NOVA IMS Information Management School

MGI

Master Degree Program in
Information Management

DATA GOVERNANCE VALUATION

A model for assessing the impact on Organisations' business

João Silva Pestana

Master Thesis

presented as partial requirement for obtaining the Master Degree in Information Management

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

DATA GOVERNANCE VALUATION

A model for assessing the impact on Organisations' business

by

João Silva Pestana

Master Thesis presented as partial requirement for obtaining the Master's degree in Information Management, with a specialization in Knowledge Management and Business Intelligence

Supervised by

Professor Gonçalo Baptista, PhD, NOVA IMS (Universidade Nova de Lisboa)

NOVEMBER, 2023

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledged the Rules of Conduct and Code of Honour from the NOVA Information Management School.

João Pestana

[Lisbon, 22nd November 2023]

ABSTRACT

In an era of exponential data growth driven by technology, organisations struggle to harness data for competitive advantage, whilst ensuring compliance, and maintaining quality. Recognising data's critical role in organisational success, many organisations invest in Data Governance programs to secure the confidentiality, integrity, quality, and availability of their data. Despite these efforts, establishing a formal Data Governance program remains a challenging task, marked by a lack of understanding regarding tangible benefits and a limited practical guidance in existing literature, hindering organisations from fully unlocking the value of their Data Governance initiatives. This study explores Data Governance's organisational impacts on Business Value, proposing a conceptual model supported in IS business value theories. To empirically assess this model, the study performs a multi-organisational questionnaire across industry sectors, targeting IT or data-related roles. Employing the PLS-SEM technique, the results demonstrate that, when integrated into a Data Governance program, Data-Driven Culture, Data Competencies, Data Processes and Data Stewardship generate value across various stages of the value chain, especially in terms of Data Quality, Business Value and Regulatory Compliance. However, Tools and Technologies show no direct impact in the value-generation process, challenging prior research. Additionally, the research recognises the influence of contextual factors in shaping the organisational outcomes of a Data Governance Program. Ultimately, this research contributes to the evolving field of Data Governance by providing organisations tools to evaluate the success of their governance practices, enabling informed decision-making and fostering a proactive approach to Data Governance.

KEYWORDS

Data Governance; Data Governance Program; Data Quality; Business Value; Regulatory Compliance

Sustainable Development Goals (SDG):

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AVE	Average Variance Extracted
BCBS	Basel Committee on Banking Supervision
CB-SEM	Covariance-Based Structural Equation Modelling
CCPA	California Consumer Privacy Act
CMB	Common Method Bias
CSF	Critical Success Factor
DAMA	Data Management Association
DDC	Data-Driven Culture
DG	Data Governance
DGI	Data Governance Institute
DGO	Data Governance Office
DQM	Data Quality Management
GDPR	General Data Protection Regulation
HTMT	Heterotrait-monotrait ratio
IS	Information Systems
ISO	International Organisation for Standardization
IT	Information Technology
KBV	Knowledge Based View
KPI	Key Performance Indicator
MDM	Master Data Management
PLS-SEM	Partial Least Squares Structural Equation Modelling
ROI	Return of Investment
SD	Standard Deviation
SEM	Structural Equation Modelling
VIF	Variance Inflation Factor

1. INTRODUCTION

1.1. BACKGROUND

With the evolution of technology, our world has become more digital. The Internet and mobile devices have revolutionised consumer needs, fuelling a demand for fast, practical, and easily accessible products/services. To comply with these emerging demands, organisations turned to digital channels, creating constant streams of data, which are transforming business processes.

Organisations are now recognising the abundance of data sources in commercial, production, value-chain, and marketing processes. Fuelled by internal and external data, these sources have grown exponentially and will continue to do so for the foreseeable future (Vasconcelos & Barão, 2017). Thus, to derive value and competitive advantages, organisations must systematically analyse this expanding volume of data essential to their operations.

In this context, data has transcended into an ownable, controllable asset capable of creating and holding value, standing alongside tangible resources like financial and human capital (Al-Ruithe et al., 2019). Nowadays, organisations globally recognise that to stay competitive, their business decisions must no longer be based on instincts but instead on the actionable insights extracted from data analysis (DAMA International, 2017). This shift in decision-making gave businesses a competitive edge, enabling a deeper understanding of customers, fostering innovation in products and services, and enhancing operational efficiency for improved financial performance and risk control (Attard & Brennan, 2018).

Furthermore, as today's fast-paced global market makes digital disruption the norm and more organisations identify themselves as "data-driven", data is now a form of Organic Growth (Ladley, 2020) and a pillar in supporting organisational businesses (Bhansali, 2014). However, with this reality came several challenges with high-risks, high-impact and high-costs for organisations:

- **Data Silos -** Today's organisations use data that they collect internally, as well as data that they acquire from external sources, which requires managing ever-increasing volumes of data (Cai & Zhu, 2015). In addition, data is rarely static. Throughout its lifecycle, data moves from department to department due to the interconnection of operations within an organisation, being cleansed, transformed, merged, enhanced, or aggregated according to the necessities of each area (DAMA International, 2017). All these factors create deep data inconsistencies, silos and even redundancies across an organisation (Abraham et al., 2019), making the process of retrieving the required data for business purposes or regulatory needs complex and error-prone.

- **Data Quality -** Data is a key asset in any organisation. However, with such tremendous data volumes, abundant data types and complex data structures that come from a diversity of data sources, allied with a rapid transformation and passage of data through the organisation, it becomes challenging to ensure its accuracy, completeness and timeliness (Cai & Zhu, 2015). If organisations cannot rely on data to meet business needs, the effort to collect, store, secure, and enable access to it is wasted. In other words, poor quality data is costly to any organisation, whether directly in the form of customer dissatisfaction, organisational inefficiencies and low productivity or indirectly as workarounds, rework and reputational costs (DAMA International, 2017).

- **Dependence on Skills and Technology -** While working with data is crucial to sustain a market position, it is also a complex process in terms of technology and collaboration. Beginning with the design and administration of systems to support the systematically collection of data, as well as its storage and integration, followed by the analysis in order to understand possible issues and problems, and finishing with the analytics to interpret data and bring strategic thinking, all of these phases have a high dependency of different skills and technologies. The challenge here is to get the people in an organisation with these skills and perspectives to build a common understanding of data so that they collaborate towards the same goals (DAMA International, 2017). Additionally, technology has a big impact on data, making decisions about it without proper organisation, preparation and alignment with data requirements a risk for organisations (Kim & Cho, 2018).

- **Data Privacy and Protection -** As consumers become more aware of the amount of personal data collected by organisations and how it is being used, they expect a better understanding of their pain points and unmet needs (Anant et al., 2020), but also protection of their information and respect for their privacy (Dommeyer & Gross, 2003). These increasing concerns, allied with multiple consumer-data breaches, have led governments to adopt new regulations, such as the General Data Protection Regulation (GDPR) in Europe and the California Consumer Privacy Act (CCPA) (Anant et al., 2020). For organisations, these regulatory requirements completely transformed their approach to data, forcing them to have strict controls over the data they have, where it came from and how it is stored, what it represents, its usages and processing, who has access to it and, if necessary, its disposal (Abraham et al., 2019). Furthermore, implementing and understanding these regulations made organisations re-plan their data strategy, including assigning new roles and responsibilities and acquiring new expertise (Shao & Oinas-Kukkonen, 2019).

- **Data Valuation -** In any organisation, the value of its physical and financial assets is accounted in the balanced sheet. However, data is different from other assets since it does not obey the same economic laws and it has some unique properties, which make it challenging to measure its value (Moody & Walsh, 1999). Unlike other assets, data can be infinitely shared and replicated. Its value does not strictly depend on volume but instead increases with quality and use. However, due to its temporal trait, data tends to depreciate over time, as older data points may no longer be relevant today (Moody & Walsh, 1999). As such, it can be challenging for organisations to calculate the monetary value of data, as its costs and benefits do not conform to a standard, thus making it difficult to measure how data contributes to organisational success (DAMA International, 2017).

With so many challenges when leveraging data assets, it becomes crucial for organisations to manage their data lifecycle needs, by developing and implementing architectures, policies, and procedures that control, protect, deliver, and enhance the value of their data assets (DAMA International, 2017). However, for data management to be successful, it cannot be restricted and unique to each department. Instead, it requires an organisation-wide approach (Weber et al., 2009). As such, Data

Governance (DG) importance emerges alongside data management, providing a framework and guidelines to ensure that all data management initiatives are consistent with the organisation's mission, strategy, values, norms, and culture (Weber et al., 2009), while also considering the impact of people and processes on data (DAMA International, 2017).

In recent years, data governance has rapidly gained popularity in both literature and practical applications, being considered an emerging field in Information Systems (IS) (Alhassan et al., 2018). Recognised by professionals and academics, data governance programs offer a promising solution for organisational data challenges in data-driven businesses (Alhassan et al., 2018). A good program helps organisations create a clear mission, increase confidence in using the organisational data, establish accountabilities, maintain scope, build standards, and define measurable successes (Al-Ruithe et al., 2019).

Moreover, data governance assumes an even more prominent role in financial and insurance services. Although every organisation is subject to governmental and industry regulations, banks and insurance are heavily regulated industries. Some regulations, such as the BCBS 239 (Basel Committee on Banking Supervision) and Basel II, for banks, and the Solvency II, for insurance, establish core principles for data management (DAMA International, 2017). Therefore, organisations need a data governance program to guide the implementation of adequate controls to monitor and document compliance with these regulations (DAMA International, 2017). In this current context, data governance has become both a necessity and an imperative.

1.2. PROBLEM DEFINITION

Many organisations are becoming aware of the increasing importance of governing their data to ensure its confidentiality, integrity, quality, and availability (Al-Ruithe et al., 2019), prompting investments in data governance programs. Solutions implemented for this purpose range from formal structures connecting business, IT, and data management functions, to processes and procedures for decision-making and monitoring, to practices that support stakeholder collaboration (Abraham et al., 2019). However, in practice, establishing a formal data governance program has proven to be a challenge with no "one-size-fits-all" solution (Weber et al., 2009). Since every organisation has its own business strategy while also needing to comply with strict regulatory requirements involving data and data practices, professionals can struggle to understand how to leverage data governance to respond to the complex reality of their organisations. This factor, allied with few practical implementations addressed in the existing data governance literature (Benfeldt et al., 2020), impacts the value-creating potential of governance processes.

Furthermore, some studies (Begg & Caira, 2012) revealed that practitioners have difficulties comprehending the value of a data governance program, not being able to grasp its impact on the organisation, perceiving the costs to be greater than the benefits. As a result, practitioners are left with a poor understanding of data governance, which remains problematic since awareness is a primary success factor when developing and implementing a data governance program (Benfeldt et al., 2020). Without awareness of how data governance contributes to organisational goals, mobilising employees to implement and adopt a data governance program becomes an impossible challenge (Benfeldt et al., 2020). In the absence of this investment, its significance becomes apparent only when an organisation experiences significant regulatory pressure, a data breach, or other severe data-related problems (DalleMule & Davenport, 2017). By then, it is too late to prevent reputational and financial damages.

