

**Diffusion of data governance and data-driven decision making**

A case study of data governance implementation in a selected retail company in  
South Africa.

**Student Number:** 19386754

A research project submitted to the Gordon Institute of Business Science, University of Pretoria, in partial fulfilment of the requirements for the degree of Master of Business Administration.

1 December 2022

## **Abstract**

The amount data generated and collected by organisations is growing at an accelerated rate. Retail organisations are in need of interventions to better manage their data to leverage it for a competitive advantage. Data governance has become vital for organisations that are looking for derive the maximum strategic value from their data assets. Organisations especially in retail are in need of guidance on the practical implementation of data governance.

The study leverages the diffusion of innovation theory to explore the implementation of data governance within a selected retail organisation in South Africa. Through the review of the literature, it was established that there has been limited research in the implementation of data governance and the implementation of data governance in practice still remains a challenge for many organisations. This study was therefore conducted to understand elements that can facilitate the implementation of data governance using the diffusion of innovation theory.

A qualitative research methodology was adopted by the study due to its exploratory nature. Semi-structured interviews were conducted with senior employees at the selected retail organisation in the process of gathering data for the study.

The results of the study revealed that the diffusion of innovation theory can be used to facilitate the implementation of data governance in a retail organisation. Additionally, a data governance implementation framework adapted from the diffusion of innovation theory was proposed which can support the implementation of data governance.

## **Keywords**

Big Data, Data governance, data-driven decision-making, Diffusion of Innovation theory, Data governance implementation

## **Declaration**

I declare that this research project is my own work. It is submitted in partial fulfilment of the requirements for the degree of Master of Business Administration at the Gordon Institute of Business Science, University of Pretoria. It has not been submitted before for any degree or examination in any other University. I further declare that I have obtained the necessary authorisation and consent to carry out this research.

2 December 2022

---

Busisiwe Tshabalala

# Table of Contents

<b>CHAPTER 1: RESEARCH PROBLEM AND PURPOSE STATEMENT .....</b>	<b>7</b>
1.1 BACKGROUND OF THE STUDY.....	7
1.2 CONTEXT OF THE STUDY .....	8
1.3 PROBLEM STATEMENT .....	10
1.4 SIGNIFICANCE OF THE STUDY .....	12
1.5 DELIMITATIONS .....	13
1.6 CONCLUSION .....	14
<b>CHAPTER 2: THEORY AND LITERATURE REVIEW .....</b>	<b>15</b>
2.1 BIG DATA AND ANALYTICS .....	15
2.2 APPLICATION OF BIG DATA IN THE RETAIL INDUSTRY.....	16
2.3 DATA-DRIVEN DECISION MAKING.....	17
2.4 DATA GOVERNANCE.....	18
2.4.1 <i>What is data governance?</i> .....	18
2.4.2 <i>Why organisations need Data Governance?</i> .....	20
2.4.3 <i>Benefits of Effective Data Governance</i> .....	20
2.4.4 <i>Data Governance Frameworks</i> .....	21
2.5 THEORETICAL LENS - DIFFUSION OF INNOVATION THEORY .....	28
2.5.1 <i>Innovation</i> .....	28
2.5.2 <i>Communication Channel</i> .....	30
2.5.3 <i>Time</i> .....	31
2.5.4 <i>Social System</i> .....	32
2.5.5 <i>Innovation Process in Organisations</i> .....	33
2.6 CONCLUSION .....	35
<b>CHAPTER 3: RESEARCH QUESTIONS .....</b>	<b>37</b>
3.1 RESEARCH QUESTIONS.....	37
3.2 CONCLUSION .....	39
<b>CHAPTER 4: RESEARCH DESIGN AND METHODOLOGY .....</b>	<b>40</b>
4.1 RESEARCH DESIGN .....	40
<i>Purpose of research design</i> .....	40
<i>Research philosophy (paradigm)</i> .....	40
<i>Research strategy</i> .....	41
<i>Methodological choice</i> .....	42
<i>Time horizon</i> .....	43
4.2 PROPOSED RESEARCH METHODOLOGY.....	43

4.2.2 Population .....	43
4.2.3 Unit of analysis.....	43
4.2.4 Sampling method and size .....	44
4.2.5 Measurement Instrument .....	44
3.2.6 Data gathering process .....	45
4.2.7 Analysis approach .....	45
4.2.7 Ethical considerations .....	47
4.2.8 Quality controls.....	47
4.2.8 Limitations .....	48
4.3 CONCLUSION .....	49
<b>CHAPTER 5: RESULTS .....</b>	<b>50</b>
5.1 INTRODUCTION .....	50
5.2 OVERVIEW OF PARTICIPANTS.....	50
5.3 INTERVIEW PROCESS .....	51
5.4 SATURATION ANALYSIS.....	51
5.5 RESEARCH QUESTION 1: PERCEIVED NEEDS THAT DRIVE THE IMPLEMENTATION OF DATA GOVERNANCE .....	52
5.5.1 The role of data in the organisation.....	52
5.5.2 The perception of data governance .....	53
5.5.3 The perceived need for data governance.....	55
5.5 RESEARCH QUESTION 2: PERCEIVED INNOVATION CHARACTERISTICS OF DATA GOVERNANCE .....	56
5.6 RESEARCH QUESTION 3: COMMUNICATION CHANNELS USED TO DISSEMINATE DATA GOVERNANCE. ....	62
5.7 RESEARCH QUESTION 4: IMPACT OF THE SOCIAL SYSTEM ON THE IMPLEMENTATION OF DATA GOVERNANCE.....	63
5.8 RESEARCH QUESTION 5: IMPACT OF DATA GOVERNANCE ON DATA-DRIVEN DECISION-MAKING .....	65
<b>CHAPTER 6: DISCUSSION OF FINDINGS .....</b>	<b>67</b>
6.1 INTRODUCTION .....	67
6.2 RESEARCH QUESTION 1: PERCEIVED NEED FOR DATA GOVERNANCE .....	67
6.3 RESEARCH QUESTION 2: PERCEIVED INNOVATION CHARACTERISTICS OF DATA GOVERNANCE .....	68
6.3.1 Relative advantage .....	69
6.3.2 Compatibility.....	69
6.3.3 Complexity .....	70
6.3.4 Trialability .....	70
6.3.5 Observability .....	71
6.4 RESEARCH QUESTION 3: COMMUNICATION CHANNELS USED TO DISSEMINATE DATA GOVERNANCE .....	71
6.5 RESEARCH QUESTION 4: IMPACT OF SOCIAL SYSTEM ON THE IMPLEMENTATION OF DATA GOVERNANCE .....	72
6.7 RESEARCH QUESTION 5: IMPACT OF DATA GOVERNANCE ON DATA-DRIVEN DECISION-MAKING .....	73
6.8 CONCLUSION .....	74

<b>CHAPTER 7: CONCLUSION .....</b>	<b>76</b>
7.1 RESEARCH BACKGROUND AND OBJECTIVES .....	76
7.2 PRINCIPAL FINDINGS .....	76
7.2.1 <i>Perceived needs for data governance</i> .....	76
7.2.2 <i>Perceived innovation characteristics for data governance</i> .....	77
7.2.3 <i>Communication channels used to disseminate data governance</i> .....	78
7.2.4 <i>Impact of social system on the implementation of data governance</i> .....	78
7.2.5 <i>Impact of data governance on data-driven decision-making</i> .....	78
7.3 PROPOSED FRAMEWORK FOR DATA GOVERNANCE IMPLEMENTATION .....	79
7.4 IMPLICATIONS FOR MANAGEMENT AND OTHER RELEVANT STAKEHOLDERS .....	81
7.5 LIMITATIONS OF THE RESEARCH .....	81
7.6 CONTRIBUTIONS AND SUGGESTIONS FOR FUTURE RESEARCH .....	82
<b>REFERENCES .....</b>	<b>83</b>
<b>APPENDIX 1: INFORMED CONSENT LETTER.....</b>	<b>94</b>
<b>APPENDIX 2: INTERVIEW GUIDE .....</b>	<b>94</b>
<b>APPENDIX 3: ETHICAL CLEARANCE LETTER.....</b>	<b>96</b>
<b>APPENDIX 4: CODE BOOK DEVELOPED FROM LITERATURE .....</b>	<b>96</b>

## **Chapter 1: Research problem and purpose statement**

This section discusses the background of the study and outlines the research problem as well as the purpose statement.

### **1.1 Background of the study**

In today's business world, leaders can no longer ignore the strategic value of data as an asset in gaining a competitive advantage (Alhassan et al., 2019a; Benfeldt et al., 2020; Santoro et al., 2019). The use of data has grown more sophisticated, which has brought about new and complex data demands necessitating different ways of governing and managing the data (Al-Ruithe et al., 2019; Al-Badi et al., 2018a). Consequently, data governance has become vital in deriving value from data as an asset (Alhassan et al., 2019b; Al-Ruithe et al., 2019; Abraham et al., 2019), and increasing the decision-making effectiveness of management in organisations (Ababneh & Aga, 2019; Jimenez et al., 2019).

Data governance is a broad concept with various definitions in academia and practice (Alhassan et al., 2018; DAMA International, 2017; Zhang et al., 2022). However, there is consensus among multiple scholars of data governance that its main objective is to drive the consideration of data as a strategic asset for the business (Alhassan et al., 2018; Zhang et al., 2022). The report defines data governance as the use of decision-making authority and control over data, which involves the preparation, monitoring, and application of standards and procedures to turn data into a strategic asset. (Brous et al., 2020; Zhang et al., 2022). Therefore, data governance is essential for creating business value from data and curbing risks that come with data-driven initiatives, as it can help organisations view data as a competitive asset, thus maximising the value that can be derived from their data (Brous & Janssen, 2020). It has also been demonstrated that data-driven decision-making can provide multiple benefits for organisations (Pereira & Frazzon, 2021), especially those operating in the retail industry (Ying et al., 2021). The purpose of the study is to explore the implementation of data governance in a retail organisation using the Diffusion of Innovation theory stages.

The Diffusion of Innovation (DOI) theory has been used to study the implementation of various new interventions and innovations. According to (Dearing & Cox, 2018), “diffusion is a social process that occurs among people in response to learning about an innovation” (p. 183). Moreover, an innovation based on the DOI theory is an idea, practice, product or service which is perceived as new to the people in the particular social system (Rogers, 2003). For this study, innovation refers to the introduction and implementation of data governance in the organisation, and diffusion is the process of communicating the new data governance standards and practices to the organisation’s social system. For example, the implementation of complex healthcare innovations (Dryden-Palmer et al., 2020), the introduction of mobile devices in nursing schools to enhance the learning process (Doyle et al., 2014), the implementation of a pharmacy residency program (Wang et al., 2021) and the implementation of an Enterprise Resource Planning Systems (Akca & Ozer, 2014). Therefore, the above studies are similar in nature as they are investigating the implementation of a new idea into an organisation. Which makes the DOI theory relevant and useful in understanding complex interventions into existing systems and structures which is the case with the implementation of data governance. The DOI theory will be further discussed in the following chapter.

## **1.2 Context of the study**

Retail businesses have been impacted by rapid technological advancements and a spike in digitalisation and the big data (Kaur et al., 2020; Santoro et al., 2019; Shamim et al., 2020). This has drawn attention to the value of data within an organisation (Seetharaman et al., 2016). As a result, the expansion of e-commerce has fuelled drastically altering consumer purchasing patterns and the direction of the retail sector (Ratchford et al., 2022). With the rise of omnichannel retailing, retailers are faced with new challenges and opportunities in the fulfilment services and information delivery to support consumers’ end-to-end shopping experiences using the various channels which include online and offline touchpoints (Momen & Torabi, 2021).

According to Griva et al. (2018) retailers have acknowledged the importance of data-driven decision-making in improving their decision quality and better serving their customers. Furthermore, retailers frequently employ data analytics to improve customers’ convenience and personalization to expand their customer base both



physically and digitally (Kaur et al., 2020). Pereira and Frazzon (2021) add that retailers who employ data-driven decision-making can make quality decisions faster based on data insights, which is imperative in the rapidly changing retail market. However, to reap the benefits that come with data-driven decision-making, organisations need to have solid data management capabilities (Hossain et al., 2020a) which are influenced by data governance (Alhassan et al., 2018). Therefore, the organisations' ability to access, analyse, manage and govern large data volumes becomes increasingly important, as retailers look to advance their business performance imperatives (Ying et al., 2021).

Despite the increased awareness of the importance of data governance, it remains an under-researched topic (Abraham et al., 2019; Al-Ruithe et al., 2019). This necessitates more attention from academics and practitioners as the implementation and monitoring of data governance remains a challenge for many organisations today (Alhassan et al., 2019a, 2019b; Brous & Janssen, 2020). In addition, Abraham et al. (2019) states that organisations today still grapple with providing compelling use cases that can be used to link data governance to organisation value creation which is one of the challenges that deters organisations from prioritising data governance.

The context of this study is the retail industry in South Africa. Retail sales in the country decreased by 2.5 percent in June 2022 (Statistics South Africa, 2022; Trading Economics, 2022) which was the largest decline in retail activity since January 2021 demonstrating the impact of higher inflation and interest rates on consumer spending (Trading Economics, 2022). Therefore, the retail industry in South Africa has been severely impacted by the current economic climate requiring retailers to be dynamic in their bid to remain competitive. Data-driven business decisions play a critical and strategic role in sustaining a firm's competitive advantage (Pereira and Frazzon, 2021). The South African retail industry has seen a surge in the use of loyalty cards which has enabled personalisation and target marketing. However, this has raised concerns about how the data will be processed, and stored and who has access to it (Mnoney & Van Belle, 2016), since the legislation of the Protection of Personal Information Act (POPIA) in July 2020 (POPIA, 2022). Therefore, data governance plays a critical role in ensuring that decisions are made

using data that is of high quality and accurate as well as governing compliance with legislation.

### **1.3 Problem Statement**

Making decisions that are accurate, timeous and of good quality is no longer just essential but a matter of survival for many organisations as they navigate the complex and competitive business environments they operate in (Akter et al., 2019). Scholars have argued that data-driven decision-making impacts firm performance positively (Shamim et al., 2020, Yu et al., 2021) and organisations that are data-driven are more productive (Shamim et al., 2020). Retail businesses in South Africa are faced with a myriad of challenges which include the lack of local economic growth, the consequent pressure on the consumers and business profits, globalisation which exposes them to global competitors as well as an unstable political and changing regulatory environment (PWC, 2021). In addition to the lingering effects of Covid-19, the crisis in Ukraine has led to increased fuel prices globally which have negatively impacted South African consumers resulting in increasing prices and the implementation of higher interest rates by the South African Reserve Bank (NielsenIQ, 2022). Therefore, South African retail businesses must make drastic shifts to survive and remain competitive by leveraging their data assets in making data-informed decisions. Aloysius et al. (2018) state that potential benefits of leveraging data and data-driven decision-making in retail include the improvement of sales and the capability to plan operations, the improvement of supply chain efficiencies, and the capability to dynamically respond to fast-changing environments.

Scholars have found that data-driven decision-making and performance rely on the information being of good quality and accurate (Shamim et al., 2020, Chatterjee, Chaudhuri & Vrontis., 2021). Martijn et al. (2015) have demonstrated that data quality deteriorates when there are no clear data governance policies, rules and controls defined within an organisation about who is accountable and responsible for which data. Therefore, data that is poorly governed may lead to profit losses as incomplete and incorrect information may lead to strategic disadvantages, inefficiently organised business processes as well as pose a risk of being non-compliant with data regulations which may have severe financial implications (Martijn et al., 2015; Mikalef

et al., 2020; Zhang et al., 2022). Data governance has become central to running a successful business, as data increasingly influences operational and strategic decisions (Alhassan, Sammon & Daly., 2019). Studies on the implementation of data governance are scarce, studies include one on the critical success factors of data governance using a telecommunication case study (Alhassan et al., 2019a). There was also a study found on the management of big data in the retail industry in Singapore (Ying et al., 2021) however the focus of this study was more on the management of big data as opposed to data governance.

Research on data governance as a means of deriving value from data resources is increasingly gaining momentum, especially with the emphasis on its strategic relevance in supporting business functions in the organisation (Al-Ruithe et al., 2019; Janssen et al., 2020; Lis et al., 2022). However, there is a gap in understanding the organisation and industry-specific factors that influence the implementation of data governance in the organisation (Abraham et al., 2019; Al-Ruithe et al., 2019; Lis et al., 2022). Furthermore, there is a gap in research on the practical implementation of data governance that provides actionable direction for organisations (Alhassan et al., 2018, 2019a; Benfeldt et al., 2020; Karkošková, 2022). Therefore, the implementation of data governance remains a challenge for many organisations (Alhassan et al., 2019a).

There is no single approach to the implementation of data governance for all organisations (Al-Ruithe et al., 2018). The successful implementation of data governance requires a business perspective and should contribute to organisational goals (Alhassan et al., 2019a; Benfeldt et al., 2020; Jimenez et al., 2019). Effective data governance can assist organisations in establishing accountabilities, keeping scope and focus, achieving clarity, increasing trust in using organisational data, and defining measurable success (Al-Ruithe et al., 2018; Alhassan et al., 2019b). Furthermore, the implementation of a data governance programme is a continuous operational discipline which requires consistent collaboration between the organisation's top management and the employees (Janssen et al., 2020; Jimenez et al., 2019).

The introduction of data privacy and protection laws has increased the urgency for organisations to firmly manage their data: knowing what data is stored, where it is

stored, ensuring the safety of the data and how the data is being used throughout the organisation (Abraham et al., 2019). South Africa legislated the Protection of Personal Information Act (POPIA) on the first of July 2020 indicating the increased importance of governing how data is collected, processed and used by organisations (POPIA, 2022). However, it is important to note that data governance has evolved from its perceived predominant role of assuring compliance and ensuring data quality to a strategic asset for competitive advantage and data-driven decision-making and innovation (Lis et al., 2022).

Data is ubiquitous, and businesses are employing a variety of channels to gather it fast and broadly (Abraham et al., 2019; Janssen et al., 2020) which makes the data inconsistent, drastically degrading data quality and subsequently leads to issues with data management risks (Zhang et al., 2022). These risks create a distrust of the data among decision-makers and business managers (Brous & Janssen, 2020; Zhang et al., 2022). Data governance can assist in addressing these challenges by ensuring that the data is fit for use by the data consumers and is available at the right time in the correct format (Benfeldt et al., 2020). It is also important to note that in 2020, the State Council of China officially defined data as the fifth factor of production, in addition to the four (capital, land, labour and technology) (Chen, 2022; Shijia & Jia, 2020). This further demonstrates that data governance has started to receive significant attention in national policy, management practice and theoretical study (Benfeldt et al., 2020; Zhang et al., 2022). Therefore, understanding the diffusion of data governance in organisations and exploring the practical implementation for retailers in South Africa in the era of Big Data will better enable them in understanding the data governance mechanisms that are needed to ensure that they harvest the full benefits of data-driven decision-making while ensuring that the data is governed effectively.

#### **1.4 Significance of the study**

Using Big Data, organisations have discovered a way to learn insightful information about themselves and the goods and services they offer. In addition, this information has assisted organisations in realising prospects for better and novel products and services, as well as enhanced decision support (Mneney & Van Belle, 2016). Decisions made using data analytics in business are now crucially important from a

strategic perspective for the company to maintain its competitive advantage. Businesses now have opportunities to take swift and thorough decisions based on insights about their customers and operations, thanks to big data and analytics (Pereira & Frazzon, 2021). The decisions made should be based on high-quality data that is well organised. Therefore, the companies' willingness to adopt data-driven decision-making also increases their need for data governance (Jia et al., 2015). However, the majority of organisations continue to struggle with the implementation and adoption of data governance (Alhassan et al., 2019a). This study aims to explore the implementation of data governance through a case study in one of the big retail organisations in South Africa. The study uses the diffusion of innovation theory to understand the characteristics of data governance as an innovation in the organisation, the communication channels used to communicate data governance, the process of adoption (innovation process) as well as the influence that the social system has on the implementation.

