UNIVERSITY OF THE WESTERN CAPE

Understanding the factors that influence the adoption of Big Data at a government department in the Western Cape.

A thesis submitted in fulfilment of the requirements for the
Master's degree in Information Management in the
Faculty of Economic and Management Sciences
of the University of the Western Cape

by

Andre Bruintjies

Supervisor: Prof. James Njenga

August 2022

Abstract:

Purpose / Problem Statement

Big Data, henceforth (BD), enables governments to produce and accumulate huge amounts of structured and unstructured data through their everyday activities. Due to the types of data produced and the volume, velocity and complexity of the data, government organisations need to find innovative ways to analyse it. Government organisations have since realised the potential to derive value from BD and, therefore, a need to adopt BD into their data activities. Currently, South African governmental organisations have not fully committed to adopting BD because they are unsure if they are adequately equipped, and additional empirical research is required to understand the salient factors that influence BD adoption.

Design methodology / Approach

This study investigated the factors influencing BD adoption in a governmental organisation through an explorative case study. Data was collected through open-ended interviews and analysed through a thematic analysis approach to find themes and relationships amongst them. The Technology-Organization-Environment (TOE) framework was used to identify and investigate these factors.

Key findings

The results indicate that most of the factors that influence BD adoption were discovered in the organisational context of the TOE framework and that the successful adoption of BD is dependent on understanding how these factors influence the adoption process and how these factors can be mitigated. Without properly understanding these factors, government organisations may struggle to adopt BD and use its full potential for efficient service delivery to citizens.

Research limitations

This research was conducted at one government organisation in the Western Cape, South Africa, because of time and budget constraints. A contributing factor for conducting this study at one government organisation was due to the effects of the COVID-19 (infectious disease caused by the SARS-CoV-2 virus) pandemic and access to study participants. For example, the organisation and university had a moratorium that prevented face-to-face access to study participants, limiting access to the number of participants that the research would have liked to interview.

Originality / Value

There is scarcity in research that focuses specifically on the factors that influence the adoption of Big Data (BD) at a government organisation in the Western Cape, South Africa. Therefore, this research provides an overview of the factors influencing the adoption of Big Data. In addition, this study provides

government executives with the factors that need to be considered for BD adoption and the recommendations on how to adequately address these factors to ensure successful adoption.

Keywords:

Adoption

Data

Big Data

Government

Technology

Information Systems

Information Technology

Service Delivery

Factors

Technology-Organisation-Environment (TOE) framework

Declaration:

I declare that this study, with the title "**Factors Influencing Big Data Adoption in a Government Institution**", is my own work, and that it has not been submitted before for any degree or examination at any other university, and that all the sources I have used or quoted have been indicated and acknowledged as references.

Andre Bruintjies

Date:

29 August 2022

Acknowledgements:

Firstly, I would like to thank the **Almighty God** for giving me the wisdom, strength and determination to complete this research study.

Secondly, I would like to express my sincere appreciation to my supervisor, Dr James Njenga, for his effective and unceasing support during my study. Your patience, motivation and immense knowledge assisted me tremendously.

Thirdly, I would like to thank my wife and children for their sacrifices during my studies. Thank you for being patient and affording me the time needed to complete my studies.

Fourthly, I would like to thank my family (mother, brother and sisters), friends and colleagues. Your continued belief and encouragement kept me focused.

Lastly, I thank the government organisation and its participants who contributed to my research study. Without your approval and participation, this study would not have been a success.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1. Background

Merendino *et al.* (2018) claim that data provides an understanding of environmental trends that creates greater insight into people's behaviour and provide actions through data. Gale (2013, p. 32) emphasises the significance of data to the government, stating that "*Governments live and die by data*". Gale further argues that for governments to serve their citizens truly, they need to transform their data into knowledge, giving them a better understanding of what citizens need and how they behave. Grover *et al.* (2018) argue that government organisations are changing their view on data, realizing that data is an important strategic asset that needs to be protected, leveraged, and analysed to meet the requirements of their citizens. Due to these changes in the data landscape, Wang *et al.* (2018) argue that Big Data (BD) can help transform data and the decision-making process by providing enhanced visibility of organisational operations and improved performance measurement mechanisms.

BD has the potential to expose people's hidden behavioural patterns and shed light on their intentions (Omoyiola, 2022). More precisely, it can explain what people want to do, what they do, and how they interact with others and their environment (Michael & Miller, 2013). Yan (2015) claims that BD can improve predictions, save money, boost efficiency, and improve decision-making in traffic control, weather forecasting, disaster prevention, finance, fraud control, business transaction, national security, education, and health care.

Government agencies have realised BD's significance and started integrating it into their data programs, for example, by integrating BD and e-government (Liu & Ma, 2019). Citizens produce large amounts of data every day in e-government systems (Liu & Ma, 2019). Ogbuju *et al.* (2019) state that e-government is data-driven because data moves in and out government channels from public and private domains. Sometimes processed data becomes the input for other processes (Ogbuju et al., 2019).

Liu and Ma (2019) emphasise that incorporating BD in e-government delivers i) information and data sharing, ii) statistical analysis and visualisation of the results of analysis, and iii) predictive analysis. The information produced through analysing DB is useful to government agencies in supporting decision making in areas such as law enforcement, social services, homeland security and, where situational awareness and anticipatory approaches are required for near-real-time decision making (Michael & Miller, 2013).

However, despite the excitement and interest in BD and its high operational and strategic potential, little is known about the factors involved in adopting BD in government organisations (Riggins & Wamba, 2015). With the potential to provide powerful competitive advantages, government organisations

struggle to establish effective governance and privacy in connection with BD initiatives (Shin & Choi, 2015). While the potential of BD technology is real, the adoption in government is lagging because of fundamental concerns about high expectations and excessive financial investment (Eynon, 2013). As Boyd and Crawford (2012) critically note, most discussions concerning BD have been technologically biased and industry-oriented, leaning towards the technical aspects of its design and not adoption.

Government agencies are still lagging in adopting BD despite research and business reports showing how government can use BD to serve citizens and overcome national challenges such as rising health care costs, job creation, natural disasters prevention, and terrorism (Kim *et al.*, 2014). However additional empirical studies are required to assess BD's potential and the adoption process (Wamba et al., 2015). This study explores the factors influencing the adoption of BD in a government organisation. The conceptual framework developed in this research could serve as a guideline to government organisations that experience challenges in adopting BD or those interested in adopting such technology.

1.2. Problem statement

Data has become a critical asset for government agencies to enhance societal needs (Gutierrez, 2017; Jewell *et al.,* 2014). Governments produce and accumulate huge amounts of data through everyday activities, such as managing pensions and allowance payments, tax collection, national health system records, recording traffic data, and issuing official documents (Cavanillas *et al.*, 2016). The knowledge hidden in large datasets is increasing, as is the pressure on governments to turn the data into actionable information (Gutierrez, 2017).

The effective use of data could be the missing link between good governance and capacity building, where insights through data is used to improve service delivery (Gutierrez, 2017). Government institutions have realised that accurate, up-to-date data is required to improve service delivery, accountability, transparency, efficiency and productivity (Arpaci et al., 2012; Hoti, 2015). Furthermore, BD directly affects the economy because public sector income is collected through taxes and social contributions, efficient services relates to direct cost savings (Cavanillas *et al.* 2016). For example, the more efficient the public sector is, the better-off citizens are, as fewer resources (taxes) are needed to provide the same level of service (Cavanillas *et al.*, 2016).

However, because of the volume, velocity, and variety of data produced, some organisations struggle to support their data requirements: capturing, storing, analysing, and using their data. (Han *et al*. 2014). These data requirements demand that organisations re-examine their current data strategies and have appropriate cultural, policy, legal, regulatory, institutional, organisational, and technical environments in place, which is required to control, manage, share, protect, and extract data value (OECD, 2019).

Furthermore, organisations often experience legacy challenges that is inherited from analogue business models, that stems from outdated data infrastructure, data silos, skill gaps, regulatory barriers, the lack of leadership and accountability, and an organisational culture which is not prone to digital innovation and change (OECD, 2019).

As an example of BD adoption, the British HM Revenue and Customs Office successfully adopted the British Connect system at forty-five million pounds (£) to detect tax fraud and other irregularities (Maciejewski, 2017; Fenwick & Vermeulen, 2019). In its first year of existence, the system brought in an additional one point four billion pounds for the treasury (Maciejewski, 2017). Because of the system's success, the British government granted an extra one hundred and fifty million pounds to develop further tax avoidance and evasion methods and reduce fraud, error, and debt in the tax credits systems (Maciejewski, 2017). This resulted in the recovery of thirty-five billion pounds in unpaid taxes in 2014 (Maciejewski, 2017).

Also, various governments have established digital or e-government programs to introduce data science technologies into the public sector (Engin & Treleaven, 2019). A leading example is Government as a Platform (GaaP) which is used in Estonia (Engin & Treleaven, 2019). The conception of GaaP offers to encapsulate the use of digital technologies to support the resolution of common problems at various levels, including city, county, national and regional, through shared software, data and services, and in this manner, improving the efficiency and effectiveness of government and governance by doing more for less (Margetts & Naumann, 2017).

Even though government agencies of other countries are adopting BD, South African government departments are lagging when it comes to adopting BD (Gang-Hoon *et al.* 2014). In general, South African government departments recognise the potential of BD, but seem hesitant about accepting its introduction, and do not show enough conviction of being adequately equipped for its use (Klievink *et al*., 2016).

Government departments are under pressure to use data for greater responsiveness and efficiency in various areas, like identifying fraud, improving regulatory oversight, optimizing business travel, or providing more effective direct services (Morabito, 2015). South African government organisations tend to have disjointed data, making it difficult to use as it is not integrated. Furthermore, the data is not readily available for use by management as it sits in varied data systems, hindering analysis and decision making. Since data is contained in disjointed systems, management often receives inconsistent data and late reports, which impacts on their decisions and prioritising.

As Klievink *et al*. (2016) add, BD would offer great opportunities for South African government departments to structurally improve and transform. Thus, South African government departments should not wait any longer before putting strategies in place to adopt BD. Failing to do so deprives the government and the citizens of the opportunity to find speedy resolution on issues such as efficiency in grant program management, improved services based on better insight into citizen demands and needs, solving persistent social problems (such as transport congestion, healthcare provision, sustainable energy production), identifying more opportunities for savings, and better understanding of their own operations as well as constituents' needs (Klievink *et al.*, 2016).

1.3. Research approach

This study investigated the factors influencing BD adoption in a Western Province governmental organisation. The study used qualitative methods in the form of an explorative case study to understand the interactions between information technology (IT) related innovations and organisational contexts. Case studies are empirical inquiries that investigate a current phenomenon within its natural context, allowing the researcher to understand the phenomenon under study (Andrade, 2007). Creswell (2013) states that this type of research inquiry seeks to establish the meaning of a phenomenon from participants' views through in-depth accounts of occurrence.

First, as motivation for selecting a case study design for this research study, Andrade (2007) explains that case studies aim to answer research questions regarding 'how' or 'why' a phenomenon exists. Second, the objective of a case study is to do intensive research on a specific person or population, such as a group, organisation, or community (Rashid *et al*., 2019). This applies to a study population identified as a government organisation in the Western Cape. Third, case studies allow the researcher to focus on behaviours, attributes, actions, and interactions, and it is the preferred strategy when the researcher has little control over events and when the focus is on current phenomena within some real-life context (Rashid *et al*., 2019).

This study adopted the Technology-Organization-Environment (TOE) framework to understand the factors that influence BD adoption at a government organisation. Various other theories exist, like the technology acceptance model (TAM), theory of planned behaviour (TPB) and the unified theory of acceptance and use of technology (UTAUT) (Oliveira & Martins, 2011; Hoti, 2015). The motivation for using the TOE framework is that it specifically focusses on studying a phenomena at an organisational level, which none of the other theories do (Oliveira & Martins, 2011; Hoti, 2015).

The interpretivist's approach was used in this study to guide the construction of reality in a social world about the adoption of BD in a Western Cape government organisation. Interpretivism tries to understand phenomena by assessing the meaning that people (participants) assign to them and focusing on their

cultural and historical context (Thomas, 2010). The emphasis is on performing research amongst people instead of objects (Sheppard, 2020). Walshan (1995) argues that interpretive research methods assume that human actors socially construct people's knowledge of reality. Also, the reality is derived from sense-making of the environment and shared meanings from subjectivity rather than objectivity (Iyamu & Mgudlwa, 2018).

An interpretive approach provides a deep insight into the complex world of lived experiences from the viewpoint of those who live it (Andrade, 2007). Similarly, this approach is consistent with constructing the environment characterised by the interaction between the researcher and the participants (Scotland, 2012). Interpretive research assumes that reality is socially constructed, and the researcher becomes the vehicle by which this reality is revealed (Andrade, 2007).

This study aims to obtain a rich set of data through structured interviews on the specific factors that affect the adoption of BD in a governmental organisation while capturing the contextual complexity (Benbasat *et al*., 1987). This is in line with Baxter and Jack's (2008) assertion that case study research may have multiple sources in documentation, archival records, interviews, physical artefacts, direct observations, and participant observation.

During this research inquiry, open-ended interviews were conducted, because they enable participants to provide in-depth accounts of the research problem being studied (Creswell, 2008; Walsham, 1995). Qualitative interviews are an excellent way to gather detailed information, not only because participants may elaborate on their answers or statements relating to BD adoption in a way that is not possible with other methods, but because they can share information with researchers in their own words and from their perspectives, rather than attempting to fit those perspectives into the perhaps limited response opinions (Sheppard, 2020).

Babbie (2005) suggests that this approach's inherent flexibility is one of its major advantages. Therefore, there is an expectation that the participants' views are more freely expressed when the interview format is more flexible and open-ended than when the interview style is regulated and confined to a standard set of questions or a survey (Flick, 2002). In order to fully cover the research questions, however, it is important for the researcher to have a set interview schedule that serves as a guide to the discussion and mitigates the chance of important topics being neglected.

1.4. Key findings

A key finding of the research study shows that the employees of the organisation under study, which this study names OrganisationXYZ, had a good understanding of BD as a technology which stems from the fact that the organisation is becoming data-driven, for example, grounding decisions on data instead

of on (mere) sentiments. The organisation's process for technology adoption was evaluated and amended to incorporate the major phases of innovation adoption, namely, initiation and implementation, as described by Rogers (1995). Additional information regarding the process was discovered and is presented in Chapter 4.

Also, an initial list of eleven factors that influence the adoption of BD at the organisation were identified and described in Chapter 2. Six additional factors that influence the adoption of BD were discovered during the research. These factors are presented and discussed in Chapter 5.

1.5. Research benefits / Key finding implications

The purpose of the study was to provide insight into the factors that influence the adoption of BD at a government organisation in the Western Cape. This was accomplished by providing evidence of how the participants in this research study perceived the influence of the identified factors on the BD adoption process within the organisation.

The study presents the factors identified through the use of the Technology-Organisation-Environment (TOE) framework and gives recommendations on how to address these factors. These factors were discovered through applying the TOE framework, presented in figure 1 below and discussed in Chapter 5.

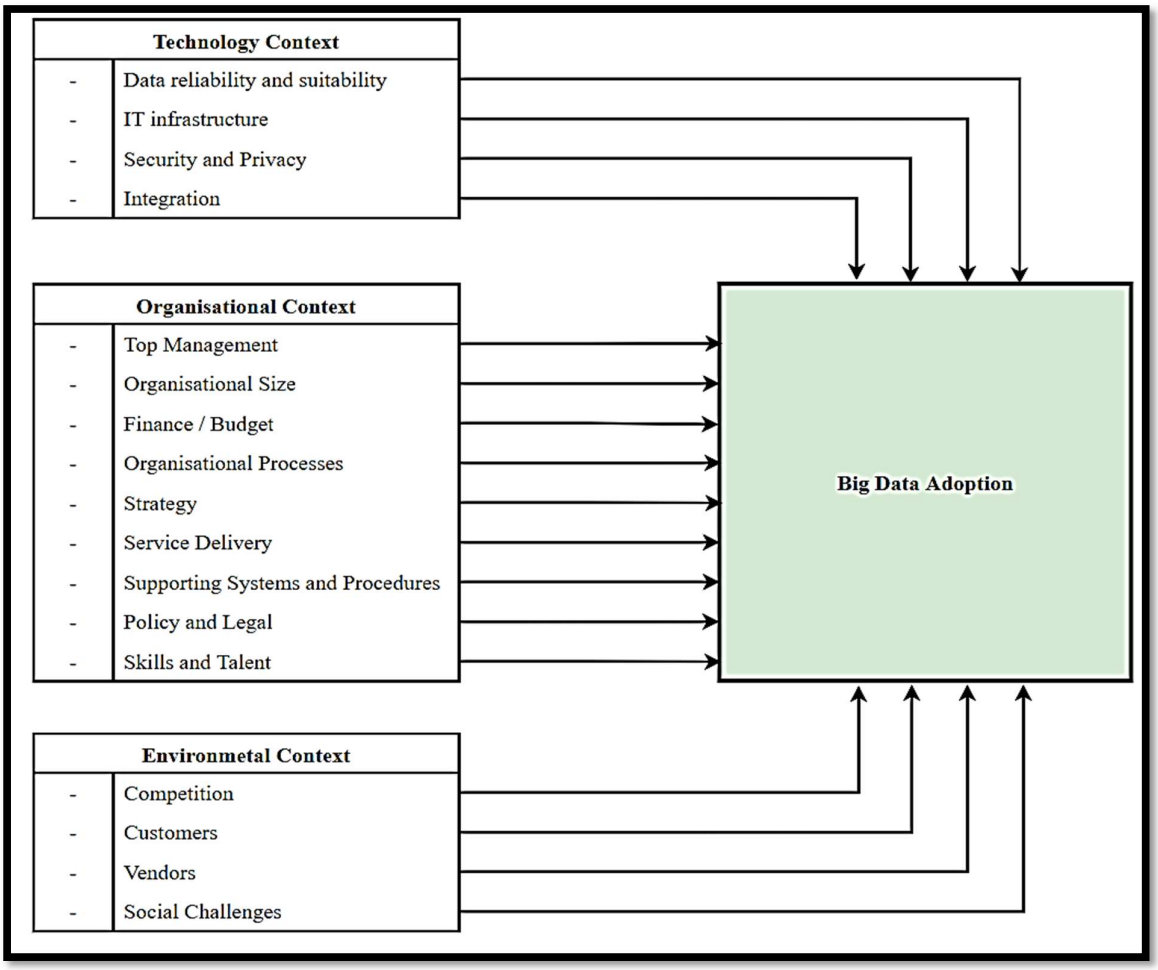


Figure 1: Government Big Data Adoption (Adapted from Sun *et al.*, 2018; Almoqren & Altayar, 2016; Thabet & Soomro, 2015; Engin & Treleaven, 2019).

It is hoped that the outcome of this research study enables government staff to understand the factors that influence the adoption of BD, ultimately incorporate the technology into effective decision making, and overcome potential deficiencies in the use of information.

1.6. Research questions and objectives

It is crucial for organisations to understand the benefits and challenges of BD adoption because it leads to new data-driven services that improve processes and enable innovative products and services (Bremser, 2018). Bremser (2018) claims that only 14% of organisations that start with BD implementation projects have done so successfully. Therefore, it is important to study the BD adoption process for academic, science, and practical interest.

Based on the research statement, the main research question has been identified, as stated below. This study looks at the main research question as constituting three specific questions, which are also stated below, immediately after the main research question.

What factors affect Big Data adoption at a government department in the Western Cape?

- What are the challenges and benefits of adopting BD in the public sector?
- How does BD adoption impact the strategy of Western Cape Governmental organisations?
- How is BD adoption affecting the constituents of OrganizationXYZ the Western Cape?

To answer these questions, the research objectives of this study are to:

a. Assess the current state of literature relating to the concept of BD, the use of BD in government organisations, and the factors that influence BD adoption.
b. Identify how the organizations technology adoption process can be aligned for Big Data.
c. Explain if/how the identified factors influence Big Data adoption in government organisations.

1.7. Background research

The unlimited potential of a data-driven economy has been widely recognised, leading to a growing interest in BD (Chen *et al.,* 2012). The data that government agencies collect can reveal gaps and inefficiencies that would have been hidden, helping government agencies save money, increase transparency, and improve services (Gale, 2013). However, to uncover the full potential, the data have to be gathered and analysed on time, and many government agencies cannot process the data surge (Gale, 2013).

Government departments collect and produce significant amounts of data while performing their business processes and activities (Hossain & Chan, 2015). Therefore, data has become a vital asset for government agencies because it has the potential to unravel problems through its unique ability to provide deeper insight from masses of data that are continuously being produced and collected (Braganza *et al.*, 2017; Gutierrez, 2017). Government agencies realise the importance of data, and the need to collect and store large data sets and use that data when making critical decisions (Gutierrez, 2017). For example, the United States Internal Revenue Service used BD analytics to fix errors early in the tax filing process for 2012, saving $100 million in erroneous claims (Yan, 2015).

Entire government systems can benefit from using BD (Kumar *et al.*, 2016). However, government agencies need to devote time and allocate budget and resources to adoption projects (Kumar *et al.*, 2016). In turn, successful adoption leads to effective citizen care, through data-driven services (Kumar *et al.*, 2016). Sun *et al.* (2016) elaborates by stating that BD creates business value through its unique

analytic ability and predictive decision support capabilities, allowing it to handle data that could not be processed with traditional methods.

Many authors have defined BD differently, but all definitions refer to it as advanced techniques and technologies that enable the capture, storage, distribution, management, and analysis of large, complex, structured and unstructured datasets to derive facts (Gandomi & Haider, 2015; Dumbill, 2013; Yan, 2015; Eynon, 2013; Agrawal *et al.*, 2011; Sun *et al.*, 2016). Madden (2012) defines BD as data that is "too big" (in terms of volume), "too fast" (velocity) and "too hard" (variety) for earlier technologies to process (Madden, 2012). Rossouw (2012) refers to it as datasets that are large and difficult to manage using traditional database tools. Yan (2015) defines it as innovative methods and technologies to collect, store, disseminate, manage, and analyse larger-sized datasets with high-velocity and diverse structures that conventional data management methods cannot handle. For this study, the following definition has been adopted:

"Big Data is a term that describes large volumes of high-velocity, complex, and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the information" (Yan, 2015, p. 3).

Technology adoption has been studied for many decades and has produced many theoretical models. Researchers have studied the process of BD adoption using the Technology-Organisation-Environment (TOE) framework (Almoqren & Altayar, 2016; Nam *et al.*, 2015; Sun *et al.*, 2018). The focus of these studies was to determine the factors that influence the adoption of technology at an organisational level instead of an individual level (Almoqren & Altayar, 2016). Almoqren and Altayer (2016) suggest that the TOE framework is suitable for these types of studies because it measures the success of information systems adoption in terms of human, organisational, and environmental factors.

Technological factors play a significant role in ensuring successful adoption of information systems, including technology integration, IT infrastructure, technology readiness, and technology resources (Almoqren & Altayar, 2016). The term 'organisational context' refers to top management, human resources, IT training programs and information management (Almoqren & Altayar, 2016). Sun *et al.* (2018) state that it (organisational context) is the resources and other characteristics of the organisation, such as organisational size, organisational structure, managerial structure, human resources, and employees' skills. The environmental context describes how the organisation conducts business and includes competitors, information intensity, and regulatory environment (Almoqren & Altayar, 2016). Nam *et al.* (2015) claim that it includes the size and structure of the industry, the organisation's competitors, and dealings with the government.

1.8. Delineation and limitation of the research

This research was conducted at one government organisation in the Western Cape, South Africa, because of time and budget constraints. The main department that was targeted is the organisation's Information Communications Technology (ICT) department whose strategic plans, operational plans, policies, and procedures were reviewed, and interviews conducted. The motivation for conducting the research in these departments was that their staff would be the most familiar with technology adoption within the organisation. No other governmental institution was included in this study. This approach ensures that only primary data relevant to the ICT departments of this specific organisation is reviewed. However, this is not seen as a substantial limitation and does not affect the reliability and findings of this study.

1.9. Ethical considerations

The data collection and handling were completed in accordance with the ethical requirements of the university. The Humanities and Social Science Research Ethics Committee provided ethical approval through Ethics Reference Number HS 18/9/13 (Appendix 3).

The following five were the ethical considerations for this study: i) conducting the research in a manner that safeguards the identity of the participants, organisation, university, and the researcher; ii) analysing the collected data with the highest integrity to ensure that it is accurate, concise, and specific to the proposed research; iii) participants were requested to partake in the study voluntarily; iv) the researcher explained the nature of the research to which participants were being asked to contribute and avoided any unfair, prejudiced, or discriminatory practice; and v) a formal request asking for permission to conduct the research was sent to the organisation, accompanied by a letter of reference from the university (Appendix 4 and 5).

1.10. Contribution to research

This research case study aims to add value to existing theory and the body of knowledge in Information Technology adoption at government institutions. This research provides South African governmental institutions with a basis for determining the most suitable method for the adoption of Big Data. Academically, it contributes to the information technology adoption theory, which other governmental institutions and scholars can use.

1.11. Chapter outline

The chapter outline of this research study has been presented below in tabular form, with the description of each chapter on the left and the purpose on the right.

Chapter	Description	Purpose
1	Introduction	This chapter introduces the research and explains the objective and reasoning for the research.
2	Literature Review	This chapter outlines the literature findings associated with the research problem and research objectives. It outlines the current state of the literature regarding Big Data characteristics, use, and adoption.
3	Research Design and Methodology	This chapter describes the research design and methodology and motivates the research study's research methods.
4	Research Findings: Presentation and Discussions	This chapter presents findings from the case study and interviews that were conducted.
5	Conclusion and Recommendations	This chapter summarises the research results, and provides recommendations based on the research and identifies areas for further research.

Table 1: Chapter outline (Source: The researcher, 2022)

2. Literature review

2.1. Introduction

This chapter presents a critical synthesis of empirical literature describing what BD is, how it can be used in government organisations, the challenges, and benefits, and how it can promote service delivery. It also describes the factors identified in the Technology-Organisation-Environment framework that influence the adoption of BD in governmental organisations.

2.2. The concept of Big Data (BD)

BD encompasses collecting, managing, and analysing massive volumes of different data types at the right speed within the right time frame while providing meaningful output to the end-user (Runting *et al.*, 2020). Since data constantly changes, it is not easy to express how much data defines BD in specific, measurable terms. Hardware and software differ from industry to industry and organisation to organisation, so generalising technologies is also challenging (Klievink *et al.,* 2016).

De Mauro *et al.* (2016) state that BD can be classified into four groups, based on its i) attributes, ii) technological needs, iii) thresholds, and iv) social impact. Data attributes provide one of BD's most popular definitions, underpinning the expected three-dimensional increase in data: Volume, Velocity and Variety (Kitchin & McArdle, 2016). Taleb *et al.* (2018) claim that BD is more than simply a matter of size, but an opportunity to find insights in new and emerging types of data and content, to make useful decisions. Various authors extended the "3 Vs" model, adding other BD features, such as Veracity and Value (Abdullah, 2017; Walker & Brown, 2019). Technological needs focus on the processing power of large amounts of data. Microsoft describes BD as a process in which serious computing power is applied to massive and often highly complex data sets (De Mauro *et al.*, 2016). Similarly, the National Institute of Standards and Technology (NIST) emphasises the need for scalable architectures for efficient storage manipulation and BD analysis (De Mauro *et al.*, 2016).

On the threshold, BD has been compared to the past analytical power and approaches (De Mauro *et al.*, 2016). Morabito (2015) reports that BD requires processing capacities that exceed those of traditional database systems. De Mauro et al. (2016) mention that the concept of "big" in terms of size is associated with Moore's Law and consequently with the capacity of commercial storage solutions, while Oussous *et al.* (2018) suggest that the most significant challenge is related to scalability when dealing with BD analytics. The concept is often described in terms of datasets that are so large, varied and dynamic that conventional hardware and software cannot process them because of their high volume, velocity and variety (Sun *et al.*, 2018; Taleb *et al.*, 2018).

Finally, the social impact orientation view of BD as "a cultural, technological, and scholarly

phenomenon" is based on three elements, namely: i) technology, which refers to the maximisation of computational power and algorithmic accuracy; ii) analysis, which is the identification of patterns on large data sets; and iii) mythology, which refers to the belief that large data sets offer a superior form of intelligence, carrying an aura of truth, accuracy and objectivity (De Mauro, Greco & Grimaldi, 2016; Mills, 2018).

De Mauro *et al.* (2016) describe BD in three main shifts in analysing data that improve our understanding of society and its organisation. These are: i) more data, using all available data instead of samples; ii) messier, using even incomplete data instead of limiting it to complete data; and iii) correlation, which becomes more important, overtaking causality as a privileged means to decide (De Mauro *et al.*, 2016). Furthermore, Silva *et al.* (2019) mention that organisations should not only collect data, but also find appropriate ways to analyse it in order to produce actions based on statistics and trends.

BD is also regularly explored when the prevailing traditional Relational Database Management Systems (RDBMS) and file systems processing capacities of the organisation are exceeded because of high transactional volumes, velocity, responsiveness, quantity, and data (Bhadani & Jothimani, 2016; Sadineni, 2020). Datasets are currently produced through many advancements in technology. They are also large, move too fast, and often do not fit in traditional RDBMS architectures (Kalema & Mokgadi, 2017). To gain value from these datasets, organisations need to seek alternative ways to process it (Bhadani & Jothimani, 2016).

One of BD's most attractive features is its potential to be used in virtually any situation where data is available (Klievink *et al.*, 2016). Data generation has increased because of the conversion of information into digital format, the emergence of different social network platforms, blogs, sensors, the adoption of hand-held digital devices and the explosion in using the Internet (Bhadani & Jothimani, 2016; Sadineni, 2020). Organisations have used these large datasets and applied BD analytics to make informed decisions, which provides them with an edge over their competitors (Bhadani & Jothimani, 2016).

First, with the advent of BD, and transactional processing being more complex because data streams from multiple sources, it is difficult to use only Relational Database Management Systems (RDBMS) and analysis techniques like Structured Query Language (SQL) for day-to-day operations to assist with decision-making and planning (Bhadani & Jothimani, 2016). Second, Bhadani & Jothimani (2016) state that because of the increase in data size, specifically unstructured data, it has become almost impossible to process these datasets with older storage techniques and basic queries (Bhadani & Jothimani, 2016). Yafooz *et al.* (2020) reiterate this by stating that because of BD's complexity, overlapping, and being unstructured most times, it cannot be processed with a single tool like a database (Yafooz *et al.*, 2020).

As mentioned previously, an important aspect of the generated data nowadays is its expanding variety in form. The traditional data in database tables are being overtaken by the growing availability of unstructured data sources, like videos, pictures and texts produced by humans (De Mauro *et al.*, 2016). The multiplicity of data types and their co-existence is one of the major challenges associated with BD's handling (Al-Sai *et al.*, 2019).

BD's concepts and technologies allow governments to achieve various objectives that enhance levels of sustainability and governance (Al-Sai & Abualigah, 2017). First, governmental organisations can increase service delivery and offer transparency, effectiveness, efficiency, and certainty (Al-Sai & Abualigah, 2017). Second, BD has the power to transform government practices by generating added value for public services and the ability to motivate and support the digital innovations of governments (Al-Sai & Abualigah, 2017). Third, it improves the time required to process data and deliver high-quality services and products that meet citizens' demands while improving internal business decisions (Al-Sai & Abualigah, 2017).

Software developers developed applications enabling citizens to report and provide information on issues such as potholes, where they see properties being vandalised with graffiti, broken pavements, and streetlights (Morabito, 2015; Desouza & Jacob, 2017). These applications allow citizens to upload information as identified individuals or anonymously and pin the location of the incident on a street map (Morabito, 2015). The data is then sent to local councils to resolve these issues, and it provides tracking of the date of when the issues will be fixed (Morabito, 2015). Thus, as shown in this example, BD has enabled citizens to inspect their neighbourhoods and reduced the cost for local government agencies to send out inspectors.

Perhaps this is nothing new because an issue like this could be reported using other methods, like calling the council or writing a letter. However, what makes this reporting different is immediacy, transparency, and non-evasiveness (Morabito, 2015). The direct interaction between citizens and government agencies incorporates three governmental goals, which are: i) to engage the citizens into civic functions, as citizens actively engage in the process; ii) to decrease the cost of civil service, as citizen engagement is voluntary and free of charge; and iii) to improve the transparency of public services, as the issue handling process is now traceable online (Morabito, 2015).

Moreover, BD is not a single technology but a combination of technologies and processes that support organisational data analytical processes to gain insight while effectively handling data load and storage problems (Halper & Stodder, 2014; Wamba & Mishra, 2017; Yakobi *et al.*, 2020). When government organisations use BD technologies that offer new effective features that provide interactive services, government data will be more than just a big and more than just a data (Al-Sai & Abualigah, 2017).

De Mauro *et al.* (2016) suggest that some of the reasons behind the rapid expansion of BD are: i) the extensive degree at which data is created, shared and utilized currently; and ii) the increasing availability of information due to the proliferation of personal devices connected to the Internet which is equipped with digital sensors such as cameras, audio recorders and GPS locators.

2.3. Characteristics of Big Data

BD can be described as the advanced techniques and technologies that enable the capture, storage, distribution, management, and analysis of large, complex, structured and unstructured datasets to derive facts (Gandomi & Haider, 2015; Dumbill, 2013; Yan, 2015; Eynon, 2013; Agrawal *et al.*, 2011; Sun *et al.*, 2016; Yakobi, 2020). BD involves high volume, high velocity, or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimisation (De Mauro *et al.*, 2016).

The characteristics of BD were previously defined through a 3V model by Laney in 2001 as high volume, high velocity and high variety of information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making (Desai, 2018). This definition has subsequently been extended to a 4V model by adding a new "V", "*Value*", and even further to a 5V model by adding *"Veracity*" (Bello-Orgaz *et al.*, 2016).

Figure 2: The 5 V's of Big Data (Adapted from Kumar, 2015; Younas, 2019)

Younas (2019) states that BD features can be compared to traditional data and characterised by 5V's, huge volume, high velocity, high variety, low veracity, and high value as shown in Figure 2 and Table 2.

Characteristic	Description
Volume	Volume is derived from the amount of data generally from terabytes to petabytes. It is an important, distinctive BD feature that imposes specific requirements for traditional technologies and tools. BD volume includes size, scale, and amount of data collected by either data-rich processes or transactions stored in individual files or databases. All need to be accessible, searchable, processed, and manageable.
Velocity	BD is often generated at high speed and includes data generated by arrays of sensors or multiple events and needs to be processed in real-time, near real-time, in batches, or as streams.
Variety	Variety deals with the complexity of BD and the semantic models behind the data. The data can be collected as structured, unstructured, semi-structured, or mixed data. Data variety imposes new data storage and database design requirements, which should dynamically be adapted to cater to these data formats, particularly scaling up and down.
Veracity	Veracity of BD includes two aspects: data consistency, which their statistical reliability can define; and data trustworthiness, that is defined by several factors including data origin, collection and processing methods, and trusted infrastructure and facilities. BD veracity ensures that the data used is trusted, authentic and protected from unauthorised access and modification. The data must be secured during its lifecycle from the collection from trusted sources to processing in trusted compute facilities and stored in protected, trusted storage facilities.
Value	Value is an important feature of the data which is defined by the added value that the collected data can bring to the intended process, activity or predictive analysis/hypothesis. Data value is dependent on the events or processes they represent which can be stochastic (having a random probability distribution or pattern that may be analysed statistically but may not be predicted precisely), probabilistic (based on or adapted to a theory of probability; subject to or involving chance variation), regular or random. Depending on these requirements, it may be required to collect all data in a dataset and store it for long periods.

Table 2: The 5 V's of Big Data (Adapted from Arputhamary & Arockiam, 2015; Younas, 2019)

Business processes is the foundation for operations of any organisation to achieve their goals (Sakr *et al.,* 2018). It provides operational resources for organisations to fulfil their operational goals (Sakr *et al.,* 2018). Currently, organisations are able to gather massive amounts of event data through business processes that are executed and stored in transaction logs, databases, e-mail correspondences, and social media (Sakr *et al.,* 2018). Harnessing the data allows organisations to apply analytic techniques to it which aid in their decision making process (Sakr *et al.,* 2018).

Presently, the IT industry has witnessed significant advancements in BD analytics (Sakr *et al.,* 2018). Sakr *et al.* (2018) state that BD can play an important role in improving the efficiency and quality of various data-intensive business operations using a wide spectrum of emerging BD systems. By recognising Big Data Analytic (BDA) potential, organisations improve the efficiency and quality of their business processes through effective Business Process Management (BPM) (Wamba & Mishra, 2017). BPM not only improves processes but also monitors the technological advances that can be integrated into the development of efficient processes through Business Process Re-engineering (BPR) and Business Process Innovation (BPI) (Wamba & Mishra, 2017). The successful integration of BDA and business processes create a new class of economic asset that aids top-performing organisations to redefine their business and outperform their competitors (Wamba & Mishra, 2017).

Furthermore, in the public sector the use of BD differs from the use of conventional digital data. Klievink *et al.* (2016) suggest that BD has five distinguishing characteristics that can be described in terms of the 5V's, namely: i) volume, using and combining multiple, large datasets, from various sources, both external and internal to the organization; ii) variety, using and combining structured (traditional) and less structured or unstructured (non-traditional) data in analysis activities; iii) velocity using incoming data streams in real time or near real time; iv) value, development and use of advanced analytics and algorithms, distributed computing and/or advanced technologies to handle very large and complex computing tasks; and v) veracity, through innovative use of existing datasets and/or data sources for new and radically different applications than the data were gathered for.

Morabito (2015) and Younas (2019) also suggest that due to the characteristics of BD, analytics searches to identify opportunities require different application policies and protocols. Morabito (2015) describes the process as: i) analytical goals, in which recognising key priorities is very important to understand the business context and efficiently set up analytics goals, for example, some organisations seek new insight while others are evaluating their businesses models against competitors; ii) selecting data, to ensure the correct data is selected for analysis. Since BD consists of references to a variety of information, selecting the relevant data allow the analytics processes to be focused and efficient; iii) processing data, as intended to enhance data reliability by defining missing data, identifying and removing irrelevant data or outliers, and the complexity of which often depends on the size and the variety of the data and the nature of the research; iv) data mining, in which data is processed to extract meaningful signals based on the goals of the analysis; v) evaluation and pattern interpretation, in which the results of the analysis is evaluated to extract meaningful insights that can enhance the organisational performance, and iv) visualization and feedback, in which innovative visualization techniques such as tables, lists or graphs are used to represent the results in an intuitive way for non-technical staffs or management teams for better understanding of the analysis outcome. Also, to improve the analysis results, a direct feedback process can be set up to refine the prediction process (Morabito, 2015).