Aiming to provide the means for organisations to evaluate the success of their governance models and data governance practices, this work will study the problem of assessing the organisational impact of a data governance program on business value.

Problem Definition:

Assess the impact of data governance on business value

On this basis, the following research questions have been investigated:

(1) What are the dimensions of a Data Governance Program?

(2) What is the relationship between data governance practices and organisational outcomes?

(3) How can organisations assess the impact of their Data Governance Programs?

(4) Can a model be built to measure the impact of implemented governance models?

1.3. CASE STUDY

1.3.1. Description

The development of this study was done under the scope of an internship program in a service provider firm as part of a set of initiatives launched by the Technology Consulting Department to expand the value of their Data Governance services. This project was born out of the clients' need to quantify the impacts of data governance practices, demonstrating the value of their data management efforts to stakeholders in order to justify investments in data governance and secure support for ongoing initiatives. Additionally, as the data governance field remains highly theoretical and conceptual, with few studies focusing on practical aspects (Nielsen, 2017), this study aims to mitigate these literature gaps by providing actionable insights about the potential benefits to be derived from implementing data governance in a real-world context.

As such, the following objectives were undertaken as the research subject for this study:

- Review the main concepts associated with data governance and the practices implemented within the scope of an organisational data governance program
- Develop a model that enables corporate clients to identify and measure the impact of implemented governance programs and the adoption of data governance practices
- Build an instrument to identify the current state of data governance practices in organisations and evaluate their impact on various organisational aspects
- Explore the implications of any differences between the data governance impacts identified through the model and the recommendations from existing literature

This work has both theoretical and practical implications. From a theoretical viewpoint, this study extends data governance research by empirically testing a new theoretical framework that assesses the antecedents and impacts of data governance on organisations. In doing so, this study reviews the entire business value-chain of data governance in a multi-industry context, linking DG capabilities to outcomes at the process and business level, as well as clarifying the role of contextual factors in shaping these relationships. On a practical level, the impact of this work is two-fold. First, it allows organisations to better understand the impacts of data governance on their business operations, stakeholders, and regulatory compliance. Secondly, the proposed theoretical model provides added value to data governance practices by enabling organisations to measure and monitor the impacts of their data governance initiatives.

Ultimately, this model will be passed to the department managers to be incorporated as a last step in the Data Governance projects to support the teams in the client follow-up phase. Therefore, the project goal is centred around equipping corporate clients with a methodology to measure the impacts of the implemented practices, enabling them to comprehend the value associated with their data governance program.

1.3.2. Structure

The study is structured into seven chapters. Chapter 1 introduces the research topic, providing background information, the research problem, objectives, and the study's significance. Chapter 2 reviews existing data governance literature, exploring theoretical foundations, frameworks, relevance in organisational contexts, while identifying gaps in current research. Chapter 3 outlines the theoretical model, hypotheses, and research methodology, guiding the empirical analysis.

Chapter 4 details the data collection process, covering the research instrument, procedures, dataset characteristics, and addressing potential biases. In Chapter 5, the theoretical model is applied to empirical data, assessing constructs, indicators, and relationships, and analysing the structural model to test hypotheses. Chapter 6 offers an in-depth examination of findings, interpreting results within the context of organisational dynamics, aligning with established theories, and discussing practical implications.

The final chapter synthesizes main findings, provides study limitation (e.g., data constraints, methodological issues), and outlines future research areas in the field of data governance.

2. LITERATURE REVIEW

2.1. DATA GOVERNANCE CONCEPTS

Data Governance is an ongoing topic in both scientific and practical literature (Nielsen, 2017). However, its presence in the literature is not unified (Abraham et al., 2019), as the lack of an official definition continues to permeate the field (Benfeldt et al., 2020). Therefore, before conducting the work for this study, it is essential to clarify the concept of Data Governance.

Whether in scholarly literature or practitioner publications, the lack of a standard has led researchers to formulate different definitions of Data Governance, adapting existing ones to their context and reality (Abraham et al., 2019; Al-Ruithe et al., 2019). As a result, there can be an incorrect understanding of the concept of Data Governance and how it differs from others in the field.

One common mistake among researchers and practitioners is the interchangeable usage between Data Governance and Data Management (Al-Ruithe et al., 2019). According to the Data Management Association (DAMA), Data Management is "the development, execution, and supervision of plans, policies, programs, and practices that control, protect, deliver, and enhance the value of data and information assets". Whereas, Data Governance involves "the exercise of authority and control (planning, monitoring, and enforcement) over the management of data assets". Abraham et al. (2019), by reviewing 145 publications, further differentiates Data Governance as a cross-functional framework for managing data as a strategic enterprise asset, specifying the decision rights and accountabilities for decision-making, and formalising data policies, standards and procedures, as well as monitoring compliance. Weber et. Al (2009) also introduces the notion of "enterprise-wide", mentioning that Data Governance specifies a structural corporate-wide framework for decision-making rights and responsibilities regarding the use of data aligned with the organisation's mission, strategy, values, norms, and culture.

Overall, there is a consensus that Data Governance encompasses the entirety of decision rights and responsibilities concerning data assets management in organisations, guiding all data management functions (Al-Ruithe et al., 2019). Therefore, Data Governance represents a high-level plan, specifying which data management decisions need to be made and who makes these decisions (Otto, 2011a) to ensure that data is appropriately managed according to policies, best practices and corporate objectives (Ladley, 2020). In contrast, Data Management's core objective is to ensure that data produces value for the organisation, focusing on the practical implementation of the decisions and appropriate actions (Otto, 2011a). Given its completeness and theoretical support, this study adopts Abraham et al.'s (2019) definition of Data Governance.

Data Quality Management (DQM) is another concept often discussed in the context of Data Governance (Nielsen, 2017). According to Wende (2007), DQM refers to quality-oriented data management, encompassing all activities related to improving data quality. Going beyond mere reactive action (e.g. identification and correction of data issues), DQM works as a proactive and preventive concept, involving a continuous cycle of activities to define, measure, analyse, and improve data quality (Otto, 2011b). The close connection between these two concepts is a result of the need to maximise the value of data as an enterprise asset (Otto, 2011b), ensuring the management of its quality and "fitness for use" (Wang, 1998), recognised as a primary business driver for adopting a formal data governance program (DAMA International, 2017; Brous et al., 2020).

Figure 1 summarises the aforementioned fundamental concepts related to Data Governance, as well as their respective relationships with each other.

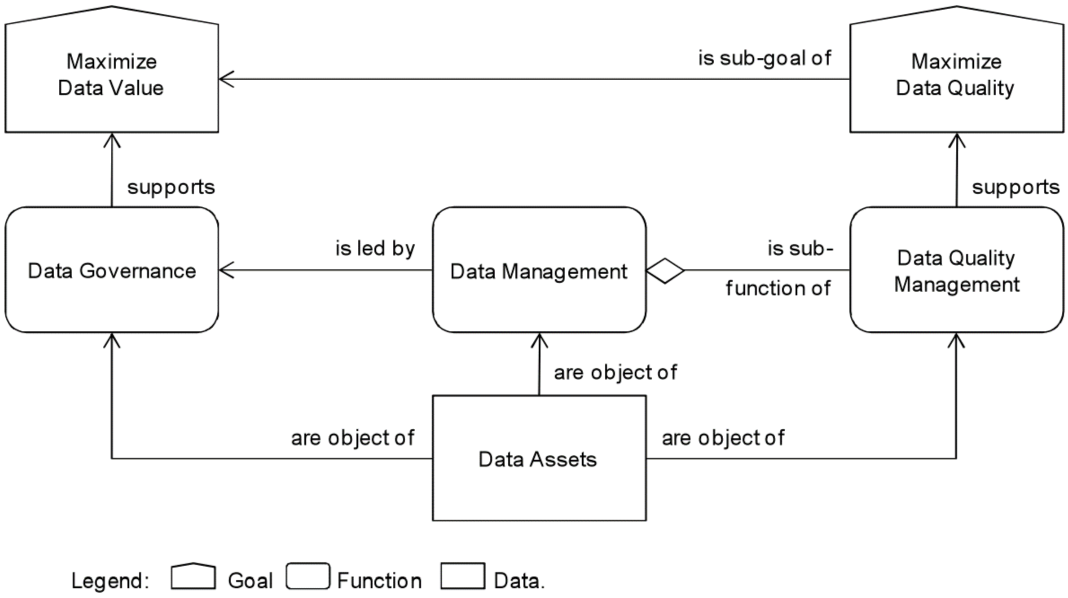


Figure 1 – Fundamental Concepts of Data Governance. (Otto, 2011a)

Nielsen's (2017) study also identifies that the link between Information Technology (IT) and Data Governance is still present in today's practices. For this reason, it is essential to emphasise that neither of these concepts should be subordinated to the other, since their alignment is critical for the successful management of both data and IT assets (Begg and Caira 2012). These two practice areas fall under the broader umbrella of Corporate Governance, encompassing all the processes, customs, policies and laws that affect organisational direction, administration, and control (Monks & Minow, 2011). Additionally, while Data Governance remains under-researched, IT Governance is a more mature area, with publications dating back four decades (Al-Ruithe et al., 2019). Industry bodies, such as the IT Governance Institute and the International Organisation for Standardization (ISO), define IT governance as the "leadership and organisational structures and processes that ensure that the organisation's IT sustains and extends the organisation's strategies and objectives", establishing detailed IT governance standards and processes, widely implemented by organisations (Panian, 2010).

Accordingly, the main difference between these two concepts lies in their practical aspects. Data Governance is designed to govern data assets in technology-agnostic environments, whilst IT Governance focus on decisions about IT investments, applications, and project portfolios (Al-Ruithe et al., 2019). Furthermore, IT Governance practices are designed primarily around an organisation's applications—not its data—clearly defining owners, processes, and policies to manage enterprise software applications (Panian, 2010).

2.2. DATA GOVERNANCE FRAMEWORKS

Although the roots of data governance research traces back to the early 1980s, the first efforts to create a framework for data governance were published in 2007 (Niemi, 2013). These frameworks consisted of Wende's scientific paper at the ACIS conference (Wende, 2007), IBM's data governance maturity model (IBM Institute, 2007), and later Radcliffe's "Gartner's Seven Building Blocks" (Radcliffe, 2009), which for many years was the primary authority in the area of Master Data Management (MDM) (Niemi, 2013). However, as the challenges faced by data governance professionals evolved and new business and technical requirements appeared, some industry associations started to develop their own data governance frameworks (Niemi, 2013).