According to (Benfeldt et al., 2020) literature on the practical implementation of data governance is limited giving organisations minimal guidance that can assist in putting data governance into practice. The study will add to the body of knowledge insights on the practical implementation of data governance and an understanding of the challenges that hinder the implementation as well as factors that enable the successful implementation of data governance. The study will provide insights into examining the implementation of data governance that practitioners can use as a guide to implement data governance successfully. Since data governance is still in its formative stages as a concept, findings from this study will add to the body of knowledge and could form a reference point for future research. There was no evidence of the diffusion of data governance found in the literature, suggesting the need for a study such as this one.

### **1.5 Delimitations**

The scope of the study is limited to one selected case study organisation in the retail industry in South Africa. The study will be exploring the implementation of data governance using the DOI theory as a theoretical lens. The study will not review other theories outside the DOI theory. The study will not be evaluating data governance frameworks but how an organisation can go about implementing data governance.

## **1.6 Conclusion**

Data is increasingly being viewed as an asset for many organisations due to the strategic value it can bring to the organisation (Benfeldt et al., 2020). Specifically, retailers have realised the value of data-driven decision-making in enhancing the quality of their decisions and better serve their customers (Griva et al., 2018) which will result in increased business performance.

However, data governance is vital in ensuring that organisations can derive value from their data assets (Brous et al., 2016; Brous & Janssen, 2020). It has been highlighted that there are limited literature studies on the implementation of data governance (Benfeldt et al., 2020) and that organisations still struggle with the implementation of data governance (Alhassan et al., 2019a). The study aims to explore the implementation of data governance within a retail organisation in South Africa.

There are a total of seven chapters in this research project with the current chapter being chapter 1. The following chapter 2 provided a review of the literature relevant to the study. Next chapter 3 will introduce the research questions formulated from the literature reviewed in chapter 2. Chapter 4 outlines the methodology that was used in the execution of the study. Chapter 5 presents the results of the study and chapter 6 discusses the results. Lastly, chapter 7 concludes the research with the key findings from the study, its contribution to academia, the recommendations to managers, the limitations of the study and suggestions for future research.

## **Chapter 2: Theory and Literature Review**

The following section will discuss data governance, data-driven decision and the theoretical lens: the diffusion of innovation theory.

### **2.1 Big Data and Analytics**

Big Data can be differentiated from regular data sets due to its distinctive features. Volume, variety and velocity were the three Vs that were initially used to describe the big data (Seetharaman et al., 2016) however this has expanded to include six Vs: volume, variety, velocity, veracity, value and variability (Akter et al., 2019). Big data's volume attribute is a representation of the sheer volume or size of the data. To improve decision-making, businesses collect enormous amounts of data from dynamic, heterogeneous and ubiquitous resources and devices (Acharya et al., 2018). These various sources make it possible to gather different types of data which include structured, semi-structured, unstructured, multi-structured and streaming data (Ying et al., 2021).

Another distinctive characteristic of big data is velocity which is the rate at which data is generated. Veracity is a term used to describe the ambiguity and unreliability of several huge data formats, large data sets' complexity, consistency, and anonymity can make the data unreliable (Ying et al., 2021). The data flow rates may vary which is brought on by the inconsistencies in the velocity of the large data sets which relates to the variability (Akter et al., 2019). And lastly, the value of big data is the undeniable benefit a business derived from the insights it gains through the big data (Akter & Wamba, 2016; Hossain et al., 2020b; Maheshwari et al., 2021).

Big Data Analytics (BDA) is a holistic process that comprises the gathering, analysing, utilising, and interpreting data for multiple functional divisions to gain actionable insights, produce business value, and establish competitive advantage (Akter et al., 2019; Akter & Wamba, 2016; Shamim et al., 2020). BDA is becoming more widely adopted and has emerged as a critical component for improving business processes and outcomes (Gawankar et al., 2020). Shamim et al. (2020) concur that BDA is a crucial aspect in improving efficiency and effectiveness with strategic and operational implications.

Akter and Wamba (2016) argue that even though big data can be used to increase business value, there is always the potential that it will produce redundant, erroneous, and duplicate data, which could make decisions more difficult to make. Schroeck et al. (2012) contend that poor data quality and insufficient data governance are major challenges in the effective use of BDA. It is therefore important to emphasize that the use of even the most sophisticated analytics would be useless if the wrong data were utilized, or if the data used were of low quality. In the era of big data, organisations need to have effective data governance in place which will assist in ensuring that the business can leverage the big data collected.

## **2.2 Application of Big Data in the Retail Industry**

Businesses are dealing with substantial changes in how they manage their operations, customers and business models as a result of the data-driven revolution in the business management (Acharya et al., 2018; Raguseo, 2018). This is caused by the emergence of new opportunities brought on by the rapid advancement in big data technologies and the enormous amount of data that businesses can collect from various sources to inform strategies and decision-making which promises an improvement in the overall business performance (Maheshwari et al., 2021).

The retail industry is recognised as one of the biggest industries in the world and is currently seeing a rapid rise in retail purchases made via mobile commerce (m-commerce) and electronic commerce (e-commerce); this is due to the introduction of high-speed internet connections, improvements in smartphones and online-related technology, improvements in the product lines of e-commerce companies, a selection of delivery options, and better payment options (Akter & Wamba, 2016; Seetharaman et al., 2016). Retail businesses have leveraged and adopted the use of big data and have seen an improvement in their decision-making performance which has resulted in increased business performance (Gawankar et al., 2020)

Data generation has accelerated due to digitisation (Shamim et al., 2020), in the retail industry there has been a rise in the demand for effective real-time systematic and serialised identification of manufactured goods example using RFID-enabled tags attached to individual items (Acharya et al., 2018). According to (Ying et al., 2021),



the value of BDA in the retail industry can be realised through the analysis of big data to assist decision-makers with: customer targeting which is useful for noting individual customer behaviour to assess customer satisfaction, inventory management to help improve stock forecasting and predict changes in demand and lastly in-store behaviour and customer sentiments to examine market demands on price or product changed, which may inform pricing decisions, store layout, shelf location as well as product mix. Additionally, e-commerce today is brimming with numerous large data sets that are being leveraged to solve business difficulties and due to the requirement to stay competitive, it is among the most rapid BDA adopters (Akter & Wamba, 2016).

### **2.3 Data-Driven Decision Making**

Data-driven decision-making involves decisions that are made using insights derived from the analysis of data instead of relying on intuition and experience (Shamim et al., 2020, Jia et al., 2015). Troisi et al. (2020) emphasize that data-driven decisions making considers data as a strategic resource and requires leadership to play an active role in fostering a data-driven culture and carefully consider data management at each step of decision-making.

Organisations in the era of big data and digitization are increasingly looking for ways to best leverage data for the decision-making (Shamim et al., 2019), as studies have shown that data-driven decision-making positively influences organisational performance (Shamim et al., 2020, Yu et al., 2021) and organisations that are data-driven are more productive and innovative (Chatterjee et al., 2021). To further enhance data-driven decision-making, organisations are introducing self-service reporting and analytics, thus creating a need to ensure that there is a consistent understanding of the data in all areas of the organisation (Abraham et al., 2019).

In contrast, Ababneh & Aga (2019) contend that the quality of information used, and the speed of the decision driven by the availability of data at the right time enhances data-driven decision-making effectiveness. Shamim et al. (2020) add that the quality of information can be assured through data governance which plays an essential role in creating trust, and support and providing knowledge resources. It has also been

found that organisations with well-developed data governance programs are better equipped to be data-driven and can adjust to new initiatives that drive data-driven decision making such as the data science (Brous & Janssen, 2020).

Shamim et al. (2020) advocates that big data enables organisations to make decisions that are data-driven which consequently enhances decision-making performance. A study conducted in India by Gawankar et al. (2020) found that the adoption of big data improved the decision-making performance of retail organisations in India which had a positive impact on their profits. Data-driven decision-making contributed to improving their supply chain processes, logistics, inventory control and the reduction of costs (Gawankar et al., 2020). Conversely, big data presents numerous challenges to organisations such as privacy and security of personal information leakages and the monitoring of customers' lives (Al-Badi et al., 2018). This is particularly challenging for South African organisations as well as the implementation of the Protection of Personal Information Act (POPIA).

## **2.4 Data Governance**

### **2.4.1 What is data governance?**

Researchers and practitioners often cannot differentiate between data governance and data management or use them interchangeably (Al-Ruithe et al., 2019). Therefore, for this study, it is important to clarify that data governance is not data management even though the two are not mutually exclusive. According to DAMA International (2017), data governance is “the exercise of authority and control (planning, monitoring, and enforcement) over the management of data assets” (p. 67). Data governance can also be seen as a framework that outlines the decision rights and accountabilities concerning an organisation's data assets (Alhassan et al., 2018, Al-Ruithe et al., 2019), which determines the decisions made about the data, in what way the decisions are made, and who within the organisation has the rights necessary to make the decisions (Abraham et al., 2019).

In contrast, data management according to DAMA International (2017) “is the development, execution and supervision of plans, policies, programs and practices that deliver, control, protect, and enhance the value of data and information assets

throughout their lifecycles” (p. 17). Therefore, data management is focused on defining the data elements and the decisions on how it is structured, stored and moved whereas data governance is the elevated strategic planning and control over how the data is managed (Al-Ruithe et al., 2019). The significance of data control is becoming increasingly clear to enterprises today and there is a rising focus on data governance to assure data security, high quality and improved data management to boost confidence in using the organisation’s data (Jimenez et al., 2019). The study is exploring the implementation of data governance in a retail business and the impact the implementation has on data-driven decision-making.

Cheong and Chang (2007) assert that as data have no will nor intent on its own, tools and people shape the data and determine how it is collected and used, therefore, data governance is the governance of people and technology. Data being perceived as an asset has been seen as the main driver for data governance in organisations (Alhassan et al., 2018, Alhassan et al., 2019a). It has also become increasingly important to use data governance as a mechanism to encourage the strategic use of data in organisations (Lis et al., 2022). To encourage desirable behaviour, data governance policies, guidelines and standards need to be developed and implemented across the organisation and aligned with the organisation's mission, strategy, values, norms and culture (Al-Ruithe et al., 2019). DAMA International (2017) adds that data governance is most effective when it aligns with the organisation's strategy and is executed as an enterprise effort.

According to Al-Ruithe et al. (2019), efforts at governing data failed historically due to being driven exclusively by Information Technology (IT) which subjected it to inflexible processes and fragmented business activities that were done on a system-by-system basis. Data governance requires support and participation from both IT and business because even though the data is stored in IT systems, the data is created and updated by the business processes (Karkošková, 2022), therefore requiring collaboration across the business for success. According to Benfeldt et al. (2020) organisations, today still battle to recognise and quantify the value-creating potential of data governance which is a result of the overarching goal of data governance remaining abstract, ambiguous and perplexing for practitioners who are then unable to mobilise the collective action of data governance. Martijn et al. (2015) add that stipulating the outcomes of data governance can be extremely complex as

data governance involves many sector-specific factors that cannot be operationalized. Mikalef et al. (2020) stress that sound effective data governance should go beyond just being a system of tactics to derive value from data however, it should be a mechanism that can influence organisational behaviour and strengthen the organisation's analytical capabilities to drive data-driven decision-making.

#### **2.4.2 Why organisations need Data Governance?**

Business and IT leaders are increasing of the opinion that data is a valuable resource and an asset to the organisation, which implies that data governance is essential for effective data management (Cheong & Chang, 2007). Alhassan et al. (2019b) argue that the governing of data has become essential in operating a business successfully as the lack of confidence in the data can result in knowledge workers losing up to 50% of their time searching and validating data as opposed to when it is trusted it gets shared which can result in higher returns on data investments.

Organisations today view data governance as a crucial means of enhancing collaboration, and decision-making legitimacy, and accelerating decision-making (Zhang et al., 2022). However, Zhang et al. (2022) also acknowledge that organisations may lose out on data-driven opportunities and waste resources due in large part to vague data governance processes. According to a McKinsey report, businesses waste 30% of their time on non-value-added tasks due to inadequate data quality and availability which limits scalability and has a major impact on the productivity of employees across the organisation (Petzold et al., 2020).

Benfeldt et al. (2020) posits that robust data governance is needed to reap the value of data. Retailers need to comprehend the fundamental levers outside of information technology that will allow them to extract the most value possible from their data (Ying et al., 2021) and data governance is one of the fundamental levers that will allow them to extract value from their data assets.

#### **2.4.3 Benefits of Effective Data Governance**

Well-formulated and successfully implemented data governance programs can generate many benefits for organisations, which include better decision-making,

reduction in operational conflict and the protection of stakeholders' data needs (Al-Ruithe et al., 2019). Additionally, the consistent approaches to resolving data issues and data standards enable business processes to be repeatable as data is standard across the organisation which reduces costs and increases effectiveness (Al-Ruithe et al., 2019). Data governance also plays a vital role in fostering trust in data initiatives especially those used for decision-making such as data science and positively influences the adoption and use of their outcomes (Brous & Janssen, 2020). Alhassan et al. (2019b) add that a critical indicator of an effective data governance program is the availability and accessibility of data at the right time and in the correct format enabling decision-makers to make better decisions timeously.

According to Varshney and Allen (2020), effective data governance promotes data-driven decision-making, innovation, growth, and compliance with corporate policies and privacy regulations from multiple jurisdictions and additionally makes it easier for IT and data teams to collaborate effectively. In addition, they add that one of the main goals of data governance is good data quality as good data quality increases how well data is used generally and enables data-driven decision-making (Varshney & Allen, 2020).

According to (Janssen et al., 2020), data governance can have both direct and indirect benefits such as efficiency gains, a rise in revenue and market share, decreased risk, and lower costs incurred are all examples of direct benefits of data governance for business operations and the improvement of expenditure on information management projects, the perception of how information initiatives perform, and the improvement of trust in information products are all examples of indirect benefits of data governance. Risks can be decreased by lowering privacy violations, boosting data security, and lowering the possibility of facing legal and/or regulatory repercussions (Ababneh & Aga, 2019; Al-Ruithe et al., 2019).

#### **2.4.4 Data Governance Frameworks**

The earlier data governance framework was introduced by the Data Governance Institute (DGI) in 2004 in response to a growing demand for a method to categorise, organise and communicate intricate processes involved in making decisions about and acting on organisations' data (Data Governance Institute, n.d.). The DGI data

governance framework consists of ten universal components which are grouped into four areas: rules and rules of engagement, people and organisational bodies and processes (Thomas, 2006). This data governance framework was closely followed by one of the first scholarly attempts to propose a data governance framework by Wende (2007).

In contrast to the DGI data governance framework, Wende (2007) presented a data governance framework with three components focused on data quality management: roles, major decision areas and assignment of accountabilities. It belongs to the philosophy that combines technical and business-driven approaches to address quality management’s strategic and operational concerns. The primary contribution was to suggest a modular model for data governance that included a framework of roles, responsibilities, and decision-making areas and create a RACI chart to suggest a list of data quality roles (Wende, 2007) (see Figure 1)

**Figure 1: Wende’s Data Governance Model (Wende, 2007)**

Roles	Executive Sponsor	Data Governance Council	Chief Steward	Business Data Steward	Technical Data Steward	...
<b>Decision Areas</b>						
Plan data quality initiatives	A	R	C	I	I	
Establish a data quality review process	I	A	R	C	C	
Define data producing processes		A	R	C	C	
Define roles and responsibilities	A	R	C	I	I	
Establish policies, procedures and standards for data quality	A	R	R	C	C	
Create a business data dictionary		A	C	C	R	
Define information systems support		I	A	C	R	
...						

R – Responsible; A – Accountable; C – Consulted; I – Informed

In the same year, Cheong and Chang (2007) proposed another data governance framework drawing on Thomas’ earlier in 2006, which stated that a lack of authority and clear roles and responsibilities are contributors to stewardship failure (Thomas, 2006). They draw on their findings from a case study of a large utility company in Australia to propose a data governance structure and framework that focuses on

data management issues such as reactive data management, data assets that are difficult to access, problems aligning on data standards, and data improvement projects that are overrun and over budget - emphasizing the importance of collaboration between business and IT departments to support organizations. The components of the data governance framework consist of organisational bodies and policies, standards and processes and the data governance technology (Cheong & Chang, 2007) (see Table 1).

Their research emphasizes the significance of data governance structures, as well as policies and processes for effective data management but make no recommendations on how to achieve this. They do, however, highlight certain crucial success elements for data governance. Business intelligence (BI) is referred to by Wende (2007) and Cheong and Chang (2007) as one of the main operating business domains where data governance can have a significant impact on providing quality data to users, however, no further details or detailed instructions are provided on how to better support the BI or Big Data Analytics systems.

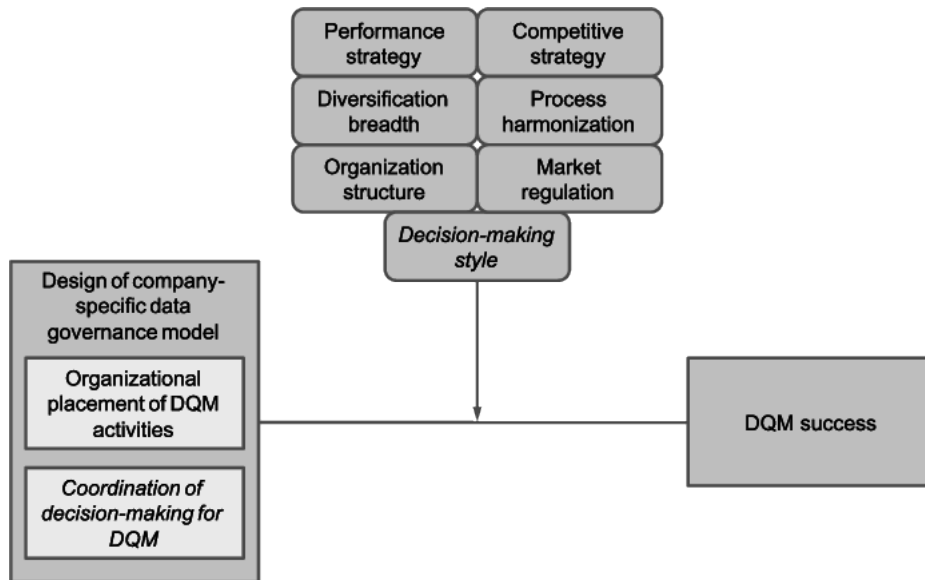
**Table 1: Cheong and Chang’s Components of Data Governance Framework (Cheong & Chang, 2007)**

Organisational Bodies and Policies	Standards and Processes	Data Governance Technology
<ul style="list-style-type: none"> <li>- Governance Structure</li> <li>- Data Custodianship</li> <li>- User Group Charter</li> <li>- Decision Rights</li> <li>- Issue Escalation Process</li> </ul>	<ul style="list-style-type: none"> <li>- Data Definition and Standard (Meta data management)</li> <li>- Third Party Data Extract</li> <li>- Metrics Development and Monitoring</li> <li>- Data Profiling</li> <li>- Data Cleansing</li> </ul>	<ul style="list-style-type: none"> <li>- Metadata Repository</li> <li>- Data Profiling tool</li> <li>- Data Cleansing tool</li> </ul>

In 2009, Weber et al. (2009) build on Wende's earlier data governance research by proposing a paradigm for enterprise-wide data quality management that addresses accountability and decision rights (see Figure 2). Nonetheless, they build on this by arguing that organisational structures are not one-size-fits-all and advocate for a flexible approach to data governance. Each business requires a unique data governance configuration that addresses a distinct set of risks. The study specifies these contingencies as firm size, structure, and decision-making style, claiming that

these variables could influence the model. They stated in their contingency approach that each organization should have a specific data governance configuration that meets a set of context considerations (Weber et al., 2009).