De Mauro *et al.* (2016) mention that some scholars adopted an implicit definition for BD through anecdotes, success stories, characteristics, technological features, trends or impact on society, firms and organisational processes. De Mauro *et al.* (2016) and Yakobi *et al.* (2020) also state that BD is a term that describes social phenomenon, information assets, data sets, storage technologies, analytical techniques, processes and infrastructures.

In conclusion, the aforementioned scholars state that the charateristics of BD include the following aspects: i) volume, velocity and variety, which describe the characteristics of information; ii) technology and analytical methods, which describe the requirements needed to make proper use of such information; and iii) value, which describes the transformation of information into insights that may create economic value for organisations and society (De Mauro *et al.*, 2016).

2.4. Big Data ecosystem

A BD ecosystem is a collection of infrastructure, analytics, and applications used to capture and analyse data. Chatfield *et al.* (2015) and Matheus & Janssen (2020) claim that BD ecosystem's components may vary from government to government and that the ecosystem produces innovative and transformative means which can be used for enhancing government transparency and accountability, innovation and productivity in public services, and active citizen engagement.

In government, the BD ecosystem actors represent diverse entities that provide data, consume data, manipulate data to offer paid services, and extend data services like data storage, hosting services to other actors (Iftikhar et al., 2020). The interrelated components that make up a BD ecosystem must work well together to support evolving data models and infrastructure in delivering meaningful facts (Iftikhar et al., 2020). Figure 3 depicts the dimensions of a BD ecosystem. Chatfield *et al.* (2015) observe that leading governments have advanced their vision of a data-driven government by converting their data into digital formats, data-driven policy agendas, and making data-driven decisions (Chatfield *et al.*, 2015).

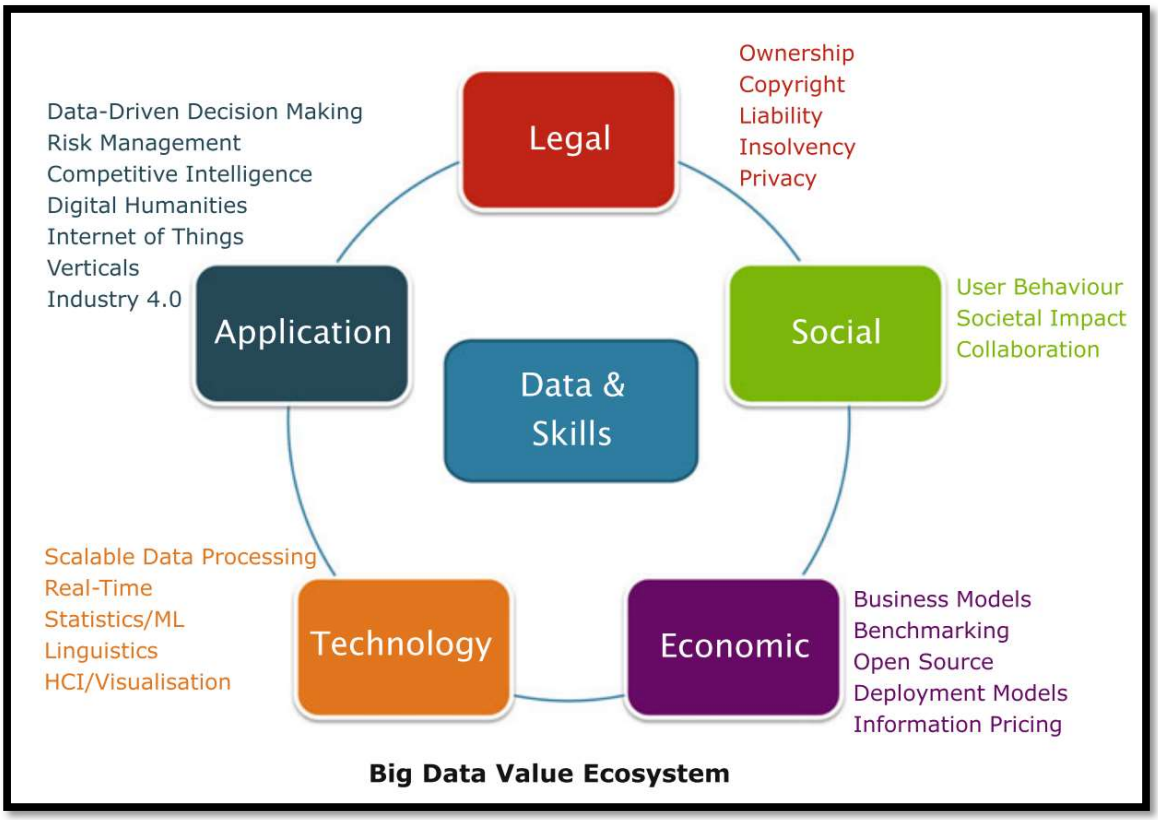


Figure 3: The dimensions of a BD Value Ecosystem (Source: Cavanillas *et al.,* 2016)

Tomar *et al.* (2016) state that two key factors influence the successful adoption of a BD ecosystem: institutional arrangement and technical aspects. Formal institutions deal with policies, governance structures, and legal, regulatory frameworks. Implications relate to external factors such as the behavioural impact of data-driven management structures on civil servants (Tomar *et al.*, 2016). The technical aspects of information systems, including infrastructure and physical capital, human capital and operational procedures, are required to efficiently function as a BD system (Tomar *et al.*, 2016).

Due to innovative devices, sensors, and social networks, BD arose to challenge practitioners to shift the way they operationalize data (Rad & Ataei, 2017). This has changed the context for many industries, like government organisations, and challenged leaders to move to adopt a BD ecosystem (Rad & Ataei, 2017). Organisations are starting to realise that the context and content of BD is well liked, due to the interplay between the user and the provider (Rad & Ataei, 2017).

As BD evolves and analytical practices change, the industry is seeing advancements and complex implementations of BD, like Mechanical Turk's (a crowdsourcing website for businesses to hire remotely located "crowdworkers" to perform discrete on-demand tasks that computers are currently

unable to do) (Rad & Ataei, 2017). This heralds a new BD ecosystem with four main role players in the relationship (Rad & Ataei, 2017). These four main role players are: i) data generators, which are the devices that generate new data about data; ii) data collectors, which collects data about users and devices with attention to their attributes and attitudes; iii) data aggregators, who collect data and draw patterns based on it; and iv) data consumers, who benefit from the data gleaned and crunched by others within the data value chain (Rad & Ataei, 2017).

The relationship amongst these four groups gives life to the BD ecosystem. Data generators using various devices produce a deluge of data that will be gleaned by data collectors (Rad & Ataei, 2017). Data collectors then pass these massive amounts of BD to data aggregators (Rad & Ataei, 2017). This is where the complex analytics are applied, and patterns revealed. Next, insights and patterns will be sold to other organisations for competitive advantage and business intelligence (Rad & Ataei, 2017).

Iftikhar *et al.* (2020) have also identified three main concepts relating to government BD ecosystems. The first concept relates to socio-technical networks, which is about the relationship of socio-technical elements such as people, processes, technology, organizations, data, and infrastructure (Iftikhar *et al*., 2020). The second concept is data functions, which involves the different phases (data collection, data integration, analysis, data storage, sharing, use, data security and protection, and data archive) to convert data/information into knowledge (Iftikhar *et al.*, 2020). The third concept is data value creation, which is about the extraction of value from the BD (Iftikhar *et al.*, 2020). These values (public service delivery, data-driven administration, transparency, data economy, and new businesses) are the product of the data functions performed on BD in a collaborative environment by the stakeholders (Iftikhar *et al.*, 2020).

2.5. Big Data lifecycle

BD lifecycle models present a structure for arranging the tasks and activities related to data management within a project or organisation (Pouchard, 2016). It also presents a way of communicating the tasks to the intended audience, including researchers, data managers, curators, repository specialists, librarians, and project managers in charge of organising data in a project or a lab (Pouchard, 2016).

Data management is becoming more difficult, particularly with the rise of the BD (Arass & Souissi, 2018). Arass and Souissi (2018) state that the best way to manage this data is to organise it into a lifecycle from creation to destruction. The BD lifecycle designs must adapt to the specific domain and the organisation's purpose to enhance value generation (Blazquez & Domenech, 2018). The design should manage the entire BD lifecycle within the organisation with the end goal being to generate added value and discover knowledge (Blazquez and Domenech, 2018; Sinaeepourfard *et al.*, 2016). The benefits of adopting a data lifecycle model includes the following: i) easing for planning and handling complexity of data management throughout all data life stages; ii) preparing data products for end-users

access, achieving the expected constraints and efficiency requirements; iii) providing a high level data quality, removing any waste and noise; iv) identifying the appropriate sequence of essential activities related to data life; and v) helping system designers to create sustainable and efficient software (Sinaeepourfard *et al.*, 2016).

Sivarajah *et al.* (2017) suggest that some challenges relating to the BD lifecycle can be grouped into three main categories, based on the BD lifecycle: data; process; and management challenges. The data challenges relate to the data itself, for example, data volume, variety, velocity, veracity and value (Sivarajah *et al.*, 2017). The process challenges relate to techniques on how to capture data, how to integrate data, how to transform data, how to select the right model for analysis and how to provide the results (Sivarajah *et al.*, 2017). Management challenges incorporate privacy, security, governance and ethical aspects (Sivarajah *et al.*, 2017).

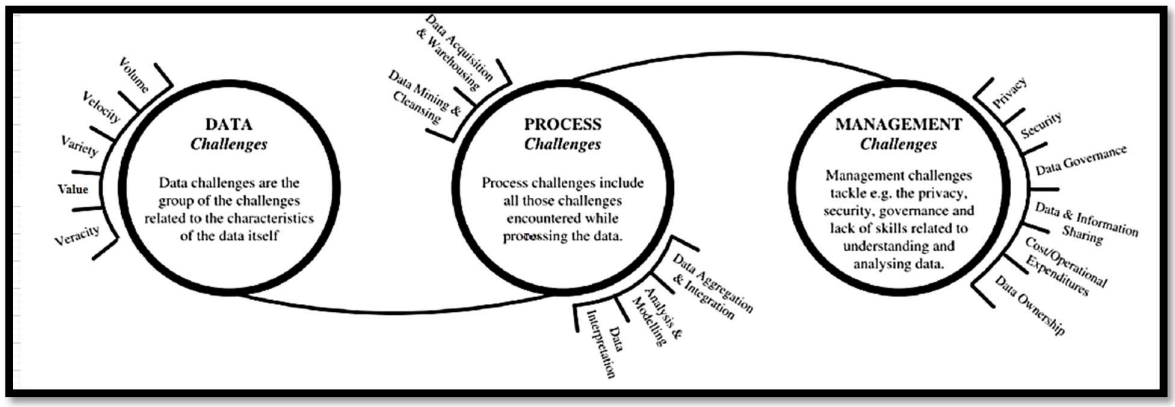


Figure 4: Conceptual BD life cycle challenges (Adapted from Sivarajah *et al.,* 2017)

BD has become a key enabler for discovering business insight, business intelligence, and service (Sun *et al.*, 2018). It has drawn an extraordinary interest from businesses, universities and governments by changing how they operate and use data analytics to provide new opportunities (Kleiner *et al.*, 2012; Sun *et al.*, 2018). Finding value through BD requires careful planning and organisation of all the processes in the data lifecycle, considering the particularities of the social and economic analyses, including the wide variety of heterogeneous sources of information and strict governance policies (Blazquez & Domenech, 2018).

The study explored natural disasters in Malaysia to illustrate the BD lifecycle. Abdullah *et al.* (2017) explain that decision making in natural disaster management has its own challenge that needs to be addressed. When a disaster occurs, government, as a response organisation, must come to timely, accurate decisions to ensure that rapid assistance and effective recovery for the victim involved can be conducted (Abdullah *et al.*, 2017). Futhermore, the scholars above suggest that BD can be used as a

tool for strategic decision making in government relating to disaster management through Big Data Analytics.

In Malaysia, seventy-six disasters were recorded between 1965 and 2016 (Abdullah *et al.*, 2017). These disasters included wildfires, storms, landslides, mudflows, epidemics, tsunamis and droughts (Abdullah *et al.*, 2017). By law, Malaysian governmental agencies must manage such events and implement measures to aid society (Abdullah *et al.*, 2017). The management of such an event involves many activities that require immediate decision making based on real-time data (Abdullah *et al.*, 2017).

The Malaysian government acknowledged BD's potential in these situations and put procedures that encourage collaboration between government agencies in place (Abdullah *et al.*, 2017). They selected four public agencies and identified five projects as Proof Of Concept (POC) (Abdullah *et al.*, 2017). Specifically focusing on disaster management, the Malaysian National Hydraulic Research Institute (NAHRIM) BD Framework for Disaster Management suggested three stages: data acquisition; data computation; and data interpretation (Abdullah *et al.*, 2017). The data acquisition phase included "Data Source", to aggregate information digitally from diverse sources for further storage and analysis (Abdullah *et al.*, 2017). This data was collected from a historical-based modelling process using high-performance computing environments and large volumes of historical and projected hydroclimate data (Abdullah *et al.*, 2017).

The data computation stage included the following three phases: Data Management, Analysis, and Data Visualisation (Abdullah *et al.*, 2017). Data cleansing was performed during the data management phase, removing incomplete data (Abdullah *et al.*, 2017). These datasets were then assigned to pre-defined categories through a specific extraction method to get the semantics and correlations of the data (Abdullah *et al.*, 2017). The collected data was then modelled and analysed. Descriptive, diagnostic, predictive, and prescriptive analysis methods were then used (Abdullah *et al.*, 2017). During descriptive analysis, hydroclimate historical data were fully utilised to quantify, track and report what might have previously occurred and how things are going in the disaster management domain (Abdullah *et al.*, 2017). Then, diagnostic analytics were used to obtain deeper insight into what transpired and to understand why these events occurred. Based on the previous analysis, the hydroclimate data presents the root cause of the disaster, allowing the subject matter expert(s) to focus on *"what happens next?"* (Abdullah *et al.*, 2017). The subject matter experts seek to answer this question using predictive analysis and statistical techniques (Abdullah *et al.*, 2017). The final analytic approach, prescriptive analytics, was used to address decision making and efficiency as soon as a good measure of accuracy on the predictive algorithm was achieved, thus justifying the prescriptive interventions (Abdullah *et al.*, 2017). They could compile a credible explanation of how disasters are likely to re-occur and understand how predictable the disaster occurrence is (Abdullah *et al.*, 2017).

The final stage in data computation included the visualisation phase. In BD frameworks, the visualisation phase is vital, as it allows organisational users to mash-up disparate data sources to create custom analytical views (Abdullah *et al.*, 2017).

The final stage is the data interpretation stage, which comprises the disaster management and the decision phase (Abdullah *et al.*, 2017). During these stages, the Subject Matter Expert (SME) and decision-makers play a vital role in understanding the information to make strategic, rational and relevant decisions based on the analysis's insights (Abdullah *et al.*, 2017). The data interpretation requires knowledge and experience by domain experts, such as hydrologists and climatologists, to help elucidate the analysis before making any decisions (Abdullah *et al.*, 2017). The stages and elements for this data lifecycle is presented in Figure 5 below.

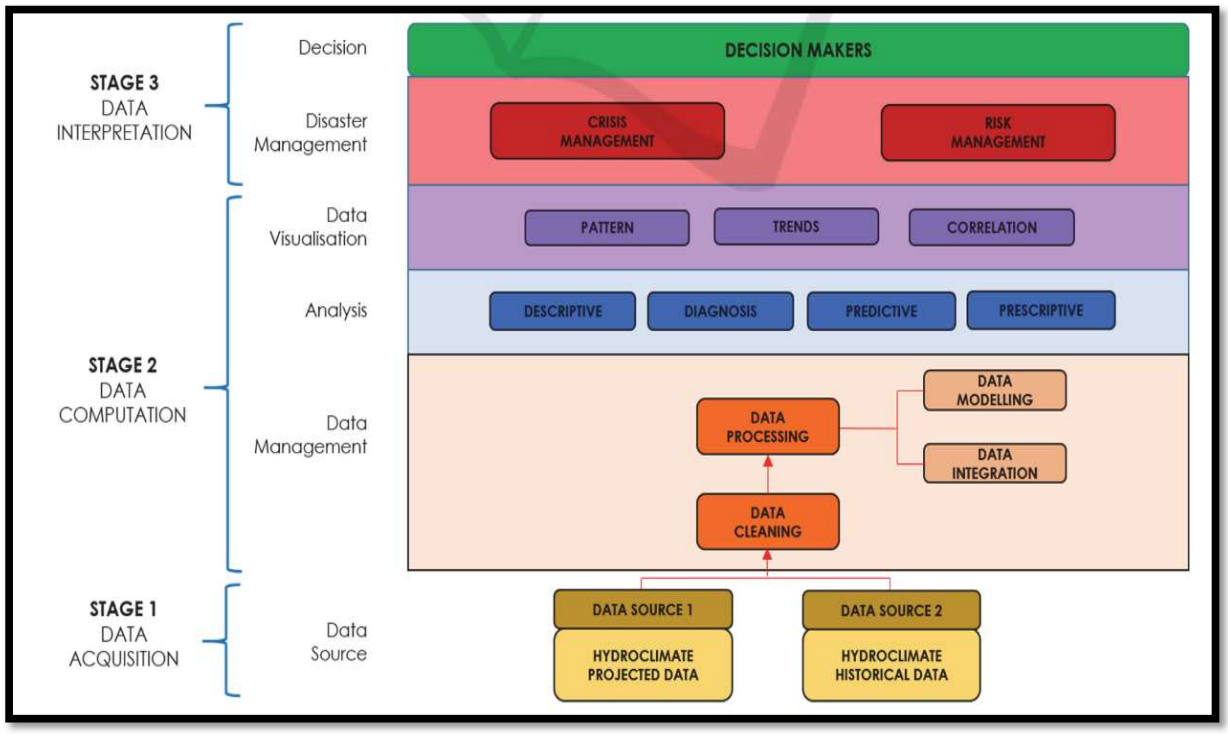


Figure 5: NAHRIM Big Data Framework for Natural Disaster Management (Abdullah *et al.*, 2017)

In conclusion, SME's and decision-makers discovered disasters could be predicted or avoided through mitigation or adaptation if decisions are made based on quality and accurate data from BDA. They could also put assessment and risk management controls to reduce the catastrophic impact of disasters based on the results. Because complexities, urgency, and uncertainty characterise disaster management, organisations need to have a fast, effective decision-making process in place.

2.6. Big Data use in world governmental organisations

Information systems have evolved from transaction recording systems to supporting business decisions at various levels (Jeble *et al.*, 2018). Business executives are challenged by customers' high expectations, high competition, rising labour costs, and materials with shorter product lifecycles (Pugna *et al.*, 2019). Aiming to solve these challenges, scholars, analysts, and entrepreneurs from a wide range of fields actively pursue approaches to mining digital traces and stashes of data that comprise BD (Desouza & Jacob, 2017). BD's use enables executives to make better decisions based on evidence rather than intuition (Jeble *et al.*, 2018).

BD promises to transform public decision making for the better by making it more responsive to actual needs (van der Voort *et al.*, 2019). In recent years, social media's acceptance has led to information being generated and continuously collected (Jeble *et al.*, 2018). This allows the public and officials to report important events (Jeble *et al.*, 2018). The data needs to be scrutinised through Big Data Analytics (BDA) which enables e-commerce, government, politics, science, technology, health, security and public safety to make better decisions (Jeble *et al.*, 2018).

BDA refers to examining massive amounts of data (BD) to uncover hidden patterns, market trends, customer preferences, and other useful information that aids in making informed decisions (Pugna *et al.*, 2019). Abdullah *et al.* (2017) describe BDA as executing algorithms on powerful computer platforms to discover hidden facts in BD. Pappas *et al.* (2018), define BDA as the capability of a data user to effectively deploy technology and talent to capture, store and analyse data, create value, change business, and bring about societal change (Pappas *et al.*, 2018). The main objective of BDA is to discover new insights and relationships which otherwise might have been unseen and to provide insights about the users who created the data (Chaha & Gulia, 2016). BDA must efficiently mine huge datasets at different levels, in real-time or near real-time (Wiseman, 2018).

Using BD can take shape in various forms relating to the government (Giest, 2017). BD can be used to design more effective or efficient policies because what the information decision-makers receive is more precise, vast, or even predictive for the issue they are tackling (Giest, 2017). However, a country's regulatory framework determines how public and private entities can use BD because privacy laws restrict collecting, sharing, or utilising personal information (Giest, 2017). Giest (2017) also claims that an alternative way of looking at BD relating to government is changing how public services are provided. This relates to citizens providing information to the government, while governments offer personalised services based on the additional data from citizens, neighbourhoods or communities (Giest, 2017).

Organisations face challenges relating to BD because they have a wealth of information. Still, they do not know how to get value out of it because it is sitting in mostly raw form or a semi-structured or unstructured format, and as a result, they do not even know whether it is worth keeping (Zikopoulos *et al.,* 2016). Similarly, public sector organisations accumulate large amounts of data of which a significant portion has never been analysed, meaning a significant amount of valuable facts remain undiscovered (Vassakis *et al.*, 2018).

Citizens worldwide are appealing for openness in government and greater civic participation in public affairs and seeking ways to make their governments more transparent, responsive, accountable, and effective (Sarker *et al.*, 2018). Governmental organisations can achieve this by analysing BD to reduce corruption, ensure swifter public service delivery, reduce public hassle, provide easy access to public services, and reduce both errors and poverty through e- services, e-management, e-democracy, and e-commerce (Sarker *et al.*, 2018).

The promise of Big Data analytics (BDA) to predict the present and the future makes BD potentially invaluable for decision-making. The literature on BD shows an abundance of these potentials. For example, BD analytics can produce valuable facts regarding customer preferences, behaviours and market trends, thus improving business intelligence and business decisions (van der Voort *et al.*, 2019). In a well-functioning society, citizens need to know what their government is doing, and, in order to do that, they must have free access to government data and information (Koznov et al., 2016). By opening up data, citizens are made much more informed and are directly involved in decision making (Koznov et al., 2016).

2.6.1. The Russian Federation

The Russian Federation created its "Open Government Strategy" to promote BD by creating many public portals and internet resources that focus on interactions (machining contact) with citizens (Anna & Nikolay, 2015). For example, the Central Bank of the Russian Federation obliged other banks to report to The Federal Tax Service of the Russian Federation about all citizens' transactions, including the opening of accounts, deposits, the flow of funds and account details changing (Anna & Nikolay, 2015). This enabled the Federal Tax Service of the Russian Federation to react immediately if suspicious activities were identified and allowed them to block these accounts at all banks immediately by using BD processing instead of the step-by-step blocking they used to do earlier (Anna & Nikolay, 2015).

The advances in the information and communications technology sector unswervingly affect the nation's business climate and, the dynamics of foreign investments (Ekimova *et al.*, 2018). In Russia, the Pension Fund of the Russian Federation implemented Systems, Applications, and Products (SAP)

High-performance ANalytic Appliance (HANA) project as a single multifunction automated accounting and cost system that controls spending and fund allocation. (Anna & Nikolay, 2015). The Russian Pension Fund is an important social institution that provides state services in pension administration and citizens' social security (Ekimova *et al.*, 2018). The measurement of information processed by the institution shows that it is rated among the top-five information-intensive and service-intensive state agencies (Ekimova *et al.*, 2018). All this data enabled the revenue service to efficiently and effectively manage compliance, risks, and to meet rising service expectations (Ekimova *et al.*, 2018).

It is noteworthy to mention that IT's proficient use has enabled the future success of revenue bodies in effectively and efficiently managing compliance risks and meeting rising service expectations (Ekimova *et al.*, 2018). Furthermore, traffic control is a major issue in the Russian Federation because of the rapid growth of vehicles on roads (Buslaev *et al.*, 2015). This is in addition to the fact that the speed at which road networks are constructed in Russia significantly lags (Buslaev *et al.*, 2015). The Federal Road Agency launched a system that forecasts traffic jams and accidents developed by Yandex Data Factory (Anna & Nikolay, 2015). This system uses data about traffic congestion, weather reports, metrics based on "Yandex Maps" application's data, and information about pavement quality and road line numbers (Anna & Nikolay, 2015).

2.6.2. The United States of America (USA)

The US government spends a large amount of money on BD research (Anna & Nikolay, 2015). There are many reports about different aspects of BD on the White House official site, which include:

- Big Data and Differential Pricing (February 2015)
- Big Data: Seizing Opportunities, Preserving Values (February 2015)
- Big Data and Privacy: A Technological Perspective (Report to the President, May 2014)

The United States of America's (USA) government (White House) also announced the BD Research and Development Initiative on March 29, 2012 (Chang & Grady, 2019). The aim of this initiative was to aid in accelerating the speed at which discovery is made in science and engineering, strengthening national security, and transforming teaching and learning by improving analysts' ability to extract knowledge and insights from large and complex collections of digital data (Chang & Grady, 2019).

The initiative led to six federal departments and their agencies investing more than $200 million in commitments which spread across more than eighty projects, with the goal of significantly improving the tools and techniques required to access, organize, and draw conclusions from huge volumes of digital data (Chang & Grady, 2019). The initiative also challenged industry, research universities, and

non-profit organisations to join with the federal government to make the most of the prospects created by BD (Chang & Grady, 2019).

Driven by the White House initiative and public suggestions, the National Institute of Standards and Technology (NIST) has accepted the challenge to encourage the collaboration among industry professionals to enhance the secure and effective adoption of BD (Chang & Grady, 2019). Subsequently, the NIST's Cloud and Big Data Forum was held on January 15–17, 2013, at which there was strong encouragement for NIST to create a public working group for the development of a BD Standards Roadmap (Chang & Grady, 2019). In the forum, participants requested that the roadmap define and prioritize BD requirements, including interoperability, portability, reusability, extensibility, data usage, analytics, and technology infrastructure because in doing so, the roadmap would accelerate the adoption of the most secure and effective BD techniques and technology (Chang & Grady, 2019).

Furthermore, on June 19, 2013, the NIST Big Data Public Working Group (NBD-PWG) was launched with widespread involvement from industry, academia, and government from across the nation (Chang & Grady, 2019). The scope of the NBD-PWG involved creating a community of interests from all sectors, including industry, academia, and government in an endeavour to develop consensus on definitions, taxonomies, secure reference architectures, security and privacy, and, from these, a standards roadmap (Chang & Grady, 2019). The consensus created a vendor-neutral, technology and infrastructure independent framework that would enable BD stakeholders to identify and use the best analytics tools for their processing and visualization requirements on the most suitable computing platform and cluster, while also allowing added value from BD service providers (Chang & Grady, 2019).

The USA is also leader in a number of BD project deployments in the public sector, from initial data gathering and consolidation projects to advance predictable analytics (Anna & Nikolay, 2015). Examples of these include collaborating with Rutgers University's School of Criminal Justice. The National Institute of Justice (NJI) awarded them with a two-year $500,000 grant to conduct risk terrain modelling research. The university's Public Securities Department worked with the police force of Arlington (Texas), Chicago (Colorado Springs), Glendale (Arizona), Kansas City (Missouri) and Newark (New Jersey) to map and analyse local crime as a part of a pilot project. This project's main focus was to help the police force to suppress crime efficiently and demonstrate how terrain risk modelling can be adopted in other cities to do the same (Anna & Nikolay, 2015).

The system integrates components that provide a platform to build safer cities (Ghosh *et al.*, 2016). It allows authorities to swiftly analyse many data sources to determine if the offenders or victims involved in an offence are related, and it also identifies prior records (Ghosh *et al.*, 2016). This enables analysts

in the command centres to promptly produce valuable tactical information of the suspects, which can be provided to law enforcement in the field (Ghosh *et al.*, 2016). The coordination of back-office operations and fieldwork yields better results (Ghosh *et al.*, 2016). It uses the connections between people, places, events, and vehicles, based on past crimes, to create new information as links between people, places and objects, which results in possible leads for investigation (Ghosh *et al.*, 2016).

The Internal Revenue Service (IRS), a bureau of the Department of the Treasury of the USA, explored ways in which BD helps streamline processes and increase revenue (Anna & Nikolay, 2015). The department collects taxes and enforces the Internal Revenue Code. Their BD project aimed to analyse financial and social information to address taxpayer error, evasion, and other lost revenue sources. Robo-audits could track citizens' online activity and flag patterns of concern (Anna & Nikolay, 2015). The data collection and analysis enable the IRS to generate and track unique attributes regarding financial transactions. Most of the data is used for research and tax enforcement to combat non-compliance. Traditionally, social data was only used when irregular tax returns required further attention (Anna & Nikolay, 2015).

2.6.3. Singapore

The Singaporean government launched a portal http://data.gov.sg/ (Kim *et al.*, 2014, p. 84) to provide access to publicly available government data gathered from over 5,000 datasets and 50 ministries and agencies (Foulonneau *et al.*, 2014; Kim *et al.*, 2014). This aimed to create value while prompting BD research, analysis and applications (Anna & Nikolay, 2015). Various agencies in Singapore collaborated to help the public and private sectors leverage innovative analytics tools. The tools developed focused on navigation and mobility (traffic, maps), Tourism, Environment, Security (Emergency contacts points), Religion (places of worship), Business (Multiple listing of companies registered in Singapore) and open jobs in public administrations (Foulonneau *et al.*, 2014).

The Singaporean government used its websites to recognise citizens' needs through a new BD analytics tool. This cloud-based tool can accurately process and understand a citizen's question and provide an answer within seconds. These applications are available on multiple mobile phone operating systems, including iOS and Android (Foulonneau *et al.*, 2014). This capability enables citizens to navigate government services better and get personalised advice when using online services. The tool also provides government agencies with insights into citizens' needs and priorities (Anna & Nikolay, 2015).

2.6.4. Africa

Healthcare systems in the developed countries are recording breakthroughs due to the application of BD, while additional research is required to understand the impact of BD in health systems in Africa

(Akinnagbe, Peiris & Akinloye, 2018). Akinnagbe et al. (2018) claims that well-organized healthcare delivery is information-driven, and that doctors and other healthcare workers need adequate and the right kind of data in order to diagnose diseases, prescribe drugs and provide appropriate counsel to patients. Similarly, healthcare managers rely on quality information in order to make right decisions pertaining to provisioning of appropriate personnel in hospitals, management of facilities and drugs inventories, and efficient management of disease outbreaks (Akinnagbe, Peiris & Akinloye, 2018).

Iyamu & Mgudlwa (2021) claims that there are gaps in patient healthcare data which hinders the progress in South African healthcare environments. They suggest that the integration of various datasets is fundamentally important and useful to healthcare service providers. because good-quality data in a timely manner is always needed due to the rapid changes that happen in the healthcare environment (Iyamu & Mgudlwa, 2021).

2.7. Delivering value to government and citizens through Big Data

The concept of BD has already been introduced earlier in this chapter. Anshari *et al.* (2018) describe it as the management and administration of large amounts of data to reveal concealed value (Anshari *et al.*, 2018). In a government setting, the data can be used at an individual level to create citizen profiles to detect fraud, fight crime, terrorism, or improve collective public endeavours and affect privacy (Anshari *et al.*, 2018).

To illustrate BD use opportunities, examples from various domains can be reviewed (Giest, 2017). Jeble *et al.* (2018) mention that BD is a great tool to use when making decisions in the medical field, specifically when used to predict the risk associated with diseases and infections. There are huge amounts of data in the medical field that human beings cannot analyse. BD analytics can be used to look for patterns that predict the onset of diseases, thus reducing patients' stay in a hospital (Jeble *et al.*, 2018).

2.7.1. Healthcare

Many developing countries' health systems cannot provide critical services to their citizens because of insufficient resourcing (Amankwah-Amoah, 2016). However, technological innovations and the development of powerful software for data analytics have created conditions and methods to analyse large datasets which can be used for the social good (Amankwah-Amoah, 2016). These innovations have helped improve the quality of decision making towards undertaking global health issues (Amankwah-Amoah, 2016). For example, we explore how these types of innovations were used to aid in bringing the Ebola outbreak in developing African countries under control (Amankwah-Amoah, 2016).

The Ebola outbreak occurred in three countries, specifically Guinea, Liberia and Sierra Leone (Amankwah-Amoah, 2016; Taylor, 2017). When the outbreak occurred, the health systems in these three countries were under-resourced, understaffed and disjointed (Amankwah-Amoah, 2016). Gathering and distributing information about the disease was beyond any of the affected countries (Amankwah-Amoah, 2016). Therefore, collective effort across the three countries was required to help map resources and expertise to potentially risky areas (Amankwah-Amoah, 2016). This approach provided a foundation for organised resources to help the wider international community be better prepared to manage the disease (Amankwah-Amoah, 2016).

Adopting an archival approach that involves examining publicly available governments and international agencies' publicly available records, which were critical and underutilised resources (Amankwah-Amoah, 2016). These were used to enhance their understanding and discover advancements in related past events (Amankwah-Amoah, 2016). These archives included information on strategies, government policies, surveillance systems and monitoring techniques from the World Health Organisation (Amankwah-Amoah, 2016). Besides, such technologies as social media, blogs, Twitter, Facebook and online forums presented huge amounts of data relating to perceptions and trends (Amankwah-Amoah, 2016). The data was used in a combination with fieldwork to monitor the outbreak and formulate public policy (Amankwah-Amoah, 2016).

Other technologies were used, including mobile mapping, the Ebola-tracking application and Trend mapping by search engine firms (Amankwah-Amoah, 2016; Taylor, 2017). First, mobile mapping data was gathered through mobile phones that are widely used in developing countries (Amankwah-Amoah, 2016; Taylor, 2017). When calls are made, calls create call data records (CDRs) that can track citizens' movement. CDRs include the caller and receiver's phone numbers, the time of the call and the tower that handled the call, which gives a rough indication of the device's location (Amankwah-Amoah, 2016; Taylor, 2017). When there is a sudden uptake in the number of calls made to the helpline in a specific area, authorities allocate more resources to that area (Amankwah-Amoah, 2016; Taylor, 2017). Also, through analysing, the CDR's authorities could determine areas where the epidemic is widespread and set up treatment centres in those areas (Amankwah-Amoah, 2016; Taylor, 2017).

Second, the Ebola-tracking app was developed through partnerships with the private sector (Amankwah-Amoah, 2016). This application was particularly helpful and aided governments in harnessing BD to assist with tracking the Ebola outbreak (Amankwah-Amoah, 2016). This application helped the affected governments to determine where to allocate resources. It also allowed governments to urge citizens to volunteer information, and an official could use the application to advise people to

remain in a certain area to limit the spread of the disease, and information could be sent to health workers in remote sites (Amankwah-Amoah, 2016).

Last, search engine firms have progressively been using technology to forecast disease outbreaks (Amankwah-Amoah, 2016). Through analysing search patterns in a particular area, they can determine if there is a possible disease outbreak (Amankwah-Amoah, 2016). For instance, Google trends that allow individuals and scientists to track the popularity of search patterns in combination with other health records for a specific area can help governments make informed decisions (Amankwah-Amoah, 2016).

2.7.2. E-government

Governments increasingly rely on computerised information systems for decision-making (Elezaj *et al.*, 2018). ICT performs an important role in achieving efficient governments' strategic goals (Elezaj *et al.*, 2018). E-government refers to the use of modern technology resources such as the internet and mobile technologies to improve the e-services provided to citizens and businesses (Elezaj *et al.*, 2018). The core benefit of using technological systems for e-government is building an open information society by providing public services on the Internet (Elezaj *et al.*, 2018).

Modern technologies promote e-government through ICT, allowing governments to connect electronically with citizens and private groups (Woodside, Amiri, 2015). Governmental organisations worldwide have started to adopt e-government strategies to achieve many benefits for themselves as providers of public services and for citizens who are the users of these services (Weerakkody *et al.*, 2016). Sarker *et al.* (2018) describe e-government as an approach in which public administration adopts an electronic process in its political and administrative process (Sarker *et al.*, 2018). It uses Information Resource Management (IRM) to identify and promote information creation, selection, development, and management to enable policy and effective decision-making (Sarker *et al.*, 2018). The main purpose of e-government is to ensure that citizens have easy access to public services (Sarker *et al.*, 2018).

Public administration, academics and research scholars highlight the importance of adopting modern technologies to properly implement public policies and deliver public services (Sarker *et al.*, 2018). Effective public administration requires facilitating customer-centred, cost-effective, and citizen-friendly public services to improve governmental functions (Sarker *et al.*, 2018). BD's potential in transforming digital government services, governments, and the interaction between governments, citizens, and the business sector are substantial (Mergel *et al.*, 2019). It can encourage collaboration, create real-time solutions to agriculture, healthcare, and transportation issues, and shepherd in a new era of policy and decision-making (Twizeyimana & Andersson, 2019).

Government organisations can use BD to increase openness, efficiency, and citizen satisfaction. Elezaj *et al.* (2018) state that government organisations aim to become service orientated. First, they can promote open government data sharing by providing accurate, fast information from public organisations to citizens for greater government transparency and promoting greater trust between citizens and government, which is in line with open data initiatives (Elezaj *et al.*, 2018). Second, they can encourage citizen sentiment analysis through social media analysis (Elezaj *et al.*, 2018). Third, BD can be used for economic analysis through correlation, pattern recognition, and rich data analytics from multiple sources to aid policymakers in making decisions based on knowledge (Elezaj *et al.*, 2018). Fourth, tax agencies can use structured and unstructured data from different sources for fraud detection (Elezaj *et al.*, 2018). Fifth, public institutions worldwide are using sensors to measure phenomena such as traffic volume and the location of vehicles, which improves urban management to improve the quality of life for their citizens (Elezaj *et al.*, 2018). Last, collecting and analysing data coming from government network computers and their logs increase institutions' ability to detect and prevent malicious cyber-attacks (Elezaj *et al.*, 2018).