In the literature, the presence of data governance frameworks is self-evident, with its benefits for organisations being well studied and identified at a macro level. These frameworks aim to help organisations develop and implement their own data governance programs, providing a range of perspectives on how to approach data management in the different decision domains (Khatri & Brown, 2010). Industry associations such as DAMA, DGI, and SAS presented some well-known examples, which Al-Ruithe et al. (2019) identified as the industry standards. Therefore, in this section, some theoretical notions behind these frameworks will be overviewed to understand the structure on which most data governance programs are based.

2.2.1. DAMA-DMBOK Functional Framework

The DAMA-DMBOK Framework, developed by DAMA International, is an industry standard that establishes a functional structure for standardising data governance practices and creating a shared vocabulary. The framework comprises eleven domains, with data governance at the core to ensure consistency and balance across functions (Figure 2).

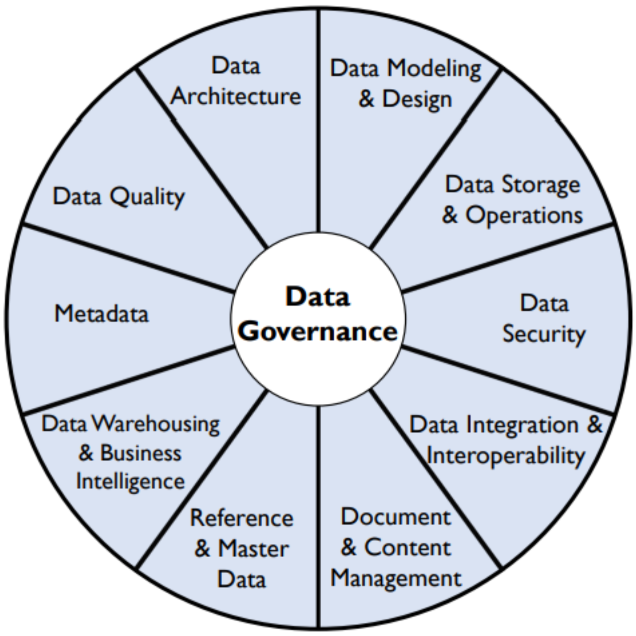


Figure 2 – DAMA-DMBOK Framework. (DAMA International, 2017)

Accordingly, the 11 domains described in this framework include (DAMA International, 2017):

Data architecture is responsible for designing and maintaining the organisation's data architecture, by defining the data models, structures, and flows that support the organisation's business requirements.

Data modelling and design is the process of creating logical and physical data models that meet the organisation's data requirements, by identifying the data entities, relationships, and attributes needed to support business processes.

Data Storage and Operations refers to the storage, retrieval, and maintenance of data to maximise its value, encompassing the management of data lifecycle operations, from planning to data processing.

Data security is responsible for protecting data from unauthorised access, use, disclosure, or destruction, implementing security controls, monitoring data usage, and ensuring compliance with data privacy regulations.

Data integration and interoperability ensures that data can be shared and integrated across different systems and applications. It involves defining data standards, developing data integration processes, and managing data exchanges.

Document and content management includes the planning, implementation and control activities to manage the lifecycle of data residing in unstructured media, especially documents required to support legal and regulatory compliance.

Reference and master data includes the ongoing reconciliation and maintenance of core-data to enable consistent cross-system use of the most accurate, up-to-date, relevant version about crucial business entities.

Data warehousing and business intelligence includes collecting, storing, and analysing data to support business decision-making. It involves designing data warehouses, developing business intelligence solutions, and providing analytical insights to business users.

Metadata includes planning, implementing, and controlling activities to provide access to high-quality metadata, including definitions, models, data flows, and other information needed to understand data and the systems that create, maintain, and retrieve them.

Data quality involves the planning and implementing quality management techniques to measure, evaluate and improve the suitability of data for use within an organisation.

2.2.2. DGI Data Governance Framework

In Thomas (2006) article, the Data Governance Institute outlines a comprehensive framework for implementing a data governance program. This framework, schematised in Figure 3, comprises 10 universal components of a data governance program, distributed among three levels: People & Organisational Bodies, Rules & Rules of Engagement, and Processes.

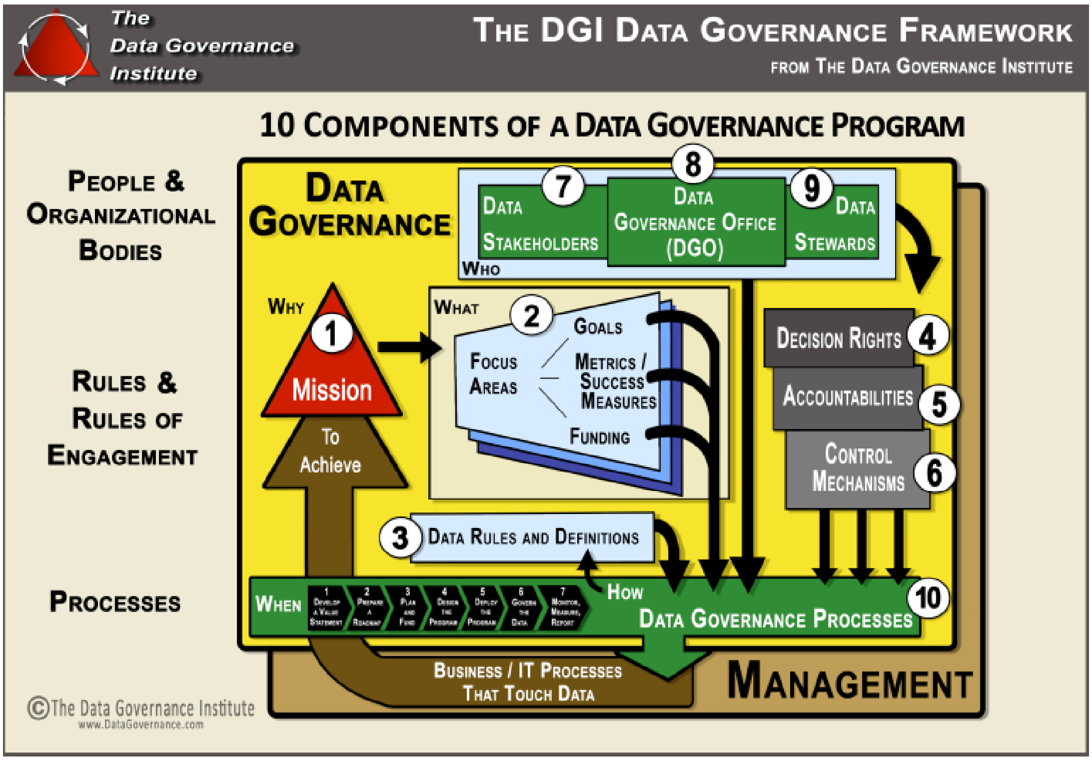


Figure 3 – DGI Data Governance Framework. (Thomas, 2006)

The Rules and Rules of Engagement form the foundational level of the framework, encompassing structural elements that define key aspects of the data governance program. This includes its mission and vision ("why"), focus areas ("what"), goals, success metrics, funding plan, data rules, decision rights, accountabilities, and control mechanisms. Accordingly, the first step is to define a clear, compelling vision of what the organisation could look like with a mature data governance program, establishing objectives, success measures, and funding strategies for each focus area. Subsequently, the framework delves into the development of data-related policies, standards, compliance requirements, business rules, and definitions. Decision rights are crucial, requiring clear documentation of responsible parties, decision-making processes, and timelines before implementing any rule or data-related decision. Following the establishment of decision rights, accountability must be defined and integrated into daily processes, mapping departmental/stakeholder responsibilities with their roles in data governance processes. To mitigate data-related risks, the program should incorporate risk management strategies through multi-level controls to uphold organisational governance goals.

At the People and Organisational Bodies level ("who"), the framework outlines three primary roles: Data Stakeholders, the Data Governance Office (DGO), and Data Stewards. Data Stakeholders contribute to data rule creation, usage, and definition, having expectations that must be addressed by the Data Governance program. This involves clarifying decision-making participants, consultative roles, and post-implementation communication. Establishing a Data Governance Board, consisting of executive stakeholders, is a recommended best practice. The Data Governance Office (DGO) is pivotal in facilitating governance activities such as data-related decision-making, defining data, ensuring compliance, and resolving issues. The DGO also oversees metric collection and reporting to stakeholders. Data Stewards form a Data Stewardship Council, which is responsible for data-related decisions, policy setting, standard specification, and making recommendations to the Data Governance

Board. This multi-tiered approach ensures comprehensive coverage and collaboration among stakeholders for effective data governance.

Lastly, the Processual level ("when" and "how") encompasses the standardisation, documentation, and repetition of data governance processes to meet regulatory and compliance requirements for Data Management, Privacy, Security, and Access Management. The DGI recommends formalising documented and repeatable processes for aligning policies, requirements, controls, decision rights, accountability, stewardship, change management, data definitions, issue resolution, data quality requirements, building governance into technology, stakeholder care, communication, and measuring and reporting value.

2.2.3. SAS Data Governance Framework

The SAS Data Governance Framework, illustrated in Figure 4, is an example of a contribution by software vendors to address data governance challenges, combating issues like a limited understanding of its value. It counters the perception of data governance as a finite project or academic exercise. This framework combines holistic and pragmatic approaches, considering both data components and organisational groups, with a phased implementation catering to short-term needs and long-term governance capability (SAS Institute Inc, 2018).

