoption of Big Data. *Journal of Computer Information Systems* , pp. 1-11.

Tabachnick, B. G. & Fidell, L. S., 2007. *Using Multivariate Statistics.* 5th ed. California: Pearson.

Tambe, P., 2014. *Big Data Investment, Skills, and Firm Value,* New York: NYU Stern.

Tarde, G., 1903. *The Laws of Imitation.* 1st ed. New York: Henry Holt and Company.

Taylor, S. & Todd, P., 1995. Decomposition and crossover effects in the theory of planned behavior: A study of consumer adoption intentions. *International Journal of Research in Marketing,* Volume 12, pp. 137-155.

Teo, T. S. H., Ranganathan, C. & Dhaliwal, J., 2006. Key Dimensions of Inhibitors for the Deployment of Web-Based Business-to-Business Electronic Commerce. *IEEE Transactions on Engineering Management,* 53(3).

Thi, L. S., 2006. *Electronic Commerce Adoption among Manufacturing SMEs in Malaysia,* Leicestershire: Loughborough University.

Thong, J. Y., 1999. An Integrated Model of Information Systems Adoption in Small Businesses. *Journal of Management Information Systems* , 15(4), pp. 187-214.

Thong, J., Yap, C. & Raman, K., 1996. Top management support, external expertise and information systems implementation in small business. *Information Systems Research,* 7(2), pp. 248-267.

Tornatzky, L. G. & Fleischer, M., 1990. *The Processes of Technological Innovation.* Lexington, MA: Lexington Books.

Tornatzky, L. G. & Klein, K., 1982. Innovation characteristics and innovation adoption-implementation: A meta-analysis of findings. *IEEE Transactions on Engineering Management,* 29(1), pp. 28-48.

Ularu, E. G., Puican, F. C., Apostu, A. & Velicanu, M., 2012. Perspectives on Big Data and Big Data Analytics. *Database Systems Journal*, 3(4), pp. 3-14.

Venkatesh, V. & Davis, F. D., 1996. A Model of the Antecedents of Perceived Ease of Use: Development and Test. *Decision sciences,* 27(3), pp. 451-481.

Venkatesh, V., Davis, F. & Morris, M., 2007. Dead or Alive? The Development, Trajectory and Future Of Technology Adoption Research. *Journal of the Association for Information Systems,* 8(4), pp. 267-286.

Venkatesh, V., Morris, M. G., Davis, G. B. & Davis, F. D., 2003. User Acceptance of Information Technology: Toward a Unified View. *MIS quarterly,* 27(3), pp. 425-478.

Watson, H., 2009. Tutorial: Business Intelligence: Past, Present and Future. *Communications of the Association for Information Systems,* 25(39), pp. 488-510.

Wejnert, B., 2002. Integrating Models of Diffusion of Innovations: A Conceptual Framework. *Annual Review of Sociology,* Issue 28, pp. 297-326.

Welsh, J. & White, J., 1981. A small buisness is not a little big business. *Harvard Business Review,* 59(4), pp. 18-32.

Winter, J. C. d. & Dodou, D., 2010. Five-Point Likert Items: t test versus Mann-Whitney-Wilcoxon. *Practical Assessment, Research & Evaluation,* 15(11), pp. 1-16.

Winter, J. & Dodou, D., 2012. Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. *Journal of Applied Statistics,* 39(4), pp. 695-710.

Wixom, B. et al., 2014. The Current State of Business Intelligence in Academia: The Arrival of Big Data. *Communications of the Association for Information Systems,* 34(1).

Worthington, R. L. & Whittaker, T. A., 2006. Scale Development Research: A content Analysis and Recommendations for Best Practices. *The Counseling Psychologist,* 34(6), pp. 806-838.

Yan, T. & Curtin, R., 2010. The Relation Between Unit Nonresponse and Item Nonresponse: A Response Continuum Perspective. *International Journal of Public Opinion Research,* 22(4), pp. 535-551.

Yeh, C. H., Lee, G. G. & Pai, J. C., 2014. Using a technology-organization-environment framework to investigate the factors influencing e-business information technology capabilities. *Information Development,* 31(5), pp. 435-450.

Ylijoki, O. & Porras, J., 2016. Perspectives to Definition of Big Data: A Mapping Study and Discussion. *Journal of Innovation Management,* 4(1), pp. 69-91.

Yong, A. G. & Pearce, S., 2013. A Beginner's Guide to Factor Analysis: Focus on Exploratory Factor Analysis. *Tutorials in Quantitative Methods for Psychology,* 9(2), pp. 79-94.

Young, R. & Jordan, E., 2008. Top management support: Mantra or necessity?. *International Journal of Project Management,* 26(7), pp. 713-725.

Zhu, K., Dong, S., Xu, S. & Kraemer, K., 2006b. Innovation diffusion in global contexts: determinants of post-adoption digital transformation of European companies. *European Journal of Information Systems,* 15(1), pp. 601-616.

Zhu, K. & Kraemer, K. L., 2005. Post-Adoption Variations in Usage and Value of E-Business by Organizations: Cross-Country Evidence from the Retail Industry. *Information Systems Research,* 16(1), pp. 61-84.

Zhu, K., Kraemer, K. L. & Dedrick, J., 2004. Information Technology Payoff in E-Business Environments: As International Perspective on Value Creation of E-Business in the Financial Services Industry. *Journal of Management Information Systems,* 21(1), pp. 17-54.

Zhu, K., Kraemer, K. & Xu, S., 2003. Electronic business adoption by European firms: a cross-country assessment of the facilitators and inhibitors. *European Journal of Information Systems,* Volume 12, pp. 251-268.

Zhu, K., Kraemer, K. & Xu, S., 2006a. The process of innovation assimilation by Firms in Different Countries: A Technology Diffusion Perspective on E-Business. *Management Science,* 52(10), pp. 1157-1576.

Appendix A: Complete Survey

Page 1 *(Introduction)*

Thank you for participating in this survey!

The purpose of this survey is to investigate the factors affecting the adoption of big data technology in Norway. To help us identify these factors, you will, on the following pages, be asked to indicate whether or not you agree with a set of statements concerning your company's relationship with big data technology. We sincerely hope that you take the time to answer all questions, and do so honestly.

The survey is designed to take approximately 5 minutes to complete, and you may at any time return to this survey by clicking on the link provided in your email. Big data is a term most people have heard of, yet very few are truly familiar with. Therefore, we begin the survey by offering a definition of big data. This is to ensure that we share a common understanding of the term.

Click the button below to begin the survey

Page 2 *(Presenting the definition of Big Data)*

What is big data?
Big data is data that is characterized by three V's: Volume, Variety and Velocity.

- Volume – The amount of available data is increasing rapidly, with data sets ranging from terabytes to zettabytes in size.
- Variety – Different types of data are available from a range of data sources, both external and internal to the firm. Variety is about handling both structured and unstructured data, from sources such as customer databases, social media, traffic cameras, and a whole host of sensor technologies.
- Velocity – The increased speed at which data is available requires real-time processing to maximize the value of the data.