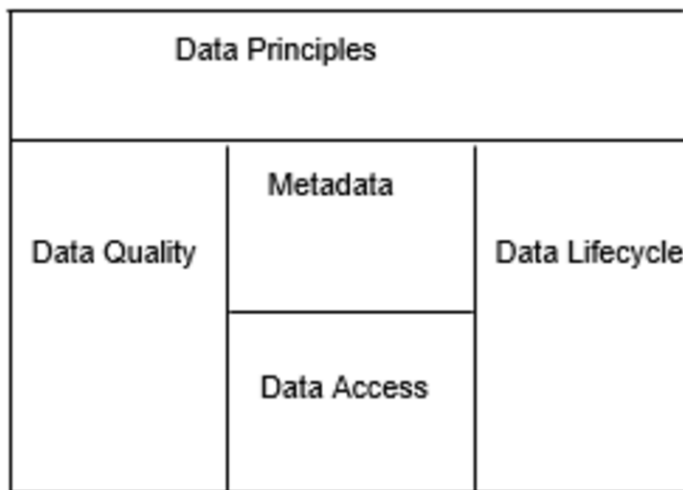
**Figure 2: Data governance contingency model (Weber et al., 2009)**



In the following year, Khatri and Brown (2010) proposed a data governance framework with five interconnected data decision domains and guidelines indicating what governance is required for each. Data principles, data quality, access, meta-data, and data lifecycle are the five decision domains (see Figure 3). Their research emphasizes the importance of organizations establishing data principles to understand what rules, standards, and guidelines are acceptable (Khatri & Brown, 2010). They argue that by first realizing data as an asset, the data becomes subject to company controls, standards, and policies (Khatri & Brown, 2010). This approach is useful since it considers the need for data life cycle decision points. This research considers this to be a critical component for the success of big data governance. Both Weber et al. (2009) and Khatri and Brown (2010) present the assumption that the realisation of data governance design alternatives is context-dependent. This is further supported by Jimenez et al. (2019) in their study of data governance in business contexts.

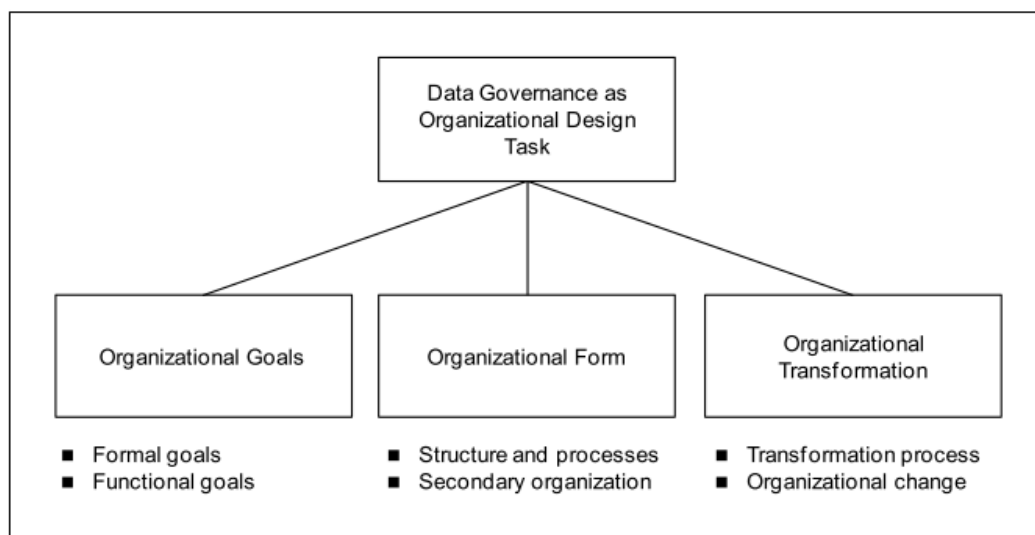


**Figure 3: Decision domains for the data governance (Khatri & Brown, 2010)**



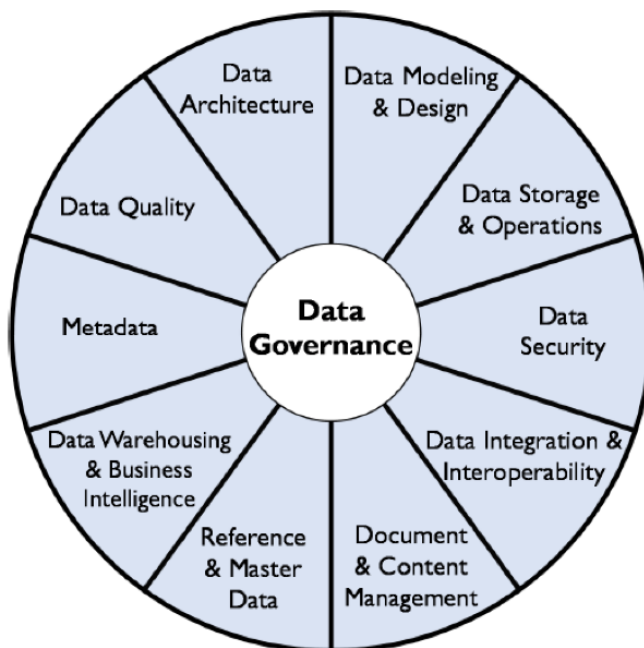
Otto (2011b) is another significant contributor to data governance research and suggested that data governance is comprised of goals and structures (see Figure 4). In his research, he states that control, organisational structure, roles and responsibilities are crucial to the goals and structures (Otto, 2011b). He further states that the organisation of data governance programs had not considered all the factors required for them to be regarded as complete (Otto, 2011b) and he illustrates this point in his study where he embarked on creating a morphology of data governance organizations (Otto, 2011a) using several mini-case studies to successfully examine the morphology of the data governance organisation.

**Figure 4: Conceptual framework of a data governance organisation**



The DAMA International industry association published the Data Management framework DM-BOK in 2009, with a second edition DM-BOK2 released in 2017. In a nutshell, it provides an industry-wide view of data management knowledge areas. It attempts to describe some of the "who," "what," and "why" of data management, but does not go into detail on the "how to" methods. It defines 11 basic knowledge areas for data management implementation, one of which is data governance and deemed the centre (DAMA International, 2017) (see Figure 5). It also identifies methodologies for data management maturity analysis.

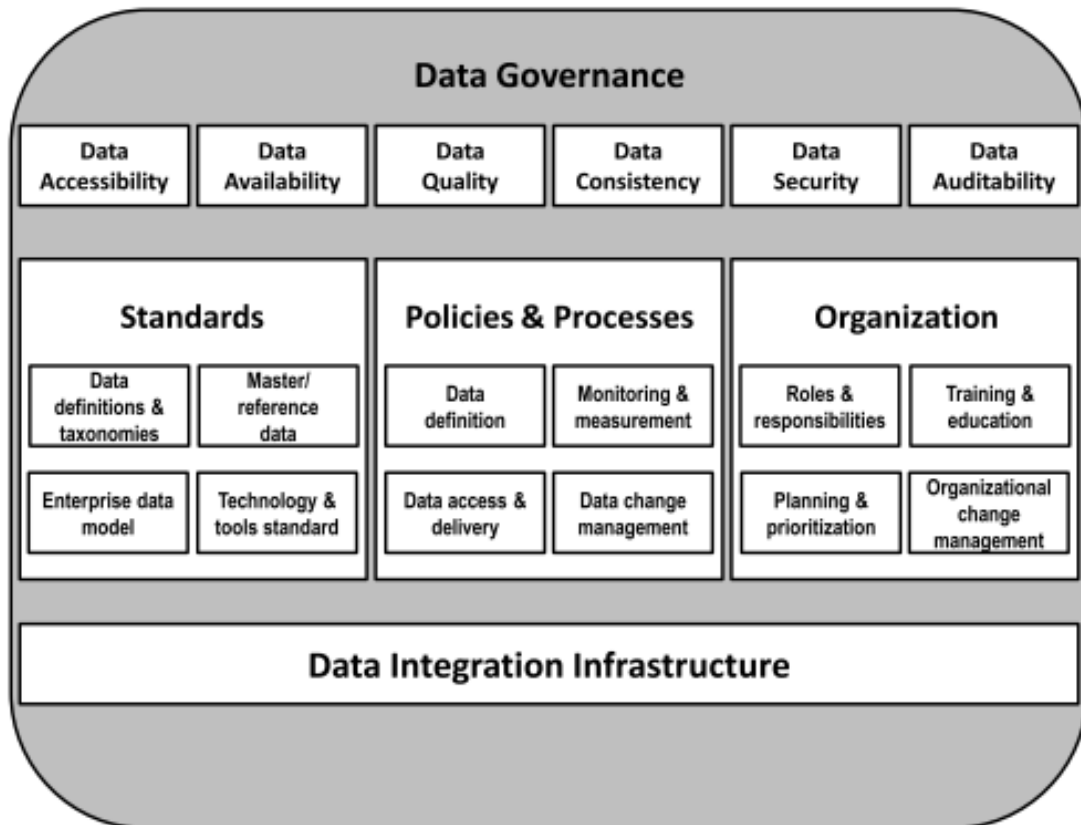
**Figure 5: DAMA-DMBOK Functional Framework (DAMA International, 2017)**



According to Panian (2010), the organizational structure, technology infrastructure, and standards necessary to ensure that data is available across the enterprise and is of high quality, consistent, auditable, and secure are all taken into consideration in the data governance framework. The core elements of the framework include standards, policies and procedures and organisation which then lead to the outcomes: accessibility, availability, quality, consistency, auditability and security ( see Figure 6). Panian (2010) argues that organisations have an integrated approach to data integration as siloed data integration can disrupt a data governance initiative, keeping it from succeeding and effectively scaling. Therefore, organisations require a consistent approach to data throughout the enterprise and throughout the data integration lifecycle. This holistic strategy methodically enforces data governance

standards and processes across the organisation and, perhaps more crucially, instils confidence in the data they rely on to make strategic business choices. However, both Panian (2010) and DAMA International (2017) have a greater technical orientation, and the framework design does not appropriately represent commercial value components.

**Figure 6: Building blocks of the data governance framework**



Although multiple data governance frameworks exist from both scholars and practitioners, the implementation of data governance remains a challenge for many organisations (Al-Ruithe et al., 2019; Alhassan et al., 2019a; Delacroix & Lawrence, 2019). The study explores the implementation of data governance in an organisation to better understand what challenges with data governance implementation. The diffusion of innovation theory has been used in multiple studies to model and understand the implementation and adoption of new systems, processes and practices in organisations operating in various industries (Akca & Ozer, 2014; Bowen et al., 2012; Doyle et al., 2014; Dryden-Palmer et al., 2020). The study will also use the diffusion of innovation theory to explore the implementation of data governance.

## **2.5 Theoretical Lens - Diffusion of Innovation Theory**

The diffusion of innovation theory is a social process through which information about an innovation or a new idea is disseminated through certain communication channels over a period within the social systems (Rogers, 2003). It is a social and psychological theory that aids in understanding and forecasting how people make decisions to adopt an innovation through the discovery of their patterns of adoption and understanding of their adoption structure (Min et al., 2019). The diffusion on innovation theory is widely used to analyse the adoption of information communications technology (ICT), digital technology and other technology innovations (Rogers, 1995, Minishi-Majanja & Kiplang'at, 2005), which makes it relevant for this current study as data governance mostly forms part of information technology implementations. Additionally, the diffusion of innovation theory is useful in understanding the adoption processes as well as the rate of adoption of an innovation (data governance) within the organization.

According to Rogers (1995), "diffusion is the process by which an innovation is communicated through channels over time among the members of a social system" (p. 5) and further adds that it is a social change where the structure and function of the social system are altered. Four major elements were identified and together contribute to the diffusion of innovation theory: innovation, communication channels, time and the social system (Lundblad, 2003).

### **2.5.1 Innovation**

Rogers (2003) defined innovation as "an idea, practice or object that is perceived as new by an individual or other unit of adoption" (p. 5). Lundblad (2003) asserts that innovation does not have to be new in terms of being newly invented; it just needs to be novel to the individual or organization adopting and executing it. As illustrated previously in the document, data governance is not a new concept and has been around for many years, however, for many organisations where it is introduced it is new. The diffusion of innovation theory suggests five characteristics of innovation namely relative advantage, compatibility, complexity, trial-ability and observability perceived by individuals to aid in explaining the rate of adoption (Rogers, 2003). According to Syahadiyanti & Subriadi (2018), the characteristics help to explain the rate of the innovation adoption process.

- I. *Relative advantage* is the level to which an innovation or new idea is perceived by the users as superior to the innovation before it (Rogers, 2003) and is economically profitable (Syahadiyanti & Subriadi, 2018). It has also been found that social reputation, convenience (ease of use) and satisfaction are also factors that are important to consider (Rogers, 2003). The adoption of the innovation is accelerated when the perceived relative advantage is greater (Minishi-Majanja & Kiplang'at, 2005; Rogers, 2003).
- II. *Compatibility* is the level to which an innovation or new idea is seen to be consistent with the prevailing values, past experiences and requirements of the potential adopters (Rogers, 2003). Therefore, innovations that are not compatible with prevailing social system features are not lightly to be adopted as quickly as compatible innovations (Syahadiyanti & Subriadi, 2018).
- III. *Complexity* is the level to which an innovation or new idea is seen to be difficult to understand and use (Rogers, 2003). Thus, the difficulty to understand and use the innovation will be an obstacle to the innovation's speed of adoption (Syahadiyanti & Subriadi, 2018).
- IV. *Trail-ability* is the level to which an innovation may be trailed and tested on a limited basis (Rogers, 2003). Innovations that can be broken up and trailed on an instalment basis are adopted faster than not divisible innovations (Syahadiyanti & Subriadi, 2018).
- V. *Observability* is the level at which innovation outcomes are observable to others (Rogers, 2003). Visibility of the results of the innovation stimulates peer discussion of the innovation which can authenticate the significance of the innovation thus accelerating the adoption (Rogers, 2003). This aligns with Min et al. (2019) who affirm that innovations are spread through people talking to other people about the adoption of the innovation.

Individual adoption rates can be explained by the characteristics of innovation. It is theorized that as each of these attributes grows, so will the rate of adoption, except complexity, where a decrease is theorized to increase the rate of adoption. Therefore, data governance as an innovation to the organisation can be evaluated in terms of the five characteristics to assist in explaining how the implementation of data governance is perceived in the organisation as well as understanding how this impact the adoption of data governance.

### **2.5.2 Communication Channel**

Diffusion is a specific type of communication where the message content is concerned with innovation or a new idea (Rogers, 2003). According to Lundblad (2003), the process of innovation requires innovation, a unit of adoption that is aware of the innovation and has used it, additional units of adoption that are not aware of the innovation and a communication channel between the two. A communication channel is a means of transmitting messages from one individual to another (Rogers, 2003). Lundblad (2003) states there is a significant relationship between the rate of adoption and the source of communication about the innovation. Wang et al. (2021) in their study of the implementation of a pharmacy residency program, further emphasizes the significance of communication channels as aspects to consider while distributing or adopting new programs

Information regarding the innovation must be communicated to create visibility and influence decisions concerning the innovation. Communication channels can include mass media which involves transmitting messages through mass mediums such as radio, television, newspapers, and newsletters (Rogers, 2003). Another communication channel involves interpersonal channels which involve face-to-face exchanges (Rogers, 2003). Zhu et al. (2021) state that a good balance of both mass media communication and interpersonal communication can encourage the diffusion of innovation. In recent times, interactive communication through the internet has become more significant for the diffusion of certain innovations (Rogers, 2003). Therefore, it is important to ensure that the right mediums of communication are implemented at the right time to facilitate the adoption of the innovation.

Thomas (2006) shared that leaders who have successfully implemented data governance initiatives revealed at the Data Governance Conference that, in their experience, data governance involved between 80 and 95 per cent of communications. She further adds that they were all taken aback by the quantity of written and verbal communication that was required to successfully bring stakeholders together to achieve their aims (Thomas, 2006). Therefore, to understand the implementation of data governance it is important to explore what communication channels are used to communicate data governance to the wider organisation and the impact this communication has had on the successful implementation of data governance.

### 2.5.3 Time

The time dimension of diffusion of innovation theory focuses on three elements namely: the innovation-decision process, innovativeness and the rate of adoption (Rogers, 2003).

**The innovation-decision process** is a process where an individual member of a social system progresses from the first knowledge of the innovation to a point where an attitude about the innovation is formed, to the decision to adopt or reject the innovation, to the implementation of the innovation and then the confirmation of the decision (Rogers, 2003).

There are five main sequential steps in the innovation-decision process: Knowledge, Persuasion, Decision, Implementation and Confirmation (Rogers, 2003). The knowledge step is achieved when an individual member of the social system learns about the existence of the innovation and understands how it functions (Rogers, 2003). Persuasion follows as individuals start forming an attitude toward the innovation (Rogers, 2003). The decision step is reached when the individuals start engaging in actions that lead them to a decision to either adopt or reject the innovation (Rogers, 2003). Implementation occurs when the innovation is used by the individual members of the system (Rogers, 2003). And lastly, the confirmation step is reached when individual members of the social system pursue reinforcement to validate an innovation decision that has already been made (Rogers, 2003). In contrast, Lundblad (2003) argues that the innovation-decision process within an organisation is more complex and comprises different five stages namely: agenda-setting, matching, redefining or restructuring, clarifying and routinising. The first two stages comprise the initiation phase where information is gathered, and planning occurs after which a choice to either adopt or reject the innovation is made (Lundblad, 2003). If the innovation is accepted, the last three stages form the implementation phase which involves actions to put the innovation into practice in the organisation (Lundblad, 2003).

**Innovativeness** is a measure of how individuals within the social system are inclined to adopt the innovation compared to other members (Lundblad, 2003). There are five classifications of innovativeness namely: Innovators, Early Adopters, Early Majority, Late Majority and Laggards (Rogers, 1995). The **adoption rate** is concerned with

the relative pace with which the innovation is adopted by individuals in the field of social systems (Rogers, 2003). It is measured in terms of the number of members in social systems that adopt the innovation over a given time which tends to result in an S-shaped curve (Rogers, 2003). According to Lundblad (2003), the characteristics of innovation outlined previously have an impact on the rate of adoption however the rate of adoption is also impacted by other factors which include the social system.

#### **2.5.4 Social System**

According to Rogers (1995), “a social system is defined as a set of interrelated units that are engaged in joint problem solving to accomplish a common goal” (p. 23). The diffusion of innovation or a new idea happens in social systems whose members may be individuals, groups, or organisations that have a common goal which binds them together as a social system (Lundblad, 2003). The structure of a social system can help or hinder the diffusion of innovations for example, in the innovation diffusion process, organisational structure qualities such as low centralization, high complexity, and low formalisation may facilitate the introduction of innovation but impede efforts to implement it (Lundberg et al., 2019).

Lundberg et al. (2019) further add that if the innovation is believed to come from within the organisation, people regard it as familiar, compatible, and easy to ascribe meaning to, making implementation easier. A social system has norms and accepted behaviours, opinion leaders and change agents that can influence the diffusion process (Rogers, 2003). Opinion leaders are members of the social system that influence other members, and their influence is based on their expertise, competence, accessibility and leadership in conforming to the system norms (Lundblad, 2003). They are central to interpersonal communication networks and serve as models to imitate when it comes to adopting the innovation (Lundblad, 2003). In contrast, change agents are external to the system however represent innovation and change and often use opinion leaders to gain acceptance within the social system to diffuse an innovation (Lundblad, 2003).

Social systems have a powerful influence on the diffusion of innovations in organisations, as the adoption of the innovation can be driven by individual members of the systems or the entire system whose decision can be made by the collective or



an authority or leaders (Rogers, 2003). The study explores the implementation of data governance in organisations, and this exercise involves many people within the organisation. It is therefore important to understand the impact that social systems have on the implementation of data governance. The study seeks to understand the impact of social systems on the implementation of data governance.