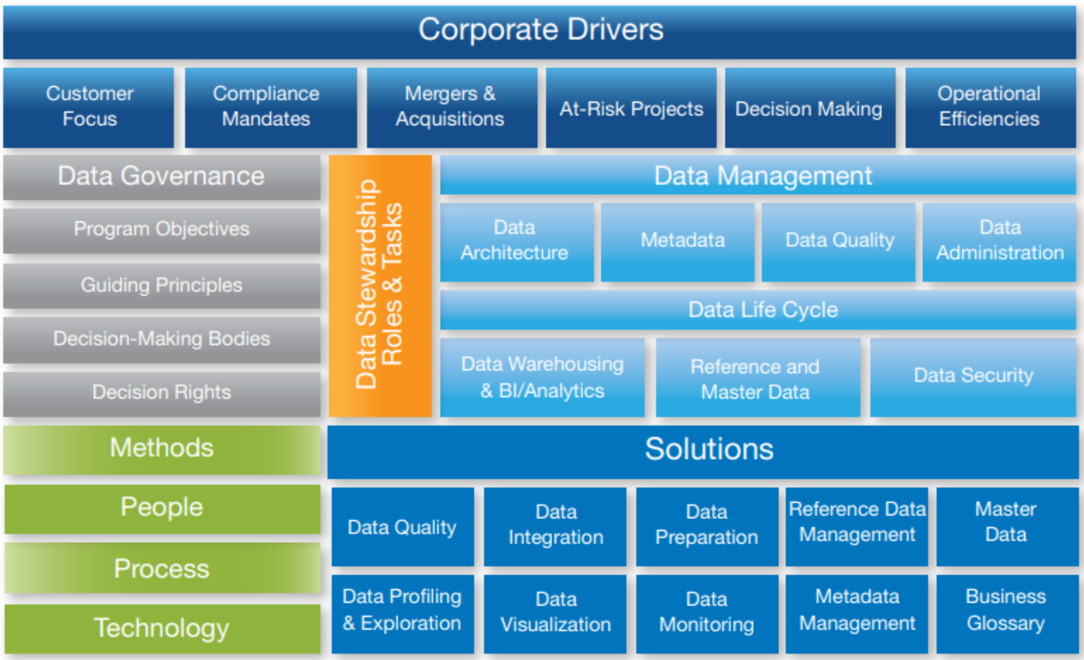


Figure 4 – SAS Data Governance Framework. (SAS Institute Inc, 2018)

The SAS framework encompasses three key facets of Data Governance: Corporate Drivers; Data Governance and Methods; as well as Data Management, Solutions and Data Stewardship. At the strategic level, Corporate Drivers address business drivers and strategies emphasising the necessity of data governance. The organising level includes the Data Governance and Methods sections, focusing on developing and monitoring policies that drive data management events (e.g. data quality, architecture and security). Lastly, Data Management, Solutions, and Data Stewardship constitute the tactical execution level, encompassing daily processes for managing data and technology.

The Corporate drivers section presents some key factors, activities, and resources impacting organisational performance. In this framework, understanding the business value and demonstrating Return of Investment (ROI) for data governance is stated as one of the main challenges. In this sense, aligning data governance activities and investments to corporate drivers will promptly associate data governance achievements with key business goals. With this strategy, data governance activities are tied to the organisation's objectives, making each activity's ROI clear to all stakeholders.

The Data Governance section outlines the planning tools that enable data governance to be implemented in a way that fits the culture and staffing – Program Objectives, Guiding Principles, Decision Bodies and Rights. All of these components comprise the initial data governance plan that needs approval by the senior leadership.

In the Data Stewardship section, this framework foresees the creation of data stewards across all data management domains to serve as contact points between the domains' governance activities and other stakeholders. Their responsibilities include defining data, ensuring usability, addressing queries, and enforcing policies, all while monitoring quality and reporting metrics. To do so, the framework specifies that granting the authority for stewards to oversee data within their domain is paramount in any organisation.

The Data Management section encompasses all domains subjected to the policies created by a data governance program. These domains have both business and IT components and consist of data architecture, metadata, quality, administration, data life cycle, data warehousing and analytics, reference data, master data management and data security. This framework recommends the incremental inclusion of these domains in the data governance program, adapted to the organisation's strategy and maturity level.

The last two sections, Methods and Solutions, are intrinsically connected as each solution used to implement and automate the various activities of a data governance program needs to consider the three main areas – People, Process and Technology – described in the Methods section.

2.3. DATA GOVERNANCE CRITICAL SUCCESS FACTORS

The Critical Success Factors (CSF) approach, initially introduced by Rockart (1979), identifies the few key areas where "things must go right" to achieve the overall goals, requiring the continual attention from management. This approach has been widely investigated and applied, especially in the IS field, remaining a valid method for identifying influential factors in communities of practice (Alhassan et al., 2019). In data governance literature, the CSF approach identifies key activity areas essential for ensuring a data governance program's success across all domains.

In Alhassan et al. (2019), the authors derive seven CSFs for data governance from interviews with over 15 industry experts in the banking sector. The first CSF, "Employee data competencies", states that the capabilities of all employees (from senior executives to entry-level) play a crucial role in a data governance program, emphasising the importance for skills in defining, implementing, and monitoring data processes and procedures, as well as awareness of data requirements and policies. In this regard, specific actions are suggested to ensure employees' data competencies, such as continuous training in data governance activities and promotion of increased awareness of critical and sensitive data.

The second CSF, "Clear data processes and procedures", comprises all activities related to data flow, integration, authorisation processes, validation, and testing. The absence of this factor leads to a lack of trust in organisational data, as there is no clear understanding of the processes that ensure the readiness of data for usage. As such, it is recommended that the processes be available in the system and the existing ones be revaluated and updated. The third CSF, "Standardised easy-to-follow data policies", stresses the need for well-known and clear policies across all domains, supporting employees in data decisions. Beyond defining policies, organisations should establish implementation methods and regular monitoring and maintenance of these policies.

The fourth CSF, "Flexible data tools and technologies", encompasses activities related to both software and hardware for data management. Seen as an enabler, this factor supports other CSFs since the investment in an appropriate IT infrastructure enables the implementation and automation of processes while also addressing data privacy, availability and future changes. The fifth CSF, "Established data roles and responsibilities", often denoted as data stewardship, has been extensively studied in the literature (Cheong & Chang, 2007; Weber et al., 2009). Most authors agree that a successful data governance program must incorporate a formal identification of roles and their respective responsibilities across all domains.

The sixth CSF, "Clear inclusive data requirements", covers all business and IT data needs. While data implementation is important, the main component of data requirements lies in ensuring compliance with data regulations, achieved through a continuous process of acquisition and maintenance. This connects with the first CSF, as understanding the proper requirements is a fundamental competency. The seventh CSF, "Focused and tangible data strategies", asserts the need to identify both short- and long-term objectives for measuring the success of a data governance program.

Another perspective on data governance CSFs is provided by Cave (2017). In this study, the author collects data from interviews in the health sector, leading to the formulation of five strategies for implementing data governance practices. However, there are some parallels between the findings of Alhassan et al. (2019) and Cave (2017), as two strategies—"Effective and strategic communications" and "Compliance with regulations"—can be directly matched to the third and sixth CSFs.

Nevertheless, Cave (2017) contributed with new CSFs, including "Benchmarking and standardisation", which advocates for program effectiveness by assessing the current state of data governance through industry benchmarking and standardising practices. Another addition consists of "Obtaining stakeholder buy-in". This entails engaging stakeholders at all levels in discussions to establish a common consensus and gather feedback on new practices, facilitating the assessment of initial outcomes. The "Structured oversight with committees and boards" is another contribution and highlights the importance of incorporating a hierarchy of committees and boards as part of the organisational structure to oversee data governance practices and policies at different levels within an organisation—local, divisional and organisation-wide. To do so, these committees must work in a coordinated way to ensure the flow of information throughout the organisation.

Al-Ruithe et al. (2019) conducted a systematic literature review on data governance, identifying 20 CSFs that present a broader perspective, with some intersections the two prior studies. As such, the following CSFs can be added to the ones specified above: "Organisational culture change" and "Develop a business case for data governance". The first one resolves around cultural change, which does not happen simply because a new policy is announced or a new system is implemented. Instead, it occurs when people behave differently because they recognise the value of doing so. This can be achieved through a careful and structured approach to managing culture change in line with the organisation's identity. The last one aims to align data governance with business needs, ensuring its continuity and constant improvement through the definition of critical terms, such as the data governance vision and mission, costs, sustainability requirements, benefits, and risks.

In summary, by reviewing three studies focused on different research areas, a set of 12 CSFs for achieving an effective data governance program were identified. This research diversity illustrates the importance of data governance across all industries and in different environments. Furthermore, the intersections between the findings of the three studies support the transversality of data governance CSFs across industries and sectors of practice.

2.4. DATA GOVERNANCE IMPACTS

The term "impacts" refers to the expected outcomes in Data Governance resulting from the implementation and application of its practices and policies. In the literature, these impacts are categorised into two types: Intermediate performance effects and Risk management (Abraham et al., 2019).

The intermediate performance effects focus on how data governance helps to strengthen an organisation's dynamic and operational capabilities, resulting in competitive performance gains (Mikalef et al., 2018). Through qualitative research, Tallon et al. (2014) illustrates this impact across various industries. In the study, the interviewees in the healthcare industry recognised the role of data governance in information provisioning (i.e., having the correct information in the right place at the right time), reducing medical errors and increasing overall efficiency. In the airline industry, data governance was associated with better decision-making in scheduling, market analysis, and ticket pricing. By adopting data governance practices, airlines were able collect, store, and analyse data of higher quality and combine it with real-time data, resulting in more accurate demand forecasts. Although Tallon et al. (2014) predominantly discusses intermediate performance effects rather than firm-level (e.g., sales growth or profitability), prior research indicates that positive intermediate-level results can aggregate to firm-level (Barua et al., 1995).

Improved data quality and a high-level understanding of its value are other intermediate outcomes commonly attributed to a Data Governance program (Abraham et al., 2019). Baker (2016) demonstrates that a well-established DG program reduces mistakes and business errors due to data inconsistencies. In this study, the author identified that organisations without a DG program to guide the implementation of data stewardship, metadata management and data processes spend more time reacting to the most recent data-related crisis, reducing the time spent on running the business and making continual process improvements. According to Randhawa (2019), promoting a data-driven culture at the organisational level, supported by an awareness of data quality requirements and the implementation of data policies, reduces data inaccuracies and operational costs by preventing later-stage data clean-up efforts.

The second type of data governance impact is the management of data-related risks. These stem from non-conformance with information policies or the absence of oversight regarding data quality (Abraham et al., 2019), significantly influencing the creation of risk reports (for internal use) and regulatory reports (submitted to external regulators) (Randhawa, 2019). These two reports require an accurate capture, analysis and reporting of organisation operations. If not performed correctly, it can increase operating costs through report regeneration and attract the attention of external regulators, leading to potential fines (Randhawa, 2019). In Randhawa's (2019) study, these issues were overcome by implementing a data governance program to ensure a complete alignment between data processes and technology. Here, it was studied the importance of having a data storage platform capable of supporting data sourcing requirements, as well as data testing and validation functions defined in the organisation processes and policies.

Further risk management impacts concern security and privacy breaches (Abraham et al., 2019), which negatively affect an organisation's clients and reputation (Randhawa, 2019). Similar to the direct relationship between regulatory reporting and data quality, data protection and privacy laws also drive security and privacy practices, imposing specific data retention, access and usage requirements (Mahanti, 2021). Through risk-mitigating processes that guide critical data handling and controls for monitoring compliance, data governance helps organisations in their compliance journey, ensuring that data is securely maintained throughout its lifecycle (Khatri & Brown, 2010). Moreover, Mahanti (2021) states that compliance requires the ability to locate and understand data associated with regulations, achievable through a data governance program that establishes data stewardship roles and standards around data elements and entities. Data governance ensures that metadata and data lineage information is up-to-date, thus facilitating the data discovery process by data stewards, who understand the business meaning and usage purposes of data (Mahanti, 2021).