Big data provides no value by itself. The value of big data comes from the ability to analyse data that was not previously available, or were too expensive to store or process, to provide new insights and improve the basis of decision making. As such data is either difficult or impossible to manage using traditional database or analytics tools, companies have begun exploring new technologies.

Big data technologies refer to software and hardware that enables the collection, processing and analysis of data of high volume, variety and velocity. Examples of technologies that are big data capable include Hadoop, CouchDB, Cassandra, Pig, Hive, MongoDB and AsterData. Although these technologies are not exclusively used for big data, their application on datasets that fit the definition of big data allow us to refer to them as big data technologies.

Q1 *Referring to the above definition, please select the statement that best describe your company*

- My company is not familiar with big data technology (1)
- My company is familiar with big data technology and/or has considered using it (2)
- My company is planning to use big data technology within the next 24 months (3)
- My company has launched pilot projects or initiatives for evaluating and/or trialling big data technology (4)
- The acquisition of specific big data technologies are planned, in progress, implemented or canceled (5)
- My company has big data technology, but we have yet to establish a program of regular use (6)
- My company has big data technology, and we have established a program of regular use (7)

Page 3 (*Demographics information)*

Q2 *Please select the option that best describe your role at your company*

- CEO/President/VP/Managing director (1)
- CIO/IT director/Technology director (2)
- CFO/Treasurer/Controller (3)
- Other C-level executive (4)
- Non-executive position (5)

Q3 *What is your company's primary industry?*

- Banking and insurance (1)
- Manufacturing (2)
- Construction, agriculture and materials (3)
- Telecommunications (4)
- Transport, logistics and post (5)
- Energy and utilities (6)
- Retail and wholesale (7)
- Services (8)
- Public sector and healthcare (9)
- Information technology (10)
- Entertainment, media and tourism (11)
- Education and scientific research (12)
- Other (13)

Q4 *What is your company size, by annual revenue?*

- Less than 85 million NOK (1)
- 85-150 million NOK (2)
- 150-300 million NOK (3)
- 300-500 million NOK (4)
- 500-1000 million NOK (5)
- More than 1000 million NOK (6)

Q5 *What is your company size, by employees?*
- Less than 50 (1)
- 50-100 (2)
- 101-150 (3)
- 151-250 (4)
- 251-400 (5)
- More than 400 (6)

Q6 *How long has your company been in business?*
- Less than 5 years (1)
- 5-10 years (2)
- 11-20 years (3)
- 21-30 years (4)
- Longer than 30 years (5)

On page 4 to 7, respondents were asked to rate the extent they agreed or disagreed with the presented statements on a Likert scale from 1 (Strongly disagree) to 7 (Strongly agree).

1 = Strongly disagree
2 = Disagree
3 = Slightly disagree
4 = Neutral
5 = Slightly agree
6 = Agree
7 = Strongly agree

Page 4

Q7 *To what extent do you agree or disagree with each of the following statements regarding* ***your company****? Please select one answer per row*
- Our IT employees have equal or better technical knowledge than our competitors (1)
- Our IT employees have the ability to quickly learn and apply new information technologies (2)
- Our IT employees have the skills and knowledge to manage IT projects in the current business environment (3)

Q8 *To what extent do you agree or disagree with each of the following statements regarding* ***big data technology****? Please select one answer per row*
- Top management believe that investment and expenditure in this technology is worthwhile (1)
- Top management believe that this technology has potential strategic value (2)
- Top management support is important to provide the resources for my company to use this technology (3)
- My company has the technological resources to adopt this technology (4)
- My company has the financial resources to adopt this technology (5)
- My company has no difficulties in finding all the necessary resources (e.g. funding, people, time) to adopt this technology (6)

Page 5

Q9 *To what extent do you agree or disagree with each of the following statements* ***regarding big data technology****? Please select one answer per row*
- This technology improves my company's performance (1)
- This technology improves my company's productivity (2)
- This technology improves the effectiveness of my company’s operations (3)
- This technology provides my company with valuable information for decision making (4)
- My company finds it easy to get this technology to do what we want it to do (5)
- My company’s interaction with this technology is clear and understandable (6)
- My company finds this technology easy to use (7)
- It is easy for my company to become skillful at using this technology (8)
- There are businesses in the community which provide support for use of this technology (9)
- There are agencies in the community who provide training on this technology (10)
- Technology agencies actively market this technology by providing incentives for adoption (11)

Page 6

Q10 *To what extent do you agree or disagree with each of the following statements regarding* ***big data technology****? Please select one answer per row*
- This technology is compatible with the data captured in my company (1)
- This technology fits well with my company’s existing operating practices (2)
- This technology is compatible with my company’s IT infrastructure (3)
- Using this technology is consistent with my company’s values and beliefs (4)
- My company has adequate tools and mechanisms to provide effective data-protection when using this technology (5)
- My company has security capabilities to adopt this technology (6)
- My company has security policies that suits the different types of data in the company when using this technology (7)
- The skills required to ensure data security when using this technology are easy for my company (8)
- It would be easy for my company to integrate security policies for this technology (9)

Page 7

Q11 *To what extent do you agree or disagree with each of the following statements regarding* ***big data technology****? Please select one answer per row*
- We believe we would lose our customers to our competitors if we did not adopt this technology (1)
- We feel it is a strategic necessity to use this technology to compete in the marketplace (2)
- We believe that our competitors get many advantages through adopting this technology (3)
- Many of our competitors are going to adopt this technology in the near future (4)
- My company’s use of this technology is limited by data protection acts, rules and regulations in Norway (5)
- My company would find it challenging to protect data privacy when using this technology (6)
- My company would find it difficult to comply with privacy related regulation when using this technology (7)
- My company would find it difficult to meet legal expectations concerning the use of big data without compromising our business goals (8)

Appendix B: Cover Letter

NORGES HANDELSHØYSKOLE

Dear [FirstName] [LastName],

Are you making good use of your data? An increasing number of companies find themselves in a position where traditional methods of collecting, storing and processing data no longer suffice. They are facing the challenges of big data!

As master students at the Norwegian School of Economics, we hereby invite you to partake in a nationwide survey about big data, as part of our master thesis. According to a recent study, 98% of Norwegian IT managers agree that analysing more and new types of data could give their company a competitive advantage. *But what factors are driving or inhibiting the use of big data technology in Norway?* This is the question we seek to address in a new study of Norwegian companies' adoption of big data technology.

Big data is characterized by data sets – both structured and unstructured – so large or complex that it becomes difficult to process using traditional tools and applications. While many companies wish to utilize big data to improve decision-making, gain new insight and optimize processes, doing so requires both new technology and methods.

Why should you participate?
People in leadership positions in Norwegian companies have been chosen to participate in this study. To get a clear picture of Norwegian companies' use of big data, it is crucial that as many people as possible participate. Your answers will contribute to identifying which technological, organizational and environmental factors facilitate or inhibit the adoption of big data technology in Norway. Our findings can give your company a new and better understanding of the areas to prioritize when adopting big data technology.

Participation from your company is essential to our research project, whether you have big data technology or not.