Albania's government created a similar portal, administered by the National Agency for Information Society (NAIS) as a multifunctional portal where citizens and businesses access public electronic services (Elezaj *et al.*, 2018). The portal began as a European Union investment in 2009 and provided service to citizens twenty-four hours a day, seven days a week. Initially, it provided six electronic services and had four systems linked to the government interoperability platform (Elezaj *et al.*, 2018). The governmental interoperability platform connects hardware, software, and services permitting collaboration between all the government institutions' connected systems (Elezaj *et al.*, 2018).

At the beginning of 2018, the system had around five hundred users registered, which increased by one-hundred and fourteen thousand in the first quarter (Elezaj e*t al.*, 2018). This is a clear indication that Albania's citizens perceived that portal as useful (Elezaj *et al.*, 2018). Currently, the portal provides more than five hundred and seventy-one services to citizens and businesses. These services include generating certificates and health cards, saving citizens time and cutting costs (Elezaj *et al.*, 2018).
The various public institutions responsible for offering online services interact and exchange real-time data by using the interoperability framework (Elezaj *et al.,* 2018). During 2016-2017, over 50 million transactions were recorded amongst the interconnected systems (Elezaj *et al.*, 2018). Today, the platform allows Albanians to download 30 different documents directly from the portal, and the documents are certified with a digital seal. This advancement has been tangible and significant in reducing paper documents (Elezaj *et al.*, 2018).

2.8. Benefits of Big Data in government

BD's efficient use and understanding as an economic asset carry great potential for economies and

society (Cavanillas *et al.*, 2015; Wamba & Mishra, 2017). It is expected that BD can impact various sectors, including healthcare, media, energy and retail (Cavanillas *et al.*, 2015). BD's constructive transformational potential has already been identified in many key sectors in Europe (Cavanillas *et al.*, 2015).

2.8.1. Healthcare

The efficiency of healthcare systems around the world needs to be improved especially because (among others) the ageing citizens are placing a significant demand on the system (Pastorino *et al.*, 2019). BD's application has the estimated potential to save this sector around ninety billion Euros from the national healthcare budget (Cavanillas *et al.*, 2015). This can be done through medical applications of BD, including comparative effectiveness research where the clinical and financial effectiveness of interventions is compared to the next generation of clinical decision support systems that make use of comprehensive heterogeneous health datasets and advanced analytics of clinical operations (Cavanillas *et al.*, 2015; Pastorino *et al.*, 2019). Research and development relating to the healthcare sector include predictive modelling, statistical tools and algorithms to improve clinical trial design, personalised medicine, and analysis of disease patterns (Cavanillas *et al.*, 2015; Pastorino *et al.*, 2019).

Retrieving meaningful BD patterns, such as patient diagnostic information, is a vital problem in the health sector, and critical applications have been developed that relay sensor data in wearable devices (Manogaran *et al.*, 2018). Wearable sensor devices are used to monitor activities and dietary habits (Manogaran *et al.*, 2018). When these devices are worn, sensors would continuously observe and store the patient's health, download and analyse it (Manogaran *et al.*, 2018). This aids doctors in diagnosing patients' health conditions and provides better results by relying on laboratory tests and patient health data collected from wearable body sensors (Manogaran *et al.*, 2018). Traditional data storage techniques and platforms are insufficient when dealing with sensor application domains where the volume, velocity, and variety of data are growing. This problem requires developing an efficient system for storing and processing voluminous data such as BD (Manogaran *et al.*, 2018).

2.8.2. European public sector

The public sector in Europe accounts for almost half of its Gross Domestic Product (GDP). This sector can save if they apply BD to their administrative processes (Cavanillas *et al.*, 2015). It is estimated that a saving of roughly 15% to 20% can be achieved, equivalent to over 150 billion Euros (Cavanillas *et al.*, 2015). Additional benefits in the public sector include improved transparency via open government and open data, improved public procurement, enhanced allocation of funding into programmes, higher quality services, increased public sector accountability and a better-informed citizenry (Cavanillas *et al.*, 2015; van der Voort *et al.*, 2019). To achieve this, policies must be put in place to allow data sharing

across government agencies and inform citizens about the trade-offs between privacy and security and the risks of sharing data and the benefits they can gain (Cavanillas *et al.*, 2015; van der Voort *et al.*, 2019). It is further envisioned that BD is able to change the relationship between citizens and government by empowering citizens to understand political and social issues in new transparent ways, enabling them to engage in local, regional, national, and global issues through participation (Cavanillas *et al.*, 2015; van der Voort *et al.*, 2019).

2.8.3. Energy and transportation

Cavanillas *et al.* (2015) further state that BD provides the public sector with innovative ways to monitor transportation and logistics by using various new data sources. They estimate that BD's potential in the transport sector is estimated at USD 500 billion worldwide in terms of time and fuel savings and the avoidance of 380 megatons of CO_2 emissions (Cavanillas *et al.*, 2015). Advancements can also be seen when energy systems are digitised to acquire real-time, high-resolution data via smart metres that can be leveraged within advanced analytics to improve the levels of efficiency within both the demand and supply sides of energy networks (Cavanillas *et al.*, 2015). Major benefits can be realised by creating smart buildings and smart cities in the energy sectors, which can reduce CO_2 emissions by more than 2 gigatons, an equivalent to 79 billion euros in utilities (Cavanillas *et al.*, 2015).

Other benefits of BD use in government have been realised in the United States of America (USA) (Gutierrez, 2017). First, the Central Intelligence Agency (CIA) helped fund Palantir Technologies, which developed analytical software to stop terrorism and counteract cyber fraud by identifying transactions that follow common patterns during fraudulent activities (Gutierrez, 2017). Second, law enforcement agencies at federal, state and county levels have access to Automated License Plate Recognition (ALPR) software that notifies them if any motorist in the area has outstanding arrest warrants (Gutierrez, 2017). They also have access to predictive technologies that forecast crime hotspots and correlate offences with repeat offenders (Gutierrez, 2017).

Third, the Department of Transportation uses license plate recognition software and cameras to monitor people's travel patterns by plane, train, and car, enabling them to understand where infrastructure investment is required (Gutierrez, 2017). Fourth, educational activities require online resources, or are completed online, as part of distance learning courses (Gutierrez, 2017). These activities generate massive amounts of data about the way learning can be analysed to discover patterns and verify if study content is sufficient (Gutierrez, 2017). Fifth, the Department of Agriculture supports the farming economy by researching and developing new BD technologies (Gutierrez, 2017). As a case in point, analysing genetic records increased dairy herds' yield by identifying which bulls were likely to breed high-yielding cows (Gutierrez, 2017). Last, the Centre for Disease Control (CDC) is tracking the spread of illnesses using social media, and the National Institutes of Health (NIH) launched an initiative called

Big Data to Knowledge (BD2K) to promote innovation based on data-derived insights (Gutierrez, 2017). Another project funded by the government aims to identify the early signs of suicidal thinking amongst war veterans based on their social media behaviour (Gutierrez, 2017).

Tomar *et al.* (2016) also state that BD analytics reduces the time to identify bottlenecks and inefficiencies, allowing public sector agencies to address immediate issues in a streamlined, rapid, and targeted manner. Additionally, Gutierrez (2017) mentions that the analysis of BD trends strengthens and tailors' specific policy interventions and public services by limiting manual processing time and providing more precise decision-making evidence.

2.9. Challenges of Big Data in government

Despite the benefits, there are barriers to the use of Big Data analytics for making decisions. These barriers include inadequacy of staff handling the advanced analytics for decision making, lack of business support and the problems that frequently arise with the database software (Jeble *et al.*, 2018). Cavanillas *et al.* (2015) state that there are critical challenges in establishing a BD ecosystem. The following section describes these challenges.

2.9.1. Data management challenges

Besides the technological challenges BD users face in governmental organisations, the data itself can also be a challenge, especially its management (Desouza and Smith, 2014). Significant challenge in BD management stems from collecting and managing huge volumes of data in real-time that is frequently incomplete or collected from silos or in a form that is not compatible with automated processing (Al-Sai *et al*., 2019; Desouza & Smith, 2014).

Government organisations need to evaluate the types and sources of the data being collected, stored, analysed, and consumed, including structured versus unstructured data (Gutierrez, 2017). BD data acquisition scenarios assume data is high-volume, high-velocity, high-variety (Gutierrez, 2017). However, sometimes low-value data is also gathered, making it important to have an adaptable and time-efficient gathering, filtering, and cleaning algorithms that ensure that only the data's high-value fragments are processed (Gutierrez, 2017). Over and above that, policy and regulatory problems need to be overcome, including data-sharing agreements, privacy and confidentiality of data, and forming alliance protocols amongst stakeholders investigating similar problems (Desouza and Smith, 2014).

Thabet and Soomro (2015) and Oussous *et al.* (2018) offer data management challenges related to BD's characteristics. The first challenge they mention relates to the volume of data and the increased availability. The second challenge, variety, deals with the various formats in which data is produced

nowadays. The third is velocity, a challenge which is linked to BD tools being unable to handle data stream generated continuously and constantly. The fourth challenge, veracity, refers to the biases, uncertainties, impressions, untruths and missing values in the data. Volatility, the fifth challenge, questions the lifespan of data. The sixth, and last challenge, quality, measures how reliable the data is. Quality concerns are further sub-dived into the following four parameters: completeness, which measures if all required data are available; accuracy, which determines if data is free of spelling mistakes; availability, which measures whether data is available when requested and easy to find; and, lastly, timeliness, which measures whether data is up to date and ready to support decision.

2.9.2. BD skills challenges

It is important to retain capable employees in the BD ecosystems. The BD ecosystem requires employees like data scientists and engineers who have expertise in analytics, statistics, machine learning, data mining, and data management. Technical experts need to be combined with data-savvy business experts with strong domain knowledge and apply their data knowledge within organisations for value creation.

A skilled workforce is required to implement BD's technical aspect (Rajaraman, 2016; Tomar *et al.*, 2016). This skilled workforce should be well informed about the strategic vision of the data usage, have a background in data science, and be familiar with the tools employed (Tomar *et al.*, 2016). Once the strategic and technical foundation is laid, establishing common procedures and training civil servants to interact efficiently with the data are fundamental steps to ensure that the data is integrated into everyday tasks (Tomar *et al.*, 2016).

Morabito (2015) and Desouza and Smith (2014) suggest that critical skills required for BD can be obtained from scholars. Desouza and Smith (2014) add that students learn critical skills for data organisation, preservation, visualisation, search, and retrieval, which they can use during their employment. Morabito (2015) suggests that data scientists are often PhD graduates who combine mathematical and programming skills to interrogate databases to uncover trends and build predictive models.

Vassakis *et al.* (2018) offer that if organisations intend to leverage data through BDA, they need human capital with a high level of technical skills, and who also understand how to exploit these systems to achieve consumable knowledge for end users. These skills include statistics, BD mining, master visualization tools, a business oriented mindset and machine learning (Vassakis *et al.*, 2018). However, these skill sets (data scientists, data analysts et cetera) are etremely rare because few people are proficiently qualified in them, thus making such individuals high in demand (Vassakis *et al.*, 2018).

Pencheva *et al.* (2018) also suggest that there is a requirement to embrace BD and technology in the public sector because it reshapes the workforce. Furthermore, the aforesaid scholars mention that Sims and Sossei conducted a survey in 2012 on public managers relating to the barriers to the adoption of BDA, in which they found that the key challenge identified was inadequate budget resources, closely followed by lack of appropriate staff.

Staff with appropriate skills/talent is required to better use citizen data to improve public services, which, by definition, should be tailored more accurately (Pencheva *et al.*, 2018). However, on the cognitive level, adopting a data-driven mindset may be challenging and profound because most people base their decisions on a combination of facts, experience, and guesswork (Pencheva *et al.*, 2018). Despite this, as Pencheva *et al.* (2018) advise, it is essential for public managers to adopt a BD mindset, which will enable them to constantly search for opportunities to unlock new forms of value from data.

In addition, government organizations use BDA to explore new opportunities that can bring benefits to them (Abdullah *et al.*, 2017). However, this goal is difficult to achieve if the wrong understanding (skills) and techniques are used to perform BDA (Abdullah *et al.*, 2017). These technical or theoretical misconceptions will have a major impact on the overall analysis process, results and outcome (Abdullah *et al.*, 2017). Equally unhelpful is when public managers cannot acknowledge that business owners' main concern is providing the sufficient tools and highly trained personnel to work with BDA (Abdullah *et al.*, 2017).

2.9.3. BD legal challenges

There needs to be a well-defined regulatory environment to facilitate BD ecosystems' implementation stipulating data ownership, usage, protection, privacy, security, liability, cybercrime, intellectual property rights, and the implications of insolvencies and bankruptcy (Sun *et al.*, 2018). Since data protection regulations are continuously changing, BD organisations need lawyers or proactive compliance officers who monitor legal changes and recognise necessary posterior changes to existing regulations (Hummel *et al.*, 2018).

Pagallo (2017) states that significant consideration has been given to studying the legal challenges associated with BD in recent years. The main emphasis of these studies often includes data protection, for example the role and content of the primary rules of the BD system (Pagallo, 2017). This is acceptable as it places the emphasis on the standards that should govern social and individual behaviour in terms of individual consent and data minimisation, accuracy and purpose limitation, integrity and confidentiality, as well as the principles of lawfulness, fairness and transparency (Pagallo, 2017).

Furthermore, creating a data-driven society in the public sector requires balancing protection against

data misuse and data sharing regulations (Tomar *et al.*, 2016). Suitable open data policies and inter-government data exchange policies are key objectives for exploring BD laws (Tomar *et al.*, 2016). La Torre *et al.* (2018) suggest that while organizations use competitive intelligence to shape their strategic planning, they sometimes cross ethical and legal boundaries.

Hummel *et al.* (2018) state that BD analytics that includes personal data must comply with many data protection guidelines and laws, for example, using encryption, partitioning, anonymisation, de-personalisation or pseudonymisation techniques, and it may also be subject to negotiations with various bodies. Furthermore, because BD systems must be scalable, on-demand and are often built in cloud environments, BD analytics must meet privacy, legal, and licensing requirements for commercial components or data (Hummel *et al.*, 2018).

2.9.4. BD technology challenges

An Accenture (2017) report examined the adoption of emerging technologies across government agencies that interact with citizens or handle citizen-facing services such as health, and social services, policing and justice, revenue, border services, pension and social security, and administration. The survey findings across nine countries and nearly eight hundred IT leaders indicated that over two-thirds of government agencies evaluate the potential of emerging technologies, including BD. However, only one fourth has moved beyond the pilot stage (Gutierrez, 2017).

The infrastructure used for BD must support the entire data lifecycle (Demchenko *et al.*, 2014). Many key technical challenges exist for government organisations wanting to adopt BD. Government organisations need to invest in such technologies as cloud computing and data warehouses, and attain or guarantee large-scale heterogeneous dataset acquisition, efficient data storage, real-time data processing and data analysis, data curation, advanced data retrieval and visualisation, intuitive user interfaces, interoperability and linking data, information, and content (Cavanillas *et al.*, 2015; Tomar *et al.*, 2016). These technological investments and technical measures need to be put in place to develop and sustain competitive advantages (Cavanillas *et al.*, 2015; Yafooz *et al.*, 2020).

2.9.5. BD software applications challenges

BD can transform many sectors, including the health, finance, energy, and transport sectors. Innovative applications and solutions must be developed, validated, and delivered in the BD ecosystems (Tomar *et al.*, 2016). BD only creates value if data consumers can interpret the facts produced through these applications (Cavanillas *et al.*, 2015). These applications' data are a significant source of valuable insights and ultimately help to make more precise decisions (Wadhwani & Wang, 2017). However, BD

has special qualities that traditional software applications cannot manage and process, which becomes a real problem (Thabet & Soomro, 2015).

First, Wadhwani and Wang (2017) mention that processing BD is complex and that conventional applications or traditional database management tools do not perform well for processing BD (Wadhwani & Wang, 2017). The second issue is that data sizes can vary from Petabytes to Zettabytes, and BD is normally in structured, semi-structured, and unstructured formats, contributing to the challenges of processing BD with traditional data analytics tools (Wadhwani & Wang, 2017).

Moreover, Sivarajah *et al.* (2017) suggest that BDA is still a labour-intensive undertaking because solutions for analytics are often based on proprietary appliances or software systems built for general purposes. Consequently, organizations need to add significant efforts to customize such BDA solutions for their individual needs, which requires integrating different data sources and setting up the software on the organization's hardware (Sivarajah et al., 2017).

2.9.6. BD social challenges

It is critical to increase awareness of BD's benefits for business, the public sector, and citizens (Cavanillas *et al.*, 2015). BD's acceptance can provide solutions for major societal challenges, such as improved efficiency in healthcare, increased liveability of cities, enhanced transparency in government, and improved sustainability (Cavanillas *et al.*, 2015; Maass *et al.*, 2018).

Adopting BD and BDA can transform what we do with these technologies and our attitude towards them, with implications for private individuals and society (Morabito, 2015). Although BD activists advise about DB's organisational advantages and opportunities, critics warn about the societal risks, specifically those that infringe on our privacy and abuse our civil rights (Morabito, 2015; La Torre *et al.*, 2018). In addition to this, data governance is still lacking, which prevents BD's use for social good (Desouza and Smith, 2014).

Halper and Krishan (2014) also observe that cultural and political issues can prevent BD from becoming more pervasive in society (Halper & Krishnan, 2014). Societal barriers have always been seen as obstacles to major changes (Morabito, 2015; La Torre *et al.*, 2018). Transforming societal beliefs to be more data-centric is a major challenge, especially when the data being used by governmental organisations are produced by them (Morabito, 2015). These changes require that citizens observe how the data is used and discover the benefits of transparency and public service delivery (Morabito, 2015).

2.9.7. BD leadership challenges

Common practices relating to business processes and communication in a governmental organisation

are usually top-down, for example, when organisational change and innovations are not driven by top management, it negatively influences government initiatives (Koo, 2019). Adopting BD technologies can have a higher failure rate if leaders do not promote them (Koo, 2019). Management support ensures that leaders accept and support the adoption of these innovations (Koo, 2019).

When there is a commitment from leadership to accept BD, successful adoption follows. Organisational leadership contributes to providing a clear, comprehensive vision for the use of data that falls within a larger development plan and includes accessible procedures and incentive alignment for creators, analysers, and users of BD (Tomar *et al.*, 2016). The engagement of these key stakeholders within government organisations is an important factor that ensures the following: i) access to relevant and timely data; and ii) the cultural transformations required within the organisations to transform the data-driven decision-making processes (Tomar *et al.*, 2016).

Organisations that strive to be successful in the data-driven era have leadership teams that govern aims, modulate achievements and ask the correct questions to be answered by data insights (Vassakis *et al.*, 2018). Regardless of its technological nature, BD's power cannot be exploited without vision or human insight (Vassakis *et al.*, 2018). Leaders of organisations with vision and the ability to reveal future trends and opportunities can be innovative and motivate their teams to work efficiently to achieve their targets (Vassakis *et al.*, 2018).

2.9.8. BD privacy challenges

Privacy concerns relating to BD are probably what most ordinary citizens are concerned about (Moreno *et al.*, 2016). Nowadays, everywhere, data is produced by everybody through the use of mobile devices, digital services, and the internet of things (IoT), which leaves digital traces of citizens (La Torre *et al.*, 2018). The collection of these traces by unauthorised individuals or groups can lead to privacy infringements. Data access and collection should occur with the awareness and permission from citizens or groups, although this is not always the case (La Torre *et al.*, 2018). The wide-ranging availability of data raises serious concerns about privacy and the use of personal and sensitive data (La Torre *et al.*, 2018).

Privacy concerns are frequently regarded as a major obstacle when organisations plan to adopt BD (Demchenko *et al.*, 2014; Moreno *et al.*, 2016). Examples of this have been seen during the adoption of BD in the USA, where the following infringements were identified: inappropriate use of data; including rights to the secrecy of personal information; and rights to remain anonymous while exercising free speech (Gutierrez, 2017a). In the research by Arp *et al.* (2017) about mobile phone tracking, it was found that embedding ultrasonic beacons into audio signals creates the ability to track users using the microphones in mobile devices (Arp *et al*., 2017).

Moreno *et al.* (2016) report that privacy can be protected with various techniques and mechanisms. The aforesaid scholars suggest the use of cryptography, anonymisation, and differential privacy. Cryptography uses algorithms to encrypt the data and access control to restrict the access of the non-desirable user. Anonymisation removes sensitive (identifiable) information from the data while differential privacy aims to provide a method to maximise the value of analysis of a set of data while minimising the chances of identifying users' identities (Moreno *et al.*, 2016).

2.9.9. BD security challenges

Security-related issues are a big concern for RDBMS approaches and are exhibited more complexly in BD because of BD's volume and velocity (Salehnia, 2015). BD's availability and accessibility to organisations introduce more opportunities for hackers and criminals to have faster access to governmental data (Salehnia, 2015; Yakobi *et al.*, 2020). In current times, more data is stored and analysed by organisations and governments, which requires regulations that meet BD security concerns (Salehnia, 2015). Sufficiently addressing security concerns relating to BD has become an important barrier that could slow down the adoption of the technology if it is not met with adequate security guarantees (Salehnia, 2015). Therefore, the adoption of BD is inevitably pegged to a high level of trust.

An important consideration for BD adoption is infrastructure guarantee data security, integrity, availability, accountability and ownership (Demchenko *et al.*, 2014). Organisations are putting policies in place to encourage collaboration. Different user groups can process data in third-party facilities, but they must comply with security requirements (Demchenko *et al.*, 2014). Citizens must also trust that the organisations processing their data have put measures in place to prevent unauthorised access (Demchenko *et al.*, 2014; Yakobi *et al.*, 2020).

Salehnia (2015) and Moreno (2016) recommend that organisations implement encryption, data anonymisation, tokenisation, and cloud database controls to address some of these security concerns. Encryption can be implemented at the software and hardware levels that operate on selected data on the fly or across an entire disk (Salehnia, 2015). However, encryption increases the central processing unit (CPU) load of computing resources, increasing computing resources' cost and ultimately increasing overall complexity (Salehnia, 2015). Data anonymisation deals with removing all data that can be uniquely related to an individual. While this technique can protect personal identification, organisations need to be cautious about the amount of information they remove (Salehnia, 2015). Tokenisation protects sensitive data by replacing data with random tokens or alias values that will be incomprehensible to someone who gains unauthorised access to this data (Salehnia, 2015). This technique decreases the chance that individuals gaining unauthorised access to data can do anything with it (Salehnia, 2015). Lastly, cloud database controls relate to access controls built into the database to protect the entire database so that each piece of data does not need to be encrypted (Salehnia, 2015).

2.10. Technology-Organisation-Environment (TOE) framework and Big Data adoption in the public sector

Rogers (1995) shows that the innovation adoption process is described by two major phases, initiation and implementation, and is separated by an adoption decision (Bremser, 2018). The initiation phase comprises the agenda-setting and matching stages. The agenda-setting is triggered by an organisational problem or the perceived usefulness of the innovation. The matching stage is used when the organisations' members explore the capabilities of the innovation to predict its potential for specific application scenarios (Bremser, 2018). The implementation phase is triggered when advantages are seen, including all the activities and decisions required to assimilate the innovation (Bremser, 2018).

New IT innovations are widely regarded as essential in ensuring the competitiveness of a country or organisation (Oliveira & Martins, 2011). However, these innovations' benefits can only be realised if and when the innovations are widely spread and used (Oliveira & Martins, 2011). Tornatzky and Fleischer (1990) developed the Technology-Organisation-Environment (TOE) framework to explain the decision to adopt a technological innovation by an organisation based on the technological, organisational and environmental contexts (Kalema & Mokgadi, 2017). The TOE framework has been extensively used in technology adoption studies and provides a useful approach to understanding the factors that influence technology adoption in organisations (Pudjianto *et al.*, 2011). The TOE framework categorises three aspects: the technological; the organisational; and the environmental contexts that influence the adoption process at an organisation level and are depicted in Figure 4 below (Almoqren and Altayar, 2016; Alqahtani, 2016; Pudjianto *et al.*, 2011).

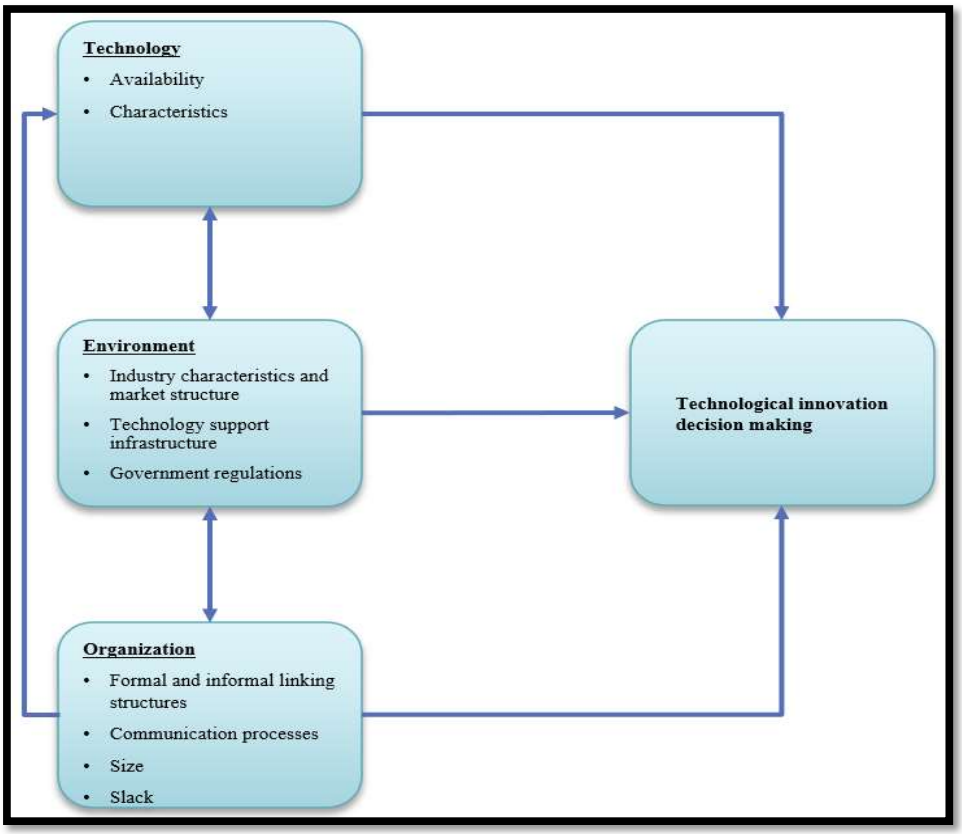


Figure 6: Technology, organisation, and environment framework (Adapted from Rahman & Aydın, (2019)

The TOE framework is suitable for studying the factors influencing BD adoption in government organisations because it allows us to evaluate the importance of different factors that affect the inclination to adopt IT (Ismail & Mokhtar, 2013). To determine the factors that influence the adoption of technology at an organisational level instead of an individual level, various researchers have studied the adoption process through the TOE framework (Almoqren and Altayar, 2016; Nam *et al.*, 2015; Sun *et al.*, 2018). Almoqren and Altayer (2016) suggest that the TOE framework is suitable for these types of studies because it measures the success of information system adoption in terms of human, organisational, and environmental factors.

The TOE framework evaluates the concept of implementation that provides identification and contributions to adoption and the need to understand innovation to provide in-depth insights to researchers and practitioners (Defitri *et al.*, 2020). The TOE framework also assists in understanding the factors that exist in technology, organisation, and environment that influence the adoption of technological innovation (Defitri *et al.*, 2020).

However, merely concentrating on one characteristic of the TOE framework may not adequately explain the successful adoption of BD within an organisation. Government organisations that adopt BD may

need to change other areas, such as streamlining business processes to realise BD's full potential and advantages. Using the TOE framework, the study aims to investigate these by, firstly, looking at the technological aspects to determine whether the organisations' present infrastructure/systems is/are appropriate to function with BD. The study looks at an organisation's IT infrastructure, reliability, and security of the data they collect. Secondly, within the organisational aspect, the study looks at whether sufficient support from top management exists within the organisation to adopt BD. The study also considers the organisational size, finances, processes, structure, and strategies. Lastly, within the environment aspect, the study investigates external factors that may influence the adoption of BD, such as customers, vendors, and competitors.

2.10.1. Technological context

Technological factors play a significant role in ensuring the successful adoption of information systems, and they necessarily include technology integration, IT infrastructure, technology readiness, and technology resources (Almoqren & Altayar, 2016). Nam *et al.* (2015) state that internal and external technologies relevant to the organisation should be examined to understand how they affect the adoption and implementation of technologies. Sun *et al.* (2018) argue that the perceived benefits gained from technology are acquired through technological innovation and support from organisational resources and other organisational characteristics. The main focus of this context is on how technology characteristics influence the adoption process and include current practices and equipment internal to the firm and the pool of technologies external to the firm (Ismail & Mokhtar, 2013).

2.10.1.1. Reliability / Suitability

Government organisations must access quality data supporting their decision-making process because their decisions affect many constituents. Thabet and Soomro (2015) state that data reliability, or suitability of data, is described by data quality and that it has four components or features (Thabet & Soomro, 2015).

First, there is completeness, which measures if all the relevant data is available. For example, when a citizen submits a service request, the record should contain details like name, account number, address, and the nature of the request (Thabet & Soomro, 2015). Accuracy is the second component, and it confirms that the data is free of misspellings, typos, wrong terms, and abbreviations (Thabet & Soomro, 2015). The third component is that of availability, which confirms that the data is available when required and is easily searchable (Thabet & Soomro, 2015). The last is that of timeliness, which indicates that the data is up to date and ready for use, to support the decision process (Thabet & Soomro, 2015). Without reliable data, government organisations cannot make crucial decisions, and, sometimes, this may have a huge financial impact.

2.10.1.2. IT Infrastructure

Engin and Treleaven (2019) believe that the effects of the 'smartification' of public services through BD substantially impacts government organisations' IT infrastructure because the services or functions they perform impact every institution or individual. IT infrastructure that provides access to these services is a major driver for innovation (Engin & Treleaven, 2019).

Kalema and Mokgadi (2017) believe that IT infrastructure is vital for government organisations that want to adopt BD because BD analytics applications frequently rely on a growing number of internal and external data sources comprising structured and unstructured data (Kalema & Mokgadi, 2017). This implies that the organisation's readiness for BD adoption highly relies on its current IT infrastructure (Kalema & Mokgadi, 2017). For the organisation to access and combine data from various sources in different formats, it needs good IT infrastructure to handle such data. If not, new systems need to be procured (Kalema & Mokgadi, 2017).

In some government systems where BD emerges, data cannot be processed or analysed because of its complexity, heterogeneity or size (Suh *et al.*, 2015). Suh *et al.* (2015) elaborate by stating that IT infrastructure plays an important role in creating, storing, processing, and managing BD, which requires adequate compute, network, and storage infrastructure to meet scalability, security, and availability requirements. Due to the attributes of BD, successful adoption of BD might not be possible if compatible IT infrastructure does not exist.

2.10.1.3. Security / Privacy

As mentioned in the IT infrastructure section above, processing and storing BD requires adequate IT infrastructure. However, it should also have the ability to enforce data policies that allow for data to be processed on trusted systems that comply with organisational requirements in terms of security and privacy (Shin & Choi, 2015). When there is a collaborative initiative between the public and private sectors, the infrastructure ensures that data portals are secure and data is ownership protected (Yakobi *et al*., 2020).

Due to the sensitive nature of government data and penalties imposed by regulators when data breaches occur, data security is the top priority for government IT executives (Lockner & February, 2012). BD provides both advantages and disadvantages in terms of security. Its unique analytic features aid organisational security. The availability of large amounts of real-time data from dispersed sources creates new risks that can lead to system intrusions, cybersecurity attacks and fraud (Kalema and Mokgadi, 2017; Yan, 2013).

2.10.2. Organisational context

The organisational context refers to elements such as top management, human resources, IT training programs and information management (Almoqren & Altayar, 2016). Sun *et al.* (2018) state that it is the resources and other characteristics of the organisation, such as organisational size, organisational structure, managerial structure, human resources, and employees' skills. Nam *et al.* (2015) propose that this element should not be studied as technology use or user adoption, but rather as how extensively the innovation is used and how deeply the organisations' use of the technology alters their processes, structures, and organisational culture (Nam *et al.*, 2015).

2.10.2.1. Top management

Management support is consistently found as a critical factor that influences the adoption of modern technologies (Hameed, 2017). Support could be in the form of finance, recruitment of skilled staff, training or aiding in the alignment of IT strategies with those of the organisation (Kalema & Mokgadi, 2017). Additionally, when employees need to embrace organisational changes, they should be confident that the organisation's management provides tangible support for change in the form of resources and information (Kalema & Mokgadi, 2017). The belief that top management is ready to support employees is also crucial for individual development and leads to a sense of self-efficacy and mind-set change towards organisational change (Kalema & Mokgadi, 2017).

Sun *et al.* (2016) state that executive attitude towards technology adoption influences the adoption decision because a faster rate of adoption is seen when executives sponsor the projects. Additionally, when executives have extensive IT knowledge, they can exert more influence on an organisation's adoption approach of BD, fostering a supportive climate towards user acceptance of innovations (Hameed, 2017; Sun *et al.*, 2016).

2.10.2.2. Organisational size

Government agencies and departments have traditionally been separated as internal rivals and even competed for superiority, which has led to an embedded silo mentality and culture (Morabito, 2015). This prevents the sharing of data and collaboration on projects (Desouza and Smith, 2014). BD promotes the sharing and re-use of data, enabling interactions between government and citizens, government and businesses, and inter-agency relationships, thus forming more effective governance and transparency (Mellouli *et al.*, 2014).

Since the government organisation under study have many departments, they receive huge volumes of data from varied sources. This suggests that the need to analyse BD within these departments could be higher than in smaller organisations (Kalema & Mokgadi, 2017). Furthermore, organisations that

generate or have been dealing with large data volumes from different sources for decades are more likely to be early adopters of BD, as this would aid in the analyses of their data (Kalema & Mokgadi, 2017).

The approach of sharing and re-use of data is being adopted by the South African government, and this is confirmed in a strategic plan released by Statistics South Africa, which states that Human Resources (HR), Finance, ICT and all other support units are re-positioned as both strategic and support partners to the organisations' core business.

2.10.2.3. Finance

Bryant *et al.* (2008) suggest that sizeable government financial investment could greatly accelerate the adoption of BD. Günther *et al.* (2017) argue that organisations that rely on BD to guide organisational strategies and day-to-day operations are expected to perform better financially than organisations that do not use DB because BD enables the creation of innovative products, services, and business opportunities. However, according to a 2012 Gartner survey, more than seventy percent of government respondents indicated that their organisations have not made any investment in BD, and thirty percent have no plan to do so (Yan, 2013). Morrison (2016) claims that this is due to the fact that government leaders do not have an awareness of BD or how it helps to shape the competitiveness of their country's businesses endeavours.

Furthermore, since BD analytics deals with increasing user needs that exceed the traditional approach to analytics, organisations may require more resources (Kalema & Mokgadi, 2017). These resources could be in the form of new technological devices, applications and tools and the employees experienced in data science that necessitate the availability of financial resources (Kalema & Mokgadi, 2017). Furthermore, analysing BD warrants that organisations upgrade from conventional Online Transaction Processing (OLTP) data stores and Structured Query Language (SQL) analysis tools to flat horizontally scalable databases with distinctive query tools (Kalema & Mokgadi, 2017). Changing from traditional to advanced BD analytical tools that process real-time data also requires that organisations are in a sound financial position to procure these tools.

2.10.2.4. Process

Chen *et al.* (2015) state that BD has the ability to fundamentally change organisational processes, business models and strategies, and even entire industries and markets. Furthermore, they claim that BD adoption decisions include layered complexity and high levels of uncertainly and risk, which potentially involve drastic structural and process changes in the organisation (Chen *et al.*, 2015; Yakobi *et al.*, 2020).

Lamba and Dubey (2015) mention that BD strategy is about discovering insights and knowledge from BD and taking the required steps to drive knowledge into organisational processes that creates new innovative products and services. The aforesaid scholars add that BD strategy should be aligned with organisational strategy.

2.10.2.5. Strategy

Kalema and Mokgadi (2017) suggest that sound organisational strategies positively influence an organisation's readiness for BD adoption. They also argue that since BD has been available for a while, organisations ought to put more emphasis on tools and other technological developments for analytics rather than organisational strategies to process BD (Kalema & Mokgadi, 2017).

Günther *et al.* (2017) state that, to use BD effectively, organisations need to exchange data with their network of partners and engage in data disclosure and screening practices. Organisations require intuitive tools that integrate data into day-to-day processes and translate modelling outputs into tangible business actions (Kalema & Mokgadi, 2017). Michael and Miller (2013) also suggest that organisations should be concerned with having experienced employees to deal with BD analytics to support managers in making real-time decisions. Since government organisations are some of the largest collectors of data, extracting the maximum value from data by building on the intelligent use of data sources is critical.

2.10.3. Environmental context

The environment context describes the environment in which the organisation conducts business and includes competitors, information intensity, and regulatory environment (Almoqren and Altayar, 2016). Nam *et al.* (2015) state that the environment includes the industry's size and structure, the organisation's competitors, and dealings with the government. Sun *et al.* (2018) refer to it as aspects that include the organisation's partners and competitors and the regulatory environment (Sun *et al.*, 2018).

2.10.3.1. Competition

Nam *et al.* (2015) state when business partners and competitors influence an organisation's willingness to adopt new technology, it is referred to as 'industrial pressure.' They explain that there might be two cases in which an organisation might feel the need to adopt new technologies due to industrial pressure, namely, (a) partner(s) requesting or recommending it to do so, or when organisations see their competitors doing so (Nam *et al.*, 2015).

Furthermore, when BD is used to its full potential, organisations can obtain better insights required to gain a competitive advantage over their rivals (Kalema & Mokgadi, 2017). Intra-governmental-

organisation competition might not be at the top of their agenda because it is more service orientated than profit oriented (Kalema & Mokgadi, 2017).

2.10.3.2. Customers

Customers' demands influence the adoption of BD since organisations need to provide real-time answers to customer queries (Kalema & Mokgadi, 2017). Besides, more people are becoming IT literate and, thus, will tend to do more of their activities online, including requesting services, which increases the amount of data that organisations receive (Kalema & Mokgadi, 2017). Due to this, organisations are compelled to have better analytical tools that can process queries in the shortest time possible (Kalema & Mokgadi, 2017).