Although some studies have investigated the topic, research on the impact of data governance on organisational outcomes is still limited (Mikalef & Krogstie, 2018), providing only qualitative evidence of its intermediate performance effects and risk management impacts (Randhawa, 2019; Mikalef et al., 2018; Baker, 2016; Tallon et al., 2014). While qualitative indicators offer directional insights into data governance impacts, quantitative metrics allow for the accurate monitoring, control, and improvement of data governance practices (Dai et al., 2016). This is particularly valuable for organisations seeking a compelling use case that links data governance with value generation (Nielsen et al., 2019).

In this context, Abraham et al. (2019) proposes that to comprehend Data Governance, it is necessary to conduct a deeper research on the intermediate performance effects and risk management, as well as their impact on strategic business outcomes. In doing so, the authors reinforce the importance of identifying the links between intermediate-level and firm-level performance effects, enabling organisations to quantify the impacts of data governance and design a business case.

2.5. VALUATION RESEARCH MODELS

Business value research explores how and to what extent the organisational practices lead to improved performance (Melville et al., 2004), encompassing efficiency and competitive impacts (Gregor et al., 2006). Particularly, internal capabilities, also termed "intangibles", are critical when theorising on the business value of IS investments, playing a fundamental role in understanding their overall value (Schryen, 2011). Data governance is an emerging IS investment (Alhassan et al., 2018), thus allowing insights from IS research to be applied and adapted to comprehend its business value.

To study the process of creating internal and competitive value, Schryen (2013) proposes a research model that encompasses three dimensions: IS capabilities, Socio-organisational capabilities and IS assets. Regarding IS capabilities, the author defines this dimension as an interlocking system of practices composed by the culture of IS use when performing organisational tasks, and competencies, which include employee and management skills. In the field of data governance, these dimension components pertain to a data-driven culture and data competencies, which for the purpose of this study no distinction is made between staff and managers. Socio-organisational capabilities pertain to the ability to align and adapt resources and activities towards operations optimisation, with the author deeming (intra- and inter-organisational) processes and structures as essential components. In data governance, structures encompass the roles and respective responsibilities defined within the data stewardship of a governance program. Lastly, IS assets refers to the investments allocated for strategic purposes, being classified as infrastructure, transactional, informational or strategic assets. In data governance, the transactional (i.e., automates processes) and the informational (i.e., provides information for managing data operations) technologies are the focal assets.

Martijn et al. (2015) presents another research model that focuses on the creation of business value. This model studies how data governance improves data quality, increasing business value and compliance. Addressing a research gap, the model establishes a direct link between data governance and data quality, derived from the various interrelated elements that compose a data governance program. These elements include data cleansing and validation activities, roles and responsibilities to establish clear ownership, data standards that define the processes and uses of organisational data, and the ability of organisations to consult, share and monitor their data. In turn, data quality determines the effectiveness of business processes and influences reporting quality, affecting business value and regulatory compliance. First, data is used in business processes and strategic reports (e.g., forecasting or budgeting) and, as such, errors directly affect business value. Second, poor-quality reports may fail to produce the necessary evidence to verify compliance with regulations.

The last research model analysed was developed by Melville et al. (2004) and describes the relationship between business value and organisational performance. Accordingly, within an organisation, the business value generation process starts with the deployment of IT resources, namely technological infrastructure and technical skills, as well as organisational resources, including structures, policies and rules, workplace practices and culture. Combining these resources enables business processes and improves their performance, leading to an increase in operational efficiency of specific business processes, such as quality improvement and enhanced cycle time. Ultimately, this generates business value, impacting organisational performance (i.e., performance profitability, market value and competitive advantage).

Contextual factors significantly shape the impacts of data governance practices, with scholars recognising these factors as Industry, Organisational and Macro-economic (Schryen, 2013). At the Industry level, differences in regulatory and technological standards that depend on the activity sector, affect an organisation's data governance capabilities (Srinivasan & Swink, 2018). Organisational factors pertain to both organisation's characteristics and operational capabilities (Schryen, 2013). Regarding organisation characteristics, the size greatly affects organisational processes (Damanpour, 1996). While larger organisations may have more resources to develop governance practices, they are also more bureaucratic and less flexible when it comes to change (Srinivasan & Swink, 2018; Damanpour, 1996). Within operational capabilities, the knowledge and experience that organisations gain over time by working with data governance, as well as the sophistication of their programs, further affect the performance and scope of data governance implementation (Elbashir et al., 2013). Lastly, Macro-economic factors, such as a country's laws and information infrastructure, can pose constraints based on their development degree (Schryen, 2013).

3. METHODOLOGY

3.1. Methodological Approach

Within the field of science, March & Smith (1995) defend the existence of design and natural science, two disciplines with different goals and approaches. Design science focuses on designing and developing new artefacts and systems to serve human purposes, producing and applying knowledge of tasks or situations to create effective artefacts. In contrast, natural science seeks to understand reality by explaining how and why things are through observation and experimentation, with new knowledge, theories, and models being its main outputs (March & Smith, 1995). Therefore, the natural science approach is the one most aligned with the goals previously defined.

Accordingly, the methodological approach chosen for the development of this study was the natural science research methodology (Figure 5). This study follows the principles, practices and procedures established by this methodology in order to conduct the research process in a rigorous, systematic, and transparent manner, ensuring the scientific validity and reliability of the research findings.

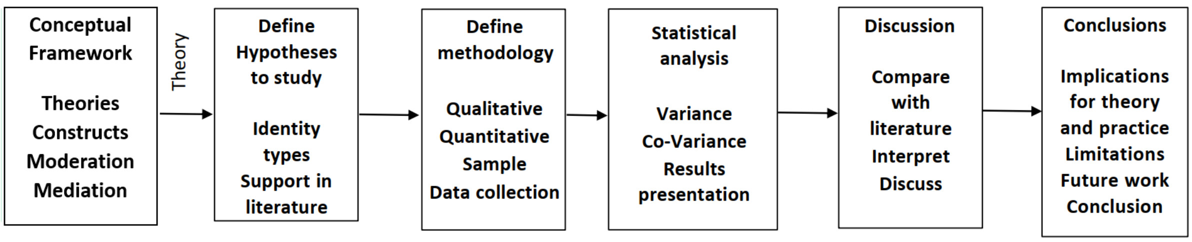


Figure 5 – Natural Science Research Methodology

Considering the stages of the adopted methodology, the work plan of this study is composed by following activities:

1. **Conceptual Framework Definition:** The research model is defined to represent the relationships among variables and concepts in the area of investigation. This model serves as a theoretical framework for conducting research, being used to guide the study's design, data collection and analysis, and results interpretation.

2. **Hypotheses Formulation:** The research hypotheses are formulated based on the literature, providing a potential explanation or prediction about the relationship between variables. The hypotheses serve as a starting point for the study research.

3. **Methodology Definition:** The methodology is established, particularly the data collection method and the type of observations, qualitative (descriptive) or quantitative (numeric), determining the subsequent analysis.

4. **Data Analysis:** The collected data is analysed to determine whether the hypotheses are supported. This stage involves the use of statistical techniques to determine the relationship between variables and to make inferences about the population based on the collected data.

5. **Results and Discussion:** The study results are interpreted to determine the meaning and significance of the findings. This involves the evaluation of the hypotheses supported by the data and the formulation of a conclusion about the relationship between variables.

6. **Conclusion and Communication:** This stage involves drawing conclusions based on the results and communicating the findings through various channels. In this case, this task includes the development of the study and its presentation at the master's degree defence.

3.2. RESEARCH MODEL

Drawing on the literature review presented in Sections 2.3 and 2.5, this study proposes the research model shown in Figure 6. This model assesses the value generation process driven by the implementation of data governance CSFs. In doing so, it is hypothesised the realisation of positive effects on Business Value (H1, H3, H5, H7, H9 and H11), Data Quality (H2, H4, H6, H8 and H10) and Regulatory Compliance (H14). Additionally, to account for some differentiating factors between organisations, the industry and country in which an organisation operates, its size in terms of employees, the DG framework used to develop a program and time since its adoption were employed as control variables. These controls aim to offer some additional insights in the explanation of the dependent variables (Data Quality, Business Value and Regulatory Compliance).

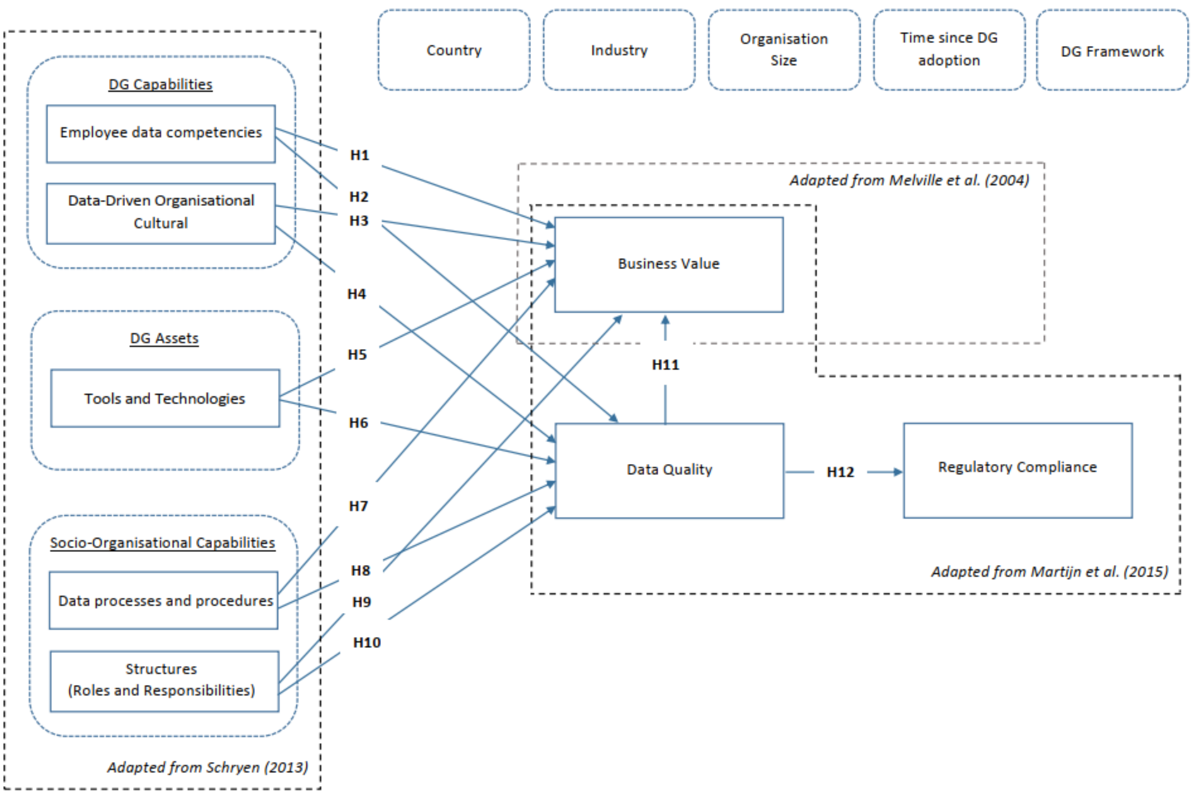


Figure 6 – Theoretical Research Model

3.3. HYPOTHESES

3.3.1. Data Governance Capabilities

Data-drivenness, as defined by Ladley (2020), involves cultivating a culture that treats data as a valuable asset. In essence, a Data-Driven Culture (DDC) fosters behaviours and practices where data plays a critical factor in organisational success, driving decision-making processes and actions to gain competitive advantages (Cao & Duan, 2014).