As students, we are dependent on your participation as well as the quality of your responses. Therefore, we hope that you take the time to answer the survey. In return, we will gladly share our findings with your company. If you wish to receive a summary of our findings, you may register your email address after completing the survey.

What does participation in the study entail?
Participation is voluntary and will take approximately **5 minutes** to complete. Your response is anonymous and all data is treated confidentially. The survey is open for participation until the 26th of April.

Click here to take the survey

Or copy and paste the link in your browser:
https://nhh.eu.qualtrics.com/jfe/form/linktosurvey

Thank you in advance for your participation!

With best regards,

Truls Petersen / truls.petersen@student.nhh.no
Truc Nguyen / truc.nguyen@student.nhh.no

Follow the link to opt out of future emails:

Click here to unsubscribe

Appendix C: Literature Review

C.1 Technology Context

	Common definition	Significant	Source
Perceived ease of use/Complexity	The degree to which an innovation is perceived to be relatively difficult to understand and use	Yes	Thong (1999), Premkumar & Roberts (1999), Art et al. (2011), Boonsiritomachai (2014), Agrawal (2015), Le et al. (2012), Alrousan (2014), Hung et al. (2016)
		No	Jeon et al. (2006), Ramdani & Kawalex (2009), Ifinedo (2011), Li et al. (2011), Hung et al. (2016)
Perceived usefulness/Relative advantage	The degree to which an innovation is perceived as being better than the idea it supersedes	Yes	Davis (1989), Iacovou et al. (1995), Igbaria et al. (1997), Thong (1999), Premkumar & Roberts (1999), Chwelos et al. (2001), Grandon & Pearson (2004), Gibbs & Kraemer (2004), Zhu et al. (2006), Jeon et al. (2006), Quaddus & Hofmeyer (2007), Oliveira & Martins (2009), Alam & Noor (2009), Scupola (2009), Ramdani & Kawalex (2009), Ifinedo (2011), Li et al. (2011), Boonsiritomachai (2014), Alrousan (2014), Park et al. (2015), Hung et al. (2016), Al-Isma'ili et al. (2016)
		No	Le et al. (2012), Nam et al. (2015), Agrawal (2015)
Compatibility	The degree to which innovations are perceived as being consistent with existing methods for executing their mission	Yes	Thong (1999), Premkumar & Roberts (1999), Grandon & Pearson (2004), Zhu et al. (2006), Le et al. (2012), Agrawal (2015), Al-Isma'ili et al. (2016)
		No	Gibbs & Kraemer (2004), Jeon et al. (2006), Ramdani & Kawalex (2009), Ifinedo (2011), Alrousan (2014), Boonsiritomachai (2014), Park et al. (2015), Hung et al. (2016)
Observability	The degree that potential adopters of an innovation can perceive the results of using that innovation from users who have already adopted it	Yes	Plouffe et al. (2001), Boonsiritomachai (2014), Alrousan (2014)
		No	Ramdani & Kawalex (2009)
Trialability	The extent to which potential adopters have the opportunity to	Yes	Randani & Kawalex (2009), Al-Isma'ili et al. (2016)

<table>
<tr><td rowspan="1"></td><td>experiment with an innovation</td><td>No</td><td>Boonsiritomachai (2014), Alrousan (2014)</td></tr>
<tr><td rowspan="2">Technology infrastructure</td><td rowspan="2">The internal technology ability to adopt new technology or the degree to which a firm has necessary technology infrastructure to adopt</td><td>Yes</td><td>Teo et al. (2006), Soares-Aguiar & Palma-dos-Reis (2008), Malladi & Krishnan (2013), Yeh et al. (2014)</td></tr>
<tr><td>No</td><td>Pan & Jang (2008)</td></tr>
<tr><td rowspan="2">Technology readiness/ maturity</td><td rowspan="2">The maturity of the information technology within an organization and its information technology capabilities encourage the organization to apply information technology to achieve its strategic goals.</td><td>Yes</td><td>Zhu et al. (2004), Pan & Jang (2008), Oliveira & Martins (2009), Yeh et al. (2014)</td></tr>
<tr><td>No</td><td></td></tr>
</table>

C.2 Organization Context

<table>
<tr><th></th><th>Common definition</th><th>Significant</th><th>Source</th></tr>
<tr><td rowspan="2">Size</td><td rowspan="2">The size of the firm (i.e., the number of employees and annual revenue)</td><td>Yes</td><td>Thong (1999), Premkumar & Roberts (1999), Zhu et al. (2003), Zhu et al. (2004), Zhu & Kraemer (2005), Buonanno et al (2005), Zhu et al. (2006), Pan & Jang (2008), Soares-Aguiar & Palma-dos-Reis (2008), Oliveira & Martins (2009), Ramdani & Kawalex (2009), Hameed et al. (2012), Le et al. (2012), Malladi & Krishnan (2013), Puklavec et al (2014), Agrawal (2015), Al-Isma'ili et al. (2016), Hung et al (2016)</td></tr>
<tr><td>No</td><td>Gibbs & Kraemer (2004), Jeon et al. (2006), Alrousan (2014)</td></tr>
<tr><td rowspan="2">Organizational absorptive capacity</td><td rowspan="2">Absorptive capacity of an organization is the ability of its members to utilize existing or pre-existing IT knowledge.</td><td>Yes</td><td>Hung et al. (2016)</td></tr>
<tr><td>No</td><td>Boonsiritomachai (2014), Agrawal (2015)</td></tr>
<tr><td rowspan="2">Organizational culture</td><td rowspan="2">Culture at various levels (national, organizational, group) can affect success of IT</td><td>Yes</td><td>Puklavec et al. (2014)</td></tr>
<tr><td>No</td><td></td></tr>
<tr><td rowspan="2">Organizational innovativeness</td><td rowspan="2">Innovativeness is the willingness degree of taking a risk and trying new solutions that not been tried or tested before</td><td>Yes</td><td>Al-isma'ili et al. (2016)</td></tr>
<tr><td>No</td><td></td></tr>
<tr><td>Centralization</td><td>The degree to which power and</td><td>Yes</td><td></td></tr>
</table>