### 2.5.5 Innovation Process in Organisations

The innovation process in organisations identifies the key decision-making steps, actions and events in the process of implementing an innovation (Rogers, 2003). According to Rogers (2003), the recallable perceptions of significant innovation process participants, the organisation's written records of the adoption decision, and other data sources are used to gather information on the innovation process. There are five stages in the innovation process: two are in the initiation subprocess and three are in the execution subprocess (Rogers, 2003). The innovation process's latter phases cannot be started until the process's earlier stages are complete (Rogers, 2003).

Agenda-setting and matching, the first two of the innovation process' five stages, together make up the initiation subprocess, which is referred to as all information collecting, conceptualization, and planning for the adoption of the innovation (Rogers, 2003). The implementation subprocess, which entails all the events, activities, and choices necessary to put the invention to use, is made up of the final three phases, redefining/restructuring, clarifying, and routinizing (Rogers, 2003). Below the phases are explained in detail:

**Agenda-setting** takes place when a widespread organisational issue is discovered and a perceived need for innovation arises, (Rogers, 2003). The agenda-setting step of the innovation process in an organisation entails identifying and prioritizing needs and challenges, as well as searching the organisation's surroundings for innovations with the potential to be effective in addressing these organisational difficulties (Rogers, 2003).

At the **matching** stage of the innovation process, a problem from the organisation's agenda is matched with an innovation and this match is planned and constructed

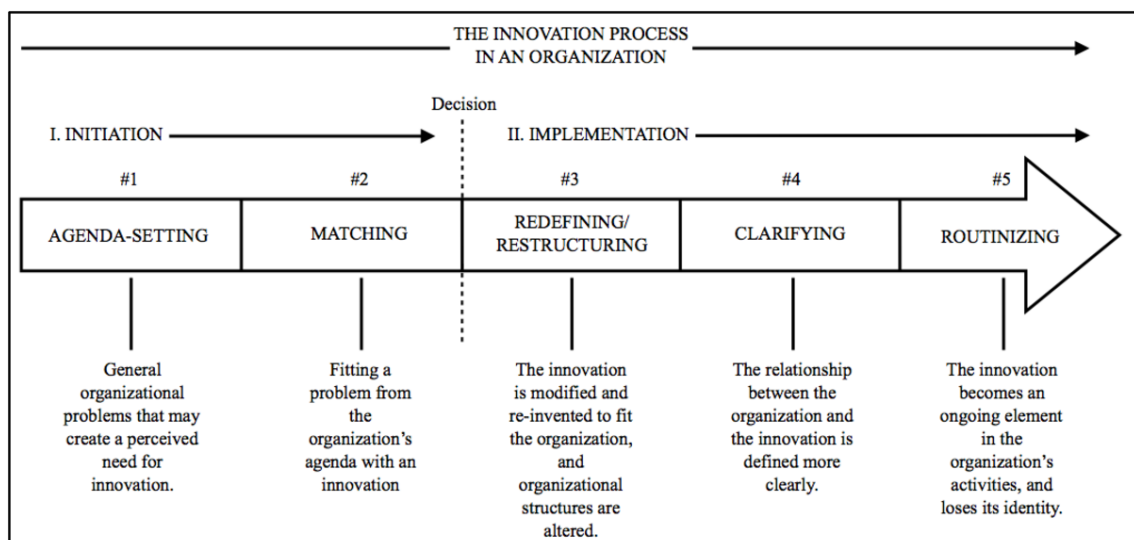
(Rogers, 2003). To determine how well they fit, the problem and the invention are conceptually matched at this second stage of the innovation process. This leads to the decision to either accept or reject the innovation before the implementation (Rogers, 2003).

**Redefining/restructuring** occurs when an innovation is revised to better suit the needs and structure of the organisation, as well as when the structure of the organization is altered to accommodate the innovation (Rogers, 2003). During this stage, both the innovation and the organisation are anticipated to change somewhat.

**Clarifying** takes place as the innovation is implemented more widely in an organization, allowing the people of the organization to increasingly understand the new invention's purpose (Rogers, 2003). Implementing innovations too quickly at the clarification stage can have disastrous effects (Rogers, 2003).

**Routinizing** occurs when an innovation loses its unique identity and is absorbed into the organization's routine tasks (Rogers, 2003). This is when the innovation is institutionalised and seen as part of the business day to day activities.

**Figure 7: Five stages in the innovation process in organisations (Rogers, 2003)**



Lundberg et al. (2019) state that implementing an idea in an organisation is difficult since many people are usually involved, including both champions and opponents,

additionally, people are normally hesitant to change existing organisational programs and processes because they have invested in and made commitments to them. The implementation of data governance is no exception to this claim as it involves people changing behaviour and how they interact with data from creation all the way to and disposal. Therefore, it is important to explore the implementation of data governance and how the diffusion of innovation theory can be used to facilitate the implementation.

## **2.6 Conclusion**

Businesses are dealing with substantial changes in how they manage their operations, customers and business models, caused by the emergence of new opportunities brought on by the rapid advancement in big data technologies (Raguseo, 2018). The enormous amounts of data that businesses can collect today from various sources to inform strategies and decision-making promises significant improvement in the overall business performance (Gandomi & Haider, 2015). However, without effective data governance and data quality management, the business cannot leverage the benefits of data (Schroeck et al., 2012).

Literature has revealed that data governance is most effective when it aligns with the organisation's strategy and its implementation enterprise-wide (Al-Ruithe et al., 2019; Cheong & Chang, 2007; DAMA International, 2017). According to Rogers (2003), many factors influence the diffusion of innovation in an organisation. From the diffusion of innovation theory, a perspective is gained that the successful implementation and the rate of adoption for data governance would be influenced by its relative advantage which is the superior advantage it is perceived to bring compared to previous ideas; compatibility which is the level to which it aligns with the norms and values of the current social system.

Data governance policies and guidelines need to align with the organisation's strategy, mission, norms and values and this according to the diffusion of innovation accelerates adoption. Another is complexity, in the literature, it was found that data governance outcomes and overarching goals are still abstract and ambiguous for most practitioners today this links to the observability characteristics as organisations are still battling to demonstrate the value of data governance. From the data

governance frameworks discussed, data governance can be broken down into smaller chunks and it is possible to trail, as it is an ongoing program.

The diffusion of innovation theory highlights the importance of communicating the innovation to the social systems and using the right mix of communication channels in communicating the innovation. Literature also revealed that the implementation of data governance is heavily reliant on the effective communication of data governance. It is therefore important to understand how data governance initiatives are communicated and the mediums used to communicate.

The innovation process in organisations indicates the main decision-making phases, actions and events involved in the implementation of an innovation. Through the five stages, the innovation process demonstrates the stages and their sequence that can be followed to implement a new idea or innovation in an organisation. This process can be used to map the implementation of data governance in the organisation.

Diffusion of data governance occurs within an organisation which is a social system. Lundberg et al. (2019) state that it can be challenging to put an idea into practice because there are frequently many stakeholders—both supporters and detractors—involved. In addition, because people have typically invested in and committed to existing organizational programs and processes, they are typically reluctant to change them. Therefore, it is important to understand the impact that the social system has on the implementation of data governance in the organisation. Literature also revealed that a data-driven culture enables data-driven decision-making in an organisation. Therefore, it is imperative to understand the influence it will have on data-driven decision-making as one of the outcomes of effective data governance.

## Chapter 3: Research questions

The previous chapter reviewed the literature on data governance and the diffusion of innovation theory which is the theoretical lens selected to guide the exploration into the implementation of data governance in the selected case organisation. The literature also reviewed big data which has been revealed as one of the drivers of the demand for data governance. The literature also reviewed data-driven decision making which has been highlighted as one of the outcomes of data governance. This has led to the formulation of four research questions that this study seeks to answer. This chapter will in detail describe the research questions in which the researcher aims to understand the implementation of data governance and its impact on data-driven decision-making.

### 3.1 Research Questions

According to McGregor (2018), research questions indicate what the author wants to know about the phenomenon. The following sections will outline the research questions the researcher intends to answer through the study.

#### *What perceived needs drive the implementation of data governance?*

There are always justifications for the implementation of innovation within an organisation. According to Rogers (2003) innovation process in organisations, agenda setting is the first step in the innovation process, where people within the organisation recognise a significant problem and seek innovation as a solution. With this question, the researcher seeks to understand the problems and drivers that prompted the need to implement data governance in the selected retail organisation.

#### *What are the perceived innovation characteristics of data governance?*

The overarching goal of data governance is still abstract, ambiguous and confusing for practitioners, which prevents them from mobilising collective action in the data governance (Benfeldt et al., 2020). As a result, organisations still struggle to recognise and quantify the value-creating potential of data governance. According to the diffusion of innovation theory, stakeholders must be able to see a relative advantage of adopting and perceive that doing so will benefit them. It is revealed in the literature that if the five characteristics of innovation increase except for

complexity which is the inverse, so will the rate of adoption. With this question, the researcher seeks to understand the innovation characteristics of data governance in the organisation to assist in matching the organisation's perceived needs and the innovation.

*What communication channels can be used to disseminate data governance?*

Literature has revealed that communication is essential in the implementation of data governance. Wang et al. (2021) emphasize the significance of communication channels as factors to take into account when disseminating or promoting new initiatives. Rogers (2003) posits that various communication channels play distinct roles at different stages in the innovation process. With this question, the researcher seeks to establish in the process of data governance diffusion in the selected retail organisation.

*What influence does the social system have on the implementation of data governance?*

Rogers (2003) observed that a system's social structure can facilitate or hinder innovation's ability to spread throughout the system. The literature revealed that data governance is the governance of people and technology as data on its own has no will nor intent and people shape it. Therefore, the implementation of data governance requires people to change their behaviour and the way they interact with data from creation through to consumption and disposal. From the literature, it is revealed that people are often reluctant to change existing organisational programs and processes because they have typically committed to and invested in them. Additionally, introducing a change in the organisation is complex as there are often many people involved with consists of promoters and detractors. With this question, the researcher seeks to understand what influence the social system has on the implementation of data governance.

*What impact does the implementation of data governance have on data-driven decision-making?*

Literature has revealed data-driven decision-making as one of the main outcomes of data governance. This, as highlighted in the literature is in part due to the increased confidence in the data-driven by the achievement of good quality data as well as the

availability of the data. With this question, the researcher seeks to understand the impact that the implementation of data governance had on data-driven decision-making in the selected retail organisation.

### **3.2 Conclusion**

The purpose of the research questions outlined in this chapter is to enable the researcher to comprehend the problem outlined in Chapter 1. The research methods and strategies that were used to explore and answer these questions during this study are described in detail in the next chapter.

## **Chapter 4: Research Design and Methodology**

The following section outlines the study's research design and methodology choices.

### **4.1 Research Design**

#### **Purpose of research design**

According to Cantrell (2011), research design can be described as an overall plan for answering the research questions and constitutes the research strategy(s) adopted by the researcher to collect information that is accurate, unprejudiced and interpretable. Therefore, the purpose of the research design is to provide a blueprint of how the research was conducted.

#### **Research philosophy (paradigm)**

Interpretivism relates to studying a social phenomenon in its natural environment (Saunders & Lewis, 2018a). An interpretivism paradigm aims to understand the subjective meanings of the people in the studied domain (Goldkuhl, 2017), with the underlying assumption that there is a greater opportunity for understanding the people's perceptions of the studied phenomena in their social context (Kelliher, 2005). The study followed the interpretivism paradigm, best suited as the researcher sought to explore the diffusion of data governance in an organisation using the diffusion of innovation theory as a framework to explore the implementation of data governance in a Retail organisation.

The study adopted an abductive approach in the research process which combines both the deductive (top-down) and the inductive approach (bottom-up). According to (McGregor, 2018), a deductive research approach entails using an existing theory to test or modify a theory. The study uses the diffusion of innovation theory as the basis to study the implementation of data governance in an organisation which speaks to the deductive nature of the study. According to Saunders and Lewis (2018a), the inductive research approach emphasizes an intimate understanding of the research context and is useful when trying to gain an understanding of the meanings that humans attach to events. Given the abstract nature of data governance, it was expected that the data collection method would yield to new results within the selected sample, specifically in the perceived characteristics of innovation and the



impact of the social system on the implementation of data governance. This speaks to the inductive nature of the study, which has led to the use of the abductive approach.

### **Research strategy**

The case study strategy enables a deep understanding of constructs or phenomena in their natural context (Wahyuni, 2015). Saunders and Lewis (2018a) add that case studies are good at facilitating a detailed understanding of the context of the research and the activity taking place within the context. The case study research strategy can include either single or multiple cases with numerous levels of analysis (Eisenhardt, 1989). Data governance is a complex phenomenon and often a challenge to fully comprehend in an organisation as it goes beyond governing the creation, collection, management, and usage of the data but also involves people and technology. For data governance to be successful, people and applications (systems) need to be considered to stimulate the desired behaviours (Janssen et al., 2020).

According to Wahyuni (2015), it is ideal for case study research to use a multiple case study design which will involve having multiple case sites and using multiple methods to collect data. In contrast, Mariotto et al. (2014) argue that single cases enable the creation of a more complicated theory than multiple cases since the researcher can fit the theory exactly to the many details of a particular case. The study will adopt a single case study research which is appropriate for exploring the implementation of data governance using the diffusion of innovation theory, as it will allow the researcher to consider all the factors that influence implementation using the chosen theoretical lens.

Multi-method qualitative research is using more than one data collection method and uses quality procedures to answer the research question (Wahyuni, 2015). According to Saunders and Lewis (2018a), a case study research strategy may include a combination of interviews, observations, and documentary analysis to enable triangulation. This study used a multi-method qualitative approach by collecting data from more than one source, in the context of this study the data was collected using semi-structured interviews and external documents. This approach was taken following Bhattacharjee (2019) who states that qualitative data in a case

study strategy may be collected using a combination of interviews, personal observations, and looking at internal or external documents.

### *Case Selection*

Selection of the case or cases is an important step in building theory from the case studies (Eisenhardt, 1989). It is an integral step of a good research strategy that aims to achieve well-defined research objectives (Ruffa, 2020). The selection of the case for the study was done through purposive sampling. Purposive case selection is strategically selecting a case or cases based on its hypothesized characteristics about the broader population of the cases (Ruffa, 2020). Flick (2018) argues that selecting an information-rich entity is crucial, especially for case study research. Guided by Flick, a JSE-listed organisation was selected in the chosen industry context retail. Additionally, the researcher was granted access to the selected case organisation to conduct a detailed exploration of the implementation of data governance in this organisation.

### **Methodological choice**

Data governance is an under-researched subject, (Al-Ruithe et al., 2019) and the impact data governance has on its users has not yet been defined (Abraham et al., 2019). Given the growing significance of data governance in practice, it is therefore imperative to gain further exploratory insights and a deep understanding of data governance and how it can be successfully implemented in the retail sector. According to Saunders and Lewis (2018a), an exploratory study aims to pursue new insights on a topic by asking new questions and assessing topics in a new light.

According to Mohajan (2018), qualitative research can be described as a social action that emphasizes the understanding of the way people interpret and understand their lived experiences to make sense of the social reality of the individual. It is exploratory and attempts to explain “how” and “why” a specific social phenomenon operates the way it does in a specific context (Mohajan, 2018).

The study sought to understand how the diffusion of innovation theory can be used as a framework to implement data governance to drive the adoption of data governance practices. It was previously highlighted that stipulating the outcomes of data governance is complex as data governance involves many industry-specific

factors that cannot be operationalized. Consequently, the qualitative research methodology is the most appropriate research methodology to deepen the understanding and draw further insights.

### **Time horizon**

According to Saunders and Lewis (2018a), cross-sectional research is a study of a specific topic at a specific time. It is most applicable to research projects that are constrained by time. The study was cross-sectional as it was the most achievable considering the time constraints on the submission to the research project.

## **4.2 Proposed research methodology**

### **4.2.2 Population**

According to Saunders and Lewis (2018b), a population is the complete collection of the group being researched. The selected research strategy for the study is the case study research strategy, Flick (2018) states that case study research requires complex decisions at multiple levels the first level being the entity as discussed in the case selection section. The second level is the selection of participants within the case site.

The population of the study is the employees (permanent and contract) of the selected case organisation. Al-Ruithe et al. (2018) suggest that data governance influences a combination of various stakeholders that are involved in data decisions and activities that affect the organisation's data assets. According to Al-Ruithe et al. (2018), the stakeholders include data governance officers, executive leadership, data managers, data consumers, data stewards, compliance officers and legal officers. The population is all the different functions the employees may perform in the organisation related to data.

### **4.2.3 Unit of analysis**

According to (Dolma, 2010), the unit of analysis is the entity that the research study attempts to investigate. The level of analysis for this study was at an individual level, therefore appropriate for the unit of analysis to be the senior employees who were involved or impacted in the implementation of the data governance programme.

#### **4.2.4 Sampling method and size**

There are various methods of selecting samples. The choice of the sampling method depends on the research questions (Saunders & Lewis, 2018b). The sampling method adopted for this study is the non-probability purposive sampling method, which operates on the premise that the best information is obtained by zooming in on a small number of deliberately selected participants based on their known attributes (Bhattacharjee, 2019). Saunders and Lewis (2018b) state that purposive sampling is used when the researcher needs to understand what is happening to be able to make logical generalisations. Given the purpose of the study, it was important to select a sample that is relevant to the topic and has substantial knowledge of data governance with a sound understanding of the potential impact it will have on the organisation. The criteria for selecting the sample included their relevance and knowledge of data governance as well as their participation in the implementation of data governance in the organisation. The selection was based on the title and roles the participants functioned in within the organisation.

According to Bhattacharjee (2019) saturation is reached when additional reiteration in the understanding of the phenomenon and observations or the collection of additional information no longer yields any new additional insight. Guest et al. (2006) assert that data saturation for qualitative interviews is reached after twelve interviews. Saunders & Lewis (2018b) also add that for a homogeneous population the sample size should be between 4 and 12 while heterogeneous populations would require a sample size between 12 and 30. Therefore guided by these scholars and the sample of the population chosen, the study used a sample size of 15 participants who were selected based on the criteria mentioned previously.

#### **4.2.5 Measurement Instrument**

The study collected qualitative data using semi-structured interviews which are a combination of structured and in-depth interviews (Wahyuni, 2015). Semi-structured interviews use a list of predetermined open-ended questions and topics in the form of an interview guide. However, the interview guide also retains the flexibility for the interviewee to speak freely about any other topic that might not be covered in the interview guide (Wahyuni, 2015). The interview guides are a fundamental tool in collecting qualitative data as they introduce a basic structure for the interviews which ensures that the researcher stays focused and is consistent throughout the research

process (Iyamu, 2018). The questions and topics in the interview guide for this study were derived from the research questions to ensure the research questions are fully addressed.

### **3.2.6 Data gathering process**

According to Bhattacharjee (2019), qualitative data in a case study strategy may be collected using a combination of interviews, personal observations, and looking at internal or external documents to improve the credibility of the study, this is also known as triangulation. Saunders and Lewis (2018b) describe triangulation as the method of using two or more sources that are independent for information collection within a single study to ensure consistency in the interpretation of the data. Data for this study was collected using semi-structured interviews and secondary data from external whitepapers published on the implementation of data governance. The researcher obtained written approval from the organisation to conduct the study.

The appointments were set up with the selected participants through email. The informed consent letter was emailed to the participants before the interview and received signed consent to conduct the interviews. The interviews lasted between 45 and 60 minutes. The majority of the interviews were conducted virtually, this was due to the team being geographically dispersed and working from three different provinces. It was also more convenient for the participants who were then able to better schedule their time. The virtual meeting and the physical meeting were conducted via Zoom and the meeting was recorded using the software as it was more reliable and secure. The audio recording was then transcribed using the software Otter and then proofread and edited by the researcher for quality testing and to ensure that the transcripts were accurate.