When fostering a DDC, a key component is the employee mindset towards data value. In many organisations, there is a lack of knowledge regarding how and what data is being valued, especially in decision-making, leading employees to perceive it as a cost rather than an asset (Berntsson Svensson & Taghavianfar, 2020). A DDC thrives on openness about data-related activities, encouraging different teams to communicate their work and how they are leveraging corporate data (Berntsson Svensson & Taghavianfar, 2020). It is also promoted a common understanding and trust in data by communicating the procedures, tools and technologies involved in data collection and analytics, communicating as well the existing data quality and data validation processes (Berntsson Svensson & Taghavianfar, 2020). Therefore, in a DDC, employees gain a clearer awareness of the impact and value-creation potential of data and its processes, influencing their support

In a DDC, employee data competencies are crucial, encompassing abilities like understanding the data-relevant processes, guidelines and standards; leveraging the available tools and technologies; employing analytical methods; and effectively communicating data insights (Klee et al., 2021). Since employees are the starting point for daily and operative analytical work, which is the foundation for using data and generating business value (Klee et al., 2021), their competencies directly impact the performance of data processes. Scholars also emphasise the importance of complementing data processes with human knowledge to produce better results (Günther et al., 2017). Accordingly, having qualified employees ensures a more conscious judgment to identify opportunities and solve problems, essential when implementing and executing data processes.

To explore the effect of a Data-Driven Culture in an organisation's daily processes, namely the impact it has on performance and data quality standards, the following hypotheses were proposed:

H1. Developing employee data competencies will allow for more efficient data processes implementation and performance

H2. Employees with higher levels of data competencies will produce and ensure higher quality outputs

H3. Having a data-driven culture promotes widespread support for the improvement of business processes' performance with more accurate and actionable insights, generating business value

H4. Establish a data-driven culture supported by data awareness will contribute to wider organisational support towards maintaining data assets' quality

3.3.2. Data Governance Assets

Although compliance with legal and ethical regulations drives the adoption of data governance programs (Mahanti, 2021), their impacts go beyond just compliance. Organisations increasingly see data governance as a means for controlling data quality and optimising data-related operations, ensuring the delivery of trustworthy decisions. To effectively address challenges and fulfil organisational goals, data governance must encompass the systems through which data is collected, managed and used (Janssen et al., 2020). As such, technology emerges as an impactful component in every data governance domain, supporting data requirements and business operations.

A central domain within a data governance program is metadata management. Over the last decade, its importance has been rapidly increasing as metadata is transitioning from a technical aspect to a business necessity (van Helvoirt & Weigand, 2015). A clear understanding of data and business objects (e.g., materials, customers, and suppliers) is imperative for organisations to implement and benefit from integrated, automated business processes (Hüner et al., 2011). A repository that stores and maintains high-quality metadata helps to create a shared understanding of data that can be consumed by both business/IT users (van Helvoirt & Weigand, 2015), ensuring smooth and efficient business processes performance.

Another data governance domain often supported by technology is data quality. Effective data management through quality assurance produces reliable information and fosters its strategic and tactical utilisation. Mechanisms addressing data quality issues, whether single-source (e.g., flawed schema design, data input errors, duplicated values) or multi-source (e.g., inconsistent aggregation), are crucial to maintaining trust in organisational data (Sidi et al., 2012). Poor data quality discourages decision-makers from data-driven decisions, forcing them to rely on subjective knowledge, not taking advantage of data insights and not accumulating experience with corporate data usage (Kwon et al., 2014). Implementing robust data quality mechanisms encourages organisations to unlock the full potential of available data, fostering greater benefits from successful utilisation.

To understand how tools and technologies support processes and operations, as well as the efficiency of data teams, it is hypothesised that:

H5. Having an intuitive metadata management tool and ensuring its continuous maintenance impacts the performance of all tasks involving data assets, influencing business value

H6. An alignment between data processes and technology, where the employed platforms fully support the data requirements, leads to higher-quality data outputs, providing a stronger base for decision-making

3.3.3. Data Governance Processes

Like other assets, data also has a lifecycle involving operations like data collection, storage and maintenance in systems, user retrieval and manipulation, and archival or disposal (Lee & Strong, 2003). In a data governance program, the goal is to integrate all data phases into the organisation (Koltay, 2016). This is achieved through DG processes that establish activities, procedures, responsibilities and necessary systems (Koltay, 2016). These processes ensure that data is ready to be consumed by other organisational operations, detailing "what is to be accomplished" and "how", specifying inputs, outputs and timings (McGilvray, 2021).

Data governance processes establish operational routines and guidelines, contributing to organisational knowledge, a key source of business value. These processes play a critical role in proficiently managing data and delivering it to end-users, supporting business processes (Côrte-Real et al., 2017). Therefore, DG processes represent a knowledge management source, allowing information provision and helping knowledge flow to achieve business excellence and operational efficiency.

To study the impact of having data processes and procedures on organisational efficiency, the following hypotheses were proposed:

H7. Data Governance processes lead to a better information provisioning within an organisation, contributing to the efficient performance of business processes

H8. Effective Data Governance processes standardise data operations to produce higher quality data outputs, reducing the risk of data errors and inconsistencies

3.3.4. Data Stewardship

Data is vital for organisational success. As a result, to make data reliable, available, and usable across an organisation, a data governance program is often essential. When properly executed, it manages data collection, including structure, processes, and key data resources (Plotkin, 2014). Data stewardship is an integral component in a data governance program, establishing accountability and responsibility for both data and processes.

Stewardship activities depend on an organisation's specific needs, with the creation and management of core metadata as primary focus (DAMA International, 2017). Additionally, having stewards responsible for data, along with procedures that require steward consultation on data-related decisions, ensures that the decisions are based on knowledge and are made in the best interests of all data users (Plotkin, 2014). The combination of designated stewards, processes and a mission to manage data in the best interests of all enhances the quality of data assets, driving business value and regulatory compliance (Plotkin, 2014).

To assess the impact of clearly established data roles and responsibilities on the quality of an organisation's data, the following hypotheses were created:

H9. The existence of data stewards accountable for ensuring the quality of data assets, and their respective metadata, will help provide high-quality data in a consistent manner

H10. A well-defined data stewardship, and thus a clear ownership and accountability of data assets, positively influence business value

3.3.5. Data Quality

Data quality reflects the capability of data to meet business, system and technical requirements, accomplished through the use of complete, consistent, accurate, and timely data (Mahanti, 2019). Known as data quality dimensions, completeness refers to the level of missing data and data sufficiency; consistency pertains to how well data is presented; accuracy focuses on data correctness and reliability; and timeliness centres on data currency (Batini et al., 2009).

Understanding the diverse applications of data is crucial for studying the impacts of data quality. For organisations, data is no longer a system by-product, but a valuable resource driving various operations (Mahanti, 2019). For instance, product data influences sales, marketing, inventory and financial forecasting. By ensuring data quality, organisations mitigate operational inefficiencies and reduce costs associated with verification and correction resulting from redundant, incomplete, or incorrect data (Mahanti, 2019). Furthermore, good data quality means that the analysis and the KPIs used to guide the organisation's strategy and decisions are strongly supported, enhancing business value (Ji-fan Ren et al., 2017). Additionally, quality data is particularly vital for meeting and complying with regulatory controls and reporting requirements (Mahanti, 2021).

To study the multiple impacts of data quality in organisations the following hypotheses were created:

H11. High data quality supports value-generating processes that positively affect business value

H12. High quality data enables the production of regulatory reports that ensure the organisation's compliance with regulatory requirements

3.4. METHODOLOGY DEFINITION

The research model aims to establish and explain cause-effect relationships between data governance practices (independent variables) and organisational impacts (dependent variables). As such, the methodology combines confirmatory and explanatory research (Benitez et al., 2020). Following the recommendations of Abraham et al. (2019) and due to the predominance of qualitative research in the study of data governance impacts, the collection of quantitative data was considered the most added-value approach. To test the research model and associated hypotheses, a multi-organisational questionnaire was the method chosen to collect data on organisational data governance programs across industry sectors. The questionnaire strategy enables the collection of quantitative data that can be used to feed into a model aiming to explain these relationships (Saunders et al., 2009).

Regarding data collection methodology, the questionnaire will be conducted using the online surveying platform Qualtrics and distributed among the organisations in the firm's client database. To ensure data quality and reliability, employees with knowledge of data governance practices and their impacts on business processes and data quality will be the primary targets. As such, the respondent profile is defined by three criteria: working in an organisation with data governance initiatives, deep knowledge of the organisation's data operations, and holding an IT (or data-related role), executive or management position. Previous studies also employed a similar methodology (Ndamase, 2014; Côrte-

Real et al., 2017; Chaudhuri et al., 2021), demonstrating the effectiveness of purposive sampling in targeting respondents with knowledge expertise.

When defining the methodology, the sample size must also be considered. A baseline of the minimum sample size required to ensure the result's representability can be established by considering statistical power tables developed in previous studies (Kock & Hadaya, 2018). This approach, which builds on Cohen's (1992a) power tables, consists of obtaining the minimum required sample sizes based on three parameters (Kock & Hadaya, 2018): the maximum number of arrows pointing at a construct, the significance level and the effect size. By defining these parameters as 6, 5% and 0.1, the threshold is set at 157 valid responses (Appendix A).

4. DATA COLLECTION

4.1. MEASUREMENT CONSTRUCT

To ensure theoretical grounding in research and practical relevance in the industry, the measurement items used in the questionnaire were developed by combining support from literature and industry experts, resulting in an enhanced content validity and reliability of the instrument (Yu et al., 2021). Regarding content validity, the questionnaire draws upon a comprehensive literature review, where its measurement items were adapted from prior research. Moreover, a multi-item approach was employed for each construct since the combination of multiple items covers a larger number of distinct construct facets, offering a higher degree of construct validity (Ongena, 2023).