	control are concentrated in the hands of relatively few individuals in an organization	No	Hameed et al. (2012)
Formalization	The degree to which an organization follows the rules and procedures on the role of performance of its members	Yes	
		No	Hameed et al. (2012)
Top management support	The degree to which top management understands the importance of the technology and the extent to which it is involved in related initiatives	Yes	Premkumar & Roberts (1999), Teo et al. (2006), Scupola (2009), Ramdani & Kawalex (2009), Le et al. (2012), Hameed et al. (2012), Yeh et al. (2014), Puklavec et al. (2014), Park et al. (2015), Al-Isma'ili et al. (2016), Hung et al. (2016)
		No	Alrousan (2014)
Project champion	An individual who performs the task of spreading knowledge of new technology within the organization.	Yes	Puklavec et al. (2014)
		No	Hameed et al. (2012)
IT expertise	The prior experience of IT employees in terms of skill and knowledge	Yes	Thong (1999), Zhu et al. (2003), Zhu & Kraemer (2005), Zhu et al. (2006), Soares-Aguiar & Palma-dos-Reis (2008), Alam & Noor (2009), Scupola (2009), Li et al. (2011), Hameed et al. (2012), Le et al. (2012), Yeh et al. (2014), Nam et al. (2015), Agrawal (2015), Al-Isma'ili et al. (2016)
		No	Jeon et al. (2006), Alrousan (2014)
Organizational readiness	The degree to which an organization has the awareness, resources, commitment and governance to adopt IT	Yes	Chwelos et al. (2001), Thi (2006), Chong et al. (2009), Ramdani & Kawalex (2009), Hameed et al. (2012), Malladi & Krishnan (2013), Puklavec et al. (2014), Park et al. (2015)
		No	Iacovou, et al. (1995)
Slack/ organizational resources	Those resources an organization has acquired which are not committed to an existing business operation, and subsequently can be used in a discretionary manner	Yes	Zhu et al. (2004), Gibbs & Kraemer (2004), Li et al (2011), Hameed et al. (2012), Le et al. (2012), Boonsiritomachai (2014), Nam et al. (2015), Park et al. (2015)
		No	Alrousan (2014)

C.3 Environment Context

	Common definition	Significant	Source
Competitive pressure	The degree of pressure that the company faces from competitors within the industry	Yes	Zhu et al. (2003), Zhu & Kraemer (2005), Zhu et al. (2006), Soares-Aguiar & Palma-dos-Reis (2008), Oliveira & Martins (2009), Chong et al. (2009), Le et al. (2012), Malladi & Krishnan (2013), Boonsiritomachai (2014), Nam et al. (2015), Agrawal (2015)
		No	Thong (1999), Zhu et al. (2004), Jeon et al. (2006), Pan & Jang (2008), Ramdani & Kawalex (2009), Park et al. (2015), Alrousan (2014)
Industry & market complexity	The degree and instability of change in a firm's environment	Yes	Agrawal (2015), Al-Isma'ili et al (2016)
		No	Malladi & Krishnan (2013)
Partners	Enacted trading partner power measures the strength of the influence strategy (e.g., rewards and threats) used to exercise that potential power.	Yes	Chwelos et al. (2001), Zhu et al. (2003), Zhu et al. (2006), Yeh et al. (2014), Alrousan (2014), Park et al. (2015),
		No	
Regulatory environment	The adequacy of institutional frameworks and business laws governing the use of innovations/technology	Yes	Zhu et al. (2004), Zhu & Kraemer (2005), Jeon et al. (2006), Alam & Noor (2009), Le et al. (2012), Alrousan (2014), Nam et al. (2015), Agrawal (2015)
		No	Pan & Jang (2008), Park, et al. (2015)
External support	Availability of support for implementing and using an information system	Yes	Thi (2006), Scupola (2009), Le et al. (2012), Al-Isma'ili et al. (2016), Hung et al. (2016)
		No	Premkumar & Roberts (1999), Ramdani & Kawalex (2009)
External pressure	External pressure applied by suppliers and customers	Yes	Iacovou et al (1995), Premkumar & Roberts (1999), Gibbs & Kraemer (2004)
		No	Alam & Noor (2009)
Legislation barriers	Government policy, inadequate legal protection or business laws	Yes	Gibbs & Kraemer (2004)
		No	

Appendix D: Assessments of Biases

D.1 Chi-Square Tests for Demographical Differences

	Early respondents (N=69)	Late respondents (N=69)	Pearson Chi-Square Asymptotic Sig. (2-sided)	Fisher's Exact Test* Sig. (2-sided)
Role				
CEO/President/VP/Managing director	50	39	0.093	
CIO/IT director/Technology director	14	18		
Other C-level executive	5	12		
Industry				
Banking and insurance	1	7		0.494
Manufacturing	14	12		
Construction, agriculture and materials	5	6		
Telecommunications	0	1		
Transport, logistics and post	6	5		
Energy and utilities	6	6		
Retail and wholesale	10	6		
Services	7	8		
Public sector and healthcare	7	2		
Information technology	4	2		
Entertainment, media and tourism	3	4		
Education and scientific research	2	4		
Other	4	6		
Annual revenue				
85-150 million NOK	13	9	0.586	
150-300 million NOK	15	14		
300-500 million NOK	11	16		
500-1000 million NOK	13	9		
More than 1000 million NOK	17	21		
Number of employees				
50-100	23	17	0.386	
101-150	14	9		
151-250	12	13		
251-400	7	12		
More than 400	13	18		
Years in business				
Less than 5 years	5	1		0.577
5-10 years	2	2		
11-20 years	10	9		
21-30 years	11	12		
Longer than 30 years	41	45		

*Fisher's Exact test was used when the expected count for each cell was less than 5

D.2 Mann-Whitney U Tests for Response Differences

		Ranks			Test Statistics			
		N	Mean Rank	Sum of Ranks	Mann-Whitney U	Wilcoxon W	Z	Asymp. Sig. (2-tailed)
MS1	Early respondents	69	67.15	4633.5	2218.5	4633.5	-.709	.478
	Late respondents	69	71.85	4957.5				
	Total	138						
MS2	Early respondents	69	69.36	4785.5	2370.5	4785.5	-.044	.965
	Late respondents	69	69.64	4805.5				
	Total	138						
MS3	Early responses	69	72.91	5031.0	2145.0	4560.0	-1.088	.277
	Late responses	69	66.09	4560.0				
	Total	138						
IE1	Early respondents	69	68.89	4753.5	2338.5	4753.5	-.185	.853
	Late respondents	69	70.11	4837.5				
	Total	138						
IE2	Early respondents	69	69.04	4763.5	2348.5	4763.5	-.140	.888
	Late respondents	69	69.96	4827.5				
	Total	138						
IE3	Early respondents	69	69.86	4820.5	2355.5	4770.0	-.110	.912
	Late respondents	69	69.14	4770.5				
	Total	138						
OR1	Early respondents	69	70.04	4832.5	2343.5	4758.5	-.160	.873
	Late respondents	69	68.96	4758.5				
	Total	138						
OR2	Early respondents	69	67.65	4668.0	2253.0	4668.0	-.563	.573
	Late respondents	69	71.35	4923.0				
	Total	138						
OR3	Early respondents	69	62.46	4309.5	1894.5	4309.5	-2.109	**.035***
	Late respondents	69	76.54	5281.5				
	Total	138						
ES1	Early respondents	69	72.49	5002.0	2174.0	4589.0	-.931	.352
	Late respondents	69	66.51	4589.0				
	Total	138						
ES2	Early respondents	69	72.09	4974.5	2201.5	4616.5	-.796	.426
	Late respondents	69	66.91	4616.5				
	Total	138						
ES3	Early respondents	69	65.25	4502.5	2087.5	4502.5	-1.350	.177
	Late respondents	69	73.75	5088.5				
	Total	138						