### **4.2.7 Analysis approach**

According to Clarke and Braun (2017), thematic analysis is a technique for systematically generating codes that assist in finding, grouping, analysing and interpreting themes (patterns of meaning) in qualitative data. Themes within the qualitative data can be identified in two primary approaches: inductive and deductive (Braun & Clarke, 2006). This study adopted both the inductive and deductive thematic analysis approach by deductively using themes identified from the literature

review to code the data and applying the inductive approach for emerging codes that were not identified in the literature review.

All interviews will be audio-recorded and converted into text assisted by software. The researcher conducted quality checks on all the transcripts to ensure that everything was translated and resolve any gaps identified. The transcripts were then loaded into the qualitative data analysis software Atlas. ti 22. The researcher also loaded the documents gathered from the organisations in the software. All the data collected was stored in a secured public cloud instance namely One Drive and only the researcher had access to the folders. The researcher followed the following phases of the thematic analysis presented by (Braun & Clarke, 2006).

**Phase 1:** The researcher read all the transcripts after the quality tests and noted all the initial ideas and information coming from the data these are also called quotations which are the essential sections in each transcript. The quotations were then indexed to enable ease of tracking.

**Phase 2:** Interesting and relevant features of the data were systematically coded across the data sets. All relevant data identified in phase 1 was organised into meaning groups linked to the codes.

**Phase 3:** The researcher searched for the themes by sorting and grouping the formulated codes and studying how the various codes could be combined to form a theme.

**Phase 4:** The researcher reviewed all the themes identified in phase 3 to refine the themes and consider their validity of the themes. The researcher used the thematic map to test if there is a logical flow.

**Phase 5:** The researcher further reviewed, defined and enhanced the themes. The researcher analysed the data within the themes to discover the crux of each theme and determine the aspects of the data captured by each theme.

**Phase 6:** The researcher then compiled a report of the thematic analysis findings which will be presented in the following chapter.

The report was compiled after the analysis of the data reported on the aggregated information. The data was stored in a secured cloud platform (Microsoft One-Drive) with no identifiers to ensure the confidentiality of the organisation. The study used an industry whitepaper to triangulate the data gathered from the participants.

#### **4.2.7 Ethical considerations**

It was stressed by Creswell and Cresswell (2018) that a researcher must be aware of any ethical dilemmas during their research study. Moreover, the research process involves collecting data from participants about the internal processes of a listed company in the competitive retail industry in South Africa. Gontcharov (2017) also highlighted the ethical concerns associated with human research. Therefore, it was critical to adhere to the approved GIBS Ethical clearance process. The participants' identities were protected using codes in line with the ethical consideration to protect participants from negative impacts. While conducting the research, the researcher established and upheld professional relationships and avoided improper behaviour.

Before conducting this study, the researcher sought and was granted permission from an executive of the selected organisation to conduct the study. This was followed by permission from the GIBS research committee through the ethical clearance process to conduct this study in partial fulfilment of the requirements of the Master of Business Administration degree. Permission was also politely sought from the participants by signing the consent forms which also outlines the purpose of the study where they were also given the option to opt-out of the study at any time. The names of the participants were not added to the forms to protect their identity as well as the identity of their organisation. The researcher assured the participants that their participation in the study would be used for academic reasons and that the results would be presented in an aggregated form.

#### **4.2.8 Quality controls**

According to Mariotto et al. (2014), reliability refers to the condition that research must be designed in such a way that if other researchers chose to follow the same research processes as the first researcher, they would obtain the same insights. Creswell and Cresswell (2018) emphasised that the reliability of findings needs to be consistent throughout the research study. Therefore, if the research instrument is not reliable, it will result in an inaccurate research study and reporting that is not trustworthy. To test the reliability of the research instrument, the interview guide was piloted with a participant chosen from the selected retail organisation. The researcher established during the interviews that not all the questions were clear and well-understood. The interview guide was amended and finetuned using the feedback

from the participant which was then used for the final data collection process. The researcher's approach was consistent throughout the data collection and analysis process which assured the reliability of the results.

To build solid triangulated measures, Eisenhardt (1989) proposes that the researcher begin with a wide topic, establish systematic data collecting, and assure case access. In this method, qualitative research may be reinforced by integrating participant observation, interviews, and documentary sources in a single case study (Kelliher, 2005). Subsequently, credibility is the degree to which the methods of collecting data and the analysis procedures produce consistent and dependable findings (Saunders & Lewis, 2018b, Bhattacharjee, 2019). Data triangulation is very useful in single-case study research to strengthen validation in the absence of a cross-case comparison (Kelliher, 2005). This study followed a multi-method approach to collecting data through semi-structured interviews as well as document sources.

Validity refers to the extent to which the methods of data collection adequately measure the underlying constructs that were intended to be measured and that the findings of the research explain what they were supposed to explain (Saunders & Lewis, 2018b, Bhattacharjee, 2019). Using numerous data sources, constructing an identifiable chain of evidence, and having the draft evaluated by key informants all contribute to the validity of the study (Kelliher, 2005). To improve the study's validity, an experienced, qualified, and credible supervisor reviewed the research study.

#### **4.2.8 Limitations**

According to Denscombe (2010), qualitative research is subject to the researcher's bias which comes from the researcher's self-identity, beliefs and background, in creating and analysing the qualitative data. Bhattacharjee (2019) adds that the interpretation of findings from the qualitative analysis is also dependent on the researcher's observational and integrative ability. Therefore, the researcher's unconscious bias toward the subject studied can be deemed as a limitation of the study.

A single case study may be subject to case selection bias, which occurs when a case is chosen that is previously known to have particular features (Ruffa, 2020). The case study was selected using the purposive sampling method where the case was



selected because of certain characteristics, therefore making it prone to selection bias. Furthermore, a single case study cannot demonstrate whether the phenomena studied is indicative of the population of instances. As a result, there is insufficient data to establish firm generalisations about the phenomena researched. Generalisability is the extent to which the findings of the research are more generally applicable outside of the research context. Additionally, the case organisation selected operates in South Africa and in the retail industry, which introduces the limitation that the study might not be generalizable to other country contexts and industries. The case selected is also a large-sized organisation which means that the findings of the study might not apply to small to medium organisations.

The study's sample size and location pose a limitation in the generalisability of the findings to the whole case population. Additionally, the qualitative nature of the study also introduces the limitations in the replicability of the study in other countries and contexts. Schedule constraints in the delivery of the project and interviews may limit the level of insight that would have been acquired in the absence of time constraints.

The study applied the diffusion of innovation theory as a lens to study the implementation of data governance in a retail organisation in South Africa. This may limit the study in viewing other lenses that may present different outcomes.

### **4.3 Conclusion**

In conclusion, this chapter presented the research techniques used in this study. In this study, the interpretivist philosophical perspective was selected and used. The case study research strategy was used in the study, which employed a qualitative research methodology. Using these techniques, the researcher was able to contextualize the data in its natural environment. 15 senior employees of the chosen retail organization served as the study's sample. The majority of interviews were performed electronically to accommodate participant preferences and geographic areas.

## Chapter 5: Results

### 5.1 Introduction

The purpose of this chapter is to present the key findings of the qualitative data collection process which was guided by the research questions discussed in chapter 3. The findings from the data collection process have been arranged by the research questions as well as the concepts and theoretical categories introduced in chapter 2.

A total of 16 interviews were conducted including the pilot interview with a total duration of 700 minutes excluding the pilot interview. 15 of the total interviews were transcribed with a total of 147 pages and 80 103 words.

The chapter begins with a summary of the interviewees, the interview process employed in the research and saturation analysis. The remainder of the chapter will be organised around the research topics outlined in Chapter 3.

### 5.2 Overview of participants

As outlined in chapter 4, a single case study research design was adopted which means that all the participants were either permanently employed by the organisation or consulting for the organisation on a full-time basis. Table 1 outlines the participants for the study with their titles in the organisation. For ease of reference, the participant has been coded from P1 to P15 based on the order in which the interviews were conducted. The study differentiates between consultants and employees to establish if there are differences in the responses provided.

**Table 2: List of Interviewees**

Participants	Code	Title	Consultant/Employee
1	P1	Retail Programme Specialist: Master Data Management	Consultant
2	P2	Senior Executive: Master Data Management	Employee
3	P3	Master Data Consultant	Consultant
4	P4	Manager: Business Analysis	Employee
5	P5	Master Data Systems Analyst	Employee
6	P6	Technical Manager: Master Data	Employee
7	P7	Senior Business Analyst: Master Data	Employee
8	P8	Product Master Data Operations Manager	Employee
9	P9	Data Privacy SME	Employee

10	P10	Master Data Domain Lead	Employee
11	P11	Master Data and Systems Analyst	Consultant
12	P12	Data and Analytics Manager	Employee
13	P13	Executive: Customer Insights and Analytics	Employee
14	P14	Senior Manager: Data Maturity	Employee
15	P15	Master Data Domain Lead	Employee

Source: Author (2022)

### 5.3 Interview Process

Before the interviews were conducted, the participants were sent the interview guide (Appendix 2) and the informed consent letter (Appendix 2) to provide context and ensure that the participants were able to provide substantial insights into the research topic. All 12 out of the 15 interviews were conducted using online using Zoom which was due to the team being dispersed between Durban, Cape Town and Johannesburg. Additionally, it was convenient for the participants as the majority worked from home.

A pilot interview was conducted to test the reliability of the interview guide in getting adequate data to answer the research questions. The participant was from the organisation and was part of the data team that was tasked with the implementation of data governance. It was found that the initial questions were too broad and misunderstood. The interview guide was then revised to ensure that the questions were clear.

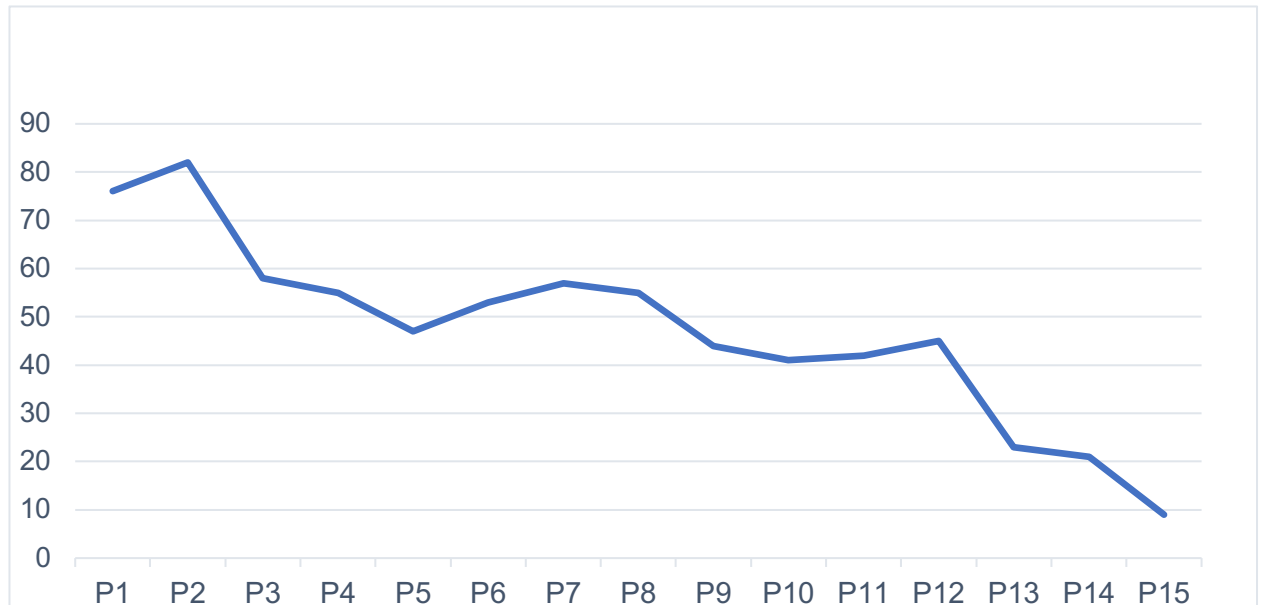
### 5.4 Saturation analysis

A total of 708 first-level codes were identified across the 15 interview transcripts collected for the research project. The codes were then organised according to their relevant characteristics guided by the code book derived from the Diffusion of Innovation theory (Appendix 4). The codes were according to the main codes and subcodes that were identified from the literature as well as new sub-codes that emerged from the data analysis (Appendix 5).

Figure 1 demonstrates that the number of new codes that were discovered from the participants declined. The analysis shows that from participant twelve there was a

significant decline in the number of new codes (topics) that emerged from the analysis.

**Figure 8: Total codes per participant**



Source: Author (2022)

## **5.5 Research Question 1: Perceived needs that drive the implementation of data governance**

As a foundation to unpack the implementation of data governance, it was important to understand the perceived needs that drove the implementation of data governance in the organisation. The participants were initially asked to provide their views on what they believed the role of data was in the organisation. It was hoped that this would provide context to understanding their perception of the importance of data in the organisation.

### **5.5.1 The role of data in the organisation**

The consensus was that data is crucial in driving decisions given the fast-paced retail environment. Data is at the core of all business operations, and it is important to have good-quality data to ensure that the business can derive value from it. The key message was well described by the following participants.

*“As retailers, we are dependent on data for virtually all our decision-making, our sales forecasts are based on historical data, our replenishments are based on sales forecasts, and we promote based on our understanding of customer data. There is an art and a science to retail, the science is data-driven, and the art is what an experienced merchant or operator will bring” (P 2)*

*“The role of data is fundamental to help the business make decisions ... No decision should be made without looking at the facts ... data should always inform the why behind every decision made” (P 13)*

*“Data is key to understanding where to improve, what to do more efficiently and where to invest more. Data is also the gathering of all actions and decisions for all business processes, the outcomes of the business process and needs to be reviewed to understand where the opportunities are.” (P 10)*

### **5.5.2 The perception of data governance**

When the participants were asked about their perception of data governance and perceived need to implement data governance in the organisation there were varying sentiments as to what data governance is and why the organisation need to implement data governance. A subset of the respondents perceived data governance to be rules, standards and policies that inform how data is managed in the organisation.

*“I believe data governance defines standards and business rules that ensure that we do not collect data for the sake of collecting it, but that it makes sense, and it has to be effective... It ensures that data drives our processes in such a way that the processes are efficient and lead to business benefit.” (P 7)*

*“Data governance defines the rules and standards of how the data is used, how it is manipulated and how the data needs to be structured.” (P 6)*

*“Data governance rules that are implemented to ensure that the data is auditable... It also sets the standard as to how the data is created and maintained.” (P 11)*

Another subset of respondents perceived data governance as a means to drive data quality in the organisation and ensure that the information derived from the data is consistent across the organisation.

*“Data governance ensures that there is consistency in terms of the information that is derived from the data across the organisation. Data governance has to do with data quality ensuring that quality is maintained centrally... Data governance also provides oversight to ensure that quality is captured at the source as well as visibility to who owns the data.” (P 3)*

*“Data governance will ensure that the business has a better chance of data accuracy.” (P 2)*

*“Data governance is a process that helps in managing the availability, usability, and integrity as well as the security of the data in the business.  
” (P 12)*

There was one participant who stressed that not all data in the organisation is equal and that it was important to consider the types of data, as the different types of data have different levels of importance and implications to the business citing personal data and company sensitive data.

*“Data has become critical, touching every point of the business... It is however important to consider the different types of data because they each have their weight or their own relevance and importance to the organisation ... Data governance involves the processes of planning and the implementation of how we manage our data, all the different data types of data from identifying what we have, how we use it, and for what purpose we use it for. Look at how we can strategically manage our data and identify the ideal management controls based on the policies.  
” (P 9)*

### 5.5.3 The perceived need for data governance

When asked what the perceived need was for implementing data governance, the shared sentiments were around the urgency to improve the data quality in the organisation. The organisation has just gone through a system migration, and this revealed how bad the quality of the data was in the organisation.

*“In any business without data, the business doesn’t exist. Not only without data but without good quality data, we have first-hand experience of businesses that were run with really poor data quality ... it a can literally bring a business to its knees” (P 15)*

*“When I joined the organisation, I was consulting on getting some of the master data issues sorted out when they migrated to the new system. It became apparent that there was no data governance ... and the business has major master data quality issues that were impacting the business big time” (P 1).*

*“We realised the importance of having a data governance organisation given the complexity of the environment and also the implementation of a new system, we had to better manage our data.” (P 3)*

Another perceived need to implement data governance was the lack of trust in the data and no standardisation. Given a business that has multiple subsidiaries it was very important to ensure that the data was consistent across the organisation.

*“The data can get quite messy because there was no standardisation from how the data was collected at the source and so down the line in the database, we find that there are a lot of gaps in the data... You really cannot trust some of the fields because the majority are incomplete” (P 13)*

Another participant also highlighted that data governance will also assist the organisation in resolving the lack of data ownership in the organisation as well as create visibility into the lineage of data.

*“Currently, there is no owner to the end-to-end process in terms of the data from source to consumption or reporting... Data governance will help us*

*understand in terms of data lineage by accounting where data comes from, who owns it and if there are any discrepancies who we can talk to.” (P14)*

In this section data, data governance and the perceived need for the implementation were discussed. Most of the respondents stated that data is fundamental to the functioning of the organisation and drives decision-making within the organisation. One participant felt that it was important to consider the different types of data and their varying importance to the organisation, in doing this the organisation can determine the right level of control to ensure that data governance yields benefit to the organisation. Additionally, the felt need for data governance was driven by poor data quality in the organisation which was having a massive impact on the business leading to the lack of trust in the data. There is also a lack of data ownership in the organisation, and it is perceived that data governance would assist in embedding this in the organisation.

## **5.5 Research Question 2: Perceived innovation characteristics of data governance**

According to Rogers (2003), there are five characteristics of innovation that help to explain the rate of adoption of an innovation. These characteristics determine whether an innovation will be accepted in an organisation or not. The researcher asked the participants to reflect on the five characteristics of innovation and share their perceptions on what these are for data governance.

### *5.5.1 Relative Advantage*

The participants were asked to share what they believed was the relative advantage of having data governance implemented in the organisation. Multiple themes emerged from the participants and the section will outline the various themes.

#### *Control*

The majority of the participants believed that data governance would ensure that the organisation has better control of their data. The organisation will have a better chance at protecting its data assets and controlling who has access to the data and what they can do with the data.



*“Data governance is going to give us better control of our data. We going to have a better understanding of who can access the data, who is able to delete the data and how long we need to keep it. It will give us a better view of our data quality through measuring our data against the business rules.” (P 2)*

*“It ensures that data is protected and used in the right way in the different areas.” (P 1)*

### *Improved Data Quality*

Another theme that emerged when the participants were asked about the relative advantage was improved data quality. Most of the participants believed that data governance would improve the quality of data in the organisation as the rules and standards put in place would assist in ensuring that validations are put in place to govern the data coming into the systems.

*“I honestly believe that any business can benefit from data governance because the better quality of data is going to improve the business all around in so many different ways.*

*” (P 5)*

*“Provides business rules that can be applied to ensure that data is accurate and aligned across the organisation and across all the various parties that looked at the data. This will improve the quality of the data and the accuracy of the data” (P11)*

The organisation was now able to track and manage its data quality using technology.

*“We had clear visibility of the information and its quality. We mapped the data field into a data quality management software, and we were able to give the data scores we were now able to proactively identify data problems and resolve them.*

*” (P 3)*

### *Improved efficiency and better decision-making*

The participants also shared that the business would be more efficient and would be able to make better decisions faster.

*“Data governance will produce better efficiencies because there will be a streamlined process through which data is acknowledged and transacted on.”*  
(P 1)

*“Data governance will result in fewer issues downstream from a transactional perspective and the way stock flows to the business as well as the way in which the business operates and interacts with customers... It is evident that it will keep consistency and efficiency in the operations of the business.”* (P 4)

*“It will help us provide insights and analytics at a faster rate because there won’t be a that long process of checking and double checking. It will remove all that time and effort spent in checking the data and have them solely focus on providing insights for the business. Therefore, will increase efficiencies in processes within the analytics team... just more trust in the data which will give the business a greater ability to innovate faster”* (P13)

### *5.5.2 Compatibility*

The participants were asked if they believed that data governance should be compatible with the organisation, the business strategy, and the way in which the business functions. There was consensus with most of the participants that the data governance put in place had to be compatible with the business. It must be fit for the business and focus on data that is important to the business.