Customer queries can take on many forms (1 to 1 or many to many, et cetera) which increases the velocity of data. This may lead to organisations being overwhelmed by the flow of data and, thus, becoming inefficient. In order to avoid this, the organisations tend to acquire new IT infrastructures and employ experienced personnel to handle these new analytical tools (Kalema & Mokgadi, 2017). Furthermore, careful consideration should be given to the trade-offs in terms of the challenges and benefits customers receive when BD is adopted.

2.10.3.3. Vendors

Having sufficient expertise within government organisations to support BD would be beneficial since this will alleviate any outsourcing costs. However, if government organisations do not have skilled experts, expertise from vendors is required. Vendors can offload tasks that are outside the organisation's core competencies and help strengthen its existing capabilities of handling BD (Kalema & Mokgadi, 2017).

Yafooz *et al.* (2020) state that various vendors offer ready-made solutions that provide BD technologies. However, industry experts advise that the development of internal solutions designed to manage and solve an organisation's existing problems is more appropriate if the organisation has suitable technical capabilities to do so (Yafooz *et al.*, 2020). Rahman (2016) notes that vendors support is very important since many BD tools are developed under an opensource licensing model. Additionally, Kalema and Mokgadi, (2017) suggest that organisations need to ensure that vendors that supply BD technologies, or the support thereof, can provide consistent support.

2.11. Chapter conclusion

The study is underpinned by Technology-Organisation-Environment (TOE) framework, which is used to understand the factors that influence BD adoption within a government organisation. Based on the

literature reviewed, eleven identified factors (reliability / suitability, infrastructure, security / privacy, top management, organisational size, finance, process, strategy, competition, customers and vendors) which may affect BD adoption within a government organisation were drawn, as depicted in figure 7 below

The strength of the TOE framework is, first, in measuring the success of information systems adoption is complex because the factors of information technology are influenced by human, organisational, and environmental factors, and the TOE provides a framework to measure these factors (Almoqren & Altayar, 2016). Second, various researchers have used the TOE framework to understand different IT adoption efforts such as Enterprise Resource Planning (ERP), Knowledge Management Systems (KMS) and BD (Oliveira & Martins, 2011). Last, this framework can be used at organisational level which is in-line with this study (Almoqren & Altayar, 2016).

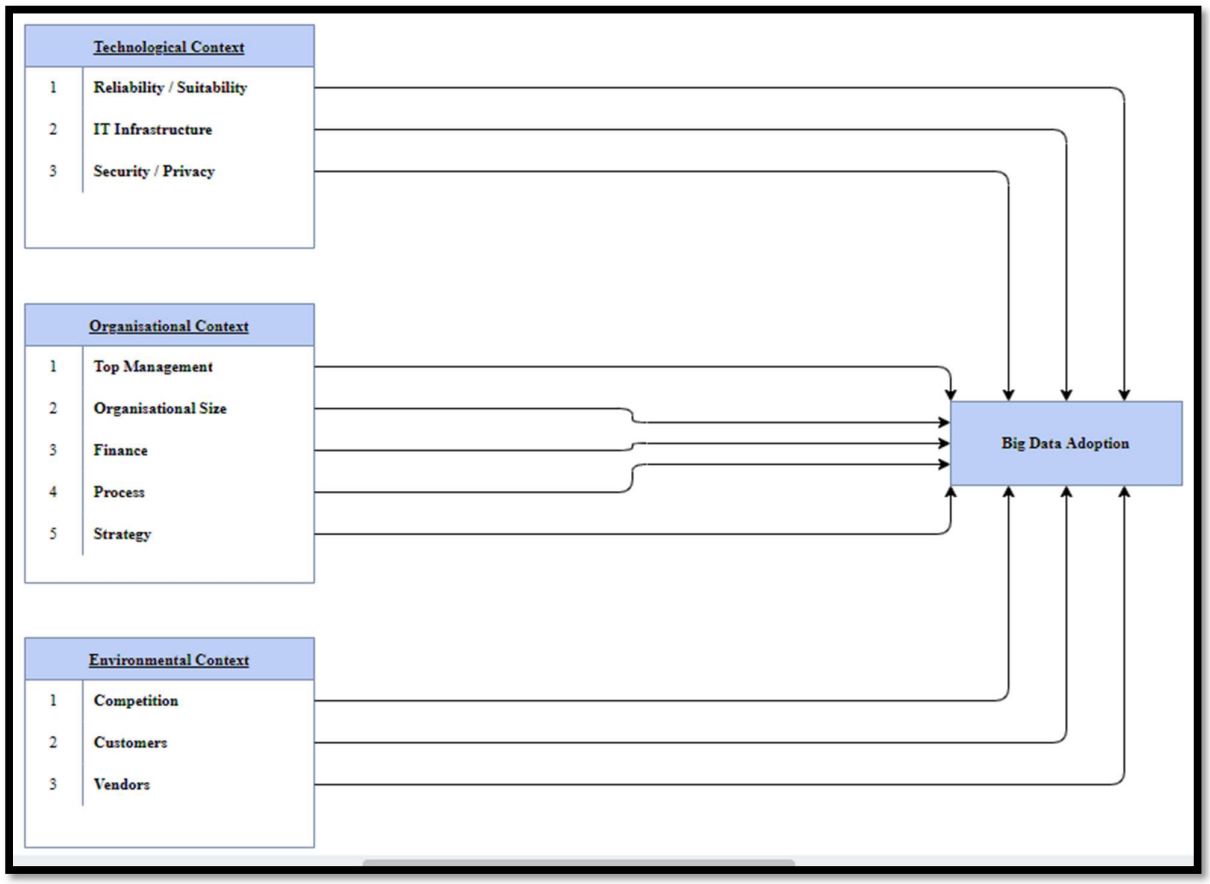


Figure 7: Adoption Factors (Source: The Researcher, 2022)

3. Research design and methodology

3.1. Introduction

Research, in its simplest form, is searching for knowledge and truth. This chapter describes the research methodology and process used to conduct this empirical case study at a governmental institution in the Western Cape. It positions the study within specific methodological tradition, details the research methodologies, data collection method, data processing and data analysis techniques used to interpret the data gathered from open-ended interviews. The importance of this chapter is to detail all processes used to conduct this research, including the methodological collection of data and its analysis, in order to answer the main research question, as stated in Chapter one.

3.2. Research design

It is important to emphasise the difference between research design and research methods. Research design has been widely described as the steps the researcher follows to find answers to the research question. Abutabenjeh and Jaradat (2018) refer to research design as a blueprint, or how a study is structured. Broadhurst *et al.* (2012) state that the function of research design is to ensure that the evidence collected allows the researcher to answer the research questions as explicitly as possible. De Vaus (2006) concurs by stating that research design is the overall strategy to integrate the various components of a study clearly and logically, ensuring that the research problem is addressed effectively.

Welman *et al.* (2005) mention that the research design guides the researcher in obtaining data relating to the research phenomenon through the data collection instrument. They continue to state that research design is the conceptual structure that details how the research would be conducted to collect relevant information (Welman *et al.*, 2005).

3.3. Research methods

Research methods are specific techniques for collecting and analysing data, including questionnaires and open-ended interviews (Rehman & Alharthi, 2016). The method used for a research project depends on the design of the project and the researcher's theoretical mindset (Rehman & Alharthi, 2016). Abutabenjeh & Jaradat (2018) mentions that research methods involve three elements, namely, i) forms of data collection, ii) analysis, and iii) interpretation of collected data. Furthermore, they state that research methods can take on three forms, qualitative, quantitative and mixed methods (Abutabenjeh & Jaradat, 2018).

3.3.1. Research approach (paradigm)

The selection of a research approach is underpinned by the philosophical assumptions (paradigms) that the researcher brings to the study, the procedures of inquiry (called research designs), and the specific

research methods for data collection, analysis, and interpretation (Creswell, 2013). Selecting a philosophical outlook to investigate a phenomenon should be guided by the necessities and requirements of the research study rather than the stubborn insistence of adhering to one particular philosophical outlook and the exclusion of others (Rehman & Alharthi, 2016).

As researchers, it is important that we have a clear understanding and be able to articulate beliefs about the nature of reality, what can be known about it and how we go about attaining this knowledge (Rehman & Alharthi, 2016). Rehman and Alharthi (2016) describe research paradigms as basic belief systems and theoretical frameworks with assumptions about ontology, epistemology, methodology and methods.

3.3.2. Ontology

Rehman and Alharthi (2016) state that ontology refers to the nature of our beliefs about reality. Mack (2010) describes it as one's view of reality and being. Similarly, Scotland (2012) states that ontological assumptions are concerned with what constitutes reality. When researchers adopt a paradigm, they need to take a position regarding their perceptions of how things really are and how things really work. Rehman and Alharthi (2016) elaborate by stating that researchers have assumptions about reality, how it exists and what can be known about it. Ontological questions, therefore, lead researchers to inquire what kind of reality exists.

The interpretivist ontological perspective was used in this study because the researcher aims to understand the social constructs of the issues under investigation. Interpretivism attempts to understand phenomena through the meanings that people assign to them, focusing on the complexity of human sense-making (Twala & Kekwaletswe, 2019). This philosophical paradigm allows the researcher to gain a deeper understanding of the phenomenon that otherwise would not be possible using the positivist paradigm (Twala & Kekwaletswe, 2019). Therefore, with the combination of the TOE framework and the identified factors, this study aims to understand the contextual elements that influence the decisions to adopt BD at a government organisation in the Western Cape.

3.3.3. Epistemology

Epistemology refers to studying the nature of knowledge and the process by which knowledge is acquired and validated (Rehman & Alharthi, 2016). Scotland (2012) writes that epistemology asks the question, *"what is the nature of the relationship between the would-be knower and what can be known?"* It is concerned with the nature and forms of knowledge, how it can be acquired, and how it can be communicated to other human beings (Rehman & Alharthi, 2016). Mack (2010) simply states that it is the view of how one acquires knowledge.

The interpretive epistemology is one of subjectivism based on real-world phenomena (Scotland, 2012). Scotland (2012) states that different people may construct different meanings to the same phenomenon due to their own way of thinking (mental state). Through subjectivism the researcher believes socially constructed multiple realities lead researchers to reject the notion that people should be studied like objects of natural sciences. Still, researchers should rather get involved with the subjects and try to understand phenomena in their contexts (Rehman & Alharthi, 2016).

Interpretivist researchers adopt a relativist ontology in which a single phenomenon may have multiple interpretations rather than truthfulness that can be verified by a measurement process (Pham, 2018). Within the interpretivist perspective, researchers obtain a more in-depth understanding of the phenomenon and its complexity within a unique context instead of generalising the foundation of understanding for the whole population (Pham, 2018).

Interpretivism has two main advantages. First, interpretivism investigates a diversity of phenomena. For example, interpretivist researchers describe not only objects, human or events, but also deeply understand them in a social context (Pham, 2018). Second, they leverage key interactive interviewing methods that allow researchers to investigate and prompt for things that we cannot observe, which allows researchers to probe an interviewee's thoughts, values, prejudices, perceptions, views, feelings, and perspectives (Pham, 2018).

Walshan (1995) argues that interpretive research methods adopt the stance that human actors socially construct people's knowledge of reality. Bhattacherjee (2017) mentions that when social order is studied through the subjective interpretation of participants, interviewing diverse participants and merging their differences among their responses using their own subjective perspectives, we are employing an interpretivist paradigm.

Based on the above ontology, interpretivism is appropriate for this study because the researcher interacts with the study participants to obtain the real meaning of the topic under study, which subsequently changes the perceptions of both parties (Chowdhury, 2014). Second, the researcher uses his own preconceptions to guide the inquiry process. Chowdhury (2014) continues to explain that interpretivist researchers look for the existence or absence of causal relationships, the specific ways in which it is manifested, and the context in which it occurs. Furthermore, because the researcher aims to understand why the participants have constructed their specific views on the subject under study, subjectivism as epistemology is also appropriate.

3.3.4. Methodology

Creswell (2014) describes research methodologies as general philosophical orientations about the world and the nature of research that a researcher brings to a study. Other researchers have labelled it 'worldviews' or 'paradigms' (Creswell, 2014). Methodologies stem from discipline orientations, student advisors' (mentors') inclinations, and past research experiences. The beliefs held by individual researchers based on these factors often lead to adopting a qualitative, quantitative, or mixed-methods approach in their research (Creswell, 2014).

Grover (2015) refers to Creswell's (2008) research regarding the basic components of a research approach. Grover (2015) claims that research approaches comprise three basic components: philosophical worldview, research design, and research methods. From their work, it is clear that the research approach has a corresponding research design that calls for possible methods that provide a range of techniques to support the method (Grover, 2015) as depicted in Figure 8.

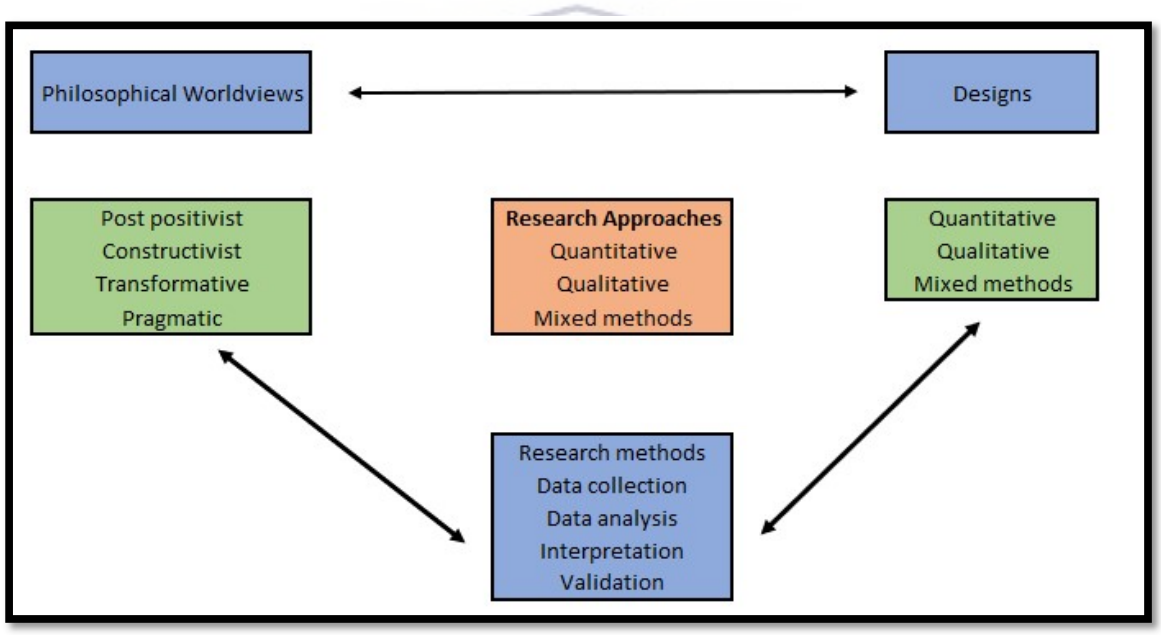


Figure 8: Interconnection - Paradigm, Design and Research (Source: Grover, 2015)

The evidence obtained from case studies is typically qualitative and focuses on developing an in-depth instead of a generalisable broad understanding (Ellis & Levy, 2017). Case studies are also useful when used to explore, describe, or explain phenomena by an exhaustive study within its natural setting (Ellis & Levy, 2017).

For this study, a qualitative case study methodology was used. First, the study is exploratory due to the relative newness of the phenomenon and because the researcher aims to address the "how" and "why"

questions as stated in Chapter 1 (Chen *et al.*, 2015; McBride *et al.*, 2019). The research aims to understand the participants' perceptions, experiences, and attitudes toward the phenomenon (McCusker & Gunaydin, 2015). Second, this research uses phenomenology because the specific topic has received little attention from scholars, and our understanding of its impact is unclear (Lester, 2006). Pietkiewicz *et al.*, (2014) describes phenomenology as an eidetic method, which is concerned with attending to the way things appear to individuals in their experience. Third, data is collected through open-ended interviews, allowing the researcher to understand the phenomenon under study because participants may respond in their own words describing their own sentiments (Rachlitz, 2017). Lastly, the type of data collected is in narrative form, which enables the researcher to obtain a deep understanding of the phenomenon (Vinet & Zhedanov, 2011).

3.4. Research inquiry strategies

3.4.1. Qualitative strategies

The three most common research methodologies are qualitative, quantitative, and mixed-method research. Qualitative research approaches are associated with the interpretive social sciences paradigm, where forms of investigation are based on the significance of human beings' subjective, experiential realm (Khomba, 2007). Qualitative research also provides avenues that lead to the discovery of the deeper levels of meaning and understanding of the participants. Qualitative research methods enable researchers to comprehend social events from the participants' perspective or understanding (Polit & Beck, 2010). This research endeavour is in the form of a case study.

3.4.2. Case study

Case studies study people, events, decisions, periods, projects, policies, institutions, or other systems that are studied holistically by one or more methods (Kumar, 2013). Kumar (2013) explains that a case study can describe an entity that forms a unit, such as a person, organisation, or institution. As an example, researchers may want to study the dynamics within different groups or individuals, for instance, the social interactions within different neighbourhoods or the processes within different production systems (Kumar, 2013).

3.5. Rationale for selecting case study research

This study aims to identify factors that influence BD adoption in a Western Province governmental organisation. The study is qualitative and, specifically, an exploratory case study. It seeks to obtain the interactions between information technology (IT) related innovations and organisational contexts (Darke *et al.*, 1998). Eisenhardt (1989) states that case study research focuses on understanding the dynamics present within a single setting. It should be used to obtain an in-dept understanding of a phenomenon when there is a lack of theories that address a specific research area. Creswell (2013)

mentions that this type of research inquiry seeks to establish the meaning of a phenomenon from participants' views through in-depth accounts of occurrence. Darke *et al.* (1998) claim that case study research is the most widely used qualitative research method in information systems research because it is well suited to understanding the interactions between information technology-related innovations and organisational contexts (Darke *et al.*, 1998).

Case study research has three strengths specific to research in information systems. First, the researcher can study information systems in a natural setting to generate theories from practice (Benbasat *et al.*, 1987). Similarly, this has proven valuable for exploring the relationship between BD adoption factors and users' perceptions and behaviours (Pozzebon *et al.*, 1998). Second, this method allows the researcher to understand the nature and complexity of the processes taking place (Benbasat *et al.*, 1987). Since the method for data collection is semi-structured interviewing, the researcher studies the nature and complexity of the processes that influence BD adoption. Last, valuable insights can be gained into new topics emerging in the rapidly changing information systems field (Benbasat *et al.*, 1987). This is identified during the study and substantiated by the data collected. These new findings are discussed in the Conclusion section. Another contributing factor for case study research was that very few theories and empirical studies existed to address the stated research problem (Eisenhardt & Eisenhardt, 1989).

The interpretivist paradigm used during this research inquiry aims to examine real-life phenomena in-depth. In this study, real-life phenomena refer to the research participants, events or situations that occur within the organisation that contribute to the participants' sentiments about the factors that influence BD's adoption within their organisation (Guest *et al.*, 1978). Interpretivist approaches try to understand phenomena by assessing the meaning participants assign to them and focusing on their cultural and historical context (Thomas, 2010).

Organisations and individuals have their common and unique features. Case study researchers aim to identify such features as well as the various interactive processes at work and show how these affect the adoption of systems and influence the way an organisation functions. Similarly, in this study, the researcher aims to understand the factors that influence BD adoption within the governmental organisation through the use of the research instrument (see appendix 1). These processes may remain hidden in a large-scale survey but could be crucial to the success or failure of systems or organisations (Bell, 2010).

This research study necessitated the adoption of the interpretivist paradigm because the data collected was analysed to identify common beliefs, themes, and meanings to address the specified research questions. To add to the usefulness of the interpretivist paradigm, the phenomenon under this study

constituted the views, experiences, beliefs and motives which the researcher sought to understand, as opposed to the technology per se (Ferreira *et al.*, 2012).

3.6. Rationale for selecting qualitative methodology

Qualitative research is characteristically exploratory, fluid, flexible, data-driven, and context-sensitive (Mason, 2002). Kumar (2013) describes qualitative research as a systematic scientific inquiry that seeks to build a holistic, primarily narrative, description to strengthen the researcher's understanding of a social or cultural phenomenon. Mason (2002) describes it as a way to explore a wide array of dimensions of the social world, including the characteristics and patterns of everyday life, the experiences, understandings and imaginings of our research participants, the ways that social processes, institutions, discourses or relationships work, and the significance of the meanings that they generate. Ritchie and Lewis (2003) argue that how people's social realities are studied, understood and interpreted is one of the central motifs of qualitative research (Ritchie & Lewis, 2003).

Furthermore, Mason (2002) describes qualitative research approaches as all having the following in common:

- Being grounded in an interpretivist position for example, they are concerned with how the phenomenon of interest is interpreted, understood, experienced, produced or constituted.
- Are based on research methods that are flexible and sensitive to social context, and
- Analytic methods which consider complexity, detail and context (Mason, 2002).

Bell (2010) suggests that researchers adopting a qualitative perspective are concerned with understanding individuals' or groups' perceptions of the world and doubt whether social facts exist and question whether a scientific approach can be used when dealing with human beings (Bell, 2010). Therefore, quantitative methods were less suitable to address the main- and sub-questions identified in this research study because numeric responses to close-ended questions would not provide the in-depth understanding of the phenomenon that the researcher is trying to obtain.

This qualitative study aims not to generalise but rather to provide a rich, contextualised understanding of some aspect of human experience through the intensive study (Polit & Beck, 2010). The researcher believes that in-depth qualitative research is especially well suited for revealing higher-level concepts and theories unique to a specific participant or setting (Polit & Beck, 2010). In this view, qualitative findings' rich, highly detailed, and potentially insightful nature make them especially suitable for extrapolation (Polit & Beck, 2010).

The above sentiments enforce the researcher's opinion to conduct qualitative research to answer the research questions for this study. A communality in the definitions is that researchers focus on obtaining an in-dept understanding of phenomena and experiences rather than focusing on a predetermined hypothesis. By conducting interviews at a government organisation, the researcher can obtain first-hand experience of the participants and environment under study.

3.7. Research techniques

Collecting and analysing data is a step-by-step process to find insight into the research questions. Researchers can use structured or unstructured techniques when conducting research studies (Eisenhardt, 1989). Structured techniques usually employ methods that use closed-ended questions, while unstructured techniques employ methods that allow open responses from the research study participants.

In this qualitative research case study, data was collected through a literature review and semi-structured interviews. The literature review provides a starting point to grasp empirical evidence about the research topic. In addition, the interviews provide a means to address specific aspects that the researcher wants to address (Lester, 2006). Also, the open-ended nature of the interview questions provides both the interviewer and interviewee the opportunity to discuss some topics in more detail (Lester, 2006). When the participant struggles to answer a specific question or responds only briefly, the interviewer can use cues or prompts to encourage the interviewee further to consider the question (Lester, 2006). This method is particularly valuable in this study because the research aims to obtain an in-depth understanding of the research topic. Using this method of interviewing was appropriate to collect such data.

Interviews were performed at the government organisation intended for this research study and kept confidential in keeping with the university's ethics requirements. This study aims to obtain a rich set of data through semi-structured interviews surrounding the specific issues that affect the adoption of BD in a governmental organisation while also capturing the contextual complexity (Benbasat *et al.*, 1987).

3.8. Research instrument design: interviews

During this research inquiry, semi-structured interviews were conducted because they enabled participants to provide in-depth accounts of the research problem being studied (Creswell, 2008; Walsham, 1995). Babbie (2005) suggests the inherent flexibility of this approach as one of its major advantages. Therefore, there is an expectation that the participants' views are more freely expressed when the format of the interview is more flexible and open-ended than where the interview style is regulated and confined to a standard set of questions or a survey (Flick, 2002).

Conforming to the above, the interview questions (appendix 1 below) were designed to answer the main- and sub-question as stated in Chapter One. In order to fully cover the research questions, it is important for the researcher to have a set interview schedule which serves as a guide to the discussion and mitigates the chance of important topics being neglected.

3.9. Sampling

The sample selection method for this research project is non-probability sampling, in the form of snowball sampling. The researcher deliberately selected staff such as directors and senior managers because they could refer the researcher to other potential participants (Rahi, 2017). Webley (2012) describes snowball sampling as a sampling method that begins with a group of research participants known to the researcher or participants that were otherwise identified in advance. Next, the researcher requests that these participants provide the details of individuals they consider to be good research subjects for a particular study. In that way, they gradually build up a larger sample of participants (Webley, 2012).

The research purposely selected its participants because these individuals are involved in deciding which technologies are put forward for review or adoption. The researcher concluded that the sample is sufficient when the data collected held similar information and enabled the research to address some of the research questions (Walliman, 2010).

3.10. Data analysis

Analysis for case study research follows the same principles as any other qualitative study where data collection and analysis occur concurrently (Baxter & Jack, 2008). When dealing with philosophical perspectives, Broadbent *et al.* (1998) state that data analysis for qualitative research case study methods such as an interpretivist approach must be used when analysing case study data (Darke *et al.*, 1998). The goal of analysis in interpretive studies in information systems is to produce an understanding of the contexts of information systems and the interactions between these systems and their contexts, and the strength of analysis in interpretive studies is derived from the strength of the explanation of the phenomena based on interpreting data (Darke e*t al.*, 1998). In this research case study, the same approach is adopted during data analysis.

Austin and Sutton (2015) state that copious amounts of data are collected and analysed irrespective of the research paradigm that the inquiry adopts. The data collected during this inquiry was processed and analysed under the interpretivism paradigm. Kothari (2004) claims that processing suggests editing, coding, classification, and tabulation of collected data to be appropriate for analysis. Creswell (2008)

suggests that researchers do the following: (a) organise the data, (b) be acquainted with the data, (c) code the data, (d) interpret the data, and (e) present the data when using the interpretivist approach.

This research approach while analysing the data is like the approach suggested by O'Conner and Gibson (2003) and O'Leary (2004). The approach is described as i) conducting and recording the interviews ii) transcribing the interviews in which interviews are transferred onto a word processing package, and the researcher is required to judge what level of detail to choose, for example, distinguishing between 'I don't, no' from 'I don't know', iii) coding the data into themes in which the researcher looks for similar words or phrases mentioned by the interviewees, after which the researcher puts them into categories/themes, iv) generalizing the data from the themes, by grouping similar kinds of information together in categories, and v) drawing conclusions, by relating different ideas and themes to one another (O'Connor & Gibson, 2003; O'Leary, 2004).

3.11. Unit of analysis

The unit of analysis selected for this research is based on content analysis: all the participants of the government organisation in question. The major entity is therefore in the form of a group. The researcher assessed all the responses to create common themes which were used to answer the main- and sub-questions.

3.12. Data management

The interviews were recorded with an telephonic conferencing system and transferred to a computer for record / safe keeping. The interviews were then transcribed and saved in a word processing application for reading. Transcribed interviews were not labelled with the actual names of the participants but rather with a generic code. The data was stored in a safe following the university's ethics requirements. A copy of the data and consent forms was handed to my supervisor to prove reliability and validate that the interviews were conducted.

3.13. Data confidentiality

Data was collected following the university's ethics requirements. Interviews did not contain any identifying information of any participants. Interviews were pseudonymised and labelled with the following generic code: P1, identifying participant 1, P2, identifying participant 2, et cetera. The specific government department where the interviews were conducted is not mentioned.

4. Results, presentation, and findings

4.1. Introduction

This chapter presents the findings from the analysis of 11 semi-structured, in-depth interviews conducted with staff from a governmental organisation in the Western Cape, South Africa. Thematic analysis of the qualitative interview data provided an understanding of the adoption process of technology (BD). These factors influence Big Data adoption and how data is used within the organisation to make decisions that influence service delivery. The following sections summarise study participants and present the main research question results and three sub-questions. These research questions provide the framework for the presentation of the results.

4.2. Description of the study participants

The researcher conducted 11 interviews over a 4-month period, which started on the 14th of July 2021 and concluded on the 14th of October 2021. It was difficult to get participants for this study because of participant work commitments. The interviews were all one-on-one interviews and were conducted via a remote conferencing system. The table below represents the participant, general role within the organisation, and the interview duration.

Participant ID	Role	Duration
P1	Project Manager	42:28 min
P2	Professional Officer	35:09 min
P3	Director	21:54 min
P4	Manager	35:24 min
P5	Professional Officer	33:28 min
P6	Professional Officer	54:54 min
P7	Professional Officer	71:54 min
P8	Information Coordinator	39:19 min
P9	Director	30:24 min
P10	Manager	36:24 min
P11	Technology Officer	56:30 min

Table 3: Participant organisational role and interview duration (Source: The researcher, 2022)

Participant names and the name of the governmental organisation were anonymised to conform with the ethical stance of this research study, as listed in table 3. The sample comprised executives, managers and analysts with expertise and involvement with data and data analytics within various organisational roles. The data collected provides a holistic view of the adoption process, factors that influence the BD adoption process, and how data can influence service delivery. These participants were interviewed

because they have the relevant knowledge to provide the data required to investigate the identified research questions. Table 4, below, describes the participants' roles in the organisation.

Participant ID	Role	Responsibilities
P1	Project Manager	A project manager guides projects from conception to completion. Within the organisation, the role of this project manager is to develop, coordinate, and report on the projects that he/she manages and ensure that the projects are completed on time and within budget. The participant is a project manager/senior officer with a demonstrated history of working in the government administration industry. Skilled in business process management, analytics, and leadership.
P2, P5, P6, P7	Professional Officer	Professional Officers are employees who hold relevant tertiary qualifications and whose primary responsibilities require applying professional knowledge, experience and judgement in the development, management, and implementation of services for the organisation. These participants are data scientists concerned with implementing systems that are used to manage, store, and analyse organisational data.
P3, P9	Director	Directors within the organisation need to comply with statutory controls while performing their duties. Their duties include acting within their powers according to the organisation's constitution and using those powers only for the purposes for which they were conferred, and exercising reasonable care, skill, and diligence. These participants are responsible for driving the strategy of the organisation, more specifically as it relates to the use of data and IT within the organisation.
P4, P10	Manager	The four primary functions of managers within the organisation are planning, organising, leading, and controlling. Through these four functions, managers work to increase the efficiency and effectiveness of their employees, processes, projects, within the organisations.
P8	Information Coordinator	An information coordinator collects, analyses, processes, and distributes information according to an organisation's policies. Within the organisation the participant is responsible for analysing data statistics and market trends.
P11	Technology Officer	The technology officer oversees the development and dissemination of technology for internal and external customers, vendors, and other clients to help improve and increase business. The technology officer is also an important decision maker in the adoption process. Within the organisation the participant has huge influence on the IT strategy of the organisation and the technologies that are adopted.

Table 4: Description of the participants (Source: The researcher, 2022)

4.3. Factors affect Big Data adoption at a government organization in the Western Cape

This research question aimed to identify the factors that influence BD adoption at a government organisation in the Western Cape. While forming the themes for the research study, it was deemed important to discover how familiar the participants are with BD and if the organisation is actively adopting new technologies. The below section describes these findings.

4.3.1. Big Data defined

Within this theme, the research focused on determining how familiar the participants were with the definition of Big Data.

The participants had a good understanding of what BD is. First, they described it as, "*looking at the Organisation XYZ data, but as looking at the Organisation XYZ's data as a whole and then analysing that data with various new types of techniques that are available now like NoSQL, ……. then bringing all of this data together into one spot, like a data lake or data house ……. then analysing that data to produce dashboards, report some analytic results that will help either predicting events or describing events that has [have] happened".* [P1]

Second, participant 2 refers to the characteristics of BD: volume, variety, and velocity. The participant mentions that he thinks the definition of BD has changed over time but defines it as large volumes of data in different formats generated at high speeds by the use of technologies to host, store, manage, access, analyse and integrate the data. However, the participant does not explicitly indicate how BD can be used within Organisation XYZ. This definition aligns with the work of Gandomi and Haider, 2015 and De Mauro *et al.,* 2016, who describe BD as high volume, high velocity, or high variety information assets requiring new forms of processing to enable enhanced decision-making and insight discovery and process optimisation.

Third, participant 6 defines BD as data that cannot be processed, aggregated, or transformed in a reasonable timeframe using a single machine. [P6] The participant mentioned that the "goalposts" (to change the rules while someone is trying to make it more difficult for them) for this definition shift depending on the technology you have available and what you need to do to the data and that the BD technologies they use are mostly driven by the analytical requirements they receive. [P6] Fourth, participant 7 describes BD as data that you cannot fit in the index on a single commodity server with half a Terabyte of RAM (Random Access Memory) and refers to data such as massive image datasets or massive video datasets.[P7] This is in alignment with the sentiment of Russom (2011), which mentions that traditional data in database tables are being overtaken by the growing availability of unstructured data sources, like video, pictures, and text produced by humans.

Last, participant 9 adds some contrast to the description, stating, "*so for me big data and as this is an interview for research purposes, it's actually a buzzword that nobody fully understands. So, within our organisation even, people have looked at big data as voluminous data, meaning the size all the data take up versus actually a data lake or what you can actually do with certain types of data. So, in my view big data doesn't relate to actually size. It actually relates to them, let's call it, amalgamation, or the combination of all types of information or data. So, repositories and then from there you derive*

value, and that value can be either information or to the further extreme can become knowledge, but big data, in my mind in my opinion is actually a repository of various information and systems data that must be mined in order to derive value." [P9]

In addition, some participants provided an average description of the term and focused more on the use instead of providing a concrete definition. Responses included, for example, "*I am familiar with big data, yes I do believe that there would be use cases for big data analysis… I think it's used cases would be around where we could get to more effective decision making for our citizens in terms of service delivery."* [P3] Second, "*I'd like to think basically it's data from various sources, you know so the Organisation XYZ is made up of different departments and also our customers are external, and we liaise with external departments as well. So, I'd like to think when we refer to by big data, I think it's bringing all of those data sources, you know where it comes from all those avenues together and to me that is a what big data reference, you know and how that can be used more most effective?"* [P5]

Participants 4 and 11 reference analysing different types of data but focus more on how the organisation uses data for decision making. It is important to note that even though these definitions did not define BD, they hinted that BD use would be evaluated on the premise of use cases within the organisation and how it would benefit the organisation.

Some participants gave an unsatisfactory definition that focused more on application of data. Responses combined data from various sources [P8] like social media applications, for example Facebook, Twitter, and WhatsApp. Content was gathered [P10], but no reference is made to BD regarding the V's.

The responses related to the definition of BD indicate that most participants would provide valuable data regarding BD because they are familiar with the technology. However, Participants 8 and 10 offered unsatisfactory definitions for BD but contributed valuable insight in their responses to the remaining interview questions and were therefore not excluded from the study.

4.3.2. Technology adoption within the organisation

In this study, it was important to discover if the organisation was actively adopting new technologies and the process for adopting these new technologies. This was crucial because if the organisation is not actively adopting new technologies, the possibility of adopting BD would be minimal. New technologies could lead to new data-driven services that improve processes and enable innovative products and services, which is important for governmental organisations, and BD can create these services. The (evidence on the) adoption of technologies constitutes): i) the technology adoption process, ii) the technology adoption timeframe, iii) who makes the decision to adopt technologies, iv)

how the adoption decision is communicated, v) the big data adoption factors, vi) rated general adoption factors, and vii) the Technology-Organisation-Environment factors.

4.3.2.1. The technology adoption process

This study used the TOE framework developed by Tornatzky and Fleischer (1990) to explain an organisation's decision to adopt a technological innovation based on the technological, organisational and environmental contexts (Sam & Chatwin, 2019). This framework has been extensively used in technology adoption studies and provides a useful approach to understanding the factors influencing technology adoption in organisations. Hall and Khan (2002) define technology adoption as a choice to acquire and use an invention or innovation, and Ismail (2006) describes it as a decision to use innovation as the best course of action available.

Research findings show that most participants described a similar process for adopting new technologies. Some participants mention that the process follows similar steps as those used in project management. [P2, P4, P6, P7, P11] First, the process begins with a business requirement [P1, P6, P7, P9 and P11] in which a specific need for a new technology is described. Second, the requirement gathering process is initiated to ensure that all the requirements for the new technology are captured. [P1, P2, P4, P6, P7, P11] Krishna and Gopinath (2021) describe the requirement gathering process as a set of tasks, knowledge, and techniques required to understand the business needs and identifying solutions to an organisational problem (Krishna & Gopinath, 2021).

Third, a Project Initiation Document (PID) is created from the output of the requirements gathering process. [P1, P6, P7 and P11] PID documents include the benefits of the project to business, the risks relating to implementing the project, the products that are developed from the project, activities included in the project lifecycle, the desired quality of the products and how resources will be applied throughout the project (Bishop, 2018). The PID creation is a collaborative step between the requesting department and Information Systems and Technology (IS&T) architecture in which the requirements are documented based on the evidence from the requirements gathered step [P1, P2, P7 and P11]. To add to this, Participant 4 mentions that a committee consisting of various stakeholders would be established to track/manage the progress of the project. [P4]

Fourth, the PID is reviewed by IS&T architecture to determine if there is an existing technology in the portfolio that meets the needs for this requirement. [P9] This step creates a conditional loop in which IS&T architecture would advise if existing technology can be modified to meet the business requirement or request approval from executives (board) for the technology based on the PID. [P11]

Fifth, Participant 4 mentions that this decision must be made in compliance with the Municipal Finance Management Act rules and regulations [P4]. This step also creates a conditional loop in which the request can either be denied, additional information is requested or approved. Sixth, on approval a tender document is sent to vendors to provide a solution (technology) to the business problem [P4, P6, and P9].

Seven, based on vendor feedback, their proposals are scored based on the suitability of their suggested technology [P4, P7]. Eight, shortlisted vendors are invited to build a POC of the recommended technology. Nine, the POC is evaluated to ensure that it meets all the business requirements and put forward for acceptance. Ten, Acceptance leads to additional project tasks and ultimately the adoption of the technology, while rejection leads to the suggestion of alternative technologies and the re-evaluation of those technologies. The re-evaluation re-triggers sending the tender documents to vendors for proposal and initiates another shortlist process. Figure 9 was developed to illustrate the adoption process described above.