		Ranks			Test Statistics			
		N	Mean Rank	Sum of Ranks	Mann-Whitney U	Wilcoxon W	Z	Asymp. Sig. (2-tailed)
RA1	Early respondents	69	68.46	4724.0	2309.0	4758.5	-.319	.750
	Late respondents	69	70.54	4867.0				
	Total	138						
RA2	Early respondents	69	66.26	4572.0	2157.0	4668.0	-1.008	.313
	Late respondents	69	72.74	5019.0				
	Total	138						
RA3	Early respondents	69	67.72	4673.0	2258.0	4309.5	-.551	.585
	Late respondents	69	71.28	4918.0				
	Total	138						
RA4	Early respondents	69	66.74	4605.0	2190.0	4605.0	-.860	.390
	Late respondents	69	72.26	4986.0				
	Total	138						
CX1	Early respondents	69	76.15	5254.5	1921.5	4336.5	-2.008	**.045***
	Late respondents	69	62.85	4336.5				
	Total	138						
CX2	Early respondents	69	73.30	5057.5	2118.5	4533.5	-1.168	.243
	Late respondents	69	65.70	4533.5				
	Total	138						
CX3	Early respondents	69	72.43	4997.5	2178.5	4593.5	-.889	.374
	Late respondents	69	66.57	4593.5				
	Total	138						
CX4	Early respondents	69	74.36	5130.5	2045.5	4460.5	-1.466	.143
	Late respondents	69	64.64	4460.5				
	Total	138						
SE1	Early respondents	69	69.51	4796.0	2380.0	4795.0	-.002	.998
	Late respondents	69	69.49	4795.0				
	Total	138						
SE2	Early respondents	69	70.20	4845.5	2332.5	4747.0	-.209	.834
	Late respondents	69	68.80	4747.5				
	Total	138						
SE3	Early respondents	69	69.68	4808.0	2368.0	4783.0	-.055	.956
	Late respondents	69	69.32	4783.0				
	Total	138						
SE4	Early respondents	69	71.49	4933.0	2243.0	4658.0	-.602	.547
	Late respondents	69	67.51	4658.0				
	Total	138						
SE5	Early respondents	69	72.80	5023.5	2152.5	4658.0	-.999	.318
	Late respondents	69	66.20	4567.5				
	Total	138						

		N	Ranks: Mean Rank	Ranks: Sum of Ranks	Test Statistics: Mann-Whitney U	Test Statistics: Wilcoxon W	Test Statistics: Z	Test Statistics: Asymp. Sig. (2-tailed)
CM1	Early respondents	69	69.07	4766.0	2351.0	4766.0	-.129	.897
	Late respondents	69	69.93	4825.0				
	Total	138						
CM2	Early respondents	69	67.13	4632.0	2217.0	4632.0	-.711	.477
	Late respondents	69	71.87	4959.0				
	Total	138						
CM3	Early respondents	69	63.76	4399.5	1984.5	4399.5	-1.730	.084
	Late respondents	69	75.24	5191.5				
	Total	138						
CM4	Early respondents	69	73.93	5101.5	2074.5	4489.5	-1.384	.166
	Late respondents	69	65.07	4489.5				
	Total	138						
CP1	Early respondents	69	65.72	4535.0	2120.0	4535.0	-1.132	.258
	Late respondents	69	73.28	5056.0				
	Total	138						
CP2	Early respondents	69	66.36	4578.5	2163.5	4578.5	-.958	.338
	Late respondents	69	72.64	5012.5				
	Total	138						
CP3	Early respondents	69	62.20	4291.5	1876.5	4291.5	-2.224	**.026***
	Late respondents	69	76.80	5299.5				
	Total	138						
CP4	Early respondents	69	64.68	4463.0	2048.0	4463.0	-1.450	.147
	Late respondents	69	74.32	5128.0				
	Total	138						
PR1	Early respondents	69	68.53	4728.5	2313.5	4728.5	-.292	.770
	Late respondents	69	70.47	4862.5				
	Total	138						
PR2	Early respondents	69	75.05	5178.5	1997.5	4412.5	-1.673	.094
	Late respondents	69	63.95	4412.5				
	Total	138						
PR3	Early respondents	69	71.28	4918.5	2257.5	4672.5	-.545	.585
	Late respondents	69	67.72	4672.5				
	Total	138						
PR4	Early respondents	69	69.49	4794.5	2379.5	4794.5	-.004	.996
	Late respondents	69	69.51	4796.5				
	Total	138						

*Items OR3, CX1, CP3 are significant at <0.05 level

D.3 K-S Tests for Differences between Respondents

Composite constructs*		N	Mean	Independent-Samples Kolmogorov Smirnov Test P-value	Decision
IT expertise	Non-IT manager	260	4.93	.446	Retain the null hypothesis
	IT manager	76	5.23		
Top management support	Non-IT manager	260	5.73	**.000****	**Reject the null hypothesis**
	IT manager	76	5.14		
Organizational resources	Non-IT manager	260	4.61	.973	Retain the null hypothesis
	IT manager	76	4.63		
Relative advantage	Non-IT manager	260	5.52	.915	Retain the null hypothesis
	IT manager	76	5.29		
Complexity	Non-IT manager	260	4.00	.999	Retain the null hypothesis
	IT manager	76	4.05		
External support	Non-IT manager	260	4.86	.274	Retain the null hypothesis
	IT manager	76	5.05		
Compatibility	Non-IT manager	260	4.88	.995	Retain the null hypothesis
	IT manager	76	4.90		
Security	Non-IT manager	260	3.52	.963	Retain the null hypothesis
	IT manager	76	3.51		
Competitive pressure	Non-IT manager	260	4.78	.588	Retain the null hypothesis
	IT manager	76	4.47		
Privacy	Non-IT manager	260	3.87	.757	Retain the null hypothesis
	IT manager	76	3.75		
	Total respondents	336			

*Items are averaged for each construct

**Top management support is significant at < 0.001 level

D.4 Harman's Single-Factor Test: Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Factor	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	11.724	31.687	31.687	11.089	29.971	**29.971**
2	3.681	9.949	41.636			
…	…	…	…			
36	.127	.343	99.734			
37	.098	.266	100.000			

Extraction Method: Principal Axis Factoring. Variance accounted for by a single factor: 29.97%.