*“Yes, it should be compatible as data governance is a component of the business that keeps us honest in terms of what we need to do and how we need to do it from a data analytics perspective and additionally it is a supporting component to the business strategy”* (P14)

*“Yes, it needs to align to the business strategy. Clearly, we need to be focusing on those elements of the business that have a big impact and the data that specifically has a big impact on the organisation. For example, product data*

*because as retailers we sell products and this data facilitates the sourcing, the procuring, the stock management and the sales therefore critical to the business's performance" (P15)*

However, two participants were not comfortable with the term compatible and believed that there should be an alignment but not necessarily compatibility.

*"I think an alignment is required. I hesitate to say be compatible because sometimes we must change our ways of work because we do things in a clumsy way or old-fashioned way and data governance should not adopt that way. It is important that there is an alignment and it's understood why data governance is important, but data governance should not be a slave to poor pre-existing processes" (P10)*

*"Data governance should be above that. So, you cannot take the current ways of working in a business and govern those you've got to put in the right governance to change the ways of working. So, we could have very bad ways of working, you don't just put governance in to protect that you would need to look" (P1)*

### 5.5.3 Complexity

The participants were asked to share their sentiments on the perceived complexity of implementing data governance in the organisation. The participants all agreed that data governance is extremely complex in this organisation because the organisation is still in the early stages of the journey of building data literacy in the organisation.

*"It is very complex because the organisation is still on the data journey. We still have to build data literacy across our business units, we still have to get people to understand how to use data, never mind the governance around how we use data" (P13)*

*"Understanding the business and related impact to the business... If it is not relevant to the business, properly planned, consulted, and understood it will certainly cause friction and frustrate the business" (P 9)*

*“How the organisation is currently structured is not conducive to making it easier for the implementation of data governance. I mean our architecture, so many teams, so many systems and all of these do not talk to each other. This is a journey, and we are still at the very early stages of the journey” (P14)*

*“It is complex because the systems are not common in the organisation and there are different ways of working and different ways of determining the same data information. Data governance is much easier to manage when it’s the same platform across the organisation and there is a oneness (one best way), but there is no one best way” (P1)*

#### *5.5.4 Trialability*

The participants were asked to share their perception of the trialability of data governance in the organisation and if they believe that being able to divide and trail data governance on a limited basis would aid in the implementation of data governance. The consensus was that it was important for data governance to be trailed and this was the approach that was followed by the organisation. Given that data, governance requires a significant change in the organisation it was important to first experiment in one of the subsidiaries before rolling it out to the entire organisation.

*“Yes, data governance can be trailed. The business trailed out data governance at one of its subsidiaries which advocated for the need for the business to have a centralised data governance capability to govern data for the group.” (P1)*

*“Yes, we tried it at the subsidiary, to see how people in that part of the business would accept it. This has given birth to a group-wide drive to implement data governance for the whole group.” (P3)*

*“Yes, it was trailed at the subsidiary. Even though it was not all successful however there were a lot of lessons learned to be carried forward in the implementation of data governance for the group.” (P5)*

### 5.5.5 Observability

The participants were asked to share their sentiments on the need for data governance to be observable by making the outcomes of data governance visible to the organisation. The consensus was that it is important to share the outcomes of data governance and show value to the business for the business to buy into the change. This was done by measuring data quality and showing the business the improvement in data quality.

*“The measure for data governance in the organisation was linked to the progress made in addressing the master data challenges the subsidiary had at the time. The improvement in data quality was the crucial measure that the business used to measure and track the progress of data governance.” (P1)*

*“The data quality scorecards that were developed were a way to measure the effectiveness of data governance. The business did use some of these scorecards and did see value in tracking their data quality.” (P5)*

*“Through the implemented data quality system, we designed scorecards that measured the quality of our data and identified data issues. We were able to track the data quality proactively. The merchants and planners used these scorecards to troubleshoot issues in their spaces and also track the quality of their data.” (P3)*

In this section, the five characteristics of innovation were discussed about data governance. It was found that the characteristics of innovation were also relevant in the implementation of data governance. However, there was a question on whether or not data governance has to be compatible with the organisation where there were participants that believed that data governance did not necessarily have to be compatible as it was a mechanism that should be used to facilitate change and improve business processes and not be a slave to the legacy processes that exist in the business.

### **5.6 Research Question 3: Communication channels used to disseminate data governance.**

As highlighted in the literature review, the communication channel is one of the four main elements of the diffusion of innovation theory. It refers to how the innovation is disseminated to the organisation or social system (Rogers, 2003). The participants were asked which communication channels were used to disseminate data governance to the organisation and provide their perception of the communication channels that were effective in the communication process of data governance.

There was consensus that a combination of communication channels at various phases of the implementation would have to be used to effectively communicate data governance to the organisation. Interpersonal channels in the form of training and information sessions were used extensively at the beginning of the implementation to educate the business on data governance. Those were found to be effective in the subsidiary where data governance was trailed.

*“When we introduced data governance in the subsidiary, there were many training sessions with the departments and groups of functions put together to be able to talk through what we wanted to achieve... It was very effective in the area where we did it because it was controllable it becomes a little more difficult as it becomes a group function” (P1)*

In addition to the interpersonal channel, the participants also shared the same sentiments in using existing mass media communication such as the organisation's town hall meetings, newsletters and how to guide flyers to reach a much wider group of the organisation and create awareness.

*“It starts with people being aware that this is happening, it can be a launch in one of our town hall sessions and we can also run quiz competitions in our newsletters. It is fundamental that people are aware” (P14)*

*“I think it needs to be described in a town hall. Not once, it needs to be reiterated and it needs to have a fairly regular slot to explain the process and outcomes and what is in it for the people who need to change their ways to conform” (P10)*

*“There were how-to flyers, and presentations to various stakeholders across the organisation mostly online which meant we could reach a wider audience. In-person communication as well as mass communication in the form of emails. A combination of in-person and online meetings with digital flyers sent through to reinforce the message.” (P3)*

Although digital forms of communication channels were advantageous in terms of reaching people, one participant highlighted that it might be a challenge in the retail organisation as there are users on the shop floor and warehouses who do not always have access to electronic devices stressing the importance of neglecting in-person communication.

*“In a retail organisation it is quite challenging because we have users literary on the shop floor that does not necessarily have access to PCs, you have got office staff that do which makes it a bit easier, and then you have people in warehouses that would have some challenges accessing electronic communication. There a hybrid approach with a combination of electronic and in-person communication is important.” (P15)*

This section discussed the communication channels that can be used to disseminate data governance in the organisation. It was revealed that interpersonal channels in the form of training were mostly used to communicate data governance as the organisation first needed to be educated about data governance and what impact it would have on them. The participants also stressed that it had to be a combination of mass communication channels and interpersonal communication was important to reach a much wider group and drive awareness in the organisation.

#### **5.7 Research Question 4: Impact of the social system on the implementation of data governance**

It was revealed in the literature review that data governance involves both people and technology because data by itself lacks both will and intent and is shaped by the people. As a result, the successful implementation of data governance required people to alter their behaviour and the way they interact with data from its generation to its consumption and disposal. A social system is one of the four main elements in

the diffusion of innovation theory which states that the social system can either facilitate or impede the diffusion of innovations in the system. Therefore, the participants were asked what impact they believed the social system has in the implementation of data governance. This was to provide the context for the role that the organisation's social system had in the implementation of data governance.

The majority of the participants agreed that the organisation's social system had a massive impact on the implementation of data governance and that it was important for the social system to buy into the change for it to be successful.

*“The social system has a huge impact on the implementation because at the end of the day many of the people are either providing the data or they are providing mechanisms and somehow giving input into either the data itself or the mechanisms that manage or create the data” (PM 15)*

*“It is very important ... we had to find and build soldiers across the different departments and across the different businesses to become spokespeople for data governance, but it had to be people that understood the systems used in the organisation and they can be able to train their teams and also direct them to the right people when there are queries and questions.” (PM 7)*

One participant mentioned that the leaders and colleagues in the organisation are more inclined to listen more when changes or innovations come from someone external and are being paid to provide the service.

*“And something that I have learned here is that sometimes our leaders and our colleagues tend to listen more if innovation comes from someone external that's getting paid a lot of money to say exactly what you are saying” (P13)*

Another participant mentioned that when the data governance systems were introduced, it changed the behaviour because the employees have no choice because a new system was introduced, and they had to adapt it to continue with their daily tasks.



*“The introduction of the data governance systems changed behaviour because that was the new way of doing things. So, they did not really have a choice, it came with a new system and that’s how it was” (P10)*

In this section, we discussed the impact of the social system on the implementation of data governance. It was found that there was consensus that the social system was essential to the implementation of data governance. The organisation has gone through a process of identifying change agents which they called soldiers to assist in embedding data governance in their teams and departments. There was also an interesting dynamic that came up where a participant shared that in the organisation, the voice of an external person is better heard than that of an internal person. Additionally, in the pilot implementation of data governance, a system was introduced which gave the social system no option but to adopt the change.

### **5.8 Research Question 5: Impact of data governance on data-driven decision-making**

In the literature, it was revealed that data-driven decision-making is one of the outcomes of data governance. The participants were asked about their perception of the impact that data governance has on data-driven decision-making in the organisation. Their consensus was that data governance has had an impact on data-driven decision-making in the organisation, even though the organisation is still on the journey of implementation however because the business could start trusting the data, they are now using the data even more to make decisions.

*“Decision-making very much leans on the trust in the data, we have seen it with Business Intelligence systems and reporting systems, if the business doesn’t trust the report, they are going to run five other reports to validate the current report which results in it taking five times longer to make a decision. Data governance plays a significant role in making sure that the data is structured in such a way that everything built on top of it is reliable and trustworthy” (P7)*

*“Data governance will guide decision making which will make decision making quicker, easier and more responsible” (P9)*

*“In retail, it is important to have the correct data to drive decisions. If the data is not correct stock calculation will not be correct and forecasting will be skewed, and you could make wrong decisions. However, with data governance rules are put in place to make sure that the data is as correct as possible” (P4)*

However, there was one participant that data governance does not have a significant impact on data-driven decision-making. The participant believes that people need to first understand how to use data to drive data-driven decision-making before introducing data governance.

*“I don’t necessarily think that data governance plays a heavy role in data—driven decision-making. For me, the organisation needs to first understand the data before it starts to apply any level of governance. The decision as to what the data is, what it is doing and not doing and what it should be doing first needs to be understood in terms of that decision needs to be clear before we even start to think about building governance” (P15)*

This section discussed the impact of data governance on data-driven decision-making. The participants believed that data-driven decision-making would be faster as data governance will ensure that the data is structured as well have built reliability and trust in the data. However, there was a participant that did not believe that data governance has a significant impact on data driven-decision-making but rather the understanding of the data.

## **Chapter 6: Discussion of findings**

### **6.1 Introduction**

This chapter aims to offer insight into the implementation of data governance in a retail organization through the perspective of the diffusion of innovation theory considering the interview-based findings highlighted in chapter 5's discussion. This will be accomplished by using practice reports to triangulate the findings and an interpretation of the results based on the literature reviewed in chapter 2 and other relevant sources. The results will either confirm, refute, or add to the body of literature already written in this field.

### **6.2 Research Question 1: Perceived need for data governance**

According to Rogers (2003), as highlighted in chapter 2, the initial step in the innovation process in organisations is agenda setting, where the organisation identifies an important organisation-wide problem and seeks an innovation that has the potential to effectively address the recognised problems. This implies that there is always justification for the introduction of innovation in an organisation.

In the interview discussion with participants detailed in section 5.5.1 to establish their perception of the role of data in the organisation identified data as being crucial for decision-making in the organisation. This is in support of literature highlighted in section 2.2 that argued that data especially big data is essential for decision-making in the increasingly competitive retail industry. Gawankar et al (2020) emphasised that retail organisations that leveraged their data saw an increase in their decision-making performance which yielded increased business performance.

In section 5.5.2, the interview discussion was to establish their perception of data governance a subset of the participants perceived data governance as rules, standards and policies that inform how data is managed in an organisation. While another subset perceived data governance as a means to manage data quality in the organisation and ensure that information derived from the data is consistent across the organisation. In contrast, data governance in section 2.4.1 taken from the DAMA international body of knowledge data governance “is the exercise of authority and control (planning, monitoring and enforcement) over the management of data assets”

(p 67). The perceptions provided by the participants speak to the “how” of data governance and are consistent with Cheong and Chang’s (2007) data framework components and Weber et al. (2009) data governance framework which proposes a paradigm for enterprise-wide data quality management consecutively.

In section 5.5.3, the interview discussion was to establish their perceived need for data governance. The shared consensus was the urgency to improve the data quality in the organisation which was impacting the business negatively. This also fueled the lack of trust in the data due to inconsistencies and lack of quality. This is consistent with section 2.4.2 where it is reported that poor data quality harms employee productivity and wastes the organisations’ resources. Therefore, a relevant problem that can be resolved through the implementation of data governance is highlighted by Wende (2007) in his proposed data governance framework that focuses on data quality management. Additionally, it also emerged that the lack of data ownership was a need that has the potential to be resolved by data governance and is consistent with Otto (2011b) who states that control, organisational structure and roles and responsibilities are crucial for effective data governance.

In conclusion, the perceived needs for data governance in the selected organisation were identified to be the recognition of data as a key resource for decision-making. An identified problem was the challenges with poor data quality that was harming the organisation, the lack of trust in the data due to the lack of standardization and lastly the lack of data ownership in the organisation. The participants shared the sentiments that data governance has the potential to address perceived needs.

### **6.3 Research Question 2: Perceived innovation characteristics of data governance**

The second phase of the innovation process in an organisation is matching which refers to the matching of the perceived needs and the innovation (Rogers, 2003). The innovation needs to fit the problem, and this leads to the decision to adopt or reject the innovation before implementation. This is done through the application of the five characteristics of innovation by the diffusion of innovation theory.

### **6.3.1 Relative advantage**

In section 5.5.1, the interview discussion was to establish the participants' perception of the relative advantage of data governance. Most of the participants believed that with data governance the organisation would have better control of their data and the organisation would have a better chance at protecting its data assets. This is aligned with Otto (2011b) in section 2.4.4 who states that control is vital to the goals and structures of data governance.

Another relative advantage the participants mentioned was improved, where most of the participants believed that data governance would improve the quality of data in the organisation. As highlighted in section 2.4.3 by Varshney and Allen (2020) one of the main goals of data governance is good data quality. The organisation also deployed technology to monitor data governance.

Lastly, improved efficiency and better decision-making were also highlighted as relative advantages of implementing data governance. The participants shared that the organisation would be more efficient and would be able to make better decisions faster. Al-Ruithe et al. (2019) state that data governance programs that are well-formulated and implemented successfully generate benefits for the organisation such as better decision-making and the reduction of operational conflicts which increases efficiencies. Therefore, this relative advantage is supported by the literature.

In conclusion, the identified relative advantages of the implementation of data governance as identified by the participants include better control over the data, improved data quality and improved efficiency and better decision-making.

### **6.3.2 Compatibility**

In section 5.5.2, the interview discussion was to establish if the participants believed that data governance had to be compatible with the organisation. The consensus was that it had to be compatible with the organisation. Two participants believed that although there had to be an alignment however data governance "should not be a slave" (P10) to the legacy processes and systems. However, Rogers (2003) in section 2.5.1 stresses that if an innovation is not compatible with the social system features that have the potential of delaying the adoption of the innovation. This might

also impact the decision in the innovation process to either adopt or reject the innovation before the implementation.

In conclusion, to facilitate the decision to adopt the innovation needs to fit the organisation's problem therefore the innovation needs to be compatible with the organisation.

### **6.3.3 Complexity**

In section 2.5.1, Rogers (2003) states that complexity is the level to which an innovation is perceived to be difficult to understand and use. The difficulty to understand and use can act as an obstacle to the adoption of the innovation. In section 5.5.3, the interview discussion was to establish if data governance was perceived to be complex in the organisation. The participants all agreed that data governance was complex in the organisation citing reasons for the organisation being in the early stages of the journey and the need to build data literacy in the organisation. Also understanding the business and related impact, the organisation structure not being conducive to simplifying data governance and lastly, the organisation operates using different systems that are not in sync and are not integrated. This has the potential to delay or even impede the successful implementation of data governance.

### **6.3.4 Trialability**

In section 5.5.4, the interview discussion was to establish the participants' perception of the trialability of data governance. The participants all agreed that it was important to trail data governance and implement it in a phased approach. This was indeed the approach that the organisation adopted by implementing data governance at one of their subsidiaries first before embarking on the organisation-wide implementation. Syahadiyanti & Subriadi (2018) in section 2.5.1 argue that innovations that can be trailed on an instalment basis are adopted faster than innovations that cannot be broken down into smaller pieces

### **6.3.5 Observability**

Rogers (2003) in section 2.5.1 posits that the visibility of the results and outcomes of an innovation stimulates peer discussion about the innovation which can authenticate the importance of the innovation and thus accelerate the adoption of the innovation. The interview discussion in section 5.5.5 was to establish the sentiments of the participants on the observability of data governance. The consensus was that it was important to share the outcomes of the data governance and show value to the business to facilitate buy-in from the business. This was done through data quality scorecards that were used to track the quality of data in the organisation. Given that data quality was and remains a challenge for the business being able to track the improvement of data quality could validate to the business the importance of having data governance.

In conclusion, the five characteristics of innovation have described the perceived innovation characteristics of data governance.

### **6.4 Research Question 3: Communication channels used to disseminate data governance**

In section 2.5.2, Thomas (2006) shared that at the Data Governance Conference leaders who had successfully implemented data governance revealed that 80 to 95 per cent of the implementation of data governance involved communication. Wang et al. (2021) also stress the importance of communication channels as factors to consider when disseminating a new idea. Therefore, it is important to effectively communicate during the implementation of data governance which requires the organisation to be strategic with the communication used to disseminate data governance.

In section 5.6, the interview discussion was to establish the communication channels that were used to disseminate data governance to the organisation. The consensus was that a combination of communications at various stages of the implementation would have to be used to effectively communicate data governance to the organisation. In the initiation phases, interpersonal channels such as training and information sessions were used to start educating the business about data governance.

In the later stages of the implementation, mass media was adopted to start spreading awareness about data governance and the wider organisation being brought along on the implementation journey. This was using town hall meetings, how-to guides, and company newsletters to spread the innovation. This approach is supported by Rogers (2003) in section 2.5.2 who states that different communication channels play distinct roles at the different stages in the innovation process.

Given the widespread use of virtual meetings and digital communication, one participant cautioned against only relying on digital forms of communication as in the retail organisation warehouse and shop floor workers might have access to view the digital communication.

In conclusion, a combination of communication channels was used at different stages of the implementation process. Interpersonal channels such as training and information sessions were used to educate the business about data governance. Then mass media channels were used such as the organisation's town hall meetings, how-to guides, and newsletters to drive up awareness and also re-enforce the message.

#### **6.5 Research Question 4: Impact of social system on the implementation of data governance**

The social system outlined in section 2.5.4 is one of the four main elements of the diffusion of innovation theory. Rogers (2003), in the same section, notes that a social system can facilitate or hinder the ability of innovation to spread across the organisation. In section 5.7, the interview discussion was to establish the perceptions of the participants on the impact of the social system on the implementation of data governance.