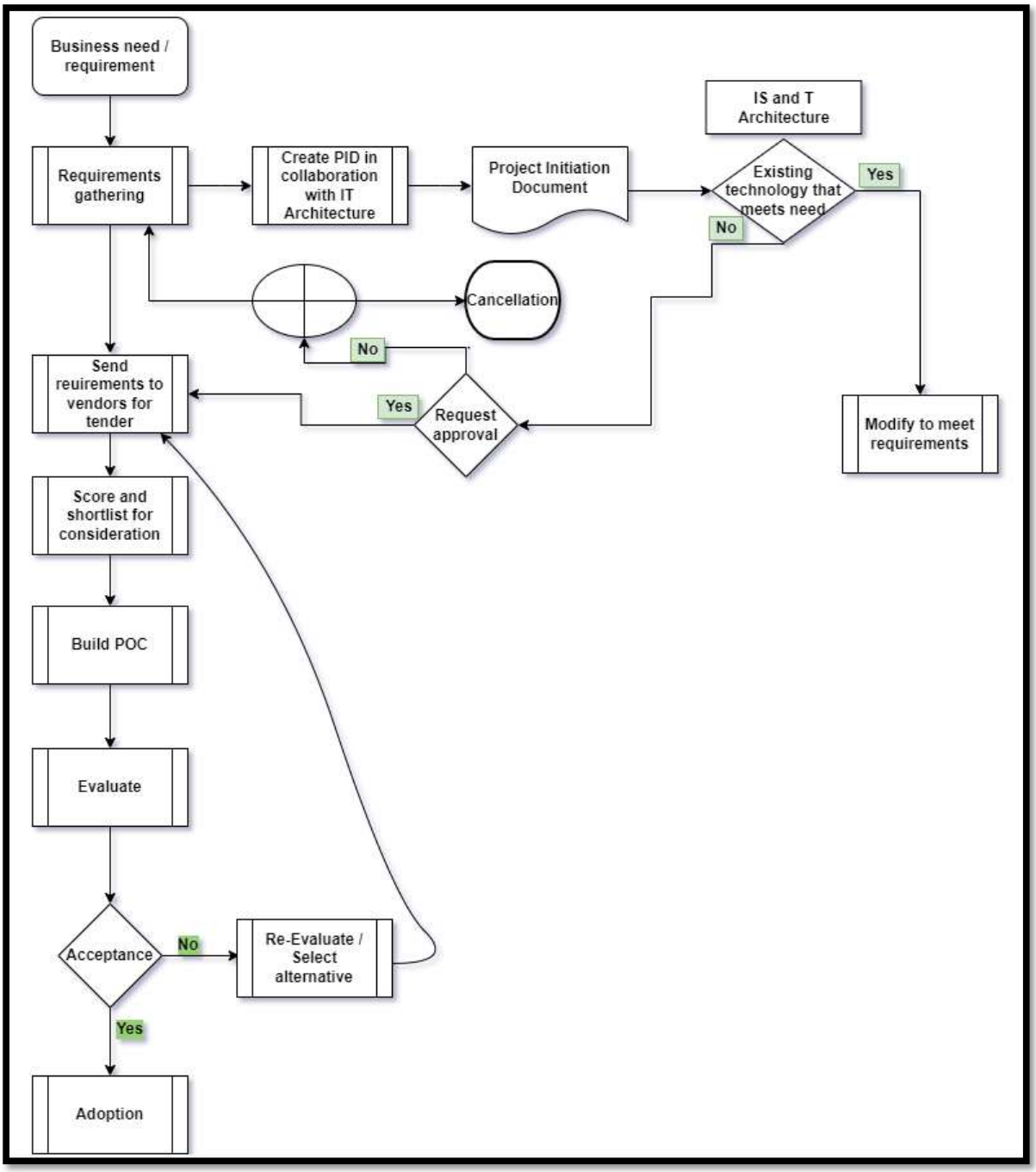


Figure 9: The organisational technology adoption process (Source: The researcher, 2022)

Corroboration for the process is provided by many participants and is best described by Participant 9, "*Typically, as an organisation, we go through what we call a project initiation phase, PID process, and in that particular process the business would put down the functional requirements, put down their actual needs analysis based on the information from business analysis done over these respective processes, or they'll then come onto IS and T, IS and T would then look at what we have in our existing portfolio that can either be used as-is or that can be adapted slightly to meet the business requirement.*

If all of those things are not meth there, we would typically allow the business to go out and find their own solution via a tender process. Obviously based on the standards that architecture and governance has defined up front relating to information systems so that that's just normal business requirements." [P9]

Another interesting aspect of the participants' responses is that the organisation is reviewing all their corporate applications (systems) through what has been noted as the CAR (Corporate Application Review) project. The response showed that the organisation is reviewing its digital landscape, confirming the adoption of new technologies. The purpose of the project was to review the organisation's digital maturity and compare it with available technologies from a market perspective and see what the best practices are for similar types of organisations. "*we are asking the market to respond to us and actually tell us what the best of breed is out there that will give us that digital life span for the next 20 years."* [P9]

The organisation also compared their future business strategies. Based on a combination of the strategy, the current landscape and where technologies are moving towards, they came up with insight into building a digital roadmap for relevant future technologies. Participant 3 was quoted as follows: "*we looked at our future business strategies and then based on a combination of our future business strategies and our current landscape, along with where technologies are moving to, we came out with insights on how we could build out a digital road map for technologies of the future."* [P3]

4.3.2.2. Adoption process alignment

Rogers (1995) shows that the process of innovation adoption is described by two major phases, namely, initiation and implementation, and is separated by an adoption decision (Bremser, 2018). The initiation phase comprises of requirement gathering and matching stages in which the requirement gathering is triggered by an organisational problem or the perceived usefulness of the innovation and the matching stage is used when the organisation's members explore the capabilities of the innovation to predict its potential for specific application scenarios (Bremser, 2018). Based on this, the organisational adoption process was separated to include the adoption stages as described by Rogers (1995). This is presented in Figure 10, as shown below.

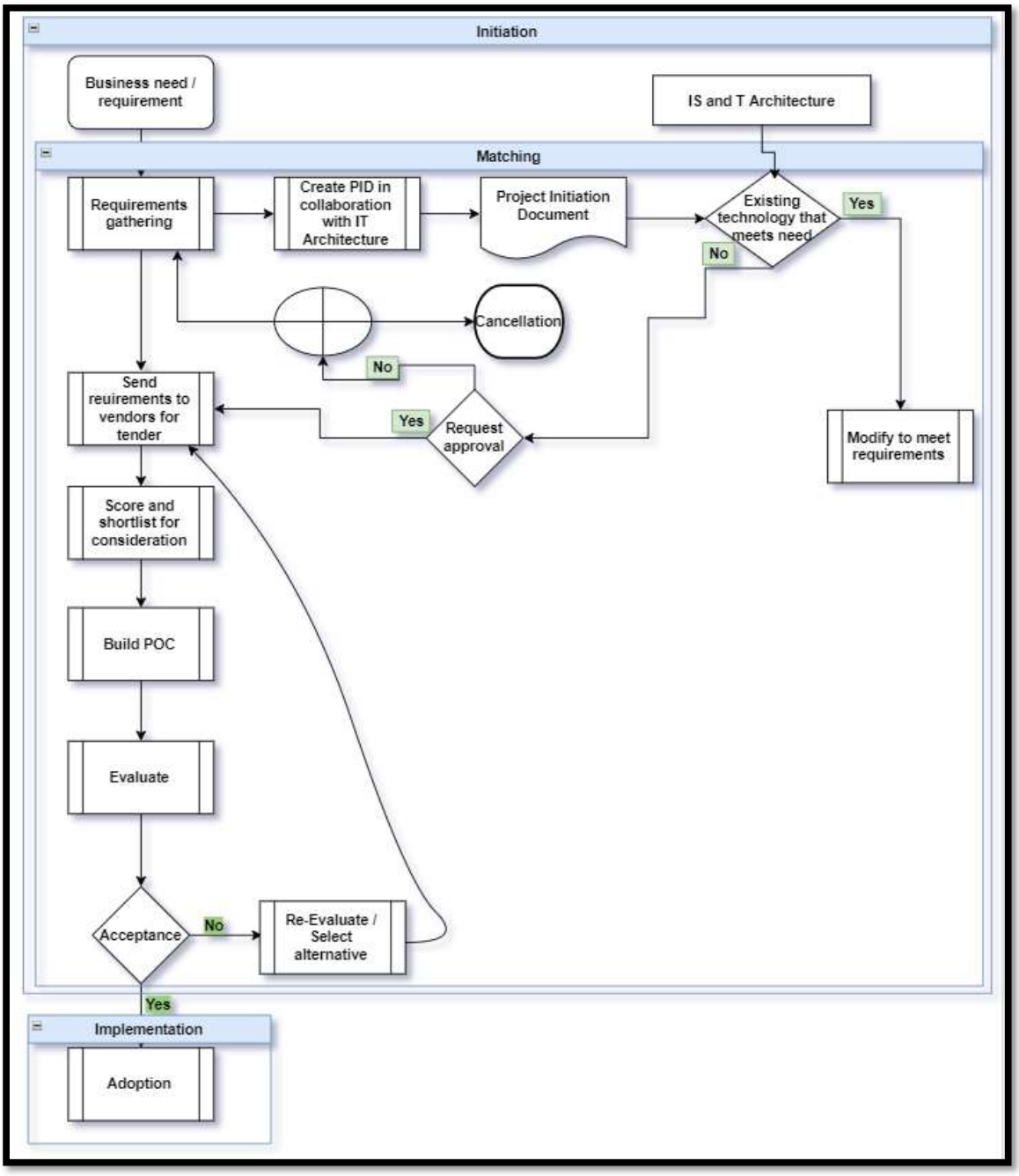


Figure 101: Organisational adoption process aligned to the innovation-decision process model of Rogers 1995 (The researcher: 2022)

4.3.2.3. The technology adoption timeframe

Evidence indicated that the technology adoption timeframe could range from six months to three years. The responses from the participants generally confirmed this. Participant 6 mentioned that it could be as fast as ten days in their department. [P6] This indicates a disparity in how technologies are adopted within different business units based on the particular need of the unit. Here is the quantitation from

Participant 6: "*From the point where we started investigating it to where a beta was deployed, it's probably about 10 days."* [P6]

Participants indicated that this disparity was because of the technology being adopted and the project's requirements. For example, enhancements to current technologies will deliver projects quicker, while new technology adoptions will have longer turnaround times.

Based on the disproportion in adoption timeframes within the organisation, different business units adopt new technologies within different timeframes. Stenberg and Nilsson (2020) suggest that this can be because of the difference between expectations and actual outcomes of the adoption process. The aforementioned authors point out that one factor that influences the adoption timeframe relates to implementation and restructuring lags, which suggests that it often takes longer than expected to fully harness the potential of new technologies. This statement supports participants' view that new organisational technologies have longer adoption timeframes.

4.3.2.4. Who makes the decision to adopt

In the process described in section 4.3.2.1, participants indicate that the adoption decision is collaborative between members of the committee started for the project. The evidence provided by Participant 5 states the following: "It's a collaboration the, the first decision is taken by the specific department, you know, the owner of the data or the owner of the systems". [P5] This often includes a Proof of Concept (POC), "Which will be a working example of their proposal" [P6], which provides the organisational leaders with the information required to base their decision on.

The following are the salient steps that constitute input for the adoption decision: i) the creation of the PID document which contains all the requirements of the technology to meet the business requirement. The PID is also used to determine if an existing technology exists in the portfolio which meets the business need, ii) approval to continue with the project, which leads to either additional requirements gathering, project cancellation or continuation, iii) the tender process, which includes sending the requirements to vendors, scoring the vendors, requesting that the identified vendors build a POC, and evaluating the POC, and iv) the committee accepting the technology. In the last step, the technology can either be accepted, which leads to adoption, or rejected, which re-initiates the tender process.

Additional information can be provided during the PID creation step in the form of a project meeting to influence the adoption decision. In addition, Participant 10 mentions that representatives from businesses will be identified and nominated to form what he refers to as a 'project steering committee'. McGrath and Whitty (2019) define a steering committee as a group or committee accountable to the organisational management for the success (or failure) of a project. In BD adoption, this is important

because steering committees possess the power (authority) to approve actions for the project on behalf of the organisation, influencing desired outcomes due to their influence within their own contributing organisation (McGrath & Whitty, 2019).

Damanpour *et al.* (2018) state that project meetings are important because they allow decision-makers to acquire additional knowledge required to make the adoption decision and reduce uncertainty. Also, organisational processes that enable cooperation and conflict resolution among groups and units, communication and interaction between members and managers, and cross-pollination of ideas across the organisation help reduce the uncertainty inherent in innovation decisions and facilitate the initiation and implementation of new programmes (Damanpour *et al.*, 2018).

Participant 4 adds that it is important for governmental organisations to obtain approval from the Mayoral Committee member that manages specific directorates in this phase because the member would be influential when obtaining approval from top management to adopt the specific technologies. [P4]

4.3.2.5. How is the adoption decision communicated

The involvement of senior management staff at various functional levels of the decision-making process promotes communication and professional discussions that contribute to the organisation's development (AlHijji *et al.,* 2018).

In this government organisation, the adoption decision is communicated through various forms, for example: i) the internal communications department, which produces and delivers messages and campaigns on behalf of executives and managers, facilitating a two-way dialogue, *"the process of communication in the Organisation XYZ is handled by a communications department"* [P1]; ii) staff newsletters which inform employees about the latest company updates, *"but when it applies across the organization we have general communications so staff newsletter"* [P11]; iii) their intranet, which is a private network used by the organisation to distribute communications exclusively to its employees, *"...process is available to the organization via the Intranet"*, [P2]; and iv) change management, which is the formal manner in which the organisation describes and implements changes within the organisation. *"So typically, through project environments with active change management strategies behind that"* [P3], *"Generally speaking, the city's quite strong on change management, so accompanying a big project, and particularly if it's a technology that's being rolled out widely there would be a change management process that would be communicated"* [P7].

Participant 1 responded that when the Executive management team makes key technology adoption decisions, such communication occurs via the organisation's communications department and staff newsletter. [P1] In addition, Participant 2 provided supplementary information which indicated that

additional information regarding the technology adopted is provided on the organisation's intranet website and that the process is the same for any technology adoption. [P2]

Participants 3, 7 and 11 indicated that a strict change control process normally accompanies the organisation's project implementation. As part of this, relevant information regarding the technology being adopted would be communicated throughout the organisation. These three participants were quoted as saying this: *"typically, through project environments with active change management strategies."* [P3] *"Generally speaking, the Organisation XYZ is quite strong on change management, so accompanying a big project, and particularly if it's a technology that's being rolled out widely there would be a change management process that would be communicated".* [P7] and *"At which point in time there is a training and change management campaign to[for] those users."* [P11]

Yue *et al.* (2019) argue that a major factor that affects change management is strategic internal communication and that, thus, initiatives which lack strategic internal communication inevitably fail. The forenamed scholars offer that through effective internal communication, change implementers provide staff with a better understanding of the content and rationale of the change. Quality internal communication reduces perceived uncertainty related to the change and decreases staff's resistance to change (Yue *et al.,* 2019).

4.4. Organisational factors for Big Data adoption

The data for this theme was collected through three interview questions. First, the researcher provided the participants with a list of factors discovered in the adoption literature, and then asked them to rate their criticality. Second, the researcher probed the participants with a general question relating to the factors that need to be investigated for technology adoption within the organisation. Last, specific factors discovered in the TOE literature were presented, and the participants were asked to provide details on how these factors would influence the adoption of BD within the organisation.

4.4.1. Rated general factors

The researcher presented the following request: *"Rate the criticality of the following factors for technology adoption."* to participants and provided a list of factors that were discovered for general technology adoption. The factors were rated from 1 to 8, with eight being the most important and one the least important. This question aimed to determine which factors the participants perceived to be important for technology adoption. However, this research question was used to setup the interviews, but did not gauge or test any construct. The data for this research question is presented in Table 6 below. The significance (presented in the total column) of each factor was determined by adding the values assigned by each participant to the factor.

Factor	A clear company strategy	Support by upper management	Talent	Training	Supporting systems and procedures	Financial budget	An organisational structure that supports multi-disciplinary projects	A sound procedure for legal and ethical issues
P1	6	6	5	4	4	5	4	5
P2	8	8	8	8	8	8	8	8
P3	1	3	4	7	6	5	2	8
P4	8	8	6	6	8	8	8	8
P5	8	8	8	8	8	8	8	8
P6	2	8	6	2	2	4	2	3
P7	4	6	7	3	5	6	4	6
P8	7	7	7	7	7	7	6	6
P9	8	4	4	4	6	7	4	5
P10	8	8	8	6	7	8	6	8
P11	6	6	6	4	4	4	4	2
Total	66	72	69	59	65	70	56	67

Table 5: Factors rated by participants (Source: The researcher, 2022)

The participants' responses indicated that support by upper management was deemed the most important factor for technology adoption, followed by financial budget and organisational talent. It is noteworthy that Participants 2, 3, 4, 9 and 11 provided ratings but no detailed descriptive content to elaborate the rating and Participants 6, 8 and 10 provided limited descriptions. Also, the response provided for *"A sound procedure for legal and ethical issues"* did not provide the research with descriptive content to elaborate on this finding. The following section describes the direct responses received from Participants 1, 5, 7 and 8.

Participants 1, 5, 7 and 8 indicate that a clear company strategy is important for adopting technologies. Participant 1 mentions that the strategy is well defined but believes there is still room for growth [P1], while Participant 7 states that *"a clear strategy becomes [an] enormously complicated thing because we have so many objectives that we're constantly balancing."* [P7] and Participant 10 believes that all the adoption factors are directly influenced by organisational strategy. [P10]

Participant 6 mentions that the organisation's maturity influences these factors: *"So for instance in the very beginning, you know just getting permission to procure servers, support by upper management*

that had to be an 8. For instance, whereas right now support for upper management is not that important because we've established the value of what we're doing, and so you know we don't need to have very high upper management support to continue driving out big data options". [P8] Furthermore, Participant 7 mentions that if the organisation wants to start using technology effectively, it has to feature in decision-making. [P7] For it to feature in decision-making, senior leadership just needs to care about that.

The general view on talent from participants is that the organisation is well resourced, and if specific skills are required, the organisation will go to the market to obtain those. Participant 7 believes it would be challenging for the new technologies to be adopted if the appropriate talent were unavailable. [P7] In addition, participants believe that training is important and also interdependent on talent. The view from Participant 8 is that without the talent and the training, you will not be able to support systems and procedures. [P8] In contrast, Participant 1 states that the organisation has identified training as key, but training programs have not caught up yet. [P1]

Participant responses indicated that the organisation would define support systems and procedures and build support mechanisms around the technology during the adoption process. The participants also pointed out that the organisation is well resourced. However, the organisation's adoption process (as described earlier in the chapter) needs to be adhered to adopt new technologies.

The rating for the factor "an organizational structure that supports multi-disciplinary projects" had the lowest score. Responses indicated that participants were not clear on how the organisation approached this factor. As support, Participant 7 provided the following statement: *"it's quite a difficult thing to reason about in the context of the Organisation XYZ".* [P7]

4.4.2. Technology-Organisation-Environment factors

The participants were asked: *"In your opinion how would the following affect Big Data adoption within the organisation?"* A list of factors from the TOE framework was provided, and the participants were asked to describe how each factor will influence the BD adoption process. The next section presents data reliability and suitability, IT infrastructure and data security and privacy as the factors for the technology context, top management, organisational size, finance, organisational process and organisational strategy as the organisation context, and competition, customers, and vendors as the environmental context of the TOE framework. Additionally, the participants were asked to extend the list of factors by including additional factors that they deemed important for BD adoption.

4.4.3. Technological context

4.4.3.1. Integration

Participant 8 mentions that many organisations' systems are integrated. Therefore, a reliable system integration should always be considered when adopting new technologies. *"Obviously I think one of the factors that you that that we always should keep in mind is that it's the systems integration in within the Organisation XYZ"* [P8].

Furthermore, Almoqren and Altayar (2016) and Sivarajah *et al.* (2017) state that a significant impediment to the adoption of BD relates to data not being pooled (integrated) for the entire organisation. Attard *et al.* (2016) mention that the purpose of integration is to enrich existing data with the intention to use it in unique ways. Based on the statement from Participant 9 and the claims from Almoqren and Altayar (2016) and Sivarajah *et al.* (2017), integration was included as one of the factors that affect BD adoption for this research study.

4.4.3.2. Data security and privacy

Sivarajah *et al.* (2017) believe organisations need to safeguard data through robust security infrastructure that enables employees of each division to access only the data they are required to have access to. Moreover, standard privacy laws must govern the use of data, and strict adherence to security regulations must be applied to the data (Sivarajah *et al.,* 2017). Abbasi *et al.* (2016) believe that security breaches would be more prevalent because employees have access to more sensitive data and there are organisational security behavioural challenges relating to compliance and insider threats.

Also, the total volume of data collected and created in a typical BD environment is one factor that raises security issues (Salleh & Janczewski, 2016). Consequently, new and improved security tools and mechanisms are required to effectively protect BD environments (Salleh & Janczewski, 2016).

Besides technology and infrastructure, security aspects should be evaluated from the organisational and environmental perspectives (Salleh *et al.*, 2015). Participants 1, 6, 7, 8 and 9 mention that BD security and privacy should be evaluated in the POPI (Protection of Personal Information) Act. The following constitute the purpose of this act: i) safeguard personal information while adhering to privacy rights and protecting the flow of information; ii) regulate how personal information may be processed; iii) provide ways to protect against the use of personal information that is inconsistent with the act; and iv) creating an Information Regulator to ensure the enforcement of the rights protected by this act (Staunton et al., 2021).

Regarding the act, first, Participant 1 states that data security and privacy are a big issue for BD adoption and that the organisation is scrambling to get aligned. [P1] The participant mentions that this might not be a top priority for the organisation. [P1] This contradicts Salleh *et al.* (2015), who claim that organisations should prioritise security-readiness assessments in adopting data-intensive technologies such as BD because of the massive amounts of data being handled by organisations. Likewise, Participant 9 claims that the act brings additional laws and policies that need to be adhered to for the adoption of BD. [P9]

Furthermore, Sun *et al.* (2016) state that when an organisation introduces BD, there arise additional security challenges. Participant 9 agrees and mentions that the organisation has been averse to adopting specific technologies because of security concerns. *"Our organisation is behind the curve when it comes to technology such as cloud adoption and the like… the Organisation XYZ has been cloud averse up until now, purely because of a security, let's say the assumptions and the perceptions around security in the cloud"* [P9].

Second, Participant 2 mentions that any data management initiatives should be important. [P2] Third, Participants 6, 7, 8 and 10 advise that security and privacy policies are set out by the IS&T department or on a national level. These need to be considered when technologies are adopted. The organisation is fortunate to have in-house information officers and legal advisors who can guide with respect to these matters. The concerned participants also mention that the use of the data must be clearly defined to avoid reprimand from legal practitioners within the organisation.

Additionally, the comments from the remaining participants relating to security and privacy do not contribute any noteworthy information. These comments are described as follows: *"It will get accessed and is not seen as a factor that will stop the adoption "*[P3], *"That as well, same thing critical"* [P6] and *"data security and privacy, and that's never been an issue for us. It doesn't stop us from implementing big data."* [P11]

4.4.3.3. Data reliability and suitability

BD involves collecting and integrating data from various sources (Walker & Brown, 2019). Decision-making is affected by the quality of the data; hence, the more relevant, timely, reliable and accurate the data, the more it positively affects decision-making (Thabet & Soomro, 2015; Walker & Brown, 2019). Some participants elaborated on the statement from Walker and Brown (2019) and provided the following as evidence:

First, Participant 1 mentions that there are some concerns regarding the quality of data and that getting the data into a reliable, suitable state will require organisational change. [P1] The participant mentions that departments want data to make the next decision and are not concerned about the data quality. [P1]

Second, Participant 2 claims that reliability and suitability of data reveal two issues that influence the adoption process. For one, the data is not accurate, and two, the data is not what you are looking for. [P2] It is not related to the problem you are investigating or the issue under investigation. [P2] While Participants 5, 6, 8, 9 and 10 mention that reliable, accurate data is very important because you would not get value from the data and build a good reputation or trust based on your data. [P5, P6. P8, P9 and P10] The response from Participants 6 and 7 is noteworthy: *"that that obviously affects it, you can't even get going unless you have confidence that you have a ground truth …. there is very little doubt that they are ground truths, and so in [the] establishment of data provenance is extremely important."* [P6], *"for a governmental organisation or a local government organisation, I don't know how we would get started without having a high degree of confidence that the data was robust."* [P7]

Third, Participant 3 believes that data reliability and suitability is a factor that will get assessed during the adoption process and should not have an effect on adoption. [P3]

Fourth, Participants 4 and 11 mention that the quality of data is influenced by the business process that generates the data. [P4 and P11] The direct citation from Participant 11 is as follows: *"if you do generate data, but it's off a bad business process, then the data is not reliable. And how do you make sense of bad data? So, for me a huge thing is getting the transactional systems right first, so they generate good quality data or else you can't do any good analytics out of it. And when I talk about good quality data, I'm not just talking about it being accurate, but also it [is] being meaningful and the only way you get [to be] meaningful data in a transactional system is if you have well designed business processes."* [P11]

Kwon *et al.* (2014) encapsulate the evidence provided by participants by stating that there may be a causal relationship between better corporate management practices when data quality provides individuals with more confidence in data-driven decision making. Hence, the heightened quality of corporate data could be a positive force in shaping an organisational culture that encourages internal and external data used for operational and strategic decision-making. This becomes a virtuous force in furthering the adoption of new data-related IT capability (for example, BD) and sustaining the organisation's competitiveness and growth (Kwon *et al.,* 2014). With the statement of Walker and Brown (2019), Kwon *et al.* (2014) and the research participants, the researcher can deduce that data quality (reliability and suitability) would influence the adoption of BD for better decision-making.

4.4.3.4. IT infrastructure

Defitri *et al.* (2020) state that IT infrastructure significantly influences the adoption process and bears a significant cost sometimes. However, this factor needs to be addressed because it can provide long-term economic effects, especially because infrastructure is required for capturing, storing, organising and analysing BD (Defitri *et al.,* 2020; Han *et al.,* 2014). Kalema and Mokgadi (2017) argue that infrastructure necessitates sufficient transmission and storage capacities and computing power. Therefore, without the availability of infrastructure, the adoption of BD will be considered an unrealistic program (Defitri *et al.,* 2020b). This sentiment is shared by Participants 5, 6, 7, 8 and 10: "*it's very important that you need to have all your hardware, software and infrastructure in place. Otherwise, you would not be able to disseminate the data or to serve the data towards your customers if it's internal or external"* [P8] and "*I mean if our infrastructure is not adequate to accommodate such. Again, it does not make sense to adopt that."* [P10]

Zulkarnain *et al.* (2019) asserts that one of the most critical issues for adopting BD in government is infrastructure. They offer that government organisations need to integrate IT infrastructure resources, especially datasets, to enable BD (Zulkarnain *et al.,* 2019). Participant 2 agrees with this statement and mentions: *"I think you need IT infrastructure that allows for integration, that allows for access to data. So, a scalable and agile infrastructure, flexible dynamic infrastructure is important as the business needs to grow".* [P2] They continue to mention that government organisations need to have clear standards for all departments relating to IT infrastructure.

Participants 3 and 4 recommend that IT infrastructure gets assessed during the adoption project, a process it has to undergo, just as any other technology which is adopted. This indicates that Participants 3 and 4 believe that strict adherence to the adoption process is vital for the adoption of BD within Organisation XYZ.

Participants 1, 9 and 11 imply that IT infrastructure should be a lesser issue for the adoption of BD, because IT infrastructure within the organisation is in a "good" state to accommodate the adoption of BD. [P1, P9 and P11] This agrees with Kalema and Mokgadi (2017), who write that BD has a high dependency on current IT infrastructure and on an organisation's ability to access and integrate data received from various sources and in different formats. Kalema and Mokgadi (2017) also state that current organisational systems need to have good potential for handling such data, but new systems also need to be acquired (Kalema & Mokgadi, 2017).

4.4.4. Organisational context

4.4.4.1. Skill and talent

Participant 2 was quoted verbatim, stating: *"I think there is always the people perspective when it comes to technology implementations. It's always important to ensure that you're resourcing and change management is catered for specifically with the technology implementations. I think sometimes we focus on the technology, but the people and process aspect of implementation is quite critical and that should be really factored in from the beginning."* [P2], while Participant 11 claims that having talented employees to address critical business issues is vital., [P11]

The sentiment of Participants 2 and 11 is shared by Argawal (2016) and Awa *et al.* (2016). Agrawal (2016) claims that technical IT skills are vital in analysing, designing, and implementing changed business processes. Awa *et al.* (2016) mention that technological competence stretches further than physical assets. Skilled individuals generate competitive advantages for innovators since skills and know-how complement physical assets and are more difficult to imitate by competitors (Awa *et al.,* 2016). Proficient individuals understand the technology's usefulness and use their experience to turn the complexities of technology into mental effortlessness (Awa *et al.,* 2016).

4.4.4.2. Financial support (Budget)

Participant 10 believes that the financial budget should be considered for organisational technology adoption. Sun *et al.* (2018) state that BD adoption costs relate to an organisation's expenses to sustain BD usage and future scalability. Also, substantial financial investment is needed for organisations to adopt an innovation that might include an extensive investment in both hardware and software and to shoulder the cost of hiring new employees as well as training the existing staff (Sun *et al.*, 2018).

Participant 10 is quoted as follows: *"Do we have the, you know the budget for that? And if we do, what will be the cost linked to the maintenance of that technology"* [P10]. This statement is supported by Sun *et al.* (2018), who assert that the investment's initial and overall budget is affected by an organisation's intention to adopt BD and the cost of adopting BD. Therefore, the adoption of BD might be a severe challenge for many organisations.

Kalema and Mokgadi (2017) state that the availability of financial resources influences an organisation's readiness for BD adoption. While Nam *et al.* (2015) highlights that financial support for new technology innovation, implementing any subsequent enhancements, and ongoing expenses are very important. Participants 1, 2, 3, 8 and 11 suggest that financial support is not a concern for BD adoption because if you have support from top management, the budget will be made available if it makes business sense to adopt BD depending on the organisational priorities. [P1, P2, P3, P8 and P11]

Participants 5 and 10 mention that financial support is critical for adopting BD. [P5 and P10] Sam and Chatwin (2019) agree and claim that BD adoption requires substantial financial support. Kalema and Mokgadi (2017) indicate that because BD deals with increasing user requirements that exceed the traditional approach to analytics, organisations might require more resources and tools that will also have financial impact. In addition, acquiring new technological devices, applications, tools, and experienced data scientists necessitate the allocation of financial resources (Kalema and Mokgadi, 2017; Motau, 2016).

Alqahtani (2016) identifies a further financial constraint as the lack of appropriate planning and budgeting, to provide and maintain system software and hardware as well as training in order to develop and enhance the technology adopted. Participants 6 and 7 believe that some financial constraints can be alleviated by applying appropriate skills. [P6 and P7] Last, Participant 9 believes that understanding organisational processes is more important and cost-effective than adopting BD. [P9]

4.4.4.3 Organisational strategy

The availability of quality, accessibility data, and well-defined rules for collecting, storing, and analysing data is critical for creating a supportive organisational strategy for BD adoption (Kumar and Krishnamoorthy, 2020). In general, organisational strategies entail setting clear objectives that organisations need to address through their resources, leading to successful adoption projects (Alqahtani, 2016). Kalema and Mokgadi (2017) concur and suggest that sound organisational strategies positively influence an organisation's readiness for BD adoption.

Argawal (2016) also writes that BD is an extensive innovation that can radically change any organisation's strategic planning and operational processes. However, doing so necessitates considerable IT and managerial capability. Technologies that enables radical improvements to the organisation requires substantial effort in matching changes to organisational structures, routines, and policies, consequently changing organisational and process adaptations (Agrawal, 2016).

Participants 3 and 7 provided responses related to organisational strategy. Participant 3 mentions that strategic business needs must be considered and technologies of the future in terms of existing offerings and bringing together those strategies and business needs. *"What are the technologies of the future in terms of offerings that exist out there and bringing together the strategies and those business needs"* [P3]. Participant 7 identifies clarity of purpose (clearly stated purpose) as a factor that needs to be considered. [P7]

First, as discovered in the data collected, Participant 1 mentions that the organisation has a data strategy that mentions the adoption of BD. [P1] However, the participant is unclear if BD adoption is specifically

mentioned as a major goal for the organisation. [P1] Participant 6 mentions that the organisation invests a lot of resources into generating a good data strategy, but he is ambivalent about whether that is of any good use to the organisation. [P6] He supports the statement with the following quote, *"Mostly because you know, data literacy is very low and data maturity is very low, so we don't know what we're trying to get to as a broader organisation".* [P6]

Second, Participants 2, 4, 5, 8, and 10 mention that organisational strategy is important for BD adoption because it enables the technology. [P2, P4, P5, P8, and P10] Kumar and Krishnamoorthy (2020) state that management's clear vision towards adopting BD is critical for aligning the BD adoption strategy with the business requirements. Lamba and Dubey (2015) believe that the organisations BD strategy enables the discovery of knowledge and takes the required steps to drive knowledge into organisational processes that create new innovative products and services.

Third, Participants 7 and 9 believe that coherence is probably the most important aspect of an organisation's strategy that drives BD adoption. [P7 and P9] Participant 9 is quoted as saying, *"you must have a common direction, and everybody need to know where they're going to"*. [P9] While participant 7 mentions in his response that, *"there must be a shared understanding in the organisation of what we are trying to do"*. [P7] Rahman and Aydin (2019) and Lamba and Dubey (2015) agree and mention that departmental strategies and goals should be aligned with organisational strategy and goals because these strategies and goals cannot be addressed in isolation.

Last, Participant 11 believes that organisation strategy is not important for BD adoption. [P11] The participant provides the following perceptive for the claim *"Organisational strategy is not important, every organisation generates data points. Whether you are in food and beverage, government, insurance or financial. So, your organisational strategy doesn't have much to do with big data, everybody is using data. I think we famously say we are in the knowledge economy, regardless of our industry. And if you believe in the Fourth Revolution hype, then every business is a digital business. So, your organisational strategy doesn't really matter because we all have to use IT to execute that organisational strategy ".* [P11]

4.4.4.4 Organisational process

Kumar and Krishnamoorthy (2020) state that organisations are less likely to use any technology if they find it problematic to both comprehend and integrate with their organisational processes. They refer to this as a 'complexity' factor. Participant 1 perceives organisational process directed explicitly at BD as immature even though the organisation has other data processes. [P1] However, they are trying to address the lack of BD processes. [P1] The participant mentions that data management happens in silos,

not at an enterprise level, and that the organisation is dealing with security, privacy, and infrastructure matters to put organisational processes in place that deal with BD. [P1]

Nam *et al.* (2015) describes organisational processes in terms of assimilation and describes them as the extent to which technology diffuses across organisational work processes and becomes routinised in the activities associated with those processes. Participants 2, 4, 5, 8 and 10 agree and indicate that the organisational process should stem from organisational strategy to influence BD adoption. [P2, P4, P5, P8 and P10]

Choi *et al.* (2018) argue that since environmental factors are complex for an organisation to control, it is perceived as critical for organisational process innovation. Participant 6 agrees and mentions that current organisational processes are almost by definition something that cannot accommodate existing organisational hierarchy and is probably incompatible with existing processes because BD did not exist when the processes were created. [P6] The participant claims that this creates resistance because established processes always conflict with something that existed before. [P6] Therefore, when organisational processes are redesigned, top management needs to be there and *"in your court"*, to advocate for tolerating the tension until you can show value. [P6] The participant continues: *"I think that these are things that make the journey perhaps harder, it might mean that your journey has more conflict than it needs to have, but I don't think that you can create a process without going through the conflict first to force the process change."* [P6]

Participants 9 and 11 also mention that organisational processes??? change should be driven by top management with the focus on using data for decision-making. [P9 and P11] Participant 3 believes organisational processes do not play a role in BD adoption, [P3] while Participant 7 says that it is difficult to link organisational processes to BD adoption. [P7] However, the same participant believes that it is important to accommodate the new way of working when BD is adopted.

4.4.4.5 Organisational size

Malak (2017) states that organisational size affects the adoption of new technologies and has, therefore, gained significant consideration as the main factor in the success of adopting technological innovations. In the responses from the participants, it was discovered that various participants consider the organisation "big" because the organisation has around 30000 employees.

Agrawal (2015), Malak (2017) and Motau (2016) claim that one reason for the relationship between IT adoption and organisational size is that larger organisations have more resources (for example, technical, financial, and personnel) and can assign more resources to the adoption of new technologies

and absorb more risk. Participants 2, 7 and 10 support this argument by claiming that organisational size influences the resources available for BD adoption projects. [P2, P7 and P10]

Participants 5, 6, 9 and 11 describe the organisation's size in terms of its complexity. They provide evidence to this in terms of, first, the number of systems the organisation uses. [P5, P6, P9 and P11] If the same critical business systems are used throughout the organisation, it reduces the organisation's complexity required to integrate BD with many systems. Second, these four participants mention that the organisational structure is complex because of the various levels of reporting within the organisation. Dooley (2002) provides a supporting narrative for both the first and second annotations by stating that the ability of organisational entities to connect depends on the number of entities and their diversity because it is easier to make connections between similar entities as opposed to larger dissimilar ones.

4.4.4.6 Top management

Top management support refers to top management's role in influencing the adoption decision in organisations (Al Mudawi *et al.,* 2019). Hence, top management support refers to the role of top managers in supporting decision-making that goes into the adoption and implementation of information technology to develop the organisation with technological investments (Al Mudawi *et al.,* 2019). Therefore, top management is crucial for any adoption project because they determine its approval (Almoqren and Altayar, 2016; Çaldağ *et al.,* 2019; Al Mudawi *et al.,* 2019). When top management views technological investments as positive, it encourages employees to adopt the new technology to improve organisational effectiveness (Al Mudawi *et al.,* 2019).