To test the reliability of the questions together with its clarity, 12 experienced practitioners in the field of data governance and data quality were invited to review and provide feedback on the questionnaire items. The main objective was to identify unclear or ambiguous questions, questions the respondent might feel uneasy about answering, and whether there were any major topic omissions. Following their feedback, redundant and ambiguous measurement items were eliminated or modified, and a deeper literature search was performed to fill identified gaps.

The questionnaire, structured into three sections—Demographics, Practices, and Impacts—was developed in line with the research model, yielding 57 items. The Demographics section details respondent and organisational profiles through five items: industry sector, role within the organisation, organisational size, DG program operationalisation, and DG framework. Regarding the country, no measurement item was created for this control variable since the Qualtrics platform automatically retrieves respondent location, considered a proxy for the organisation's operating country. The Practices section comprised 31 items designed to measure organisational adoption of data governance practices across five constructs: Data-driven Culture, Data Competencies, Tools and Technologies, Processes and Procedures and Data Stewardship. In the Impacts section, 21 items measure the perceived impact of governance practices, encompassing key indicators of Data Quality, Business Value and Regulatory Compliance. The final version of demographic and measurement items, along with their sources and constructs, is presented in Appendix B and C.

Afterwards, the questionnaire was operationalised on the Qualtrics platform, beginning with an introduction where the aim of this study is explained, and respondent anonymity and confidentiality is assured. Upon participant consent, the Demographic questions are presented to collect control information, using a multi-choice format where the respondents select the option that better represents their organisations' reality. Next, the measurement items for the Practices and Impacts sections are introduced in the form of statements, and a seven-point Likert scale (1="strongly disagree" to 7="strongly agree") was used to quantify the responses. This was preferred over the five-point scale since it offers more options for a more reliable expression of the respondent's views, while still providing a span of options that does not overwhelm the respondent's capacity to apprehend and distinguish the items (Preston & Colman, 2000). Lastly, to collect further insights regarding the respondent's profile, two control questions were presented about the degree of knowledge concerning the questionnaire theme and NOVA IMS.

4.2. Collection Process

After operationalising the questionnaire, an email was sent to the representatives of each Organisation in the firm's client database asking for consent to disseminate the questionnaire among the data and IT teams. Upon consent, a follow-up email was sent providing some additional context along with the questionnaire link.

The questionnaire remained active for a period of two months, spanning from April to May of 2023. Figure 7 depicts the progression of responses, including both complete and incomplete, throughout this period.

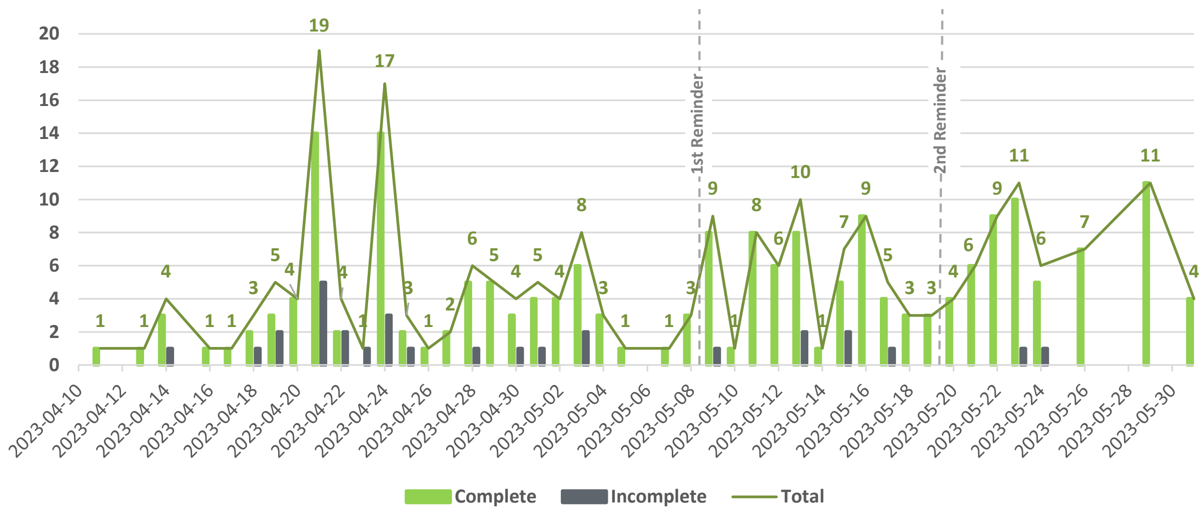


Figure 7 – Responses Progression

As presented in Figure 7, within the first month, 107 responses were received, out of which only 85 were deemed usable, with 21 being incomplete, and one participant declined participation. When assessing the response rate, the number responses only reached around half of the defined objective. Furthermore, incomplete responses accounted for approximately 20% of overall submissions, indicating a relatively high degree of incompleteness. Employing strategies from prior studies (Baptista & Oliveira, 2015; Côrte-Real et al., 2017) to enhance response rates and overcome the incompleteness issue, a multi-contact strategy was adopted. The first reminder, dispatched via email on May 8th, aimed to engage non-respondents, by reiterating the questionnaire's final deadline, highlighting the significance of completing it, and emphasising response confidentiality.

Following the first reminder, 62 additional responses were received, comprising 55 valid responses, 6 incomplete and 1 non-agreement. As illustrated in Figure 7, the reminder successfully resulted in an increase in the overall number of responses, reducing the rate of incomplete responses to 9%. These outcomes effectively demonstrate the tangible benefits derived from adopting a multiple contacts strategy. Continuing with the strategy implementation, a second reminder was sent on May 19th to notify respondents about the approaching deadline. Similar to the first reminder, the second one resulted in a new peak of responses, culminating in 58 responses, with only two remaining incomplete.

In conclusion, throughout the active period of the questionnaire, 227 responses were received. However, only 196 responses were considered valid due to either incompleteness or non-consent to participate, with the different response types presented in Figure 8. Therefore, considering the established minimum target of 157 responses, it can be concluded that, given the size of the obtained sample, the objective was not only met but exceeded by a substantial margin

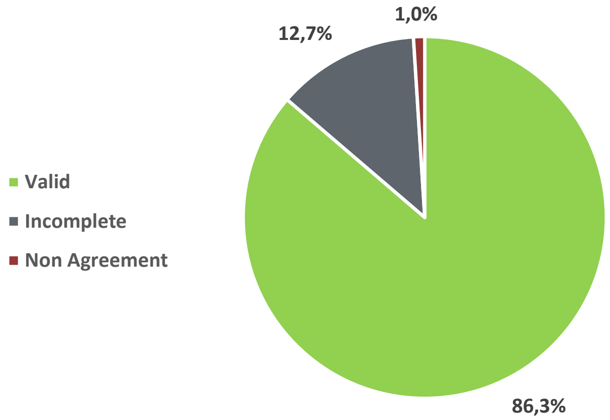


Figure 8 – Responses Received

4.3. Sample Demographics

The theoretical model integrates control variables to account for contextual factors related to industries, organisations, and macroeconomics. These variables serve two purposes: firstly, to explain the impacts of data governance practices, and secondly, to provide insights into respondent characteristics, allowing for a comprehensive characterisation of the collected sample. Table 1 displays descriptive statistics, such as frequency and percentage, for the control variables concerning country, industry, respondents' positions within the organisation, organisation size, time since the adoption of a data governance program, and the framework used.

Table 1 – Descriptive statistics of sample profile

Sample Demographics (N=196)	Frequency	Percentage (%)
Country		
Portugal	76	38.78%
United Kingdom	46	23.47%
Angola	34	17.35%
Germany	32	16.33%
Brazil	8	4.08%
Industry		
Banking/finance	31	15.82%
Business services	28	14.29%
Technological development	20	10.20%
Insurance	19	9,69%
Other	18	9,18%

Oil and Gas	11	5.61%
Education and Research	11	5.61%
E-Commerce	9	4.59%
Retail/Distribution	8	4.08%
Communications	8	4.08%
Utilities	8	4.08%
Defence	6	3.06%
Government/Municipalities	6	3.06%
Logistics	6	3.06%
Food and Beverage	3	1.53%
Manufacturing	2	1.02%
Health services	2	1.02%
Respondent Position		
Operational/Data worker	53	27.04%
Middle/first line manager	50	25.51%
Middle/knowledge worker	50	25.51%
Top/executive manager	28	14,29%
Operational/Production worker	13	6.63%
Don’t Know/Don’t want to answer	2	1.02%
Number of Employees		
>5000	66	33.67%
1001-5000	47	23.98%
501-1000	29	14.80%
251-500	23	11.73%
50-250	17	8.67%
<50	13	6.63%
Don’t Know/Don’t want to answer	1	0.51%
Data Governance Adoption		
The program has been operational for over 5 years	75	38.27%
We do not have an operational program	44	23.98%
Don’t Know/Don’t want to answer	40	14.80%
The program has been operational for 3 to 5 years	23	11.73%
The program has been operational for 1 to 3 years	13	6.63%
The program has been operational for less than 1 year	1	0.51%
Data Governance Framework		
Don’t Know/Don’t want to answer	153	78.06%
SAS Data Governance Framework	17	8,67%
Internally-developed framework	9	4,59%
DAMA-DMBOK Functional Framework	5	2.55%
DGI Data Governance Framework	5	2.55%
PwC Data Governance Framework	3	1.53%
Boston Consulting Group Framework	3	1.53%
Eckerson Modern Data Governance	1	0.51%

As seen in Table 1, the sample encompasses five countries, with a predominant focus on European nations, accounting for approximately 78.58% of the total. Among these countries, Portugal exhibits the most pronounced predominance. The sample also includes a significant representation of an African nation (17.35%), such as Angola, and despite having the lowest representation (4.08%), South America is also present in the sample through the inclusion of Brazil. To analyse the results, the Country control variable was converted into a binary variable named "Developed Country". For this variable, countries that possess developed status, like Germany, Portugal, and the United Kingdom, were assigned a value of 1. The remaining countries, categorised as developing, were assigned a value of 0.

Regarding the industry variable, the sample contains a diverse range of sectors, with banking/finance being the most prominent, accounting for 15.82%. Adhering to the approach proposed by Yu et al. (2021), dummy encoding was exclusively employed for industries with significant representation within the sample, due to the limited representativeness in certain sectors. As a result, four dummy variables were generated to represent banking/finance, business services, technological development, and insurance, each corresponding to their respective sectors.

The remaining variables focus on respondent and organisational characteristics. In terms of respondents, the sample displays a significant representation of different positions, with Operational/Production workers being the exception, constituting less than 10%. Regarding organisational size the majority are large organisations with over 1000 employees (57.65%), followed by medium-sized companies ranging from 250 to 1000 employees (26.53%), and smaller organisations with fewer than 250 employees (15.30%). Following Chen et al.'s (2014) recommendation, the ordinal encoding technique was employed for the Organisation Size control variable. Numeric codes were assigned based on natural order, with "<50" coded as 1 and ">5000" as 6, while the response "Don't Know/Don't want to answer" was coded as -1.