The perception from most of the participants was that the social system had a great impact on the implementation of data governance and that it was important for the social system to buy into the change for it to be successful. It was highlighted that during the implementation they had to build "soldiers" across the organisation so that



they would be able to train their team members. Rogers (2003) in section 2.5.4 calls these individuals opinion leaders, who are members of the system that influence other members of the system and their influence is based on expertise, competence, accessibility, and leadership in conforming to the system. Building up these individuals from the systems may facilitate the implementation of data governance within their systems.

In contrast, it was also highlighted by another participant that new ideas or innovations that come from someone external are far more accepted Lundblad (2003) in section 2.5.4 calls these individuals change agents who are external to the system, however, represent innovation and change. They often use opinion leaders to gain acceptance within the social system to be able to diffuse innovation.

Another participant highlighted that in the implementation of data governance in the pilot subsidiary, a new data governance system was introduced, and the social system must adapt to the new system as an authority decision was made to use the new systems. This refers to the innovation decisions which can be optional, collective or driven by authority. In this instance, the decision was driven by the authority.

In conclusion, the social system has a major impact on the implementation of data governance as it can facilitate or impede the diffusion of innovation. It is important to use opinion leaders and change agents as they influence the social systems, it may assist in driving the implementation of data governance.

### **6.7 Research Question 5: Impact of data governance on data-driven decision-making**

As highlighted in section 2.4.3 by Al-Ruithe et al. (2019), one of the many benefits of data governance is better decision-making. This is due to the vital role that data governance plays in promoting trust in data initiatives (Brous & Janssen, 2020). Alhassan et al. (2019b) argued that the availability and accessibility of data at the right time and in the right format enabling decision-makers to make better decisions timeously is one of the critical success factors of an effective data governance program.

In section 5.8, the interview discussion aimed to establish the perception of the participants on the impact of data governance on data-driven decision-making. The consensus on the perception was that data governance has a significant impact on data-driven decision-making. Even though the organisation is still on the implementation journey they are starting to see the potential outcomes of data governance. The confidence in the data is increasing and the business is starting to use data even more for decision-making.

In conclusion, data governance has a positive impact on data-driven decision-making. Because of more confidence in the data and with the increased availability of the data, the organisation will increasingly use data to make decisions.

## 6.8 Conclusion

Table 3 depicts an updated codebook using the literature code book initially developed (see Appendix 4) as a basis as well as incorporating the research findings based on the newly identified codes.

**Table 3: Updated code boo**

Research Question	Main Codes	Sub Codes
Research Question 1: Perceived needs from data governance	Perceived Needs	Data is crucial for decision making
		Challenges with data quality impacting business
		Data governance as rules and standards to derive value from data
Research Question 2: Perceived innovation characteristics of innovation	Relative Advantage	Control
		Data Quality
		Efficiency and better decision-making
	Compatibility	
	Complexity	
	Observability	
Research Question 3: Communication channels used to disseminate data governance	Mass communication channels	
	Interpersonal communication channels	
Research Question 4: Impact of social system	Social structure	
	System norms	Culture
		Data Literacy
	Opinion leaders and change agents	
	Innovation decisions	Optional
		Collective
Authority		

Research Question 5: Impact of data governance on data-driven decision making	Data-driven decision making	
---	-----------------------------	--

## **Chapter 7: Conclusion**

This chapter will first reiterate the research objectives outlined in chapter 1. Following that, it will summarise the main conclusions and observations about the study questions from chapter 3 considering the literature review from chapter 2, as well as the findings and discussion from chapters 5 and 6, respectively.

### **7.1 Research Background and Objectives**

Making accurate, timely, and high-quality decisions are no longer just necessary for business operations; it is a question of survival for many organisations as they traverse complex and competitive business environments. Data governance is increasingly becoming important for organisations to derive value from their data assets and leverage them in their decision-making to gain a competitive advantage. However, many organisations today still struggle with the implementation of data governance and there is a lack of literature on the pragmatic implementation of data governance.

The research objective was to study the implementation of data governance in a retail organisation in South Africa using Rogers' (2003) diffusion of innovation theory. The researcher aimed to formulate a framework that can aid the implementation of data governance in the organisation.

### **7.2 Principal Findings**

This section summarises the main findings based on interviews with the senior employees at the selected retail organisation.

#### **7.2.1 Perceived needs for data governance**

Establishing what the participants felt were the key needs driving the implementation of the organisation was important as it set a foundation which explained why the organisation was looking to implement data governance as well as the problems data governance would be addressed for the organisation.

The perceived needs identified by the participants were the recognition of data as a key resource for decision-making which drove the need to better manage the data to facilitate decision-making as supported by Cheong and Chang (2007). Identified problems were the challenges with poor data quality, which was harming the business, the lack of data ownership which implies the lack of accountability on data issues as well as the lack of trust in the data which was driven by the poor data quality and the lack of standardisation. Data governance was viewed as the potential solution to addressing the above-mentioned needs. This aligns with Rogers (2003) agenda-setting phase in the innovation process in organisations.

### **7.2.2 Perceived innovation characteristics for data governance**

The question was addressed with the use of the five innovation characteristics, namely: relative advantage, compatibility, complexity, trialability and observability. The relative advantages highlighted in the study were better control over the data, improved data quality, increased efficiency and better decision making these were supported by (Alhassan et al., 2019a, 2019b; Brous & Janssen, 2020; Jimenez et al., 2019). It was established that it was important for data governance to be compatible with the organisation however the organisational structure will also have been altered somewhat to fit data governance. This is consistent with the Redefining/ Restructuring phase in the innovation process by Rogers (2003).

It was highlighted that data governance in the selected organisation is complex due to the complex organisational structure of the organisation and the disparate legacy systems that do not talk to each other. The complexity is viewed as an impediment to the introduction of data governance. For the social system to adopt an innovation, it needs to be easy to understand and use (Rogers, 2003).

It was also established that a phased approach to data governance can aid the successful implementation of data governance. Being able to trail it reduces the perceived risk (Rogers, 2003) of data governance harming the organisation as the implementation can be divided into smaller more manageable chunks to not overwhelm the organisation. This is consistent with the clarification phase in the innovation process where Rogers (2003) cautions against rapid implementation as the organisation begins to define the relationship between the organisation and the innovation in the context of the study being data governance.

Lastly, the observability of data governance becomes important when implementation. This ensures that the organisation has visibility on the results and outcomes of data governance this creates dialogue amongst peers about data governance which can facilitate further implementation of data governance Rogers (2003). This also facilitates the routinizing of data governance.

### **7.2.3 Communication channels used to disseminate data governance**

Communication is critical to the implementation of data governance (Thomas, 2006). Rogers (2003) states that different communication channels play distinct roles at different stages in the innovation process. From the participants, it was established that a combination of both interpersonal and mass media communication channels needs to be used to disseminate data governance to the organisation. It is important that communication not only drives awareness but that it is also educational and educates the business about data governance, the objectives as well as what is in it for them.

### **7.2.4 Impact of social system on the implementation of data governance**

It was established from the participants that the social system has a significant impact on the implementation of data governance. Diffusion occurs within an organisation which is a social system. Lundberg et al. (2019) state that it can be challenging to put an idea into practice because there are frequently many stakeholders—both supporters and detractors—involved. In addition, because people have typically invested in and committed to existing organizational programs and processes, they are typically reluctant to change them (Lundberg et al., 2019). Therefore, it is critical to involve the social system in the implementation because often they are most impacted by the change.

### **7.2.5 Impact of data governance on data-driven decision-making**

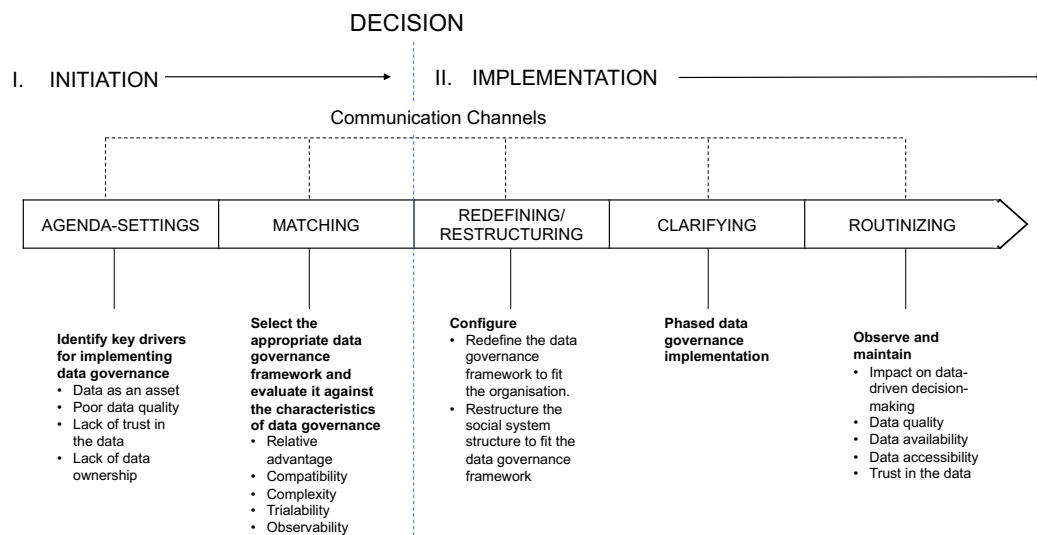
It was established from the participants that data governance has a positive impact on data-driven decision-making. This was due to the increased confidence in the data as well as the availability and accessibility of the data. This meant that the organisation could increasingly start adopting the use of data in making decisions. Various literature on data governance supports that data governance has an impact

on the data-driven decision-making (Alhassan et al., 2019a; Brous & Janssen, 2020; Jimenez et al., 2019)

### 7.3 Proposed framework for data governance implementation

Considering the results and the discussion of the results in Chapters 5 and 6 within the context of the literature reviewed in chapter 2, the researcher suggests the framework in Figure 9 to assist the retail industry in South Africa with the implementation of data governance.

**Figure 9: Proposed data governance implementation framework (Adapted from Rogers (2003))**



The framework is adapted from the five stages of the innovation process in organisations by Rogers (2003) in the diffusion of innovation theory. Similarly, to the innovation process, the framework follows the five stages sequentially and is split between initiation and implementation.

The agenda-setting stage entails the organisation identifying the key drivers for implementing data governance. The common drivers include the recognition of data as an asset and key problems such as poor data quality, the lack of trust in the data as well as the lack of data ownership in the organisation. However, the drivers are not limited to these organisations and can have more or even different key drivers.

The matching stage entails selecting or building the appropriate data governance framework. As demonstrated in the literature review, there are various data governance frameworks that exist in academia as well as in practice. And as demonstrated, the frameworks are different and may address different data requirements and challenges. It is therefore important to choose the most appropriate. The data governance framework will then be evaluated using the five characteristics of innovation namely relative advantage, compatibility complexity, trialability and observability. This will determine whether the data governance framework chosen fits the key drivers in the organisation's agenda. From there the decision to either adopt or reject the data governance will be made.

If the decision is to adopt data governance, the process will move to the redefining and restructuring phase. Here the organisation will configure the data governance framework to fit the organisation and also restructure the social system to fit the data governance. This is to ensure that the implementation of data governance is set up for success.

In the clarifying stage, a phased approach to the data governance elements is implemented. In this phase, the relationship between data governance and the organisation becomes more clearly defined as the organisation starts to implement and use the elements of data governance. Lastly, the routinizing phase which entails observing and maintain data governance against the key drivers identified in the agenda-setting stage. The objective of this phase is to embed data governance into the business processes and systems that become part of the way things are done in the organisation.

It is important to note that communication channels appear in all stages of innovation. It is important to ensure that there is constant communication in the implementation of data governance. It is also important to ensure that the appropriate communication channels are used at every stage of the innovation process.

In conclusion, the study has demonstrated that the diffusion of innovation theory innovation process can be applied in the implementation of data governance. It adds to the body of knowledge a framework that can facilitate the implementation of data governance.



## **7.4 Implications for management and other relevant stakeholders**

The following suggestions are put out for the relevant data governance stakeholders in the retail business based on the findings.

Data governance is increasingly recognised as a means to derive maximum value from data assets. However, the successful implementation of it in organisations remains a challenge. Managers must pay special attention to understanding the objectives for implementing data governance and then developing a data governance framework based on those objectives. One size does not fit all when it comes to data governance.

Communication is an essential part of implementing data governance and it is also just as important to ensure that suitable communication channels are used in the various stages of implementation. Additionally, data governance is a journey, therefore, following a phased approach to the implementation of data governance gives the organisation the time and space to trail out data governance which as stated in the literature has a significant impact on the adoption of data governance.

## **7.5 Limitations of the research**

Qualitative research is considerably more prone to bias. This is a natural result of the endeavour to probe further into the existing understanding of any field of study.

With the time restrictions in conducting the research, the researcher is aware that the sample selected might not be a full representation and may be skewed.

The study was conducted using a single case study strategy in a retail organisation, the researcher is therefore aware that the findings from the study may not be generalisable to other organisations in the same industry.

Because this is a personal interest study done in the industry where the researcher worked, they are aware of their own bias. Due to the researcher's potential tendency to overemphasize some aspects of the topic based on personal experience, the interpretations and findings of this research may be biased.

Despite all these limitations, the researcher is certain that attention was done to evaluating the literature and the interview method can be deemed sufficient to reduce these indicated biases.

### **7.6 Contributions and suggestions for future research**

Since qualitative research is exploratory in nature, it aims to learn more about the topic under investigation, albeit this does not always produce answers that can be considered conclusive but rather point to another research that needs to be done.

This research study contributes to the existing body of knowledge on data governance through the probe of the implementation of data using the diffusion of innovation theory's innovation process in a selected listed retail organisation in South Africa. Furthermore, the study presented a framework for the implementation of data governance adapted from the diffusion of innovation theory to aid in the implementation of data governance.

Future research on this subject can expand through testing the proposed framework using a larger set of case study organisations. Additionally, the study can test the framework in a different industry.

## References

- Ababneh, T. A. M., & Aga, M. (2019). The impact of sustainable financial data governance, political connections, and creative accounting practices on organizational outcomes. *Sustainability (Switzerland)*, 11(20). <https://doi.org/10.3390/su11205676>
- Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, 49, 424–438. <https://doi.org/10.1016/j.ijinfomgt.2019.07.008>
- Acharya, A., Singh, S. K., Pereira, V., & Singh, P. (2018). Big data, knowledge co-creation and decision making in fashion industry. *International Journal of Information Management*, 42, 90–101. <https://doi.org/10.1016/j.ijinfomgt.2018.06.008>
- Akca, Y., & Ozer, G. (2014). Diffusion of Innovation Theory and an implementation on Enterprise Resource Planning Systems. *International Journal of Business and Management*, 9(4), 92–114. <https://doi.org/10.5539/ijbm.v9n4p92>
- Akter, S., Bandara, R., Hani, U., Fosso Wamba, S., Foropon, C., & Papadopoulos, T. (2019). Analytics-based decision-making for service systems: A qualitative study and agenda for future research. *International Journal of Information Management*, 48, 85–95. <https://doi.org/10.1016/j.ijinfomgt.2019.01.020>
- Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: a systematic review and agenda for future research. *Electronic Markets*, 26(2), 173–194. <https://doi.org/10.1007/s12525-016-0219-0>
- Al-Badi, A., Tarhini, A., & Khan, A. I. (2018). Exploring Big Data Governance Frameworks. *Procedia Computer Science*, 141, 271–277. <https://doi.org/10.1016/J.PROCS.2018.10.181>
- Al-Ruithe, M., Benkhelifa, E., & Hameed, K. (2018). Data governance taxonomy: Cloud versus non-cloud. *Sustainability (Switzerland)*, 10(1). <https://doi.org/10.3390/su10010095>
- Al-Ruithe, M., Benkhelifa, E., & Hameed, K. (2019). A systematic literature review of data governance and cloud data governance. *Personal and Ubiquitous*

- Computing*, 23(5–6), 839–859. <https://doi.org/10.1007/s00779-017-1104-3>
- Alhassan, I., Sammon, D., & Daly, M. (2018). Data governance activities: a comparison between scientific and practice-oriented literature. *Journal of Enterprise Information Management*, 31(2), 300–316. <https://doi.org/10.1108/JEIM-01-2017-0007>
- Alhassan, I., Sammon, D., & Daly, M. (2019a). Critical success factors for data governance: a telecommunications case study. *Journal of Decision Systems*, 28(1), 41–61. <https://doi.org/10.1080/12460125.2019.1633226>
- Alhassan, I., Sammon, D., & Daly, M. (2019b). Critical Success Factors for Data Governance: A Theory Building Approach. *Information Systems Management*, 36(2), 98–110. <https://doi.org/10.1080/10580530.2019.1589670>
- Aloysius, J. A., Hoehle, H., Goodarzi, S., & Venkatesh, V. (2018). Big data initiatives in retail environments: Linking service process perceptions to shopping outcomes. *Annals of Operations Research*, 270(1–2), 25–51. <https://doi.org/10.1007/S10479-016-2276-3>
- Benfeldt, O., Persson, J. S., & Madsen, S. (2020). Data Governance as a Collective Action Problem. *Information Systems Frontiers*, 22(2), 299–313. <https://doi.org/10.1007/s10796-019-09923-z>
- Bhattacharjee, A. (2019). *Social Science Research: Principles, Methods and Practices (Revised edition)*.
- Bowen, C. M., Stanton, M., & Manno, M. (2012). Using diffusion of innovations theory to implement the confusion assessment method for the intensive care unit. *Journal of Nursing Care Quality*, 27(2), 139–145. <https://doi.org/10.1097/NCQ.0B013E3182461EAF>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706QP063OA>
- Brous, P., & Janssen, M. (2020). Trusted decision-making: Data governance for creating trust in data science decision outcomes. *Administrative Sciences*, 10(4). <https://doi.org/10.3390/admsci10040081>
- Brous, P., Janssen, M., & Krans, R. (2020). Data Governance as Success Factor for Data Science. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*,

- 12066 LNCS, 431–442. [https://doi.org/10.1007/978-3-030-44999-5\\_36](https://doi.org/10.1007/978-3-030-44999-5_36)
- Brous, P., Janssen, M., & Vilminko-Heikkinen, R. (2016). Coordinating decision-making in data management activities: A systematic review of data governance principles. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9820 LNCS, 115–125. [https://doi.org/10.1007/978-3-319-44421-5\\_9](https://doi.org/10.1007/978-3-319-44421-5_9)
- Cantrell, M. A. (2011). Demystifying the research process: understanding a descriptive comparative research design. In *Pediatric nursing* (Vol. 37, Issue 4, pp. 188–189). <https://web-s-ebSCOhost-com.uplib.idm.oclc.org/ehost/pdfviewer/pdfviewer?vid=1&sid=14dac2b8-800d-4000-b360-03e85cf1570f%40redis>
- Chatterjee, S., Chaudhuri, R., & Vrontis, D. (2021). Does data-driven culture impact innovation and performance of a firm? An empirical examination. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-020-03887-z>
- Chen, Q. (2022). *China Wants to Put Data to Work as an Economic Resource—But How?* Digichina. <https://digichina.stanford.edu/work/china-wants-to-put-data-to-work-as-an-economic-resource-but-how/>
- Cheong, L., & Chang, V. (2007). The Need for Data Governance: A Case Study. *Association for Information Systems*. <http://aisel.aisnet.org/acis2007/100>
- Clarke, V., & Braun, V. (2017). Thematic analysis. In *Journal of Positive Psychology* (Vol. 12, Issue 3, pp. 297–298). <https://doi.org/10.1080/17439760.2016.1262613>
- Creswell, J. W., & Creswell, J. D. (2018). Research Design: Qualitative, Quantitative and Mixed Methods Approaches. In *Research Design: Qualitative, Quantitative and Mixed Methods Approaches* (5th ed.). SAGE Publications, Inc.
- DAMA International. (2017). DAMA-DMBOK: Data Management Body of Knowledge (2nd Edition) July 2017. In *Technics Publications, LLC 14 Elm St Denville NJ United States* (Vol. 44, Issue 8).
- Data Governance Institute. (n.d.). *About the Data Governance Institute*. Data Governance Institute. Retrieved October 4, 2022, from <https://datagovernance.com/about-us/>

- Dearing, J. W., & Cox, J. G. (2018). Diffusion of innovations theory, principles, and practice. *Health Affairs*, 37(2), 183–190. <https://doi.org/10.1377/HLTHAFF.2017.1104>
- Delacroix, S., & Lawrence, N. D. (2019). Bottom-up data Trusts: Disturbing the 'one size fits all' approach to data governance. *International Data Privacy Law*, 9(4), 236–252. <https://doi.org/10.1093/IDPL/IPZ014>
- Denscombe, M. (2010). The Good Research Guide: For Small-Scale Social Research Projects. In *Nurse Researcher* (4th ed.). Open University Press.
- Dolma, S. (2010). The central role of the unit of analysis concept in research design. *Istanbul University Journal of the School of Business Administration*, 39(1), 169–174. <http://eds.b.ebscohost.com.uplib.idm.oclc.org/ehost/pdfviewer/pdfviewer?vid=33&sid=557e1ca9-810b-4e4f-ae0a-c29c73493ce2@sessionmgr103&hid=117>
- Doyle, G. J., Garrett, B., & Currie, L. M. (2014). Integrating mobile devices into nursing curricula: Opportunities for implementation using Rogers' Diffusion of Innovation model. *Nurse Education Today*, 34(5), 775–782. <https://doi.org/10.1016/j.nedt.2013.10.021>
- Dryden-Palmer, K. D., Parshuram, C. S., & Berta, W. B. (2020). Context, complexity and process in the implementation of evidence-based innovation: A realist informed review. *BMC Health Services Research*, 20(1). <https://doi.org/10.1186/S12913-020-4935-Y>
- Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *The Academy of Management Review*, 14(4), 532. <https://doi.org/10.2307/258557>
- Flick, U. (2018). *The SAGE Handbook of Qualitative Data Collection*. SAGE Publications Ltd.
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137–144. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- Gawankar, S. A., Gunasekaran, A., & Kamble, S. (2020). A study on investments in the big data-driven supply chain, performance measures and organisational performance in Indian retail 4.0 context. *International Journal of Production Research*, 58(5), 1574–1593.