Various participants share the above statements. For instance, Participant 1 believes that top management is critical to supporting and driving adoption, which goes hand in hand with financial support. [P1] Participant 6 mentions that *"top management is absolutely critical in getting going, you need to have strong top management support because BD is a new way of doing things"*. [P6] In addition, Participants 5 and 11 claim that if top management is not on board, the BD adoption project would be doomed from the start. [P5 and P11] Participant 8 states that buy-in is required from top management and they need to be confident that BD adoption provides a means for either better decision-making or future planning. [P8] The argument confirms the assertion of Çaldağ *et al.* (2019) that the success of technology adoption initiatives cannot be achieved without top management support, the same way they cannot be achieved without establishing the organisation's mission with the aim of innovation adoption.

Participant 7 believes that top management would positively influence BD adoption. [P7] If you have top management that is very oriented toward results and performance, they will be quite explicitly

supportive of data-driven, decision-making adoption projects. [P7] This is supported by the claim of Participant 10 who believes that BD adoption would provide top management with the means to understand data more easily, helping them to decide more easily with the supporting information that BD provides. [P10] And, if BD allows for better data-driven results, they will be more open to adoption. The participant claims that, in contrast, they can put a serious spanner in the works if they are overly concerned about things like compliance and, as a result, impede the ability to experiment and try new things. [P10]

Adnan *et al.* (2021) suggest that top management should be innovative to assist in the adoption process. The aforesaid describe the innovativeness as the willingness of managers to adopt new technologies, bring new experiences to the organisation and develop creative processes that will benefit the organisation and promote organisational performance. However, Participant 9 believes that top management is less influential in the adoption process: *"earlier I scored top management quite low as to its effectiveness and criticality on adoption, and that is merely because they themselves may not actually understand the nitty-gritty, and actually what data means and how to actually translate it. So, they are just the recipients at the end of the day in my view. But they do have the authority to push it down and actually understanding and influencing outcomes and adoptions."* [P9] In addition, Participant 1 believes that some top managers might be interested in BD adoption, but top management would be more concerned with service-related issues. [P1]

Participant 3 believes that the adoption of BD should be reviewed based on a specific business case related to the organisation's strategic goals. [P3] Adnan *et al.* (2021) mention that a strategic goal is defined as an organisation's strategic objectives relating to technology adoption initiatives. Previous research also indicates that the adoption decision is commonly influenced by the organisation's goals (Adnan *et al.,* 2021). Hence, the alignment between the goals of the organisation and the impact resulting from the adoption initiatives is an important aspect to consider with respect to the outcome of technology adoption projects (Adnan *et al.,* 2021).

4.4.4.7 Policy and legal

Participant 4 identified policies as a factor that needs to be considered for the adoption of new technologies. The following is a quotation from Participant 4: "*First, we need to comply with MFMA rules. In other words, your MFM is your municipal Finance Management act, were you not allowed to where they allow you only a certain amount of time period and to use certain service providers"* [P4]. Tomar *et al.* (2016) state that creating a data-driven society in the public sector requires a balance between protection against data misuse and data sharing regulation. Relevant open data policies and inter-government data exchange policies are key objectives for exploring BD regulations (Tomar *et al.,* 2016). It is also in this regard, thus, that Attard *et al.* (2016) point out that policies become a critical

factor because of ambiguity and out-of-date policies.

Zhang (2018) argues that government organisations should timeously introduce better rules and regulations for applying BD, as well as strengthen the legal aspects because citizens are paying more attention to privacy. Tomar *et al.* (2016) add that policies should be consistent with legal priorities, adjustable to legal and administrative frameworks, and feasible within capacity constraints.

4.4.4.8 Service delivery

One aspect of the data obtained relates to being service orientated and understanding the needs of citizens. Participant 5 explains that the organisation is a customer-facing organisation, and technology should not be adopted for the sake of staying abreast of technological advancements. [P5] In contrast, Participant 10 explains that whatever technology the organisation adopts, services for citizens should always be the highest priority. [P10] The organisation should rather try to understand how the technology will benefit society and the organisation. Participant 10 is quoted as follows: "*basically, to understand the needs of our customers… we need to understand who we serve and do our customers have access to that … Do they have even the ability to transact and talk to us … remotely or electronically*" [P5].

Kalema and Mokgadi (2017) explain that government organisations should endeavour to provide real-time services to citizens since more citizens are becoming IT literate and, thus, requesting more services online. This could cause a rise in service requests and the amount of data that requires processing (Kalema & Mokgadi, 2017). By this token, organisations will be required to use better analytics tools to serve these requests in the shortest time possible (Kalema & Mokgadi, 2017).

Anshari *et al.* (2018) state that organisations can use BD to reveal the hidden value of large amounts of data they possess. Anshari *et al.* (2018) also agree with Participants 5 and 10 when they state that by using BD analytical processes, large amounts of data can be converted into valuable information which can be used to understand customers and also enhance effective service delivery.

4.4.5 Environmental context

4.4.5.1 Competition

Empirical studies recognise that competition can pressure organisations to adopt an innovation (Ismail & Mokhtar, 2013). Kumar and Krishnamoorthy (2020) describe competitive pressure as the influence that an organisation experiences in the environment that it is trading to use technology to maintain or increase competitiveness. They claim that organisations with higher competition for market share,

revenue, market growth, and product development are more steered towards adopting technologies for business analytics.

Ismail and Mokhtar (2013) also describe competitive pressure as the extent to which an organisation perceives itself as threatened by its counterparts. Participants 1, 2, 5 and 11 believe that the organisation does not have any competition: *"The organisation does not have competition because of the nature of the services it provides."* [P2], *"I don't believe there is competition within the organisation, more collaboration and cooperation and working together with these departments."* [P5] and *"There is no competition in the public sector. You are a captive client. The only thing that you can do as a citizen or business is leave and go to city."* [P11]

Another interesting discovery in the data is the view from Participants 7, 8, 9 and 10. They claim that competition for the organisation relates to competing objectives because departments within the organisation could bid for the same funds to meet their objectives. Participant 7 argues that this type of tension needs to exist within the organisation to ensure that it adopts technologies based on its strategic goals. Last, Participant 9 mentions that competition for the organisation relates to the attractiveness of the area that the organisation governs. [P9]

4.4.5.2 Customers

Kumar and Krishnamoorthy (2020) identify customer orientation as a driver that directs organisations towards satisfying customer needs. Customer orientation pressure has been found to drive BD adoption in some adopter organisations (Kumar & Krishnamoorthy, 2020). Agrawal (2015) suggests that customer orientation might significantly impact BD adoption since BD can assist in shorter production cycles and radically improve customer services.

The data collected by the researcher indicates that all the participants believe that BD adoption will have a beneficial influence on the customers of their organisation. The general sentiment from participants is that the organisation's main goal is to provide services to its customers, and that through BD, these services will be improved. Participant 1 mentions that BD will enable the organisation to manage tariffs better, alleviate some of the citizens' social stresses, and provide a means to better forecast based on population statistics. [P1]

Alqahtani (2016) states that the diffusion and computerisation of services have helped increase customer satisfaction and also led to greater awareness of service improvement. Participant 2 supports this finding, saying: *"For us, it's about service delivery, you know. And if we can utilise big data to improve service delivery, then that will have a great effect on the citizens, and that's essentially what we truly after, so ensuring that our citizens needs are catered for and if we can leverage of big data to*

do this, then it's a very important factor. Because I mean at the end of the day, that's our mandate, service delivery". [P2]

Participants 4, 6 and 7 support the adoption of BD. However, they suggest that the technology should be introduced in a manner that does not cause friction for customers. [P4, P6 and P7] Instead, customers should be able to identify the benefits that the technology provides.

4.4.5.3 Vendors

Salleh *et al.* (2015) points out that there is inherent security and privacy risk when BD initiatives are outsourced to vendors. The aforesaid scholars state that since BD is relatively new, most organisations are still without the capabilities to build and maintain an in-house BD environment. This creates a need for organisations to resort to outsourcing practices for their BD environment or part of it (Salleh *et al.,* 2015).

Matua (2016) states that several vendors offer BD analytics solutions that organisations adopt. However, when an organisation creates a business case for BD adoption, they fail to include support for future business requirements that a vendor often provides. Some participants indicated that vendors either do not impact the adoption of BD within the organisation or they will negatively affect the adoption thereof. The most relevant support for this claim is provided by Participant 11, who claims that *"vendors do influence the public sector, but in a bad way. They are excellent at hyping up old technologies and trying to convince top management to buy from them." [P11]*

4.5. Challenges and benefits of implementing big data in the public sector

The first research sub-question was: "*What are the challenges and benefits of implementing BD in the public sector?"*

The results from the data analysis showed that the majority (over eight-two percent) of the participants believed that there are major benefits of using BD in this government organisation. Nine percent could not detail any benefits and another nine percent did not envision any benefits of using BD at this government organisation.

Furthermore, a participant that were unsure of the benefits of BD provided the following comment to corroborate: *"I'm not sure that the Organisation XYZ is just there yet in terms of being able to draw those kinds of benefits, …. even though I'm a technocrat, I'm not always sure that one should be spending the amount of money on those types of technologies if you are not sure on the return on*

investment for those types of technologies" [P1], while participant 11 believe there are no benefits and mention that skilled analysist is more important than acquiring BD as a technology. [P11]

4.5.1. Benefits

Some of the notable benefits of BD adoption included data integration, better decision making, improved service delivery and operational efficiency. Data integration and better decision making were believed to be beneficial by fifty-four percent of the participants, while improved service delivery was rated as beneficial by thirty-six percent.

First, responses from Participant 6, 7 and 10 indicated that BD benefits are seen when BD is used to construct a common view of reality though extracting data from coherent datasets of fine-grained data, then creating common views, to discover the facts on what is happening as well as the causes that can lead to forming the basis for an operational response plan. [P6. P7 and P10] Sarker *et al.* (2018) agrees with this claim and argue that BD enables fast error-free policymaking through information support systems that act as a tool for solving complex societal issues for government reform.

Second, Participant 8 claims that BD will promote service delivery, which is one of the main goals for government organisations. [P8] The response is noted as follows: *"first of all, our organisation is there for service delivery, so I think it improves service delivery. Not only for our external customers, but also within the OrganisationXYZ within different departments. Because the different departments rely on each other for that as well."* [P8]

Iftikhar *et al.* (2020) reinforce the claims by Participants 6, 7, 8 and 10 by stating that organisations use fine-grained data to create various opportunities in the public sector, such as efficient public service delivery, enablement of data-driven decision making for policymakers, enhancement of country digital economy, creation of new jobs for the youth, and promotion of civic participation to define and improve public policies.

Lastly, Participant 4 and 9 indicated that they envisage the benefits of BD as making better, quicker, and efficient decisions and improving productivity of staff. Pencheva *et al.* (2018) agree, adding that BD provides an increase in the accuracy, efficiency, and speed of processes and that governments can unlock vast amounts of unstructured data, thereby supplementing traditional decision techniques. Furthermore, the aforementioned state that BD has the ability to improve efficiency and effectiveness in the delivery of public services, thereby delivering savings and boosting productivity.

4.5.2. Challenges

The participants in this study agree with Rajaraman (2016) and Tomar *et al.* (2016) by stating that an important challenge for BD adoption is related to a skill (talent) shortage. They suggest that BD analytical skills should be more widespread across the organisation and that, sometimes, data analytical teams must be established within departments. Participant 6 provided the following statement to substantiate why BD analytical skills is a challenge in his opinion: *"Almost every analyst that I've ever encountered, and it's not just in the Organisation XYZ but everywhere, is accustomed to consuming reports and joining across reports in order to generate some new report. And all of that is kind of operating on the surface of reports. When I think about Big Data, it's about getting to the transactional stuff underneath the reports and being able to summarise or to aggregate from that very low-level data on the fly, to be able to sort of makeup and generate rapid insights. So just that that mind shift from dropping off, dropping from essentially reports consumption and wrangling into the generating of reports off of transactional data is a massive shift that's required, and there's very little skill probably globally for that, so I'd say that's probably the biggest fundamental problem that I can go on at length about". [P6]* Bremser (2018) also agrees by stating that organisational aspects, such as lack of analytical skills, are an obstacle for BD adoption.

Another important challenge mentioned by participants is removing the 'silo mentality' (a mindset present when certain departments or sectors do not wish to share information with others in the same company) from employees within the organisation, to ensure that data is sufficiently integrated within systems and departments. It is also claimed that silos result in different perspectives or views when same data is analysed. Furthermore, Chatfield *et al.* (2015) argue that government organisation infrastructure was historically built as department-specific silos, which resulted in fragmented structures, costly duplication, unnecessary complexity, an inability to share and collaborate, and an increased vulnerability to security threats. Subsequently, these 'silo'd' systems are rapidly becoming obsolete legacy systems that are also expensive and difficult to maintain (Chatfield *et al.,* 2015). In addition, Participant 2 states, *"I would say it is more related to change…. Changing the way and behaviour around how we do things", [P2]* which relates removing the silos from within the organisation.

Over and above that, some participants claim that the integration of systems and data that is interchanged between these systems are also a BD challenge within the organisation. Chen *et al.* (2015) state that the integration of BD technology with existing technologies is challenging, and organisations can seldom discard old systems and solely rely on their newly adopted BD systems. Technology selection, orchestration and integration challenges are only a few of the many system development paradigm shifts brought forward by the V's of BD that organisations will have to overcome to enable the successful adoption of BD (Chen *et al.*, 2015).

Participants 4 and 5 refer to the cost of storing BD as a challenge for the adoption of BD. Participant 4 claims that the cost of storing BD will sometimes outweigh the benefits. [P4] Sivarajah *et al.* (2017) elborate by mentioning that organisations are still experiencing challenges in terms of storing, managing and predominantly extracting value from the data in a cost-effective manner due to the high cost associated with BD infrastructure. Furthermore, the cost and benefit of BD adoption sometimes cannot explicitly be known beforehand due to the scope and scale of BD and the speed of technological changes (Chen *et al.*, 2015).

Lastly, Participant 4 claims that data quality and data standardisation across the organisation is also a challenge for BD adoption. [P4] Data quality refers to a multidimensional concept that describes the properties of the information, such as accuracy, timeliness, completeness, consistency, relevance and fitness for use (Janssen *et al.,* 2017). Data quality is essential for government organisations because it provides accurate information, which is needed to influence decisions that affect large populations (Talha *et al.,* 2019). Furthermore, Talha *et al.* (2019) argue that BD quality issues stem from the following: i) high volume, because the volume of data is vast and difficult to assess and improve the data quality within a reasonable time; ii) heterogeneity, because the data produced today are in most cases semi-structured or unstructured and, therefore, more complex to process than structured data; iii) data changing quickly and, thus, quickly becoming obsolete because organisations cannot collect the required data in real time, and may produce unnecessary or misleading conclusions, potentially leading to incorrect decisions; and iv) data security, because easy, flexible access to all data is necessary to allow the opportunity to evaluate the quality and, in turn, allow the required improvements.

4.6. BD impacting the strategy of Western Cape governmental organisations

The second research sub-question was: *How does BD impact the strategy of Western Cape Governmental organisations?*

The data collection for this section was done through participant responses and a review of strategic documents that are publicly available. The documents reviewed were titled 'the Organisation XYZ Data strategy work streams and deliverables and the Resilience Strategy'. One of the most important findings that comes to light in these documents indicates that the organisation values its data and uses it to make important decisions. The concerned participants refer to this practice of decision-making within the organisation as data-driven decision-making. Windt *et al.* (2019) define data-driven organisations as organisations in which decision-making is enabled by evidence found in data, rather than perception. Additionally, in order to become data-driven, organisations need quality data, good data management process and a leadership culture that promotes the use of data for decision making (Windt *et al.,* 2019).

4.6.1. Data quality

The interview question used by the researcher to obtain data for this theme is *"Comment on the access to timely, relevant and accurate data within your organisation".* Responses from participants indicate that all (11/11 or one hundred percent) of the respondents believe that data quality is crucial for the organisation. Carroll *et al.* (2019) also assert that governments perform numerous tasks that involve data. The significance of these tasks relates to making decisions about citizens, communities, and resources. Government organisations need accurate, relevant, and timely data to inform policies and decision-making (Carroll *et al.,* 2019).

Participant 1 claims that data quality is a big concern for most organisations, and that very few get it right. [P1] Additionally, the participant mentions that Organisation XYZ has systems that are more mature, in which data quality and timeous updating and accuracy are very good. The same participant also mentions that they have systems in which data quality is not so good, but that the organisation is attending to these. [P1] Participants 2, 8, 9, 10 and 11 substantiate this by mentioning that the organisation has become data-driven and is embarking on a data strategy to improve and standardise data quality and management/governance. [P2, P8, P9, P10 and P11] Furthermore, Participant 11 advises that leadership should avoid managing from instincts and relationships, but rather base decisions on data.

Participants 3, 4, 6 and 7 state that the organisation spent a great deal of time to transform their data into a good state. [P3, P4. P6 and P7] Participant 3 is quoted as follows: *"because the right data gets you to the right level of information to the right level of reporting, and the ability for decision-makers to have the right data in their hands to make decisions". [P3]* Participant 5 agrees and adds that data provides a measurement of performance that can be used as an indicator on how to improve or enhance the business, [P5] while Participant 10 mentions that it directly relates to providing service in a timely manner. [P10]

4.6.2. Data use for decision making

The interview question used by the researcher to obtain data for this theme is *"Comment on the ability of executives and business talent to use data to improve or transform your organisation."*

Executives find it difficult to introduce changes within the shortened timescales created by an ever-changing environment of new data. They are under pressure to adapt to the way strategic decisions are made (Merendino *et al.,* 2018). The decision-making processes need to be integrated and coherent with existing information systems in order for management to expand the level of knowledge of decision-makers of the proposed decisions that will be made (Alkatheeri *et al.,* 2020).

Participant 5 emphasised the importance of using data and mentioned that the data produced within their department is used as a metric for economic development within the region. [P5] Furthermore, in the data collected, Participants 1 and 10 believe that data relates to evidence and informs decisions within the organisation, but Participant 1 also doubts that this data is used in every decision. [P1 and P10] Additionally, Participants 1 and 9 observe that data is starting to play a major role in transforming the organisation, and they also add that there is still room for growth in making it a consistent tool for change. [P1 and P9]

The responses gathered from Participants 6, 7, 11 and 9 were supportive of this claim. Participant 6 believes that BD is part of the transformation journey, but also argues that it does not add value in itself. [P6] The value is realised when the data which is provided informs the decision-making process for a specific business problem, and when decision-makers actually follow through based on the data on-hand. [P6] Participants 7, 9 and 11 claim that *"quantitative or statistical literacy is generally a bit on the low side"[P7]* within the organisation, and that decision-makers sometimes need to be helped to come to appropriate conclusions based on the reports received. [P7, P9 and P11]

The responses from Participants 2, 4 and 8 were omitted because they were not relevant or helpful to the question. Participant 2 stated thus: *"I'm not sure, as not being an executive…"* [P2] while the response from Participant 8 focused on data quality [P8] and that from Participant 4 focused only on the organisations' information systems. [P4]

4.7. BD adoption affects the constituents of the Western Cape

The third research sub-question was: *How will BD adoption affect the constituents of the Western Cape?*

The evidence collected from participants indicated that the majority envisioned BD providing better service delivery to citizens. Iftikhar *et al.* (2020) argue that data by itself has no value, but that it has an extraordinary value-adding potential to promote the extraction of meaningful, actionable information. The value in question here is public service delivery (Iftikhar *et al.,* 2020).

The main remarks from the participants centred around the following: i) opportunities to discover new ways to answer age-old questions around societal issues regarding poverty and social economics; ii) informed decisions based on information and initiating programs or projects based on the information provided through BD; iii) considering service delivery from a spatial point of view, while building a framework around it; iv) better planning and rollout of services based on the needs of the citizens; v) planning for the future, to improve existing programs; vi) providing timely services, where citizens do not have to wait to be serviced (real time servicing); vii) having the correct information on-hand to make effective decisions for service delivery and to improve the performance of the organisation's staff;

and viii) basing decisions on data, specifically around housing projects and the services that accompany such developments.

Participants 1 and 2 claim that to benefit from BD adoption, data should be of a good quality and be handled in an ethical manner. Participant 1 states thus: *"my background is that I'm a mathematical statistics graduate … I know that you can make data say what you wanted to say. The point is you need to have that ethical desire to for it to help the citizens of Organisation XYZ"*. [P1]

Participant 3 mentions that BD will bring transformation to the broader proactive manner in which services are delivered and infrastructure maintained within Organisation XYZ's jurisdiction. [P3] Participants 4, 5, 6, 8 and 10 believe that BD contributes to better planning and roll-out of civil programs by analysing data from various sources, and that deriving knowledge from it leads to improved service delivery, which is ultimately the goal for this governmental organisation. [P4, P5, P6, P8 and P10] Additionally, Participant 6 mentions that BD provides timely data that informs business processes, allowing the organisation to respond quickly to the changing needs of the citizens. [P6]

Participant 9 describes the benefits in terms of the housing services that the organisation provides to citizens. [P9] The participant mentions that BD will provide the data required to make informed decisions relating to the resources which are required for new housing projects. Below is a quotation from Participant 9: *"if I want to put out a housing complex …, I need to have the data of the actual area, I need demographics. I need to know transport routes. … in order to decide whether this is a viable spot for me to actually do that housing development. The role of big data within the organisation from a service delivery is absolutely crucial and actually paramount, because without access to that data, my decision making will be flawed, which in essence will result in bad service delivery and not even to mention the financial wastage that could be incurred."* [P9]

5. Conclusion and recommendations

5.1. Introduction

The study aimed to do the following: i) review the current state of literature related to the adoption of IT, specifically Big Data at a governmental organisation; ii) identify ways in which general Big Data adoption processes can align with the adoption process of the organisation; and iii) explain if/how the identified factors influence Big Data adoption in government institutions.

In the previous chapter, the findings of this study were presented, analysed and discussed based on the responses from the participants and related literature. This chapter presents a summary in the form of a discussion on the thesis and answers to the research questions and recommendations. Additionally, it outlines the research limitations, suggestions for future research, study contributions, and terminates in the conclusion.

5.2. Summary of the chapters

This research examined the factors that influence the adoption and use of the BD in a government organisation in the Western Cape, South Africa.

Chapter 1 presented an overview of the research and provided the introduction and background to the study. It identified the research problem statement, research question, sub-questions, research objectives, and the outline of the study.

In Chapter 2, the literature review relating to BD adoption was presented. In this chapter, topics such as the concept, characteristics, ecosystems and lifecycle of BD were discussed, as well as the uses, benefits and challenges and some examples of how BD is used in other world government organisations. Lastly, the TOE framework was discussed, along with the view of how it relates to the adoption of BD in the public sector.

Chapter 3 presented the research methods used in this study. The research was aligned to the Interpretivist approach in the form of a qualitative case study. Research participants were identified using a snowball sampling strategy. The study used a thematic analysis technique to uncover themes and links that emerged from the collected data. Interview questions for the research were open-ended, allowing the participants to formulate their answers from their own perspectives. The TOE framework was drawn on to develop the interview questions, and data was collected from one government organisation (Organisation XYZ) in the Western Cape.

An introduction and description of the research participants was presented in Chapter 4. An analysis of the data collected in relation to the themes and sub-themes identified then followed. This penultimate chapter finished off with the discussion and scrutiny of the data gathered from the research interviews.

This concluding chapter (Chapter 5) starts with an introduction, followed by a summary discussion of the research findings. It continues with recommendations, research limitations, suggestion for future research, and then rounds off with the conclusion of the study.

5.3. Findings: Answers to Research Questions

5.3.1. Factors affecting Big Data adoption at a government organisation

The finding for this theme is discussed in the previous chapter under Sections 4.3 and 4.4. The focus of these sections was to discover the adoption process, adoption timeframe, newest technology adopted, who makes the adoption decision and how the adoption decision is communicated throughout the organisation. This section was important to meet one of the objectives of this study and to discover if the organisation is actively adopting technologies and the process applied to the adoption. Also, three questions were addressed, to identify the factors relating to the following: i) the factors that participants perceived as important (FPPI); ii) factors identified in literature that participants rated in terms of importance (FIL); and iii) specific factors identified in the TOE framework (FITOE) that participants were asked to discuss. These themes form the basis of determining the factors that influence BD adoption within the organisation and are depicted in Figure 11 below. It should be noted that the highest number of factors identified in this study were found in the organisational context.

Schüll and Maslan (2018) define the organisational context as descriptive measures of the organisation that relate to the organisation's structure, scope, size, and managerial capabilities. The successful adoption of BD is almost impossible without appropriately addressing these factors. Since these factors directly influence the adoption of BD, they play a crucial role in capacity building, by structuring and orchestrating resources within the organisation for the adoption of BD (Schüll & Maslan, 2018).

In the following section, these factors are summarised, and recommendations are provided to remediate the challenges associated with these factors as they relate to the adoption of BD in governmental organisations.

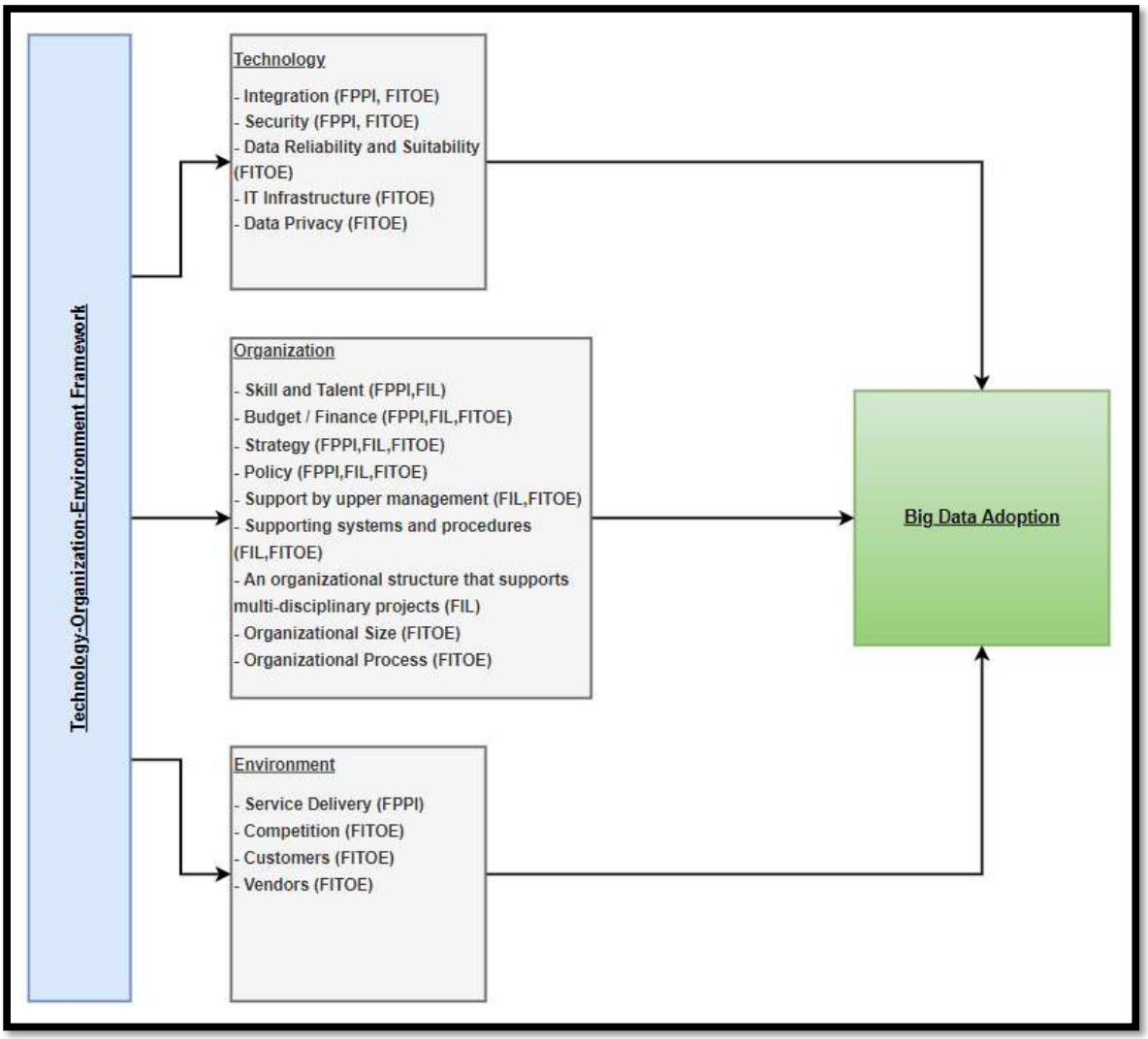


Figure 11: TOE factors discovered in the study (Source: The researcher, 2022)

5.3.2. Technology context

5.3.2.1. Data reliability and suitability

The evidence provided by participants in the analysis, in sub-section 4.4.3.3, as well as the literature by Thabet and Soomro (2015), Walker and Brown (2019) and Carroll *et al.* (2019) indicated that government organisations need a way to ensure they have accurate, relevant, and timely data to inform policies and decision-making. BD also has the ability to promote data reliability and suitability, which, in turn, influence the decision to adopt BD.

Another important consideration for BD adoption relating to data reliability and suitability is that BD infrastructure guarantees data security, integrity, availability, accountability, and ownership. The

implication of this is that government organisations evaluate the types and sources of BD being collected, stored, analysed, and consumed, for example, structured versus unstructured data.

Recommendation: BD acquisition scenarios assume that data is high-volume, high-velocity, high-variety, high-value and high-veracity. However, sometimes low-value data is also gathered, making it important to have an adaptable and time-efficient gathering, filtering, and cleaning algorithms that ensure that all data fragments are processed.

5.3.2.2. IT infrastructure

Responses from participants in section 4.4.3.4 and the literature by Engin and Treleaven (2019), Kalema and Mokgadi (2017), and Suh *et al.* (2015) claim that because of the nature and digitisation of government services, having adequate infrastructure to manage BD is an important factor that needs to be considered for BD adoption. Defitri *et al.* (2020) and the research participants emphasise the importance of IT infrastructure for BD adoption by arguing that without IT infrastructure, BD adoption projects would most likely be unsuccessful.

Recommendation: BD infrastructure needs to support the entire BD lifecycle. Government organisations need to overcome challenges like investing in technologies such as cloud computing and data warehouses, large-scale heterogeneous dataset acquisition, efficient data storage, real-time data processing and the data analysis, curation, retrieval, visualisation, and interoperability.

5.3.2.3. Security and Privacy

Evidence for this is provided in sub-sections 4.4.3.2 as well as in the work of Salleh and Janczewski (2016), who mention that new security and privacy protocol positively affect the adoption of BD. Organisations should consider implementing encryption, data anonymisation, tokenisation, and database controls to address security concerns. These techniques decrease the chances of individuals gaining un-authorised access to data and also limit the actions that can be performed on it.

Recommendation: Government organisations need to put policies in place so as to enhance trust when processing citizen data, which should result in greater synchronisation between government and the citizens.

5.3.2.4. Integration

Participants believe that system and data integration are beneficial for organisations because they can create a common view of reality though extracting coherent datasets of fine-grained data and then rolling those up into common views (sub-section 4.4.3.1). Chen *et al.* (2015) state that the integration

of BD technology with existing technologies are challenging, and that organisations can seldom discard old systems to solely rely on new adopted BD systems.

Recommendation: Integration should be considered from the start of BD adoption projects to ensure that current systems are compatible with BD. Also, BD should be adopted in a manner that ensure it is integrated into the day-to-day operations of the organisation.

5.3.3. Organisational context

5.3.3.1. Top management

In Sub-section 4.4.4.6, the responses from participants and literature from Hameed (2017), Kalema and Mokgadi (2017) and Sun *et al.* (2016) mention that top management support facilitates the availability of resources required to adopt BD and also exerts a positive influence amongst employees towards BD adoption. This is further substantiated by Almoqren and Altayar (2016), Çaldağ *et al.* (2019), Al Mudawi *et al.* (2019) and Adnan *et al.* (2021) and some of the participants who claim that when top management see the benefits of BD, they would positively affect the adoption thereof.

Recommendation: It is vital that support from top management is obtained for BD adoption project. The commitment from top management promotes the data strategy of organisations and, in turn, reveals the benefits of using BD within the organisation.

5.3.3.2. Organisational size

The literature provided a view that government organisations have many departments, and that each department generates and analyses large amounts of data specific to it (the department). Hence, the need to analyse BD within these departments could be higher than in smaller organisations. Likewise, Agrawal (2015), Malak (2017) and Motau (2016) claim that one probable reason for the relationship between IT innovation adoption and organisational size is that a larger organisation has more resources (for example, technical, financial, and personnel) and, therefore, can assign more resources to the adoption of new technologies and absorb more risk (than a smaller organisation). Thus, the size of the organisation and silos within the organisation impact on data sharing and re-use of data between departments, ultimately affecting the adoption of BD.

Recommendation: Irrespective of the size of the organisation, data sharing and the re-use of data between departments should be encouraged. The re-use of data eliminates the time required to process data.

5.3.3.3. Finance / Budget

In Sub-section 4.4.4.2, the research participants indicate that finance and budgets influence the adoption of BD. Bryant *et al.* (2008) suggest that sizeable government financial investment could greatly accelerate the adoption of BD. Kalema and Mokgadi (2017) also claim that the availability of financial resources influences an organisation's readiness for BD adoption. In addition, Nam *et al.* (2015) mention that financial support to pay for new technological innovation, the implementation of any subsequent enhancements, and ongoing expenses are very important. This is supported by the participants who claim that financial support is critical for the adoption of BD,

Recommendation: Organisations should ensure that BD projects are properly scoped and adequate budget is allocated for BD projects. When the budget is made, all the costs should be considered. These costs basically include the following: the organisation's ability to sustain BD, future scalability, hardware and software, and the cost of hiring new employees and training the existing staff.

5.3.3.4. Organisational processes

In Sub-section 4.4.4.4, participants mention that organisational data process is immature, and that data management occurs in silos. Chen *et al.* (2015) state that BD has the ability to fundamentally change organisational processes, business models and strategies, and even entire industries and markets. Lamba and Dubey (2015) claim that BD strategy is about discovering insights and knowledge from BD and taking the required steps to drive knowledge into organisational processes that create new innovative products and services.

Kumar and Krishnamoorthy (2020) suggest that organisations are less likely to use any technology if they find it problematic to comprehend and integrate it with their organisational processes. Furthermore, some of the participants mention that the organisation's current organisational processes are almost, by definition, something that cannot accommodate existing organisational hierarchy and probably incompatible with existing processes because BD did not exist when the processes were created.

Recommendation: Organisational processes should incorporate the adoption and use of BD in day-to-day tasks. Government organisations should introduce strong access and governance controls into their organisational processes.

5.3.3.5. Strategy

Participants suggest that organisational strategy influences the adoption of BD. Kalema and Mokgadi (2017) mention that sound organisational strategies positively influence an organisation's readiness for BD adoption, while Kumar and Krishnamoorthy (2020) mention that the availability of quality,

accessibility data, and well-defined rules for collecting, storing, and analysing data is critical for creating a supportive organisational strategy for BD adoption.

Recommendation: BD should be included in the strategy of government organisations. It should be well defined and clearly outlined how BD can be used to promote the vision (strategy) of the organisation. The strategy should also be well communicated to ensure that all employees are aware of their role within the organisation. In the same line of thought, Rahman and Aydın (2019) and Lamba and Dubey (2015) argue that departmental strategies and goals should be aligned with organisational strategy and goals because both the strategy and goals cannot be addressed in isolation during the process of adopting BD.

5.3.3.6. Service delivery

Participants mention that a fundamental goal for government organisations is to provide services to citizens. Thus, as the participants state, BD should be efficiently integrated within the organisation to provide better services to citizens. Arpaci *et al.* (2012) and Hoti (2015) mention that government institutions have realised that accurate, up-to-date data is required to improve service delivery, accountability, transparency, efficiency, and productivity. In Section 1.3, Bremser (2018), argues that it is crucial for organisations to understand the benefits and challenges associated with BD adoption because this leads to new data-driven services which improve processes and enable innovative products and services delivery.

Al-Sai and Abualigah (2017) mention various ways in which BD promotes service delivery, while Rajaraman (2016) states that BD promotes interaction between government and the citizens, which ultimately affects the manner in which services are delivered.

Recommendation: Service delivery is critical for government organisations. Government organisations need to be able to explain and illustrate the benefits of adopting BD, which will, in turn, increase civic participation and acceptance of the technology.

5.3.3.8. Supporting systems and procedures

The findings in the table 5 suggest that more than sixty percent of the participants indicated that supporting systems and procedures would influence the adoption of BD. Participants mentions that supporting systems and procedures will be developed during the adoption of BD to ensure that the technology can be sufficiently supported.

Recommendation: It is important that organisations have sufficient procedures in place to support BD systems. Supporting systems and procedures need to be built around BD during the adoption process. Organisations should ensure that the employees who support the technology are well trained on these systems.

5.3.3.9. Policy and Legal

Participant 4 mentions that the organisation needs to comply with the Municipal Finance Management Act (MFM) when new technologies are adopted. Similarly, Tomar *et al.* (2016) advise that suitable data regulations should be put in place for BD adoption. Zhang (2018) adds that government organisations should introduce better rules and regulations for applying BD as soon as possible and also strengthen the legal aspects because citizens are paying more attention to privacy.

Recommendation: A well-defined regulatory environment needs to be created to facilitate BD implementation. Creating a data-driven society in the public sector requires a balance between protection against data misuse and data-sharing regulations. Policies need to clearly define data ownership, usage, protection, privacy, security, liability, cybercrime, intellectual property rights, and the implications of data misuse.

5.3.3.10. Skill and talent

Participants mention that resourcing in terms of skills and talent is important for BD adoption. They suggest that the BD adoption should not only focus on the technology, but that it should also incorporate the people aspect in the adoption process. On this note, Rajaraman (2016) and Tomar *et al.* (2016) observe that an important challenge for BD adoption is related to skill (talent) shortage. Echoing this, Bremser (2018), Argawal (2016) and Awa *et al.* (2016) add that skills are vital in analysing, designing, and implementing BD.

Recommendation: Government organisations should acquire and retain skilled individuals that have the knowledge to implement, use and maintain BD. Technical experts need to be teamed up with data-savvy business experts with strong domain knowledge so that they can effectively apply their data knowledge within the organisation.