Appendix E: EFA – Initial Assessments

E.1 Normality Tests

	Mean	Std. Deviation	Skewness		Kurtosis		Shapiro-Wilk	
Item	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error	Statistic	Sig.
IE1	4.79	1.350	-.344	.133	-.184	.265	.918	.000
IE2	5.07	1.291	-.608	.133	.003	.265	.917	.000
IE3	5.15	1.273	-.680	.133	.000	.265	.904	.000
MS1	5.18	1.349	-.796	.133	.450	.265	.898	.000
MS2	5.49	1.316	-1.126	.133	1.175	.265	.856	.000
MS3	6.12	.884	-1.361	.133	**3.057***	.265	.790	.000
OR1	4.37	1.517	-.348	.133	-.722	.265	.932	.000
OR2	5.26	1.369	-.786	.133	.084	.265	.894	.000
OR3	4.22	1.509	-.237	.133	-.887	.265	.931	.000
RA1	5.41	1.116	-.843	.133	1.159	.265	.885	.000
RA2	5.33	1.157	-.921	.133	1.100	.265	.883	.000
RA3	5.40	1.077	-.753	.133	.960	.265	.890	.000
RA4	5.73	1.107	-1.110	.133	1.660	.265	.853	.000
CX1	3.96	1.303	-.064	.133	-.461	.265	.943	.000
CX2	4.25	1.178	.170	.133	.022	.265	.934	.000
CX3	3.87	1.215	-.032	.133	-.309	.265	.941	.000
CX4	3.89	1.273	.002	.133	-.524	.265	.941	.000
ES1	5.24	1.095	-.586	.133	.144	.265	.882	.000
ES2	5.16	1.143	-.409	.133	-.265	.265	.892	.000
ES3	4.30	1.172	.131	.133	.456	.265	.898	.000
CM1	5.04	1.429	-.714	.133	-.058	.265	.906	.000
CM2	4.68	1.430	-.333	.133	-.606	.265	.936	.000
CM3	4.45	1.416	-.239	.133	-.739	.265	.937	.000
CM4	5.35	1.170	-.908	.133	.784	.265	.876	.000
SE1	4.58	1.397	.345	.133	-.606	.265	.933	.000
SE2	4.73	1.423	.395	.133	-.501	.265	.934	.000
SE3	4.78	1.428	.528	.133	-.300	.265	.926	.000
SE4	4.07	1.310	.098	.133	-.586	.265	.942	.000
SE5	4.24	1.218	.395	.133	-.398	.265	.921	.000
CP1	4.55	1.609	-.390	.133	-.692	.265	.932	.000
CP2	5.12	1.509	-.806	.133	-.061	.265	.889	.000
CP3	4.35	1.368	-.245	.133	-.285	.265	.939	.000
CP4	4.80	1.453	-.383	.133	-.516	.265	.929	.000
PR1	4.24	1.474	.023	.133	-.412	.265	.935	.000
PR2	3.91	1.361	.093	.133	-.233	.265	.944	.000
PR3	3.66	1.281	.123	.133	.021	.265	.930	.000
PR4	3.55	1.240	.136	.133	.269	.265	.909	.000

*Above threshold (> 1.96)

E.2 Initial KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.894
Bartlett's Test of Sphericity	Approx. Chi-Square	8345.069
	df	666
	Sig.	.000

E.3 Initial Total Variance Explained

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings[a]
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	11.724	31.687	31.687	11.405	30.824	30.824	2.854
2	3.681	9.949	41.636	3.372	9.112	39.937	2.755
3	2.791	7.544	49.180	2.463	6.657	46.593	2.741
4	1.926	5.204	54.384	1.629	4.403	50.996	2.587
5	1.730	4.676	59.060	1.473	3.981	54.977	2.472
6	1.667	4.506	63.566	1.332	3.600	58.577	2.407
7	1.219	3.294	66.860	.917	2.477	61.054	2.317
8	1.018	2.751	69.611	.697	1.882	62.937	2.247
9	.979	2.647	72.258	.648	1.753	64.690	2.133
10	.868	2.345	74.603	.565	1.528	66.218	1.987
11	.798	2.158	76.761				
…	…	…	…				
37	.098	.266	100.000				

Extraction Method: Principal Axis Factoring.

a. When factors are correlated, sums of squared loadings cannot be added to obtain a total variance

E.4 Initial Reliability Assessment

Factors	Item	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Relative advantage	RA1	.778	.868
Cronbach's Alpha: .899	RA2	.803	.859
	RA3	.818	.854
	RA4	.702	.895
Complexity	CX1	.724	.853
Cronbach's Alpha: .880	CX2	.694	.863
	CX3	.777	.832
	CX4	.769	.834
Compatibility	CP1	.673	.748
Cronbach's Alpha: .815	CP2	.753	.707
	CP2	.549	.808
	CP4	.578	.794
IT expertise	IE1	.746	.812
Cronbach's Alpha: .866	IE2	.794	.767
	IE3	.699	.853
Top Management support	MS1	.804	.598
Cronbach's Alpha: .809	MS2	.835	.558
	MS3	.456	**.933***
Organizational resources	OR1	.508	.736
Cronbach's Alpha: .745	OR2	.588	.646
	OR3	.626	.595
External support	ES1	.697	.645
Cronbach's Alpha: .790	ES2	.761	.566
	ES3	.459	**.893***
Security	SE1	.713	.860
Cronbach's Alpha: .884	SE2	.793	.840
	SE3	.683	.868
	SE4	.741	.854
	SE5	.676	.869
Competitive pressure	CP1	.740	.845
Cronbach's Alpha: .879	CP2	.823	.810
	CP3	.685	.865
	CP4	.712	.854
Privacy	PR1	.421	**.866***
Cronbach's Alpha: .810	PR2	.762	.694
	PR3	.741	.709
	PR4	.632	.761