- <https://doi.org/10.1080/00207543.2019.1668070>
- Goldkuhl, G. (2017). *European Journal of Information Systems Pragmatism vs interpretivism in qualitative information systems research*. <https://doi.org/10.1057/ejis.2011.54>
- Gontcharov, I. (2017). Chapter 15: Qualitative Ethics in a Positivist Frame: The Canadian Experience 1998-2014. In *The SAGE Handbook of Qualitative Research Ethics* (pp. 231–247). SAGE Publications, Inc. <https://www.researchgate.net/publication/315703702>
- Griva, A., Bardaki, C., Pramataris, K., & Papakiriakopoulos, D. (2018). Retail business analytics: Customer visit segmentation using market basket data. *Expert Systems with Applications*, 100, 1–16. <https://doi.org/10.1016/j.eswa.2018.01.029>
- Guest, G., Bunce, A., & Johnson, L. (2006). How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. *Field Methods*, 18(1), 59–82. <https://doi.org/10.1177/1525822X05279903>
- Hossain, A., Akter, S., & Yanamandram, V. (2020a). Revisiting customer analytics capability for data-driven retailing. *Journal of Retailing and Consumer Services*, 56, 102187. <https://doi.org/10.1016/j.jretconser.2020.102187>
- Hossain, A., Akter, S., & Yanamandram, V. (2020b). Revisiting customer analytics capability for data-driven retailing. *Journal of Retailing and Consumer Services*, 56, 102187. <https://doi.org/10.1016/j.jretconser.2020.102187>
- Iyamu, T. (2018). Collecting qualitative data for information systems studies: The reality in practice. *Education and Information Technologies*, 23(5), 2249–2264. <https://doi.org/10.1007/s10639-018-9718-2>
- Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy Artificial Intelligence. *Government Information Quarterly*, 37(3). <https://doi.org/10.1016/j.giq.2020.101493>
- Jia, L., Hall, D., & Song, J. (2015). The conceptualization of data-driven decision making capability. *2015 Americas Conference on Information Systems, AMCIS 2015*. <https://www.researchgate.net/publication/280051851>
- Jimenez, L. M., Polo, J. A., & Duarte, N. A. (2019). Overview of Data Governance

- in Business Contexts. *IOP Conference Series: Materials Science and Engineering*, 519(1). <https://doi.org/10.1088/1757-899X/519/1/012023>
- Karkošková, S. (2022). Data Governance Model To Enhance Data Quality In Financial Institutions. *Information Systems Management*. <https://doi.org/10.1080/10580530.2022.2042628>
- Kaur, J., Arora, V., & Bali, S. (2020). Influence of technological advances and change in marketing strategies using analytics in retail industry. *International Journal of System Assurance Engineering and Management*, 11(5), 953–961. <https://doi.org/10.1007/s13198-020-01023-5>
- Kelliher, F. (2005). Interpretivism and the pursuit of research legitimisation: An integrated approach to single case design. *Electronic Journal of Business Research Methods*, 3(2), 123–132.
- Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, 53(1), 148–152. <https://doi.org/10.1145/1629175.1629210>
- Lis, D., Arbter, M., Spindler, M., & Otto, B. (2022). An Investigation of Antecedents for Data Governance Adoption in the Rail Industry—Findings From a Case Study at Thales. *IEEE Transactions on Engineering Management*, 1–18. <https://doi.org/https://doi.org/10.1109/TEM.2022.3166109>
- Lundberg, M., Engström, S., & Lidelöv, H. (2019). Diffusion of innovation in a contractor company: The impact of the social system structure on the implementation process. *Construction Innovation*, 19(4), 629–652. <https://doi.org/10.1108/CI-08-2018-0061>
- Lundblad, J. P. (2003). A Review and Critique of Rogers ' Diffusion of Innovation Theory as it Applies to Organizations. *Organization Development Journal*, 21(4), 50–64.
- Maheshwari, S., Gautam, P., & Jaggi, C. K. (2021). Role of Big Data Analytics in supply chain management: current trends and future perspectives. *International Journal of Production Research*, 59(6), 1875–1900. <https://doi.org/10.1080/00207543.2020.1793011>
- Mariotto, F. L., Zanni, P. P., & de Moraes, G. H. S. M. (2014). What is the use of a single-case study in management research? *RAE Revista de Administracao de Empresas*, 54(4), 358–369. <https://doi.org/10.1590/S0034-759020140402>



- Martijn, N., Hulstijn, J., de Bruijne, M., & Tan, Y. H. (2015). Determining the effects of data governance on the performance and compliance of enterprises in the logistics and retail sector. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9373, 454–466. [https://doi.org/10.1007/978-3-319-25013-7\\_37](https://doi.org/10.1007/978-3-319-25013-7_37)
- McGregor, S. . L. . (2018). Introduction and Research Questions. In *Understanding and evaluating research : a critical guide* (pp. 228-). SAGE Publications, Inc. <http://dx.doi.org/10.1016/j.cirp.2016.06.001><http://dx.doi.org/10.1016/j.powtec.2016.12.055><https://doi.org/10.1016/j.ijfatigue.2019.02.006><https://doi.org/10.1016/j.matlet.2019.04.024><https://doi.org/10.1016/j.matlet.2019.127252><http://dx.doi.org/10.1016/j.ijfatigue.2019.02.006>
- Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2020). The role of information governance in big data analytics driven innovation. *Information and Management*, 57(7). <https://doi.org/10.1016/j.im.2020.103361>
- Min, S., So, K. K. F., & Jeong, M. (2019). Consumer adoption of the Uber mobile application: Insights from diffusion of innovation theory and technology acceptance model. *Journal of Travel and Tourism Marketing*, 36(7), 770–783. <https://doi.org/10.1080/10548408.2018.1507866>
- Minishi-Majanja, M. K., & Kiplang'at, J. (2005). The diffusion of innovations theory as a theoretical framework in Library and Information Science research. *SA Jnl Libs & Info Sci*, 71(3).
- Mneney, J., & Van Belle, J. P. (2016). Big Data capabilities and readiness of South African retail organisations. *Proceedings of the 2016 6th International Conference - Cloud System and Big Data Engineering, Confluence 2016, January*, 279–286. <https://doi.org/10.1109/CONFLUENCE.2016.7508129>
- Mohajan, H. K. (2018). QUALITATIVE RESEARCH METHODOLOGY IN SOCIAL SCIENCES AND RELATED SUBJECTS. *Journal of Economic Development, Environment and People*, 7(1), 23. <https://doi.org/10.26458/jedep.v7i1.571>
- Momen, S., & Torabi, S. A. (2021). Omni-channel retailing: A data-driven distributionally robust approach for integrated fulfillment services under competition with traditional and online retailers. *Computers and Industrial*

- Engineering*, 157(April). <https://doi.org/10.1016/j.cie.2021.107353>
- NielsenIQ. (2022). *South African retail industry calls for a positive landing*. <https://nielseniq.com/global/en/insights/analysis/2022/south-african-retail-industry-calls-for-a-positive-landing/>
- Otto, B. (2011a). A morphology of the organisation of data governance. *19th European Conference on Information Systems, ECIS 2011*. <http://aisel.aisnet.org/ecis2011/272>
- Otto, B. (2011b). Organizing Data Governance: Findings from the telecommunications industry and consequences for large service providers. *Communications of the Association for Information Systems*, 29(1), 45–66. <https://doi.org/10.17705/1cais.02903>
- Panian, Z. (2010). Some Practical Experiences in Data Governance. *World Academy of Science, Engineering and Technology*, 62.
- Pereira, M. M., & Frazzon, E. M. (2021). A data-driven approach to adaptive synchronization of demand and supply in omni-channel retail supply chains. *International Journal of Information Management*, 57(May 2020). <https://doi.org/10.1016/j.ijinfomgt.2020.102165>
- Petzold, B., Roggendorf, M., Rowshankish, K., & Sporleder, C. (2020). Designing data governance that delivers value. *McKinsey Technology*, 1–8. <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/designing-data-governance-that-delivers-value>
- POPIA. (2022). *Protection of Personal Information Act (POPI Act)*. South African Government. <https://popia.co.za/>
- PWC. (2021). *Retail and Consumer*. <https://www.pwc.co.za/en/industries/retail-and-consumer.html>
- Raguseo, E. (2018). Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. *International Journal of Information Management*, 38(1), 187–195. <https://doi.org/10.1016/j.ijinfomgt.2017.07.008>
- Ratchford, B., Soysal, G., Zentner, A., & Gauri, D. K. (2022). Online and offline retailing: What we know and directions for future research. *Journal of Retailing*, 98(1), 152–177. <https://doi.org/10.1016/J.JRETAI.2022.02.007>
- Rogers, E. M. (2003). *Diffusion of Innovations* (5th ed.). The Free Press.
- Ruffa, C. (2020). Case Study Methods: Case Selection and Case Analysis. *The*

- SAGE Handbook of Research Methods in Political Science and International Relations*, 1133–1147. <https://doi.org/10.4135/9781526486387.N62>
- Santoro, G., Fiano, F., Bertoldi, B., & Ciampi, F. (2019). Big data for business management in the retail industry. *Management Decision*, 57(8), 1980–1992. <https://doi.org/10.1108/MD-07-2018-0829>
- Saunders, M., & Lewis, P. (2018a). Choosing your research design. In *Saunders, Doing research in business and management: An essential guide to planning your project* (2nd ed., pp. 104–136). Pearson Education Limited.
- Saunders, M., & Lewis, P. (2018b). Collecting data. In *Doing Research in Business and Management: An essential guide to planning your project* (2nd ed., pp. 134–178). Pearson Education Limited.
- Schroek, M., Shockley, R., Smart, J., Romero-Morales, D., & Tufano, P. (2012). *IBM Global Business Services Business Analytics and Optimization*. 3–5. [www.sbs.ox.ac.uk](http://www.sbs.ox.ac.uk)
- Seetharaman, A., Niranjana, I., Tandon, V., & Saravanan, A. S. (2016). IMPACT OF BIG DATA ON THE RETAIL INDUSTRY. *Corporate Ownership & Control*, 14(1).
- Shamim, S., Zeng, J., Khan, Z., & Zia, N. U. (2020). Big data analytics capability and decision making performance in emerging market firms: The role of contractual and relational governance mechanisms. *Technological Forecasting and Social Change*, 161. <https://doi.org/10.1016/j.techfore.2020.120315>
- Shamim, S., Zeng, J., Shariq, S. M., & Khan, Z. (2019). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. *Information and Management*, 56(6). <https://doi.org/10.1016/j.im.2018.12.003>
- Shijia, O., & Jia, C. (2020, April 10). New guideline to better allocate production factors. *China Daily*. <https://www.chinadaily.com.cn/a/202004/10/WS5e903fd7a3105d50a3d15620.html>
- Statistics South Africa. (2022). Retail Trade Sales (Preliminary). *Statistical Release P6242, June*, 18. [www.statssa.gov.za](http://www.statssa.gov.za), [info@statssa.gov.za](mailto:info@statssa.gov.za), Tel+27123108911
- Syahadiyanti, L., & Subriadi, A. P. (2018). International Journal of Economics and

- Financial Issues Diffusion of Innovation Theory Utilization Online Financial Transaction: Literature Review. *International Journal of Economics and Financial Issues*, 8(3), 219–226. <http://www.econjournals.com>
- Thomas, G. (2006). The Data Governance Framework. In *The Data Governance Institute*. <https://doi.org/10.1108/978-1-78973-567-320201005>
- Trading Economics. (2022). *South Africa Retail Sales YoY*. Trading Economics. <https://tradingeconomics.com/south-africa/retail-sales-annual>
- Troisi, O., Maione, G., Grimaldi, M., & Loia, F. (2020). Growth hacking: Insights on data-driven decision-making from three firms. *Industrial Marketing Management*, 90, 538–557. <https://doi.org/10.1016/j.indmarman.2019.08.005>
- Varshney, S., & Allen, K. (2020). *Implementing Data Governance*. 1–32.
- Wahyuni, D. (2015). *The Research Design Maze: Understanding Paradigms, Cases, Methods and Metho...*: EBSCOhost. Jamar. <https://web-s-ebSCOhost-com.uplib.idm.oclc.org/ehost/pdfviewer/pdfviewer?vid=1&sid=593cd0a3-6e95-4765-8747-19064e4fa09d%40redis>
- Wang, C. Y., Clavarino, A., & Luetsch, K. (2021). The implementation of a pharmacy residency program – A qualitative study on the diffusion of an innovation. *Exploratory Research in Clinical and Social Pharmacy*, 3, 100048. <https://doi.org/10.1016/j.rcsop.2021.100048>
- Weber, K., Otto, B., & Osterle, H. (2009). One Size Does Not Fit All—A Contingency Approach to Data Governance. *ACM J. Data Inform. Quality*, 4, 27. <https://doi.org/http://doi.acm.org/10.1145/1515693.1515696>
- Wende, K. (2007). *A Model for Data Governance-Organising Accountabilities for Data Quality Management*.
- Ying, S., Sindakis, S., Aggarwal, S., Chen, C., & Su, J. (2021). Managing big data in the retail industry of Singapore: Examining the impact on customer satisfaction and organizational performance. *European Management Journal*, 39(3), 390–400. <https://doi.org/10.1016/j.emj.2020.04.001>
- Yu, W., Yew Wong, C., Chavez, R., & Jacobs, M. A. (2021). *Integrating big data analytics into supply chain finance: The roles of information processing and data-driven culture*. <https://doi.org/10.1016/j.ijpe.2021.108135>
- Zhang, Q., Sun, X., & Zhang, M. (2022). Data Matters: A Strategic Action Framework for Data Governance. *Information & Management*, 59(4),

103642. <https://doi.org/10.1016/j.im.2022.103642>

Zhu, Q., Lyu, Z., Long, Y., & Wachenheim, C. J. (2021). Adoption of mobile banking in rural China: Impact of information dissemination channel. *Socio-Economic Planning Sciences*, December 2020. <https://doi.org/10.1016/j.seps.2021.101011>

## Appendix 1: Informed Consent Letter

Dear Sir\Madam

I am currently a student at the University of Pretoria's Gordon Institute of Business Science and completing my research in partial fulfilment of an MBA. I am researching the impact of data governance on data-driven decision-making in the Retail sector in South Africa. Our interview is expected to last for about an hour, and this will help us explore the impact the adoption of data governance has on data-driven decision-making in retail organisations.

Your participation is voluntary, and you can withdraw at any time without penalty. All data will be reported without identifiers and all personal data will be kept confidential. If you have any concerns, do not hesitate to get in touch with my supervisor or me. Our details are provided below.

Researcher: Busisiwe Tshabalala

Research Supervisor: Suzanne Myburgh

Email: [19386754@mygibs.co.za](mailto:19386754@mygibs.co.za)

Email: [suzanne.myburgh@hotmail.com](mailto:suzanne.myburgh@hotmail.com)

Mobile: 0765537799

Signature of participant: \_\_\_\_\_

Date: \_\_\_\_\_

Signature of researcher: \_\_\_\_\_

Date: \_\_\_\_\_

## Appendix 2: Interview Guide

### Interview Guide

#### 1. General Questions

- 1.1. Please provide a brief overview of your current role in the company
- 1.2. Are you involved in the process of implementing data governance?

1.3. What is your role in the implementation of data governance in the organisation?

## **2. Data Governance as the Innovation**

2.1. What do you think is the role of data in business operations?

2.2. What do you understand about data governance? What role does it play in the company?

2.3. What is the relative advantage of implementing data governance? (The level to which data governance will yield greater benefit than what is currently in place?)

2.4. Should data governance be consistent with the requirements of the business (business strategy) and align with the current ways of working?

2.5. How complex is it to implement data governance? What are some of the complexities?

2.5.1. How complex is it for the broader organization to understand?

2.6. Can data governance be trailed and tested on a restricted basis? What would this look like?

2.7. What are some of the measures that can be used to measure the effectiveness of data governance?

## **3. Communication Channel**

3.1. What communication channels can be used to cascade data governance to the organisation?

3.2. How effective have some of these communications been in communicating change in the organisation?

3.3. When and how often should data governance be communicated?

## **4. Innovation process**

4.1. What were some of the factors that influenced the need to implement data governance?

4.2. How is the larger organisation involved in the implementation of data governance?

4.2.1. What is the role of people in the implementation of data governance?


4.3. What do you think are some of the factors that will influence the adoption of data governance by the organisation?

## 5. Social system

5.1. What impact does the social system have on the implementation of data governance?

## 6. What influence does data governance have on data-driven decision-making?

### Appendix 3: Ethical Clearance Letter



**Gordon Institute  
of Business Science**  
University of Pretoria

**Ethical Clearance  
Approved**

Dear Busisiwe Tshabalala,

Please be advised that your application for **Ethical Clearance** has been approved.  
You are therefore allowed to continue collecting your data.  
We wish you everything of the best for the rest of the project.

[Ethical Clearance Form](#)

Kind Regards

This email has been sent from an unmonitored email account. If you have any comments or concerns, please contact the GIBS Research Admin team.

### Appendix 4: Code book developed from literature

Research Question	Innovation Process Stage	Main Codes	Sub Codes
Research Question 1: Perceived needs from data governance	Agenda-Setting	Perceived Needs	
Research Question 2: Perceived innovation characteristics of innovation	Matching	Relative Advantage	
		Compatibility	
		Complexity	
		Trialability	
		Observability	



Research Question 3: Communication channels used to disseminate data governance	All stages	Mass communication channels	
		Interpersonal communication channels	
Research Question 4: Impact of social system	Redefining\Restructuring	Social Structure	
		System norms	
		Opinion leaders and change agents	
		Innovation decisions	Optional
Collective			
Authority			
Research Question 5: Impact of data governance on data-driven decision making	Clarifying and Routinizing	Data-driven decision making	