5.3.3. Environmental context

5.3.4.1. Competition

Nam *et al.* (2015) explain that business partners and competitors influence an organisation's willingness to adopt new technology through industrial pressure when partners request or recommend it, or when they see that their competitors have adopted (or are adopting) BD. Furthermore, some of the participants

mention that competition within the organisation stems from competing objectives due to the fact that departments compete for the same funds to initiate projects.

Recommendation: In order to address this factor, government organisations should ensure that no competition exists within the organisation that would prevent the adoption of BD. There should also be intentional efforts to guard against the development or growth of silos within the organisation as this could culminate in competition between departments.

5.3.4.2. Customers

Kalema and Mokgadi (2017) state that customer demands influence the adoption of BD because government organisations need to provide near real-time answers to customer queries related to service delivery. Additionally, some of the participants mention that BD provides a means to decrease production cycles, which radically improves customer services.

Recommendation: In order to address this factor, government organisations should ensure that the services and responses to service delivery queries are of the highest quality. Because BD also influences the manner in which services are delivered, government organisations should be transparent in how these services are created and delivered.

5.3.4.3. Vendors

Some participants suggest in Sub-section 4.4.5.3 that vendors are required to assist with the adoption of BD if government organisations do not have the expertise to do so. Kalema and Mokgadi (2017) claim that vendors can offload tasks that are outside the core competencies of the organisation and help strengthen its existing capabilities for handling BD. In contrast, other participants are of the view that vendors have a negative influence on BD adoption within the organisation because of past experiences and the fact that vendors are focused on promoting their technologies.

Recommendation: Careful consideration should be given to any vendors who are involved in the adoption or maintenance of BD. Vendors should be well informed of the organisation's strategy, and, if possible, be aligned with it.

5.3.4.4. Social challenges

Some of the participants claim that BD provides a means to alleviate social stresses through better forecasting based on population statistics. Halper and Krishan (2014) and Rajaraman (2016) point out that there are also societal barriers to the adoption of technology, especially in government organisations. Transforming societal beliefs to be more data-centric is a major challenge which requires

that citizens observe how data is used and to discover and realise the benefits of transparency and public service delivery.

Recommendation: It is critical to increase the awareness of BD's benefits for business, public sector organisations, and citizens. Government organisations need to transform societal beliefs to be more data-centric. These changes require that citizens observe how the data is used and also realize and discover the benefits of transparency and public service delivery.

5.4. Challenges and benefits of adopting BD in the public sector

5.4.1. Challenges

From the data collected, the challenges for BD adoption within this organisation relate to skill shortage, removing the silo mentality from within departments, change (changing how things are done within the organisation), integrating existing systems with BD, the cost of storing data, and data quality and standardisation. From this, the researcher deduced that the overarching challenges relate to organisational strategy.

First, skills shortages relate to organisational strategy in attracting, developing and retaining individuals with the competencies required to effectively perform data analysis tasks while being aware of their role and mission within the organisation (Mahfoozi *et al.,* 2018). A participant described the skill shortage in terms of analysing data to build reports instead of consuming and joining reports to produce new evidence.

Second, the successful implementation of BD technology projects requires the coordination between multiple departments which encourages the exchange of ideas. Various participants mention this as a challenge because departments still perform their functions within silos. Removing silos from within the organisation allows employees to look at the bigger picture, which allows them to (join hands where necessary and) work towards a common goal. Other participants claim that the coordination of data activities should occur from a service delivery perspective, which will ultimately benefit citizens.

Third, changing how things are done within the organisation can be linked to organisational strategy in terms of adapting processes, and designing new business models to realise value from BD. Evidence from the participants indicates that this challenge stems from how the organisation operates, because it is managed by executives that are not always willing to collaborate on cross departmental activities.

Fourth, integrating existing systems with BD can be linked to the organisation's focus on making data-driven decisions. Within the organisation, this challenge relates to data activities or developments

happening outside of corporate systems, sometimes isolated on a user's computer. This challenge forces the duplication of work efforts that negatively affect the organisation.

Fifth, the cost of storing data relates to organisational strategy to measure whether storing all data, including historical data, is believed to be beneficial to the organisation. The opinion on this is that some datasets may never be used for analysis, therefore increasing storage costs but not contributing to any data-driven decision-making.

Last, data quality relates to the use of data to achieve the mandate of public organisations, which is to provide high-quality public services. Some participants suggest that data quality is not always of the highest quality, but that the organisation is putting measures in place to increase this.

5.4.2. Benefits

The analysis found that the benefits of BD are manifested internally and externally to the organisation. Internally, the benefits were identified as data integration, better decision-making, operational efficiency and improving productivity. This finding emphasises how BD can be used for data-driven decision-making in response to the organisation fulfilling its mandate to deliver services to citizens. The findings also indicate that BD would provide a means for better root cause analysis, allowing the organisation to create more efficient operational response plans.

Data management and the accessibility of the integrated data has also been identified as a benefit because it provides a holistic view to inform decisions. Furthermore, some participants relate the benefits to better, quicker decision-making, which reduce inefficiencies and improve productivity, service delivery and transparency.

Externally, the benefits relate to the organisation's services to its constituents. This is a common finding identified in the analysis due to the nature of the organisation. It is believed that BD would positively affect public service delivery, thereby improving the organisation's efficiency and effectiveness, ultimately resulting in savings and increased productivity.

5.5. The impact of BD on the strategy of Western Cape Governmental organisations

Based on the analysis conducted in Chapter 4, the strategic goals relating to how Big Data impacts the organisation were discovered, analysed, and interpreted. The findings indicate that one of the organisation's main strategic goals is service delivery. The organisation foresees providing better services to citizens by incorporating data into its strategic functions, thus becoming more data-driven. To achieve this goal, the organisation is putting measures in place to manage, use, and increase its data

quality. It is also perceived that in doing so, the organisation will enhance its business and provide services on time.

5.6. The effect of BD adoption on the constituents of the Western Cape

The finding based on the analysis in Chapter 4 indicates that the adoption of BD is beneficial for the Western Cape's constituents. The benefits are mostly related to better service delivery, which stems from the ability to use large amounts of data for decision-making. The findings have been articulated in Section 4.7 of Chapter 4.

None of the participants thought or said that Big Data would negatively affect the constituents of the Western Cape; instead, the participants believe(d) that Big Data will positively contribute to service delivery, decision-making, poverty alleviation, new project rollouts, planning and housing.

5.7. Contribution of the research

This research provides South African governmental institutions with a foundation to address the factors identified for adopting BD. Section 5.3 discusses the seventeen factors that influence BD adoption in government organisations. The discussion includes a summary detailing each factor identified and a recommendation on how these factors can be addressed.

This study also contributes to the body of knowledge relating to BD adoption in government organisations when the TOE framework identifies factors for BD adoption. The benefit to the government organisations is that the study creates awareness for top management regarding the factors that influence BD adoption. By considering these factors, government organisations can integrate and synergize IT and business to ensure that the core business of government, which is service delivery to the citizens, is achieved competently by harnessing the potential of BD.

This study presents empirical evidence of the factors that influence BD adoption. It can be used as the starting point for academics or researchers who want to perform similar research on other government organisations.

The main objective of this study was to discover the factors that affect BD adoption at an organisational level. This objective was met by analysing the data related to the adoption factors, as provided by the participants. However, further research is required to investigate how these factors can be incorporated in a model specifically for BD adoption at government organisations. It is also recommended that a similar research study be conducted in other Local and Provincial Government organisations to expand on the results of this study, since these spheres of government might also need to adopt BD.

5.8. Limitations of the study

This research was limited to one governmental organisation in the Western Cape, South Africa, where the interviews were conducted to gather information on the factors that influence Big Data adoption. No other governmental organisations were included in this study due to constraints of time and budget. This approach ensured that only data relevant to the IS and T departments of the specific organization was examined. One benefit of this approach is that the researcher gained a deeper insight into the phenomenon being studied in its context.

Furthermore, because the TOE framework focuses on studying how the organisation's context influences the adoption and implementation of innovations, and this case study research focuses on one government organisation, it limited the study's findings. This, however, does not influence the legitimacy of the study. Rather, this case study research serves as an indication that further research is required on similar organisations to expand the body of knowledge.

Collecting data for this study was particularly difficult and time-consuming because of the COVID-19 (Infectious Coronavirus disease caused by the SARS-CoV-2 virus) disease and the scheduling and availability of interview participants due to work commitments. Because of the disease, no physical contact was allowed for interviews by both the university and Organisation XZY. This inevitability limited the number of available participants for this research study.

5.9. Recommendations

The main objective of this research was to discover the factors that influence the adoption of BD at a government organisation, and to identify ways in which the influence of these factors can be mitigated. However, future research is required to investigate or define how these factors can be expanded to ensure that continuous improvement pertaining to the adoption of BD and organisational goals are met and to ensure that service delivery is improved through the use of BD. It is also recommended that a similar research study be conducted in other government organisations, since these government organisations may produce additional factors that influence the adoption of BD. Furthermore, some government ICT initiatives overlap, which would enable researchers to expand this type of study to more than one organisation. Lastly, it is recommended that non-IT staff be included as participant in the new study to ensure that their view is represented as well.

5.10. Conclusion

This chapter presents the study's findings and illustrates how the research objective was met and how the research questions were answered. The contribution of the study from the academic and practical perspectives was provided. Recommendations for the organisation and the government sector were

offered, and they included indicating areas that need attention in government's continued endeavour to deliver efficient services to the citizens. Furthermore, the benefits of the study to the government sector and the academic domains were presented. Finally, the chapter presented recommendations for further research.

The aim of this research was to investigate the factors influencing the adoption and use of BD in a government organisation in the Western Cape through the use of the TOE framework. The study revealed seventeen factors that employees of the organisation perceived as influential in the process of adopting BD. The majority of the factors were identified in the Organisational context of the TOE framework. These factors are as follows: top management; organisational size; finance and budget; organisational processes; strategy; service delivery; supporting systems and procedures; policy and legal; and skill and talent. These factors relate to the organisation's structure, scope, size, and managerial capabilities. It was discovered that these factors have the ability to drastically deter the adoption of BD, and that if they are not adequately addressed, they could cause the failure of BD adoption.

This chapter presents some of the key challenges and benefits of BD in the public sector. Key challenges were identified as skill shortages (talent), integration, organisational strategy, data usage, data storage and data quality. The benefits identified were better data integration, better decision-making, operational efficiency, improving productivity and better service delivery to constituents. This chapter has also given recommendations on how these factors (challenges) can be addressed. The recommendations were created based on literature and the responses from participants, as shown in Chapter 4.

This study also evaluated the impact of BD on Organisation XYZ, and the effect of BD adoption on constituents. Both of these scenarios were perceived as beneficial because it promotes service delivery to citizens by incorporating data into its strategic functions.

In conclusion, BD is often seen as an opportunity for organisations to strengthen their data capabilities, regardless of their nature. Literature indicates that traditional technologies are struggling to keep pace with BD because of the velocity, variety and volume at which BD is generated. The findings in this study provide organisations with the opportunity to leverage on the discovered factors and recommendations when they want to adopt BD. This study also contributes towards addressing the scarcity of academic literature on BD adoption in governmental organisations, and executives can use its findings to make informed decisions regarding BD adoption.

References:

Abdullah, M.F., Ibrahim, M. and Zulkifli, H. (2017a) Big data analytics framework for natural disaster management in Malaysia. In IoTBDS 2017 - Proceedings of the 2nd International Conference on Internet of Things, Big Data and Security, pp. 406–411. doi:10.5220/0006367204060411.

Abdullah, M.F., Ibrahim, M. and Zulkifli, H. (2017b) Resolving the misconceptions on big data analytics implementation through government research institute in Malaysia. In IoTBDS 2017 - Proceedings of the 2nd International Conference on Internet of Things, Big Data and Security, pp. 261–266. doi:10.5220/0006293902610266.

Abutabenjeh, S. & Jaradat, R. (2018). Clarification of research design, research methods, and research methodology: A guide for public administration researchers and practitioners. *Teaching Public Administration*. 36(3):237–258. DOI: 10.1177/0144739418775787.

Adnan, H.R., Hidayanto, A.N. and Kurnia, S. (2021) Citizens' or government's will? Exploration of why indonesia's local governments adopt technologies for open government. Sustainability (Switzerland), 13(20). doi:10.3390/su132011197.

Agrawal, D., Das, S. and El Abbadi, A. (2011) Big data and cloud computing. Proceedings of the 14th International Conference on Extending Database Technology - EDBT/ICDT '11, 443(7108), p. 530. doi:10.1145/1951365.1951432.

Agrawal, K.P. (2016) Investigating the determinants of Big Data Analytics (BDA) adoption in emerging economies. Academy of Management Proceedings, 2015(1), pp. 11290–11290. doi:10.5465/ambpp.2015.11290abstract.

Akinnagbe, A., Peiris, K.D.A. & Akinloye, O. (2018). Prospects of Big Data Analytics in Africa Healthcare System. *Global Journal of Health Science*. 10(6):114. DOI: 10.5539/gjhs.v10n6p114.

Al-Sai, Z.A., Abdullah, R. and Husin, M.H. (2019) Big Data Impacts and Challenges: A Review. 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology, JEEIT 2019 - Proceedings, (April), pp. 150–155. doi:10.1109/JEEIT.2019.8717484.

Al-Sai, Z.A. and Abualigah, L.M. (2017) Big data and E-government: A review. ICIT 2017 - 8th International Conference on Information Technology, Proceedings, (May), pp. 580–587. doi:10.1109/ICITECH.2017.8080062.

AlHijji, K.Z., AlHarrasi, N.H. and Alissaee, H.S. (2018) The Role of Information in Strategic Planning and Decision Making: Overview on Economic Organizations in Oman. Journal of Economics, Business & Accountancy Ventura, 21(1). doi:10.14414/jebav.v21i1.1278.

Alkatheeri, Y., Ameen, A., Isaac, O., Nusari, M., Duraisamy, B. & Khalifa, G.S.A. (2020) The Effect of Big Data on the Quality of Decision-Making in Abu Dhabi Government Organisations. Advances in Intelligent Systems and Computing. 1016(September):231–248. DOI: 10.1007/978-981-13-9364-8_18.

Almoqren, N. and Altayar, M. (2016) Identify Challenge Factors in The Adoption of Big Data Mining Technology in Saudi Banks. International Journal of Management and Applied Science, 2(7), pp. 33–39.

Alqahtani, F.N. (2016) Identifying the Critical Factors that Impact on the Development of Electronic Government using TOE Framework in Saudi E-Government Context: A Thematic Analysis. PQDT - UK & Ireland, (October), p. 270. Available at: https://search.proquest.com/docview/2083712405?accountid=17242.

Amankwah-Amoah, J. (2016) Emerging economies, emerging challenges: Mobilising and capturing value from big data. Technological Forecasting and Social Change, 110, pp. 167–174.

Andrade, A.D. (2007) Interpretive research aiming at theory building: Adopting and adapting the case study design. QualIT 2007 - Qualitative Research: From the Margins to the Mainstream [Preprint], (March 2009).

Anna, K. and Nikolay, K. (2015) Survey on big data analytics in public sector of Russian federation. In Procedia Computer Science. Elsevier Masson SAS, pp. 905–911. doi:10.1016/j.procs.2015.07.144.

Anshari, M., Almunawar, M.N. and Lim, S.A. (2018) Big data and open government data in public services. ACM International Conference Proceeding Series, pp. 140–144. doi:10.1145/3195106.3195172.

Arass, M. El and Souissi, N. (2018) Data Lifecycle : From Big Data to Smart Data Lifecycle : From Big Data to Smart Data. Proceedings of 5TH Edition International IEEE Congress on Information Science and Technology (CiSt'18), (November), pp. 80–87.

Arp, D. (2017) Privacy Threats through Ultrasonic Side Channels on Mobile Devices. Proceedings - 2nd IEEE European Symposium on Security and Privacy, EuroS and P 2017. IEEE, pp. 35–47. doi:10.1109/EuroSP.2017.33.

Arpaci, I. et al. (2012) Organizational Adoption of Information Technologies : a Literature Review. INTERNATIONAL JOURNAL OF eBUSINESS AND eGOVERNMENT STUDIES, 4(2), pp. 37–50.

Awa, H.O., Ukoha, O. and Emecheta, B.C. (2016) Using T-O-E theoretical framework to study the adoption of ERP solution. Cogent Business & Management, 3(1), pp. 1–23. doi:10.1080/23311975.2016.1196571.

Baxter, P. and Jack, S. (2008) Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. The Qualitative Report, 13(4), pp. 544–599. doi:citeulike-article-id:6670384.

Bell, J. (2010) Doing your research project. Fifth Edit. McGrawHill Open University Press. Available at: https://cpb-us-w2.wpmucdn.com/portfolio.newschool.edu/dist/2/14941/files/2017/06/Judith_Bell_Doing_Your_Research_Project-xhunbu.pdf.

Bello-Orgaz, G., Jung, J.J. and Camacho, D. (2016) Social big data: Recent achievements and new challenges. Information Fusion, 28, pp. 45–59. doi:10.1016/j.inffus.2015.08.005.

Benbasat, I., Goldstein, D.K. and Mead, M. (1987) The Case Research Strategy in Studies of Information Systems. MIS Quarterly, 11(3), p. 369. doi:10.2307/248684.

Bester, A.V. (2007) Efficiency in the public sector : an analysis of performance measurements employed by the Western Cape Provincial Treasury. (10677143). Available at: http://scholar.sun.ac.za/handle/10019.1/3120.

Bhadani, A.K. and Jothimani, D. (2016) Big data: Challenges, opportunities, and realities. Effective Big Data Management and Opportunities for Implementation, pp. 1–24. doi:10.4018/978-1-5225-0182-4.ch001.

Bishop, W.A. (2018) A project management framework for small- and medium-sized entities: Accounting software implementation. Journal of Economic and Financial Sciences, 11(1), pp. 1–11. doi:10.4102/jef.v11i1.183.

Blazquez, D. and Domenech, J. (2018) Big Data sources and methods for social and economic analyses. Technological Forecasting and Social Change, 130(September 2017), pp. 99–113. doi:10.1016/j.techfore.2017.07.027.

Braganza, A., Brooks, L., Nepelski, D., Ali, M. & Moro, R. (2017) Resource management in big data initiatives: Processes and dynamic capabilities. Journal of Business Research. 70:328–337. DOI: 10.1016/j.jbusres.2016.08.006.

Bremser, C. (2018) Starting with Big Data Adoption. European Conference on Information Systems [Preprint].

Buslaev, A., Zernov, A., Sokolov, P. & Yashina, M. (2015) Computer network traffic models: research, hypotheses, results. Recent Advances in Mathematics Computer. 170–175.

Çaldağ, M.T., Gökalp, E. and Alkış, N. (2019) Analyzing determinants of open government based technologies and applications adoption in the context of organizations. Proceedings of the International Conference on e-Learning, e-Business, Enterprise Information Systems, and e-Government (EEE), (August), pp. 50–56.

Carroll, S.R., Rodriguez-Lonebear, D. and Martinez, A. (2019) Indigenous data governance: Strategies from united states native nations. Data Science Journal, 18(1), p. 319. doi:10.5334/dsj-2019-031.

Cavanillas, J.M., Curry, E. and Wahlster, W. (2016) New Horizons for a Data-Driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe. New Horizons for a Data-Driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe, (January 2015), pp. 1–303. doi:10.1007/978-3-319-21569-3.

Chaha, D.H. and Gulia, P. (2016) Big Data Analytics. Research Journal of Computer and Information Technology Sciences, 4(2), pp. 1–4.

Chang, W. and Grady, N. (2019) NIST Big Data Interoperability Framework: Volume 1, Definitions. NIST Special Publication, 1, pp. 1–53. doi:https://doi.org/10.6028/NIST.SP.1500-1r2.

Chatfield, A.T., Reddick, C.G. and Al-Zubaidi, W. (2015) Capability Challenges in Transforming Government through Open and Big Data : Tales of Two Cities. Proceedings of the International Conference on Information Systems, pp. 1–21. doi:10.1109/MC.2013.155.

Chen, G., Wu, S. and Wang, Y. (2015) The Evolvement of Big Data Systems: From the Perspective of an Information Security Application. Big Data Research. Elsevier Inc., pp. 65–73. doi:10.1016/j.bdr.2015.01.002.

Chen, H.-M., Kazman, R. and Matthes, F. (2015) Demystifying big data adoption: beyond IT fashion and relative advantage. Proceedings of the Twentieth DIGIT Workshop [Preprint]. Available at: http://aisel.aisnet.org/digit2015%5Cnhttp://aisel.aisnet.org/digit2015%5Cnhttp://aisel.aisnet.org/digit2015/4.

Chen, H., Chaing, R. and Storey, V.C. (2012) Business Intelligence and Analytics: From Big Data to Big Impact. MIS Quarterly, 36(4), pp. 1165–1188. doi:10.1145/2463676.2463712.

Chowdhury, M.F. (2014) Interpretivism in Aiding Our Understanding of the Contemporary Social World. (August), pp. 432–438.

Creswell, J.W. (2013) The Selection of a Research Approach. Research design: Qualitative, quantitative, and mixed methods approaches., pp. 3–23. doi:45593:01.

Damanpour, F., Sanchez-Henriquez, F. and Chiu, H.H. (2018) Internal and External Sources and the Adoption of Innovations in Organizations. British Journal of Management, 29(4), pp. 712–730. doi:10.1111/1467-8551.12296.

Darke, P., Shanks, G. and Broadbent, M. (1998) Successfully completing case study research: combining rigour, relevance and pragmatism. Information Systems Journal, 8(4), pp. 273–289. doi:10.1046/j.1365-2575.1998.00040.x.

Defitri, S.Y., Bahari, A., Handra, H. & Febrianto, R. (2020) Determinant factors of e-government implementation and public accountability: Toe framework approach. Public Policy and Administration. 19(4):37–51. DOI: 10.13165/VPA-20-19-4-03.

Demchenko, Y., Laat, C. de & Membrey, P (2014) Defining architecture components of the Big Data Ecosystem. Collaboration Technologies and Systems (CTS). International Conference on IEE, p. 104:112. Available at: http://www.uazone.org/demch/papers/bddac2014-bd-ecosystem-archi-v05.pdf.

Desai, P. V. (2018) A survey on big data applications and challenges. Proceedings of the International Conference on Inventive Communication and Computational Technologies, ICICCT 2018, (April 2018), pp. 737–740. doi:10.1109/ICICCT.2018.8472999.

Desouza, K. and Smith, K. (2014) Big Data for Social Innovation. Stanford Social Innovation Review, Summer, p. 7. doi:10.1089/big.2015.1530.

Desouza, K.C. and Jacob, B. (2017) Big Data in the Public Sector: Lessons for Practitioners and Scholars. Administration and Society, 49(7), pp. 1043–1064. doi:10.1177/0095399714555751.

Dumbill, E. (2013) Making Sense of Big Data. Big Data, 1(1), pp. 1–2. doi:10.1089/big.2012.1503.

Eisenhardt, M. and Eisenhardt, K.M. (1989) Building Theories from Case Study Research. The Academy of Management Review, 14(4), p. 532. doi:10.2307/258557.

Ekimova, K. V., Bogdanova, O.Y., Karepina, O.I. & Kravchenco, E. V. (2018) Pension fund of the Russian Federation: Challenges and prospects of the development under modern conditions. European Research Studies Journal, 21(2), pp. 273–284.

Elezaj, O., Tole, D. and Baci, N. (2018) Big data in e-government environments: Albania as a case study. Academic Journal of Interdisciplinary Studies, pp. 117–124. doi:10.2478/ajis-2018-0052.

Ellis, T. and Levy, Y. (2017) Towards a Guide for Novice Researchers on Research Methodology: Review and Proposed Methods. Proceedings of the 2009 InSITE Conference, 6. doi:10.28945/3325.

Engin, Z. and Treleaven, P. (2019) Algorithmic Government: Automating Public Services and Supporting Civil Servants in using Data Science Technologies. Computer Journal, 62(3), pp. 448–460. doi:10.1093/comjnl/bxy082.

Eynon, R. (2013) The rise of Big Data: what does it mean for education, technology, and media research? Learning, Media and Technology, 38(3), pp. 237–240. doi:10.1080/17439884.2013.771783.

Fenwick, M. and Vermeulen, E.P.M. (2019) It Is Time for Regulators to Open the "Black Box" of Technology. SSRN Electronic Journal [Preprint], (April). doi:10.2139/ssrn.3379205.

Ferreira, I., Ferreira, S. and Ramos, I. (2012) Virtual and Networked Organizations, Emergent Technologies and Tools, Virtual and Networked Organizations, Emergent Technologies and Tools. doi:10.1007/978-3-642-31800-9.

Fosso-Wamba, S., Akter, S., Edwards, A., Chopin, G. & Gnanzou, D. (2015) How "big data" can make big impact: Findings from a systematic review and a longitudinal case study. International Journal of Production Economics, 165, pp. 234–246. doi:10.1016/j.ijpe.2014.12.031.

Fosso-Wamba, S. and Mishra, D. (2017) Big data integration with business processes: a literature review. Business Process Management Journal, pp. 477–492. doi:10.1108/BPMJ-02-2017-0047.

Foulonneau, M., Turki, S., Vidou, G., Martin, S., Henri, C. & Paris, U. (2014) Open data in Service design. 12(2), pp. 99–107.

Gale, S.F. (2013) Data Doesn't Lie. PM Network, pp. 30–39.

Gandomi, A. and Haider, M. (2015) Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management, 35(2), pp. 137–144. doi:10.1016/j.ijinfomgt.2014.10.007.

Gang-Hoon, K., Silvana, T. and Ji-Hyong, C. (2014) Big Data Applications in the Government Sector. Communications of the ACM, 57(3).

Ghosh, D., Ae Chun, S., Adam, N.R. & Shafiq, B. . (2016) Big data-based smart city platform: Real-Time crime analysis. ACM International Conference Proceeding Series, 08-10-June, pp. 58–66. doi:10.1145/2912160.2912205.

Giest, S. (2017) Big data for policymaking: fad or fasttrack? Policy Sciences, 50(3), pp. 367–382. doi:10.1007/s11077-017-9293-1.

Grover, V. (2015) Research Approach: An Overview. International MultidisciplinaryResearch Journal, 4(8).

Guest, G., Namey, E.E. and Mitchell, M.L. (1978) Qualitative Research Defining and Designing. The American journal of occupational therapy. : official publication of the American Occupational Therapy Association, p. 329. doi:10.4135/9781506374680.n1.

Günther, W.A., Rezazade Mehrizi, M.H., Huysman, M., Feldberg, F., Mehrizi, M.H.R., Huysman, M., Feldberg, F., Rezazade Mehrizi, M.H., (2017) Debating big data: A literature review on realizing value from big data. Journal of Strategic Information Systems, 26(3), pp. 191–209. doi:10.1016/j.jsis.2017.07.003.

Gutierrez, D.D. (2017) Inside Big Data: Guide to Data Analytics in Government. insideBIGDATA, pp. 1–5.

Halper, F. and Krishnan, K. (2014) Tdwi Big Data Maturity Model Guide. RDWi Resarch, 2013–2014, pp. 1–20.

Halper, F. and Stodder, D. (2014) TDWI Analytics Maturity Model Guide, TDWI Research. Available at: https://tdwi.org/pages/assessments/adv-all-tdwi-advanced-analytics-maturity-model.aspx.

Hameed, M.A. (2017) A Conceptual Model for the Organisational Adoption of Information System Security Innovations. V2(4 May 2017), pp. 1–38.

Han, H., Yonggang, W., Tat-Seng, C., Xuelong, L., Hu, H.A.N., Wen, Y., Member, S., Chua, T. (2014) Toward Scalable Systems for Big Data Analytics: A Technology Tutorial. Access, IEEE, 2, pp. 652–687. doi:0.11 09/ACCESS.2014.2332453.

Hossain, M.A. and Chan, C. (2015) Open data adoption in Australian government agencies : an exploratory study. pp. 1–12.

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, III(4), pp. 1–14. doi:10.20472/BM.2015.3.4.001.

Hummel, O., Eichelberger, H., Giloj, A., Werle, D. & Schmid, K. (2018) A collection of software engineering challenges for big data system development. Proceedings - 44th Euromicro Conference on Software Engineering and Advanced Applications, SEAA 2018, pp. 362–369. doi:10.1109/SEAA.2018.00066.

Iftikhar, S., Shah, H., Peristeras, V. & Magnisalis, I. (2020) Government (Big) Data Ecosystem : Definition , Classification of Actors , and Their Roles. International Journal of Computer and Information Engineering, 14(4), pp. 102–114.

Ismail, S. (2006) Detailed review of Roger's Diffusion of innovations theory and educational technology. The Turkish Online Journal of Educational Technology, 5(2), pp. 14–23. Available at: https://files.eric.ed.gov/fulltext/ED501453.pdf.

Ismail, W.N.S.W. and Mokhtar, M.Z. (2013) Application of Toe Framework in Examining the Factors Conceptual Model for Examining the Factors That Influence the Likelihood of Computerised Accounting Information System (Cais) Adoption Among Malaysian Smes. International Journal of Information Technology and Business Management, 15(July).

Iyamu, T. and Mgudlwa, S. (2018) Transformation of healthcare big data through the lens of actor network theory. International Journal of Healthcare Management, 11(3), pp. 182–192. doi:10.1080/20479700.2017.1397340.

Janssen, M., van der Voort, H. and Wahyudi, A. (2017) Factors influencing big data decision-making quality. Journal of Business Research, 70, pp. 338–345. doi:10.1016/j.jbusres.2016.08.007.

Jeble, S., Kumari, S. and Patil, Y. (2018) Role of big data in decision making. Operations and Supply Chain Management, 11(1), pp. 36–44. doi:10.31387/oscm0300198.

Jewell, D., Barros, R.D., Diederichs, S., Duijvestin, L.M., Hammersley, M., Hazra, A., Holban, C., Li, Y. (2014) Performance and Capacity Implications for Big Data. International Business Machines Corporation, pp. 1–36. Available at: http://www.redbooks.ibm.com/redpapers/pdfs/redp5070.pdf.

Joseph M. Woodside, Shahram Amiri, B.B. (2015) The Impact of ICT and Big Data on e-Government The Impact of ICT and Big Data on e-Government. Int'l Conf. on Advances in Big Data Analytics, (July), pp. 117–118.

Kalema, B.M. and Mokgadi, M. (2017) Developing countries organizations' readiness for Big Data analytics. Problems and Perspectives in Management, 15(1), pp. 260–270. doi:10.21511/ppm.15(1-1).2017.13.

Khomba, J.K. (2007) 7 Chapter Seven : Research Design and Methodology. pp. 238–365.

Kim, G.-H., Trimi, S. and Chung, J.-H. (2014) Big-data applications in the government sector. Communications of the ACM, 57(3), pp. 78–85. doi:10.1145/2500873.

Kitchin, R. and McArdle, G. (2016) What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. Big Data & Society, 3(1), p. 205395171663113. doi:10.1177/2053951716631130.

Kleiner, A., Talwalkar, A., Sarkar, P. & Jordan, M. (2012) The Big Data Bootstrap. Available at: http://arxiv.org/abs/1206.6415.

Klievink, B., Romijn, B.-J., Cunningham, S. & de Bruijn, H. (2016) Big data in the public sector: uncertainties and readiness. Information Systems Frontiers, pp. 1–17. doi:10.1007/s10796-016-9686-2.

Koo, E. (2019) Digital transformation of Government : from E-Government to intelligent E-Government. Massachusetts Institute of Technology. Available at: http://iibraries.mit.edu/ask.

Koznov, D., Andreeva, O., Nikula, U., Maglyas, A., Muromtsev, D. & Radchenko, I. (2016) A survey of open government data in Russian Federation. IC3K 2016 - Proceedings of the 8th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, 3(November), pp. 173–180. doi:10.5220/0006049201730180.

Krishna, V.V. and Gopinath, G. (2021) Process of Requirement Gathering and Techniques for Web Application. Webology, 18(SpecialIssue2), pp. 140–152. doi:10.14704/WEB/V18SI02/WEB18063.

Kumar, A. and Krishnamoorthy, B. (2020) Business Analytics Adoption in Firms: A Qualitative Study Elaborating TOE Framework in India. International Journal of Global Business and Competitiveness, 15(2), pp. 80–93. doi:10.1007/s42943-020-00013-5.

Kumar, G. (2015) An encyclopedic overview of "big data" analytics. International Journal of Applied Engineering Research, 10(3), pp. 5681–5705. Available at: http://www.ripublication.com.

Kumar, M., Arora, S., Pal, A. & Johri, P. (2016) A Survey of Big Data Analytics in Healthcare. INROADS- An International Journal of Jaipur National University, 5(1s), p. 239. doi:10.5958/2277-4912.2016.00046.1.

Kumar, P. (2013) Qualitative Research Designs: A Conceptual Framework. International Journal of Social Science & Interdisciplinary Research, 2(1), pp. 118–124.

Kwon, O., Lee, N. and Shin, B. (2014) Data quality management, data usage experience and acquisition intention of big data analytics. International Journal of Information Management, 34(3), pp. 387–394. doi:10.1016/j.ijinfomgt.2014.02.002.

Lamba, H.S. and Dubey, S.K. (2015) Analysis of requirements for Big Data Adoption to maximize IT Business Value. 2015 4th International Conference on Reliability, Infocom Technologies and Optimization: Trends and Future Directions, ICRITO 2015, pp. 1–6. doi:10.1109/ICRITO.2015.7359268.

Lester, S. (2006) An introduction to qualitative research. European Journal of Information Systems, 9(2), pp. 127–128. doi:10.1057/palgrave/ejis/3000350.

Liu, N. and Ma, L. (2019) Challenges and Opportunities of E-government in Big Data Era. Advances in Social Science, Education and Humanities Research, 328(Ichssd), pp. 95–97. doi:10.2991/ichssd-19.2019.19.

Lockner, B.J. and February, P. (2012) The Impact of Big Data on Data Analytics. IDC Government Insights, 14(Picciano 2012), pp. 139–145.

Maass, W., Parsons, J., Purao, S., Storey, V.C. & Woo, C. (2018) Data-driven meets theory-driven research in the era of big data: Opportunities and challenges for information systems research. Journal of the Association for Information Systems, 19(12), pp. 1253–1273. doi:10.17705/1jais.00526.

Maciejewski, M. (2017) To do more, better, faster and more cheaply: using big data in public administration. International Review of Administrative Sciences, 83(1_suppl), pp. 120–135. doi:10.1177/0020852316640058.

Madden, S. (2012) From Databases to Big Data. Internet Computing, IEEE, 16(3), pp. 4–6. doi:10.1109/MIC.2012.50.

Mahfoozi, A., Salajegheh, S., Ghorbani, M. & Sheikhi, A. (2018) Developing a talent management model using government evidence from a large-sized city, Iran. Cogent Business and Management, 5(1). doi:10.1080/23311975.2018.1449290.

Manogaran, G., Varatharajan, R., Lopez, D., Kumar, P.M., Sundarasekar, R. & Thota, C. (2018) A new architecture of Internet of Things and big data ecosystem for secured smart healthcare monitoring and alerting system. Future Generation Computer Systems, 82, pp. 375–387. doi:10.1016/j.future.2017.10.045.

Margetts, H. & Naumann, A. (2017) Government as a platform: what can Estonia show the world? Department of Politics and International Relations, University of Oxford [Preprint].

Mason, J. (2002) Qualivative Researching. 2nd edn, SAGE Publications Inc.

De Mauro, A., Greco, M. and Grimaldi, M. (2016a) A formal definition of Big Data based on its essential features. Library Review, 65(3), pp. 122–135. doi:10.1108/LR-06-2015-0061.

De Mauro, A., Greco, M. and Grimaldi, M. (2016b) A formal definition of Big Data based on its essential features. Library Review, 65(3), pp. 122–135. doi:10.1108/LR-06-2015-0061.

Mayer-Schönberger, Viktor Cukier, K., Naimi, A.I. and Westreich, D.J. (2014) Big Data: A Revolution That Will Transform How We Live, Work, and Think. American Journal of Epidemiology, 179(9), pp. 1143–1144. doi:10.1093/aje/kwu085.

McBride, K., Aavik, G., Toots, M., Kalvet, T. & Krimmer, R. (2019) How does open government data driven co-creation occur? Six factors and a “perfect storm”; insights from Chicago’s food inspection forecasting model. Government Information Quarterly, 36(1), pp. 88–97. doi:10.1016/j.giq.2018.11.006.

McCusker, K. and Gunaydin, S. (2015) Research using qualitative, quantitative or mixed methods and choice based on the research. Perfusion (United Kingdom), 30(7), pp. 537–542. doi:10.1177/0267659114559116.

McGrath, S.K. and Whitty, S.J. (2019) Do steering committees really steer? International Journal of Managing Projects in Business, 12(3), pp. 785–807. doi:10.1108/IJMPB-04-2018-0064.

Mellouli, S., Luna-Reyes, L.F. and Zhang, J. (2014) Smart government, citizen participation and open data. Information Polity, 19(1–2), pp. 1–4. doi:10.3233/IP-140334.

Merendino, A., Dibb, S., Meadows, M., Quinn, L., Wilson, D., Simkin, L. & Canhoto, A. (2018) Big data, big decisions: The impact of big data on board level decision-making. Journal of Business Research, 93(August), pp. 67–78. doi:10.1016/j.jbusres.2018.08.029.

Mergel, I., Edelmann, N. and Haug, N. (2019) Defining digital transformation: Results from expert interviews. Government Information Quarterly, 36(4), p. 101385. doi:10.1016/j.giq.2019.06.002.

Michael, K. and Miller, K. (2013) Big data: New opportunities and new challenges. Computer, 46(6), pp. 22–24. doi:10.1109/MC.2013.196.

Mills, K.A. (2018) What are the threats and potentials of big data for qualitative research? doi:10.1177/1468794117743465.

Morabito, V. (2015) Big data and analytics: Strategic and organizational impacts. Big Data and Analytics: Strategic and Organizational Impacts, pp. 1–183. doi:10.1007/978-3-319-10665-6.

Moreno, J., Serrano, M.A. and Fernández-Medina, E. (2016) Main issues in Big Data security. Future Internet, 8(3). doi:10.3390/fi8030044.

Al Mudawi, N., Beloff, N. and White, M. (2019) Cloud computing in government organizations-towards a new comprehensive model. Proceedings - 2019 IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Internet of People and Smart City Innovation, SmartWorld/UIC/ATC/SCALCOM/IOP/SCI 2019, pp. 1473–1479. doi:10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00266.