*Improved Cronbach's Alpha if item deleted

E.5 Initial Pattern Matrix[a] and Communalities

	SE	CX	RA	IE	CP	PR	MS	ES	CM	OR	Communalities
SE1	**.652**	.174	.097	.189	.080	-.047	.097	.064	.290	.119	.620
SE2	**.766**	.039	.082	.206	.025	-.033	.200	.170	.236	.185	.798
SE3	**.663**	.079	.077	.119	.068	.075	.193	.192	.198	.107	.601
SE4	**.718**	.356	.071	.188	.118	-.096	-.025	.058	.009	.136	.729
*SE5****	*.634*	*.349*	*.055*	*.173*	*.107*	*-.169*	*-.004*	*.083*	*.002*	*.125*	.619
CX1	.160	**.719**	.151	.096	.055	.049	.131	.110	.110	.135	.616
CX2	.130	**.570**	.167	.220	.114	-.020	.213	.187	.266	.182	.639
CX3	.158	**.768**	.080	.091	.091	.018	.038	.073	.165	.260	.741
CX4	.227	**.724**	.044	.106	.005	-.012	.096	.146	.146	.266	.713
RA1	.016	.106	**.671**	.059	.340	-.042	.271	.100	.210	.130	.728
RA2	.121	.141	**.775**	.012	.174	.056	.176	.075	.251	.104	.779
RA3	.090	.136	**.748**	.032	.238	.031	.211	.140	.209	.118	.766
RA4	.050	.039	**.621**	.005	.224	-.059	.233	.131	.193	.140	.571
IE1	.103	.088	.036	**.779**	.007	.014	.109	.031	.083	.139	.666
IE2	.161	.122	.030	**.865**	.030	-.002	.140	.064	.053	.077	.822
IE3	.142	.033	-.018	**.735**	.000	-.023	.030	.157	.110	.124	.616
CP1	.045	-.042	.246	.008	**.682**	.123	.299	.114	.154	.173	.701
CP2	.083	-.035	.328	.060	**.706**	.017	.283	.135	.271	.238	.846
*CP3****	*.054*	*.183*	*.242*	*-.016*	*.598*	*.099*	*.177*	*.201*	*.184*	*.059*	.572
CP4	.092	.131	.182	.025	**.602**	-.040	.236	.188	.292	.140	.619
*PR1**	*.166*	*.050*	*.046*	*.032*	*.208*	*.499*	*.162*	*.068*	*.093*	*.010*	**.365***
PR2	-.028	-.021	-.002	-.027	.054	**.841**	-.035	-.022	.007	.087	.721
PR3	-.110	-.056	-.040	-.013	-.064	**.894**	-.050	-.035	-.094	.005	.833
PR4	-.116	.053	.001	.021	-.035	**.724**	-.055	-.056	-.035	-.005	.550
MS1	.069	.150	.200	.164	.226	-.027	**.794**	.094	.195	.213	.869
MS2	.061	.089	.225	.116	.268	-.035	**.789**	.091	.191	.210	.860
*MS3**	*.065*	*.001*	*.249*	*.073*	*.188*	*.008*	*.355*	*.213*	*.125*	*.023*	**.295***
ES1	.107	.048	.110	.059	.098	-.060	.077	**.808**	.137	.080	.727
ES2	.041	.047	.065	.116	.080	-.037	.045	**.934**	.054	.037	.908
*ES3**	*.077*	*.158*	*.044*	*.056*	*.129*	*.030*	*.130*	*.444*	*.080*	*.147*	**.295***
*CM1****	*.097*	*.118*	*.313*	*.051*	*.282*	*-.033*	*.216*	*.156*	*.607*	*.103*	.655
*CM2****	*.094*	*.191*	*.247*	*.125*	*.187*	*-.006*	*.206*	*.114*	*.721*	*.180*	.764
*CM3***	*.213*	*.300*	*.041*	*.387*	*.120*	*.023*	*.039*	*.164*	*.401*	*.324*	.596
*CM4***	*.256*	*.008*	*.186*	*.144*	*.279*	*-.113*	*.201*	*.123*	*.389*	*.284*	.498
*OR1***	*.186*	*.275*	*.077*	*.396*	*.060*	*.107*	*.261*	*.122*	*.190*	*.375*	.548
*OR2****	*.038*	*.076*	*.135*	*.064*	*.084*	*.022*	*.147*	*.114*	*.126*	*.753*	.655
*OR3****	*.119*	*.304*	*.015*	*.173*	*.118*	*.072*	*.085*	*.046*	*.079*	*.655*	.600

Extraction Method: Principal Axis Factoring. Rotation Method: Equamax with Kaiser Normalization[a]
a. Rotation converged in 8 iterations.

*Items dropped due to communalities < 0.5, **Items dropped due to factor loadings < 0.4,
***Items dropped due to convergent validity: AVE < 0.5 or CR < 0.7.

Appendix F: EFA – Final Solution

F.1 Final Solution: KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.855
Bartlett's Test of Sphericity	Approx. Chi-Square	5690.589
	df	300
	Sig.	.000

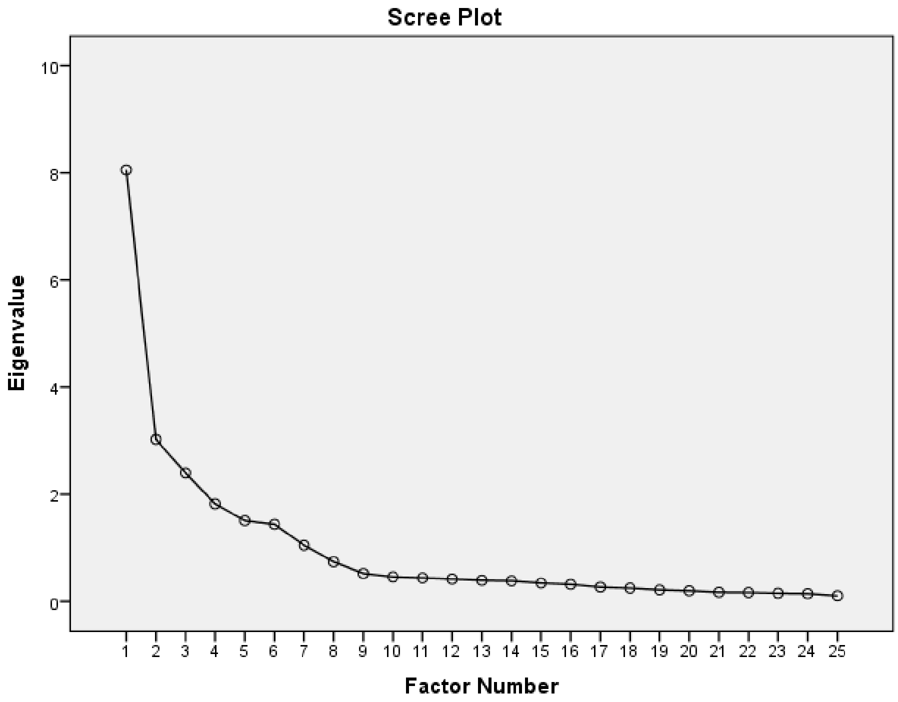
F.2 Final Solution: Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings[a]
Factor	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	8.056	32.224	32.224	7.792	31.167	31.167	2.581
2	3.026	12.102	44.326	2.745	10.979	42.146	2.534
3	2.399	9.596	53.922	2.124	8.496	50.642	2.432
4	1.822	7.287	61.209	1.533	6.134	56.776	2.213
5	1.509	6.034	67.243	1.315	5.260	62.036	2.197
6	1.440	5.760	73.004	1.158	4.632	66.668	2.153
7	1.050	4.199	77.203	.815	3.259	69.926	2.052
8	.742	2.967	80.170	.536	2.146	72.072	1.856
9	.522	2.086	82.256				
10	.457	1.830	84.086				
…	…	…	…				
25	.106	.426	100.000				

Extraction Method: Principal Axis Factoring.

a. When factors are correlated, sums of squared loadings cannot be added to obtain a total variance.

F.3 Scree Plot



F.4 Inter-Item Correlation Matrix for Two Items Constructs

	MS1	MS2	OR2	OR3	ES1	ES2	CM1	CM2
MS1	1.000							
MS2	**.874**	1.000						
OR2	.380	.369	1.000					
OR3	.341	.298	**.585***	1.000				
ES1	.239	.255	.220	.145	1.000			
ES2	.224	.206	.150	.147	**.808**	1.000		
CM1	.497	.479	.290	.228	.308	.255	1.000	
CM2	.508	.488	.363	.316	.288	.226	**.730**	1.000

*Does not satisfy r > 0.7

F.5 CR and AVE for Constructs with Two Items

	Composite Reliability (CR)	Average Variance Extracted (AVE)
Top management support	.807	.677
Organizational resources	**.664***	**.498****
External support	.866	.764
Compatibility	**.613***	**.444****

*CR < 0.7, **AVE < 0.5 → omitted constructs

Appendix G: Ordinal Logistic Regression

Case Processing Summary

		N	Marginal Percentage
Assimilation	Initiation	191	56.8%
	Adoption-decision	101	30.1%
	Implementation	44	13.1%
Size	Employees 50-250	215	64.0%
	Employees > 250	121	36.0%
Total		336	100.0%

Model Fitting Information

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	637.467			
Final	465.706	171.761	9	.000

Test of Parallel Lines

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	465.706			
General	453.274	11.898	9	.219

The null hypothesis states that the location parameters (slope coefficient) are the same across response categories.

Model Output

Parameter estimates

		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Threshold	[Initiation]	.355	.227	2.445	1	.118	-.090	.801
	[Adoption-decision]	2.752	.282	95.549	1	.000	2.200	3.304
Location	Complexity	-.084	.134	.394	1	.530	-.178	.346
	Relative advantage	.421	.158	7.063	1	**.008***	.110	.731
	Security	-.580	.140	17.253	1	**.000****	.306	.854
	IT expertise	.432	.138	9.804	1	**.002***	.162	.702
	Competitive pressure	1.389	.186	55.705	1	**.000****	1.025	1.754
	Privacy	-.156	.125	1.553	1	.213	-.400	.089
	Top management support	1.324	.186	50.440	1	**.000****	.959	1.690
	External support	-.052	.135	.149	1	.699	-.318	.213
	[Size=Employees 50-250]	-.328	.257	1.628	1	.202	-.831	.176
	[Size=Employees > 250]	0[a]	.	.	0	.	.	.

Link function: Logit. ***Pseudo R-Square: .471 (Nagelkerke)***
a. This parameter is set to zero because it is redundant.
* Significance at < 0.05 level
** Significance at < 0.01 level

Appendix H: MLR – Assumptions

H.1 Tolerance and VIF Test for Multicollinearity

	Collinearity Statistics	
	Tolerance	VIF
Relative advantage	.992	1.008
Security	.998	1.002
Complexity	.998	1.002
IT expertise	.997	1.003
Privacy	1.000	1.000
External support	.999	1.001
Top management support	.996	1.004
Competitive pressure	.991	1.009

Tolerance level > 0.2 and VIF < 10: **This assumption was met.**

H.2 Linearity of the Logit

Likelihood Ratio Tests

	Model Fitting Criteria	Likelihood Ratio Tests		
Effect	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	439.312[a]	.000	0	.
Size	440.502	1.190	2	.552
CX	440.348	1.036	2	.596
RA	440.544	1.233	2	.540
SE	440.948	1.636	2	.441
IE	440.207	.895	2	.639
CP	443.219	3.907	2	.142
PR	445.092	5.780	2	.056
MS	441.478	2.166	2	.339
ES	439.453	.141	2	.932
Ln(CX)*CX	439.450	.138	2	.933
Ln(RA)*RA	440.970	1.658	2	.437
Ln(SE)*SE	443.529	4.217	2	.121
Ln(IE)*IE	440.496	1.184	2	.553
Ln(CP)*CP	440.044	.733	2	.693
Ln(PR)*PR	444.501	5.189	2	.075
Ln(MS)*MS	442.256	2.944	2	.229
Ln(ES)*ES	439.613	.301	2	.860

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

The logit transformation does not work on negative values and zeroes. To avoid this problem, a constant k, equal to the minimum value of the sample plus 1 is added: k=min(x)+1, so that ln(x+min(x)+1). The transformation does not modify the distribution.

None of the interactions in the form xln(x) are significant ($p > 0.05$). Hence, **this assumption was met.**

Appendix I: MLR – Model Fit

Case Processing Summary

		N	Marginal Percentage
Assimilation	Initiation	191	56.8%
	Adoption-decision	101	30.1%
	Implementation	44	13.1%
Size	Medium-sized companies [Employees 50-250]	215	64.0%
	Large companies [Employees > 250]	121	36.0%
Valid		336	
Missing		0	
Total		336	100.0%

I.1 2-log Likelihood Test

Model Fitting Information

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	637.467			
Final	456.421	181.046	18	.000

I.2 Pseudo R-Square

Cox and Snell	.417
Nagelkerke	.490

Appendix J: MLR – Model Output

Model Output

Parameter estimates

		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B) Lower Bound	Upper Bound
Initiation vs. Adoption-decision[a]	Intercept	-.521	.249	4.359	1	.037			
	Complexity	.161	.158	1.035	1	.309	1.175	.861	1.602
	Relative advantage	.216	.179	1.459	1	.227	1.242	.874	1.764
	Security	-.333	.155	4.601	1	**.032***	**.717**	.528	.972
	IT expertise	.166	.158	1.102	1	.294	1.181	.866	1.610
	Competitive pressure	1.239	.207	35.815	1	**.000****	**3.453**	2.301	5.181
	Privacy	-.055	.152	.128	1	.720	.947	.702	1.277
	Top management support	1.087	.198	30.026	1	**.000****	**2.965**	2.010	4.374
	External support	-.028	.155	.033	1	.855	.972	.718	1.316
	[Size=Employees 50-250]	-.379	.303	1.571	1	.210	.684	.378	1.238
	[Size=Employees > 250]	0[b]	.	.	0	.	.	.	.
Adoption-decision vs. Implementation[b]	Intercept	-2.217	.502	19.514	1	.000			
	Complexity	-.559	.221	6.420	1	**.011***	**.572**	.371	.881
	Relative advantage	.393	.260	2.278	1	.131	1.481	.889	2.466
	Security	-.749	.258	8.414	1	**.004***	**.473**	.285	.784
	IT expertise	.655	.244	7.221	1	**.007***	**1.925**	1.194	3.104
	Competitive pressure	.723	.341	4.483	1	**.034***	**2.060**	1.055	4.021
	Privacy	-.161	.189	.718	1	.397	.852	.588	1.234
	Top management support	1.083	.382	8.051	1	**.005***	**2.954**	1.398	6.244
	External support	-.118	.224	.278	1	.598	.889	.573	1.378
	[Size=Employees 50-250]	.112	.409	.075	1	.785	1.118	.502	2.493
	[Size=Employees > 250]	0[c]	.	.	0	.	.	.	.
Initiation vs. Implementation[a]	Intercept	-2.737	.509	28.876	1	.000			
	Complexity	-.398	.236	2.850	1	.091	.672	.432	1.066
	Relative advantage	.609	.278	4.803	1	**.028***	**1.839**	1.067	3.170
	Security	-1.082	.272	15.830	1	**.000****	**.339**	.199	.577
	IT expertise	.821	.258	10.158	1	**.001***	**2.273**	1.372	3.766
	Competitive pressure	1.962	.359	29.835	1	**.000****	**7.112**	3.518	14.378
	Privacy	-.215	.209	1.058	1	.304	.806	.535	1.215
	Top management support	2.170	.396	30.078	1	**.000****	**8.759**	4.033	19.023
	External support	-.146	.235	.388	1	.534	.864	.545	1.369
	[Size=Employees 50-250]	-.268	.442	.366	1	.545	.765	.322	1.821
	[Size=Employees > 250]	0[c]	.	.	0	.	.	.	.

a. The reference category is: Initiation
b. The reference category is: Adoption-decision
c. This parameter is set to zero because it is redundant.

* Significance at < 0.05 level and ** Significance at < 0.01 level