Nam, D.W., Kang, D. and Kim, S.H. (2015) Process of big data analysis adoption: Defining big data as a new IS innovation and examining factors affecting the process. Proceedings of the Annual Hawaii International Conference on System Sciences, pp. 4792–4801. doi:10.1109/HICSS.2015.569.

O'Bien, A. (2012) The Impact of Big Data on Government. IDC Government Insights, GI237315(October), pp. 1–12.

O'Connor, H. and Gibson, N. (2003) The Quality of the Interview: A step-by-step guide to qualitative data analysis. A Journal of Aboriginal and Indigenous Community Health, 1(1), pp. 64–90. Available at: http://www.pimatisiwin.com/uploads/1289566991.pdf.

O'Leary, Z. (2004) The Essential Guide To Doing Research. The essential guide to doing research, pp. 28–41.

OECD (2019) The Path to Becoming a Data-Driven Public Sector, OECD Digital Government Studies. Available at: https://doi.org/10.1787/059814a7-en.

Ogbuju, E., Taiwo, K., Ejiofor, V. & Onyesolu, M. (2019) Big data driven e-government framework in Nigeria. Bayero Journal of Pure and Applied Sciences, 11(2), p. 252. doi:10.4314/bajopas.v11i2.35.

Oliveira, T. and Martins, M.F. (2011) Literature Review of Information Technology Adoption Models at Firm Level. Electronic Journal Information Systems Evaluation, 14(1), pp. 110–121.

Omoyiola, B.O. (2022) Overview of the Social Implications, Risks, Challenges, and Opportunities of Big Data. SSRN Electronic Journal [Preprint], (April). doi:10.2139/ssrn.4036190.

Oussous, A., Benjelloun, F.Z., Ait Lahcen, A. & Belfkih, S. (2018) Big Data technologies: A survey. Journal of King Saud University - Computer and Information Sciences. 30(4):431–448. DOI: 10.1016/j.jksuci.2017.06.001.

Pagallo, U. (2017) The Legal Challenges of Big Data. European Data Protection Law Review, 3(1), pp. 36–46. doi:10.21552/edpl/2017/1/7.

Pappas, I.O., Mikalef, P., Giannakos, M.N., Krogstie, J. & Lekakos, G. (2018) Big data and business analytics ecosystems: paving the way towards digital transformation and sustainable societies. Information Systems and e-Business Management, 16(3), pp. 479–491. doi:10.1007/s10257-018-0377-z.

Pastorino, R., De Vito, C., Migliara, G., Glocker, K., Binenbaum, I., Ricciardi, W. & Boccia, S. (2019) Benefits and challenges of Big Data in healthcare: An overview of the European initiatives. European Journal of Public Health, 29, pp. 23–27. doi:10.1093/eurpub/ckz168.

Pencheva, I., Esteve, M. and Mikhaylov, S.J. (2018) Big Data and AI – A transformational shift for government: So, what next for research? Public Policy and Administration [Preprint]. doi:10.1177/0952076718780537.

Pham, L. (2018) A Review of key paradigms: positivism, interpretivism and critical inquiry. ResearchGate, (April), pp. 1–7. doi:10.13140/RG.2.2.13995.54569.

Polit, D.F. and Beck, C.T. (2010) Generalization in quantitative and qualitative research: Myths and strategies. International Journal of Nursing Studies, 47(11), pp. 1451–1458. doi:10.1016/j.ijnurstu.2010.06.004.

Pouchard, L. (2016) Revisiting the Data Lifecycle with Big Data Curation. International Journal of Digital Curation, 10(2), pp. 176–192. doi:10.2218/ijdc.v10i2.342.

Pozzebon, M., Freitas, H. and Jenkins, M. (1998) The applicability and scientific rigor of case studies in Information Systems. (050498), pp. 1–31.

Pudjianto, B., Zo, H., Ciganek, A.P. & Rho, J.J. (2011) Determinants of E-Government Assimilation in Indonesia : An Empirical Investigation Using a TOE Framework. Asia Pacific Journal of Information Systems, 21(1), pp. 50–80.

Pugna, I.B., Duţescu, A. and Stanila, O.G. (2019) Corporate attitudes towards Big Data and its impact on performance management: A qualitative study. Sustainability (Switzerland), 11(3). doi:10.3390/su11030684.

Rachlitz, A. (2017) '3. Research design and methodology. The Political Construction of Irregularity in Germany and South Africa, pp. 62–89. doi:10.5771/9783845288284-62.

Rad, B.B. and Ataei, P. (2017) The big data Ecosystem and its Environs. International Journal of Computer Science and Network Security, 17(3), pp. 38–42.

Rahi, S. (2017) Research Design and Methods: A Systematic Review of Research Paradigms, Sampling Issues and Instruments Development. International Journal of Economics & Management Sciences, 06(02). doi:10.4172/2162-6359.1000403.

Rahman, M. and Aydın, E. (2019) Organisational Challenges And Benefits Of E- Hrm Implementations In Governmental Organisations: Theoretical Shift From Toe Model. Uluslararası İktisadi ve İdari İncelemeler Dergisi [Preprint]. doi:10.18092/ulikidince.516443.

Rajaraman, V. (2016) Big data analytics, Resonance. doi:10.1007/s12045-016-0376-7.

Rashid, Y., Rashid, A., Warraich, M.A., Sabir, S.S. & Waseem, A. (2019) Case Study Method: A Step-by-Step Guide for Business Researchers. International Journal of Qualitative Methods, 18, pp. 1–13. doi:10.1177/1609406919862424.

Rehman, A.A. and Alharthi, K. (2016) An Introduction to Research. International Journal of Educational Investigations, Vol.3(October), pp. 51–59. doi:10.4135/9780857024633.d15.

Riggins, F.J. and Wamba, S.F. (2015) Research directions on the adoption, usage, and impact of the internet of things through the use of big data analytics. Proceedings of the Annual Hawaii International Conference on System Sciences, 2015-March, pp. 1531–1540. doi:10.1109/HICSS.2015.186.

Ritchie, J. and Lewis, J. (2003) Qualitative Research Practice: A Guide for Social Science Students and Researchers, SAGE Publications. doi:10.4135/9781452230108.

Runting, R.K., Phinn, S., Xie, Z., Venter, O. & Watson, J.E.M. (2020) Opportunities for big data in conservation and sustainability. Nature Communications, 11(1), pp. 1–4. doi:10.1038/s41467-020-15870-0.

Sadineni, P.K. (2020) Sampling based join-aggregate query processing technique for big data. Indian Journal of Computer Science and Engineering, 11(5), pp. 532–546. doi:10.21817/indjcse/2020/v11i5/201105116.

Sakr, S., Maamar, Z., Awad, A., Benatallah, B. & Van Der Aalst, W.M.P. (2018) Business process analytics and big data systems: A roadmap to bridge the gap. IEEE Access, 6, pp. 77308–77320. doi:10.1109/ACCESS.2018.2881759.

Salehnia, A. (2015) Comparisons of Relational Databases with Big Data: a Teaching Approach.

Salleh, K., Janczewski, L. and Beltran, F. (2015) Association for Information Systems AIS Electronic Library (AISeL) SEC-TOE Framework: Exploring Security Determinants in Big Data Solutions

Adoption Recommended Citation “SEC-TOE Framework: Exploring Security Determinants in Big Data Solutions Adoption” S. p. 2015. Available at: http://aisel.aisnet.org/pacis2015/203.

Salleh, K.A. and Janczewski, L. (2016) Technological, Organizational and Environmental Security and Privacy Issues of Big Data: A Literature Review. Procedia Computer Science, 100(December 2016), pp. 19–28. doi:10.1016/j.procs.2016.09.119.

Sam, K.M. and Chatwin, C.R. (2019) Understanding Adoption of Big Data Analytics in China: From Organizational Users Perspective. IEEE International Conference on Industrial Engineering and Engineering Management, 2019-Decem, pp. 507–510. doi:10.1109/IEEM.2018.8607652.

Sarker, M.N.I., Hossin, M.A., Frimpong, A.N.K. & Xiaohua, Y. (2018) Promoting information resource management for E-government through big data approach. ACM International Conference Proceeding Series, pp. 99–104. doi:10.1145/3277139.3277155.

Schüll, A. and Maslan, N. (2018) On the adoption of big data analytics: Interdependencies of contextual factors. ICEIS 2018 - Proceedings of the 20th International Conference on Enterprise Information Systems, 1(Iceis 2018), pp. 425–431. doi:10.5220/0006759904250431.

Scotland, J. (2012) Exploring the philosophical underpinnings of research: Relating ontology and epistemology to the methodology and methods of the scientific, interpretive, and critical research paradigms. English Language Teaching, 5(9), pp. 9–16. doi:10.5539/elt.v5n9p9.

Sheppard, V. (2020) Research Methods for the Social Sciences: An Introduction. (December), p. 430. Available at: https://pressbooks.bccampus.ca/jibcresearchmethods/.

Shin, D.H. and Choi, M.J. (2015) Ecological views of big data: Perspectives and issues. Telematics and Informatics, 32(2), pp. 311–320. doi:10.1016/j.tele.2014.09.006.

Sinaeepourfard, A., Garcia, J., Masip-Bruin, X. & Marín-Torder, E. (2016) Towards a comprehensive data LifeCycle model for big data environments. Proceedings - 3rd IEEE/ACM International Conference on Big Data Computing, Applications and Technologies, BDCAT 2016, pp. 100–106. doi:10.1145/3006299.3006311.

Sivarajah, U., Kamal, M.M., Irani, Z. & Weerakkody, V. (2017) Critical analysis of Big Data challenges and analytical methods. Journal of Business Research, 70, pp. 263–286. doi:10.1016/j.jbusres.2016.08.001.

Staunton, C., Adams, R., Botes, M., Vries, J. de, Labuschaigne, M., Loots, G., Mahomed, S., Loideain, N.N. (2021) Enabling the use of health data for research: Developing a POPIA code of conduct for research in South Africa. South African Journal of Bioethics and Law, 14(1), p. 33. Available at: https://doi.org/10.7196/SAJBL.2021.v14i1.740.

Stenberg, L. and Nilsson, S. (2020) Factors influencing readiness of adopting AI: A qualitative study of how the TOE framework applies to AI adoption in governmental authorities.

Suh, J., Vujin, V., Barac, D., Bogdanovic, Z. & Radenkovic, B. (2015)Designing Cloud Infrastructure For Big Data In E-Government. RUO. Revija za Univerzalno Odlicnost, 4(1), pp. A26–A38. Available at: http://search.proquest.com/docview/1681254181?accountid=13031%5Cnhttp://sfx.nelliportaali.fi/nelli28b?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:journal&genre=article&sid=ProQ:ProQ%3Aabiglobal&atitle=DESIGNING+CLOUD+INFRASTRUCTURE+FOR+BIG+DATA+I.

Sun, S., Cegielski, C.G., Jia, L. & Hall, D.J. (2018) Understanding the Factors Affecting the Organizational Adoption of Big Data. Journal of Computer Information Systems, 58(3), pp. 193–203. doi:10.1080/08874417.2016.1222891.

Sun, Z., Strang, K.D., Strang, K. & Li, R. (2018) Big Data with Ten Big Characteristics Cyberprivacy and Cybersecurity in the Age of Big Data View project Intelligent Big Data Aanalytics: Intelligent Techniques for Big Data Aanalytics View project Big Data with Ten Big Characteristics. (October). doi:10.13140/RG.2.2.21798.98886.

Taleb, I., Serhani, M.A. and Dssouli, R. (2018) Big Data Quality: A Survey. Proceedings - 2018 IEEE International Congress on Big Data, BigData Congress 2018 - Part of the 2018 IEEE World Congress on Services, (August), pp. 166–173. doi:10.1109/BigDataCongress.2018.00029.

Talha, M., El Kalam, A.A. and Elmarzouqi, N. (2019) Big data: Trade-off between data quality and data security. Procedia Computer Science, 151, pp. 916–922. doi:10.1016/j.procs.2019.04.127.

Taylor, L. (2017) Safety in Numbers? Group Privacy and Big Data Analytics in the Developing World. Group Privacy, pp. 13–36. doi:10.1007/978-3-319-46608-8_2.

Thabet, N. and Soomro, T.R. (2015) Big Data Challenges. Computer Engineering & Information Technology, 04(03), p. 33. doi:10.4172/2324-9307.1000135.

Thomas, P. (2010) Chapter 3. Research methodology and study design. English as a Lingua Franca in Higher Education, pp. 291–334. doi:10.1515/9783110215519.82.

Tomar, L., Guicheney, W., Kyarisiima, H. & Zimani, T. (2016) Big Data in the Public Sector. Selected Applications and Lessons Learned. Inter-American Development Bank, (October), pp. 1–52. Available at: https://webimages.iadb.org/publications/english/document/Big-Data-in-the-Public-Sector-Selected-Applications-and-Lessons-Learned.pdf.

La Torre, M., Dumay, J. and Rea, M.A. (2018) Breaching intellectual capital: critical reflections on Big Data security. Meditari Accountancy Research, 26(3), pp. 463–482. doi:10.1108/MEDAR-06-2017-0154.

Twala, A. and Kekwaletswe, R. (2019) Strategic Cloud Computing Framework: South African Context. 12, pp. 372–361. doi:10.29007/p7br.

Twizeyimana, J.D. and Andersson, A. (2019) The public value of E-Government – A literature review. Government Information Quarterly, 36(2), pp. 167–178. doi:10.1016/j.giq.2019.01.001.

Vassakis, K., Petrakis, E. and Kopanakis, I. (2018) Big Data Analytics: Applications, Prospects and Challenges. 10(January), p. 346. doi:10.1007/978-3-319-67925-9.

Vinet, L. and Zhedanov, A. (2011) A "missing" family of classical orthogonal polynomials. Journal of Physics A: Mathematical and Theoretical, 44(8). doi:10.1088/1751-8113/44/8/085201.

van der Voort, H.G., Klievink, A.J., Arnaboldi, M. & Meijer, A.J. (2019) Rationality and politics of algorithms. Will the promise of big data survive the dynamics of public decision making? Government Information Quarterly, 36(1), pp. 27–38. doi:10.1016/j.giq.2018.10.011.

Wadhwani, K. and Wang, Y. (2017) Big Data Challenges and Solutions. International Journal of Computer Sciences and Engineering, 5(10), pp. 250–255. doi:10.26438/ijcse/v5i10.250255.

Walker, R.S. and Brown, I. (2019) Big data analytics adoption: A case study in a large South African telecommunications organisation. SA Journal of Information Management, 21(1), pp. 1–10. doi:10.4102/sajim.v21i1.1079.

Walliman, N. (2010) Research Methods: The Basics. Research Methods: The Basics [Preprint]. doi:10.4324/9780203836071.

Webley, L. (2012) Qualitative Approaches to Empirical Legal Research. The Oxford Handbook of Empirical Legal Research [Preprint], (January 2010). doi:10.1093/oxfordhb/9780199542475.013.0039.

Weerakkody, V., Irani, Z., Lee, H., Hindi, N. & Osman, I. (2016) Are U.K. Citizens Satisfied With E-Government Services? Identifying and Testing Antecedents of Satisfaction. Information Systems Management, 33(4), pp. 331–343. doi:10.1080/10580530.2016.1220216.

Welman, C., Kruger, F. & Mitchell, B. (2005) Research methodology. pp. 35–40.

Windt, B., Borgman, H. and Amrit, C. (2019) Understanding Leadership Challenges and Responses in Data-driven Transformations. Proceedings of the 52nd Hawaii International Conference on System Sciences, pp. 4987–4996. doi:10.24251/hicss.2019.599.

Wiseman, M.J. (2018) Data-driven government: The role of Chief Data Officers. IBM Center for the Business of Government [Preprint].

Yafooz, W.M.S., Binti, Z., Bakar, A., Fahad, S.K.A. & Mithun, A.M. (2020) Analytics , Data Mining and Machine Learning. pp. 217–230. doi:10.1007/978-981-13-9364-8.

Yakobi, K., Scholtz, B. and vom Berg, B. (2020) A Conceptual Model of the Challenges of Social Media Big Data for Citizen e-Participation: A Systematic Review, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer International Publishing. doi:10.1007/978-3-030-45002-1_21.

Yan, J. (2013) Big Data, Bigger Opportunities. Dataquest, 33(15), p. 72. Available at: http://www.meritalk.com/pdfs/bdx/bdx-whitepaper-090413.pdf%0Ahttp://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=108687508&site=eds-live.

Younas, M. (2019) Research challenges of big data. Service Oriented Computing and Applications, 13(2), pp. 105–107. doi:10.1007/s11761-019-00265-x.

Yue, C.A., Men, L.R. and Ferguson, M.A. (2019) Bridging transformational leadership, transparent communication, and employee openness to change: The mediating role of trust. Public Relations Review, 45(3), p. 101779. doi:10.1016/j.pubrev.2019.04.012.

Zulkarnain, N., Nizar Hidayanto, A. and Prabowo, H. (2019) Critical Success Factors for Big Data Adoption in Government. International Journal of Mechanical Engineering and Technology (IJMET), 10(3), pp. 864–875.

Appendices

Appendix 1: Interview Questionnare

Research question	Interview question	Interview question design	Rationale
PART: A	1. Please describe the process that you think your organization is following to adopt new technologies.	Open-Ended question	It is imperative to understand whether the organization have processes in place for adopting new technologies. An understanding of these processes and factors influencing technology adoption within the organization would enable the researcher to provide significant insight on how/whether these processes needs to be changed to enable Big Data adoption
What are the factors affecting Big Data technology adoption at a government department in the Western Cape?	2. In your opinion, what are the factors that your organization need to investigate for new technology adoption?	Open-Ended question	
	3. What is the newest technology is that your organization adopted, and when did it occur?	Open-Ended question	
	4. Do you know in what timeframe this project was completed?	Open-Ended question	
	5. To the best of your ability can you explain what you think Big Data is, and how it can be used within your organization.	Open-Ended question	
	6. Rate the criticality of the following factors for technology adoption.	Open-Ended question	

Factors influencing Big Data - Relates to Part A Question 6		
Interview question	**Factor**	**Assign a value from 1 to 8**
How critical are the following factors for technology adoption?	A clear company strategy	
	Support by upper management	
	Talent	
	Training	
	Supporting systems and procedures	
	Financial budget	
	An organizational structure that supports multi-disciplinary projects	
	A sound procedure for legal and ethical issues	

Research question	Interview question	Interview question design	Rationale
PART: B	1. At what level is the decision to adopt new technology made, and how is it taken?	Open-ended question	Understanding the strategic position of the organization and who it will be influenced by Big Data
How does Big Data impact the strategy of Western Cape Governmental organizations?	2. How is the process communicated throughout the organization?	Open-Ended question	
	3. Comment on the access to timely, relevant and accurate data within your organization.	Open-ended question	
	4. Comment on the ability of executives and business talent to use data to improve or transform your organization.	Open-ended question	
	5. What is the role of leadership on data utilization from IT, to or the rest of the business?	Open-Ended question	

Research question	Interview question	Interview question design	Rationale
PART: C	1. What do you think are some of the benefits for using Big Data in your organization?	Open-ended question	This would produce an understanding of the individual's perspective regarding the benefits and challenges associated with Big Data. Question 3 relates to the TOE framework
What are the challenges and benefits of implementing BD in the public sector	2. What would the biggest organizational challenges be of using Big Data?	Open-ended question	
	3. In your opinion how would the following affect Big Data adoption within the organization? a) Data Reliability/Suitability b) IT Infrastructure c) Data Security /Privacy d) Top Management e) Organizational Size f) Finance (budget to implement Big Data) g) Organizational Process h) Organizational Strategy i) Competition j) Customers k) Vendors	Open-ended question	

Research question	Interview question	Interview question design	Rationale
PART: D How will BD adoption affect the constituents of the Western Cape?	1. What would the role of BD be in your organization in terms of service delivery and new programs being rolled out by your organization	Open-Ended question	To access if the participant understands how services and programs affect the organization and citizens and how Big Data will affect these

Appendix 2: Example Interview Transcripts

00:00:00.000 --> 00:00:02.690
Andre Bruintjies
And the Transcriptionist started as well.

00:00:04.240 --> 00:00:04.630
P3:
OK.

00:00:05.900 --> 00:00:09.960
Andre Bruintjies
OK, so my first question is. Please can you describe the process your organization is following to adopt new technologies?

00:00:20.130 --> 00:00:30.870
P3:
So, in terms of new technologies, we have ran an review and to the organization over the last year where effectively in that review we looked at the digital maturity for the organization. Uh, and did compare of the digital landscape of what we have as an organization against what is up there, both from market perspective as well as best practices for cities. And we looked at our future business strategies and then based on a combination of our future business strategies and our current landscape, along with where technologies are moving to, we came out with insights on how we could build out a digital road map for technologies of the future that makes sense for us.

00:01:21.550 --> 00:01:23.370
Andre Bruintjies
OK, thank you very much.

00:01:24.960 --> 00:01:33.960
Andre Bruintjies
The following question is, in your opinion, what are the factors you organization need to investigate for new technology adoption?

00:01:41.190 --> 00:02:10.660
P3:
I think I've just spoken to that in the previous question, so you know the fact is I'm more around what of the future business strategic needs and mapped against that is. What are the technologies of the future in terms of offerings that exist out there and bringing together the strategies and those business needs and of course based on financial factors, making determination around where there are economic and financial business cases for those.

00:02:15.570 --> 00:02:17.470
Andre Bruintjies
OK, thank you very much.

00:02:18.290 --> 00:02:28.430
Andre Bruintjies
Next question is, are you aware of what the newest technology is that the organization adopted?

00:02:30.440 --> 00:02:37.250
P3:
Sure, it’s a huge organization Andre and there's been many new technologies that gets adopted. But I would say probably we could highlight new topics. New technology is in the basis of IoT as well as artificial intelligence.

00:02:54.720 --> 00:03:07.200
Andre Bruintjies
OK thank you. then just a follow up question on that, are you aware of what the time frame for that project was? Or any of those projects were that you mentioned?

00:03:09.650 --> 00:03:10.470
P3:
Well, they are ongoing, but I guess you could say about 18 months or so.

00:03:19.730 --> 00:03:21.120
Andre Bruintjies
OK, thank you very much.

00:03:22.740 --> 00:03:27.410
Andre Bruintjies
And then the next question is are you familiar with big data? And if you are, do you know if there is any benefit of using it or if there will be any benefit of using big data within the organization?

00:03:45.780 --> 00:03:47.710
P3:
I am familiar with big data, yes, I do believe that there would be use cases for big data analysis and that would be very match in this space for me, I think it's used cases would be around where we could get to more effective decision making for our citizens in terms of service delivery.

00:04:14.470 --> 00:04:21.850
Andre Bruintjies
OK, thank you very much. the next question is, so I've got factors that I that I've identified during the study, and I'd like to know or if you can identify or let me know how critical these factors or for adopting new technologies, I am going to put it in the chat window or I can share my screen, whichever you prefer.

00:04:52.750 --> 00:04:55.320
P3:
I'm happy to share your screen.

00:04:56.100 --> 00:04:57.470
Andre Bruintjies
Yeah, I can do that.

00:05:11.970 --> 00:05:14.220
Andre Bruintjies
Please let me know once you can see that.

00:05:17.020 --> 00:05:17.950
P3:
I can.

00:05:20.270 --> 00:05:24.200
Andre Bruintjies
OK, so these are the factors highlighted here.
00:05:26.030 --> 00:05:32.720
Andre Bruintjies
So let's start off so it starts off with a clear on company strategy support by upper management.

00:05:32.910 --> 00:05:39.520
Andre Bruintjies
A talent training, supporting systems and procedures. Financial budget.

00:05:40.400 --> 00:05:51.420
Andre Bruintjies
And then an organizational structure that supports multidisciplinary projects. And the last one is a sound procedure for legal and ethical issues.

00:05:56.840 --> 00:05:58.670
P3:
These are scored from high to low (1 being the most important)
So, for number 1. Yeah, definitely. I think the company strategy. And so yeah, OK, so company strategy number one. Number two, I think in organizational structure that supports multidisciplinary projects. Number three, I think he needs support by upper management. Number 4 talent. Uh, number five financial budget. #6, uh, supporting systems and procedure. #7 training. And sample rate. Yeah, #6 is. Supporting us no. No, just having this training. I'm sorry you're doing it the wrong way.

00:06:59.550 --> 00:07:03.320
P3:
So talentwise number 4 and 7 is the next one.

00:07:04.250 --> 00:07:04.850
P3:
Thanks.
00:07:05.970 --> 00:07:09.030
P3:
It's the next writing and at the bottom 8.

00:07:12.520 --> 00:07:14.450
Andre Bruintjies
OK, thank you very much.

00:07:15.420 --> 00:07:18.920
Andre Bruintjies
Uh, I'll just stop sharing again.

00:07:20.680 --> 00:07:27.960
Andre Bruintjies
Our next section deals with the impact of big data and what it can do for communities and things like that. So, the first question in in the section is at what level is this is in the decision to adopt new technologies made and how was it taken?

00:07:55.000 --> 00:08:04.540
Andre Bruintjies
Does this also tie in with the first question that you wanted to, or do you feel that there's different explanation that you can give to that to that question?

00:08:05.560 --> 00:08:07.930
P3:
I think it ties into the first, but I mean. You know, I think the one thing we must highlight is, you know they can adoption of a new technology. You have new technologies is an ongoing process, so you know. So, it's based on business case and need alongside kind of active reviews on where you sit on the landscape

00:08:29.990 --> 00:08:30.320
Andre Bruintjies
OK.

00:08:32.310 --> 00:08:33.670
Andre Bruintjies
OK, thank you.

00:08:34.510 --> 00:08:41.790
Andre Bruintjies
The next question is how is the process communicated throughout the organization?

00:08:44.680 --> 00:08:47.110
P3:
Process as in process of change.

00:08:48.180 --> 00:08:50.530
Andre Bruintjies
Uh, yeah there was in the process of change.

00:08:52.940 --> 00:08:54.080
P3:
So typically, through project environments with active change management strategies behind that.

00:09:03.290 --> 00:09:03.570
Andre Bruintjies
OK.

00:09:05.150 --> 00:09:18.510
Andre Bruintjies
OK, thank you. Then the next question is, please comment on access to timely, relevant, accurate data within the organization. Do you think it's important? And if so, how does it affect the decision-making process within the organization?

00:09:36.120 --> 00:09:37.610
P3:
It's very important. on all levels, timely, relevant and accurate. From a data perspective, we have spent many years. I think this especially with our ERP foundational systems to ensure a focus on data accuracy and getting data right. And I think that ties in know that ties into effective decision making coming through. Because the right data gets you to the right level of information to the right level of reporting. To the ability for decision makers to have the right data in their hands to make decisions.

00:10:22.920 --> 00:10:24.700
Andre Bruintjies
OK, thank you very much.

00:10:25.650 --> 00:10:31.700
Andre Bruintjies
So then I will move to part C of the interview question.

00:10:32.650 --> 00:10:40.500
Andre Bruintjies
The first question in this section is what do you think are some of the biggest benefits of using big data within your organization?

00:10:43.860 --> 00:11:00.130
P3:
I think we spoke to this one earlier. I mean, we'll look at big data use cases in spaces where you know it can really from an operational efficiency, improve our service delivery to our citizens.

00:11:01.830 --> 00:11:09.600
Andre Bruintjies
Again, thank you and then, uh, the biggest challenges with using big data and organization.

00:11:13.870 --> 00:11:23.060
P3:
Yeah, I think you know challenge wise. Probably just the cost to stores such data. So, I think it's kind of cost versus benefit and sometimes will cost outweighs the benefit.

00:11:33.670 --> 00:11:42.420
Andre Bruintjies
Thank you and then the next question is again one of those where I've got multiple factors.

00:11:56.910 --> 00:12:03.060
Andre Bruintjies
The question is, in your opinion, how would the following affect big data adoption within the organization? And the factors that I have is that reliability or suitability, IT, infrastructure, data security and privacy top management organizational size of got finance, organizational processes, organizational strategy. competition. Customers and vendors.

00:12:29.770 --> 00:12:33.220
P3:
So, so do you want me to rank it?
00:12:34.770 --> 00:12:54.760
Andre Bruintjies
We don't have to actually a bit of detail on how you think these affect the adoption of big data, but we don't have to discuss all of these.

00:12:55.820 --> 00:12:56.320
P3:
OK.

00:12:58.220 --> 00:13:28.260
P3:
OK, so I guess from my perspective in terms of big data adoption, which your kind of don't really have here. So maybe it's perhaps finance because it comes down to, you know, is there a relevant business case first of all to implement the big data solution? So why does it make business sense, so I guess finance would be probably one of the biggest factors in terms of big Data adoption. I think the second one around that I mean, to be honest with you, you know. Data reliability and suitability. For as well as things like data security and privacy. Or even IT infrastructure? I mean, all of those would get assessed. I don't think that these is a based on big data that we've seen. What the negative factors that have stopped any big data projects? So, I don't think those are huge challenges to be honest with you, I don't think the datasets are typically in such a bad state that it's not usable. Top management and organizational size., once again, I think that you know if you have a valid business case which the business can understand the value for, you'll get the buy-in and so perhaps top management and organizational sizes is business case relevant. And then you know, I think the rest of it, customers you know, and competitions probably really important. Yeah, I think you know for big data I think it is quite a bit of a fight as too. who has the access to the database today to so? So that probably doesn't play a role. And I don't think the rest to be honest.

00:15:10.540 --> 00:15:11.030
Andre Bruintjies
OK.
00:15:12.310 --> 00:15:18.920
Andre Bruintjies
OK, thank you very much. I'm the last question I have is what would the role of big data in your organization be in terms of service delivery and how new programs are being rolled out by organization?

00:15:35.560-->00:15:51.470
P3:
So I think if we think about, you know a city, I think that we'll see, probably cities in general as well as the OrganisationXYZ, move more into the IoT space into the future. And I think since the data is going to bring more and more kind of elements of big data across the organization going into the future for analysis. We, as an organization have put together a data strategy along with a data science team that I think is looking to think about that future and how it will incorporate. Big data analysis as part of kind of broader kind of proactive scenario. Thinking into the city or its service delivery and maintenance needs, and very much in the asset's spaces for us and maintenance. Besides that, I think also from a spatial point of view in terms of its delivery framework. I think more and more cities will start to use big data or for framing thinking of its kind of strategic framework as well as its spatial frameworks for the future. Yeah, so I think that's how the OrganisationXYZ as an organisation think about big data. It'll probably be part of kind of data programs that we will run into the organization, ties into kind of data analytics, artificial intelligence and big data will be part of this theming. But as I've started right from the beginning. You know, I think we'll be very business case and kind of use case relevant in terms of how our adoption of big data. So, we will adopt it in use cases where it makes sense for us, but I don't think it's a big program and just to bring data into the organization.

00:17:44.460 --> 00:17:44.860
Andre Bruintjies
OK.
00:17:46.100 --> 00:17:49.520
Andre Bruintjies
Thank you very much that is all the research questions that I have.

Appendix 3: Ethical Clearance

UNIVERSITY *of the* WESTERN CAPE

60 YEARS *of hope, action & knowledge*

14 January 2021

Mr A Bruintjies
Information Systems
Faculty of Economics and Management Sciences

Ethics Reference Number: HS 18/9/13

Project Title: Understanding the factors that influence the adoption of big data at a government department in the Western Cape

Approval Period: 26 November 2020 – 26 November 2023

I hereby certify that the Humanities and Social Science Research Ethics Committee of the University of the Western Cape approved the methodology and ethics of the above mentioned research project.

Any amendments, extension or other modifications to the protocol must be submitted to the Ethics Committee for approval.

Please remember to submit a progress report by 30 November each year for the duration of the project.

The Committee must be informed of any serious adverse event and/or termination of the study.

Ms Patricia Josias
Research Ethics Committee Officer
University of the Western Cape

Director: Research Development
University of the Western Cape
Private Bag X 17
Bellville 7535
Republic of South Africa
Tel: +27 21 959 4111
Email: research-ethics@uwc.ac.za

NHREC Registration Number: HSSREC-130416-049

FROM HOPE TO ACTION THROUGH KNOWLEDGE.

Appendix 4: Permission Letter

Andre Bruintjies
14 Strelitzia Street
Belhar, Western Cape, 7493
+27 71 068 4471
andre.bruintjies@gmail.com

15 October 2019

Dear Sir/Madam

RE: Permission to conduct research at your organization

My name is Andre Bruintjies, I am a Master's student in Information Systems at the University of the Western Cape in Cape Town. As part of my studies I will be undertaking a research project, which focus on investigating "The factors that influence Big Data (BD) adoption at a government department in the Western Cape".

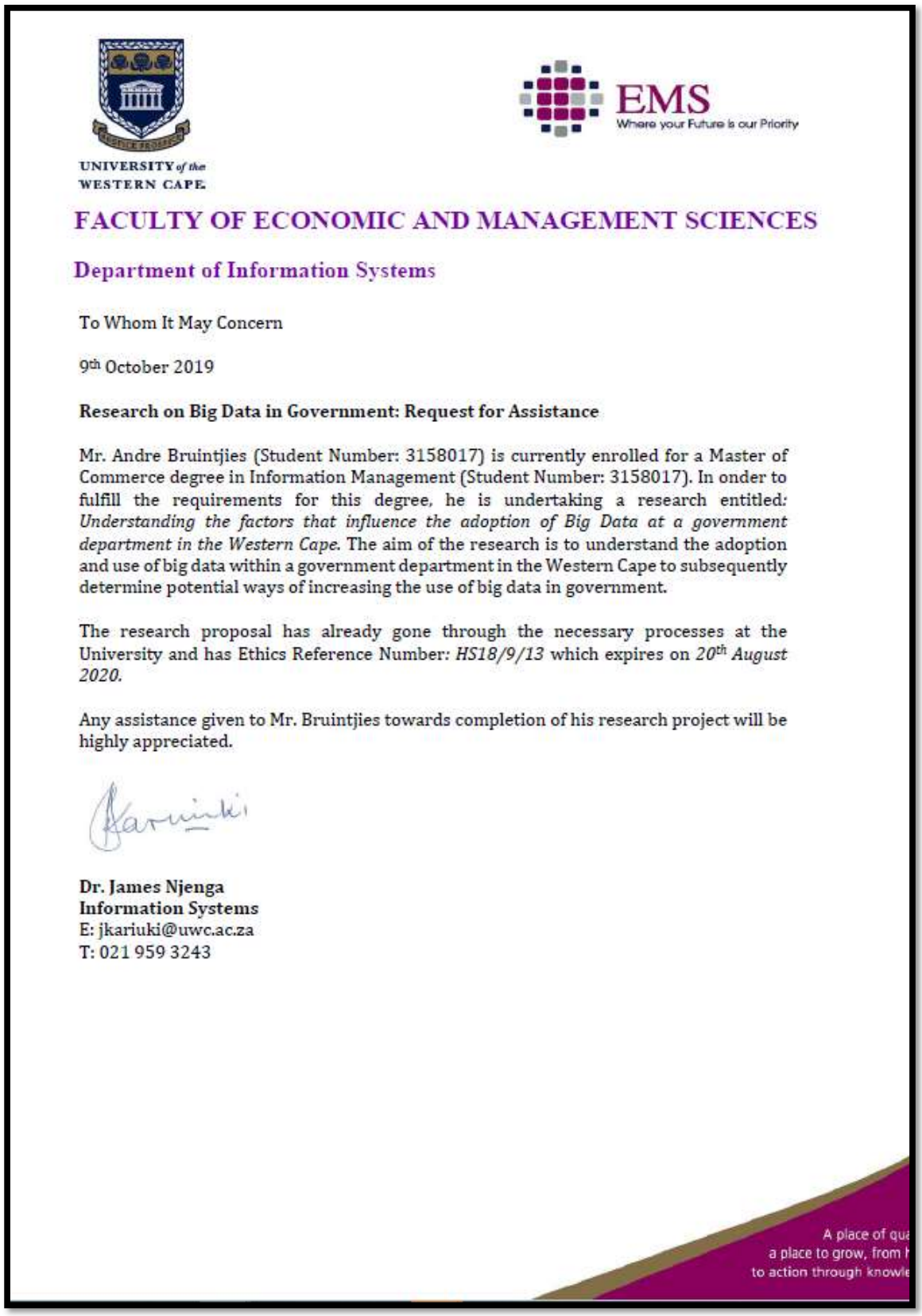
During this research project, I will be conducting conduct an interpretive case study to explore factors influencing the adoption of BD at your organization. The conceptual framework developed in this research may serve as a guideline to government departments that experience challenges in the adoption of BD or to those that are interested in the adoption of such technology.

I would therefore like to request your permission to conduct this research at your organization. The data collection and handling will be done in accordance with the ethical principles of UWC, which guarantee full confidentiality of the interviewees, including the possibility of the participants to withdraw at any point during the study.

Please contact myself or my research supervisor should any additional information be required on the details below:
Andre Bruintjies, +2771 068 4471, andre.bruintjies@gmail.com
Dr. James Njenga, +27 21 959 3680, jkariuki@uwc.ac.za

Yours sincerely
Andre Bruintjies

Appendix 5: Research Reference



UNIVERSITY of the
WESTERN CAPE

EMS
Where your Future is our Priority

FACULTY OF ECONOMIC AND MANAGEMENT SCIENCES

Department of Information Systems

To Whom It May Concern

9th October 2019

Research on Big Data in Government: Request for Assistance

Mr. Andre Bruintjies (Student Number: 3158017) is currently enrolled for a Master of Commerce degree in Information Management (Student Number: 3158017). In order to fulfill the requirements for this degree, he is undertaking a research entitled: *Understanding the factors that influence the adoption of Big Data at a government department in the Western Cape.* The aim of the research is to understand the adoption and use of big data within a government department in the Western Cape to subsequently determine potential ways of increasing the use of big data in government.

The research proposal has already gone through the necessary processes at the University and has Ethics Reference Number: *HS18/9/13* which expires on *20th August 2020.*

Any assistance given to Mr. Bruintjies towards completion of his research project will be highly appreciated.

Dr. James Njenga
Information Systems
E: jkariuki@uwc.ac.za
T: 021 959 3243

A place of qua
a place to grow, from h
to action through knowle