The operational characteristics of the organisations in the sample are reflected through the Adoption and Framework variables. Half of the sample consists of organisations with a mature data governance program, operational for more than 3 years. As for the framework used, the majority of respondents (78.06%) were unable to identify the framework employed by their organisation, which can be justified by the technical nature of this variable. In the analysis of these two control variables, the ordinal encoding strategy was also applied to the Adoption variable. The option "We do not have an operational program" was encoded as 0, while the remaining options followed a similar logic. Additionally, similar to the previous case, the "Don't Know/Don't want to answer" option was coded as -1. For the framework variable, given the lack of information collected, this variable will not be considered in the results analysis.

4.4. Bias Assessment

When conducting a questionnaire, it is crucial to test for bias to ensure the integrity and validity of the collected data. Assessing biases allows for the identification of discrepancies and the implementation of corrective measures to minimise their influence. As such, to enhance the trustworthiness of the research findings, the sample collected is examined for two types of bias: Non-response and Common Method.

Non-response Bias refers to a potential biased outcome due to systematic differences between those who respond to the questionnaire and those who do not. This bias occurs when the characteristics of non-respondents differ from those of respondents, leading to an unequal representation of certain groups (Sheikh & Mattingly, 1981). To mitigate this type of bias, specific strategies, such as purposive sampling and multiple contacts, were employed in the development of this study to ensure sample representativeness. The effectiveness of the multi-contact strategy is demonstrated in Section 4.2. As for the purposive sampling, the outcomes can be analysed by examining the control question on the respondents' level of knowledge, presented in Figure 9.

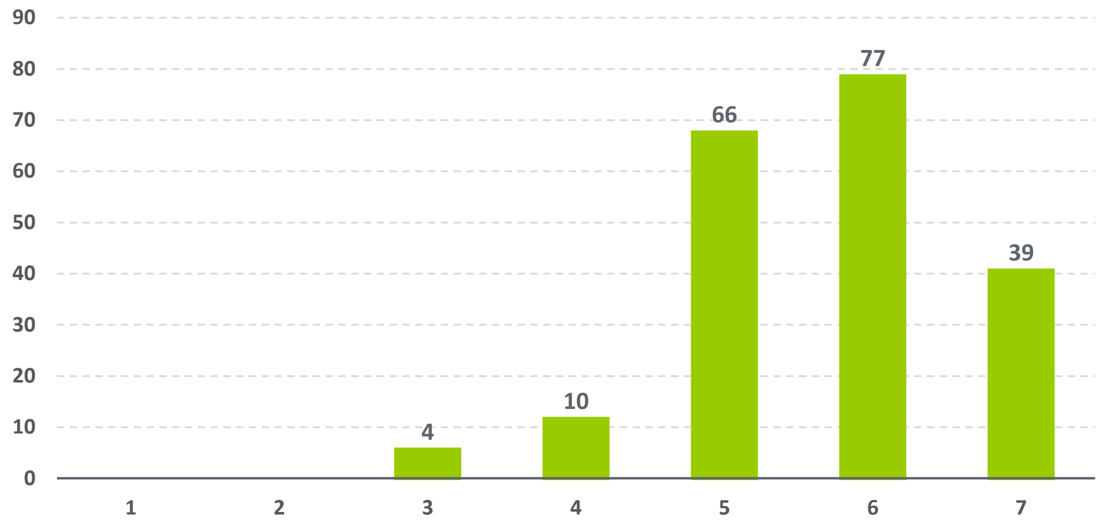


Figure 9 – Respondent's Degree of Knowledge

Upon analysing Figure 9, the majority of respondents (60%) reported a "High" or "Very High" level of knowledge regarding the data governance practices implemented in their organisations and their associated impacts. The remaining participants reported a "Neutral" or "Moderately High" level of knowledge, with only 4 respondents rating their knowledge as "Somewhat Low", suggesting a near absence of low-knowledge respondents. Consequently, considering that the average knowledge level was established at 5.7, it can be concluded that the purposive sampling strategy successfully yielded the desired results. Therefore, with the success of the two strategies confirmed, the next step involved testing the sample for the presence of non-response bias.

Based on the literature, different methods can be employed when estimating non-response bias. Due to this study nature, the Extrapolation Method was the one chosen for this purpose. This method relies on the premise that late respondents who required more time to complete the questionnaire share similarities with non-responders. By comparing early and late responders, it is possible to identify any potential differences in their characteristics and assess whether these differences will affect the study outcomes (Sheikh & Mattingly, 1981).

To conduct this analysis, the early respondents were selected as those who completed the questionnaire within the first month of its release. In contrast, the latter respondents were identified as the ones who received the follow-up reminders, requiring a longer time to respond. Using the approach of previous studies (Côrte-Real et al., 2017; Aydiner et al., 2019), a t-test was performed for each construct to compare the responses from the two groups. The results are displayed in Table 2.

Table 2 – Testing response bias: early vs. late respondents

Constructs	Full Sample N=196		Early Respondents N=85		Late Respondents N=111		T-Test	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	*t*-value	*p*-value
DD	5.66	1.10	5.44	1.14	5.83	1.05	-2.55	0.07
DC	5.81	1.11	5.66	1.21	5.92	1.02	-1.68	0.20
TT	5.67	1.08	5.45	1.13	5.84	1.01	-2.61	0.06
PP	5.50	1.13	5.27	1.24	5.68	0.99	-2.55	0.054
RR	5.32	1.22	5.11	1.28	5.49	1.15	-2.19	0.11
DQ	5.77	1.13	5.56	1.28	5.93	0.96	-2.28	0.12
BV	5.60	1.17	5.44	1.12	5.72	1.20	-1.69	0.23
RR	5.89	1.09	5.76	1.08	5.98	1.08	-1.48	0.18

Upon conducting a t-test, the results presented in Table 2 reveal that each construct demonstrates a *p*-value greater than 0.05, which was chosen as the threshold for statistical significance (Appendix A). These results indicate that the observed difference between the two groups, early and late respondents, is not statistically significant. Therefore, following Ryan's (2004) guidelines, there is insufficient evidence to conclude that the groups differ significantly based on the collected data, suggesting the absence of non-response bias.

Additionally to the respondent's profile, the measurement method used to test the relationships between constructs can also be responsible for introducing biases into the research results. Commonly known as Common Method Bias (CMB), this issue is widely regarded as a major source of measurement error that can compromise the study validity (Podsakoff et al., 2003). Given that CMB stems from a common source or measurement context, as well as the characteristics of the items themselves (Podsakoff et al., 2003), this study assesses its potential effects on the research findings. To address this concern, the recommendations of Podsakoff et al. (2003) are followed, employing both ex-ante and ex-post approaches to minimise the potential effects of CMB.

In terms of ex-ante approaches, the questionnaires were sent to multiple Organisations representing different industries. This diverse sample selection aims to incorporate data from various sources, reducing the likelihood of bias stemming from a single organisational context. Furthermore, respondents were assured that all provided information would remain completely anonymous and confidential, emphasising that any data analysis would be performed solely for research purposes.

After the data collection (ex-post), Harman's post hoc single-factor analysis was employed, recognised as a primary technique for testing CMB. The analysis aimed to identify a single factor accounting for the majority of data variance, suggesting the potential presence of CMB (Podsakoff et al., 2003). Based on the factorial analysis, it was determined that a single factor explains 42.70% of the variance in the data. This level of variance falls below the threshold of 50%, suggesting that the relationships among variables are primarily influenced by the underlying constructs rather than solely by the measurement method. Therefore, CMB is not a significant concern in this study.

5. ANALYSIS AND RESULTS

5.1. ANALYSIS APPROACH

Structural Equation Modelling (SEM) is a statistical technique used to analyse the relationships between variables, providing a framework for testing and estimating causal relationships between independent and dependent variables (Ullman & Bentler, 2012). Researchers commonly rely on SEM to assess the fit of their proposed models to the observed data and evaluate their theoretical models' validity and adequacy (Ullman & Bentler, 2012). In this study, the selection of SEM as the analytical technique enables a thorough analysis of the collected data, aligning with the research goals and providing a robust methodology for investigating the relationships among constructs.

In SEM, two primary types are recognised: Covariance-Based SEM (CB-SEM) and Partial Least Squares SEM (PLS-SEM). Considering the data characteristics and the research model, PLS-SEM was the technique chosen for its suitability in studies with limited sample sizes and complex models. Unlike CB-SEM, which requires larger samples for reliable estimates, PLS-SEM yields robust results even with small/moderate sample sizes (Hair et al., 2021). Additionally, PLS-SEM offers greater flexibility when confronted with non-normal or non-linear data distributions, since it does not assume multivariate normality, making it suitable for non-normal data and non-linear relationships (Hair et al., 2021). To conduct the analysis, the Smart PLS 4 software was used, leveraging the strengths of PLS-SEM to explore the research objectives.

In accordance with J. C. Anderson and Gerbing's (1988) guidelines, the analysis was conducted in two phases. The first phase assessed the measurement model's reliability and validity, encompassing internal consistency and reliability for each construct and its corresponding indicators. Additionally, it examined the distinctiveness between different constructs, ensuring no strong correlations were present. Next, the second phase focused on assessing the structural model and conducting hypothesis testing. In this phase, the focus shifted to examining the relationships between constructs and testing the defined hypotheses regarding the causal relationships among them.

5.2. MEASUREMENT MODEL EVALUATION

When employing the SEM technique, it is crucial to consider the model type—reflective or formative—as it significantly influences the evaluation process (Hair et al., 2021). The direction of the relationship between indicators and constructs, whether it is indicator-to-construct or construct-to-indicator, determines the model type (Freeze & Raschke, 2007). In the proposed model, the observed variables represent outcomes of a construct, reflecting various dimensions. Thus, the measurement model consists of reflective variables, supporting its classification as a reflective model.

This study builds upon the methodological framework introduced by Hair et al. (2021) to undertake a comprehensive assessment of the reflective measurement model. The evaluation encompasses the examination of indicator reliability, construct reliability, convergent validity, and discriminant validity.

To evaluate the reliability of indicators in reflective constructs, their outer loadings are often examined to assess the strength of their relationship with the latent construct they aim to measure. Outer loadings range between -1 and 1, with higher absolute values signifying a stronger association between the indicator and the construct. A common rule of thumb sets a threshold of 0.7 to identify a strong relationship. However, indicators with outer loadings between 0.40 and 0.70 should only be considered for removal if removing them improves the composite reliability or average variance extracted (AVE) of the construct. On the other hand, indicators with outer loadings below 0.40 should always be eliminated from the construct (Hair et al., 2021).

The outer loadings in Appendix D were analysed in line with the guidelines. Among the outer loadings, only the RR6 indicator of the Roles and Responsibilities (RR) construct falls within the range of 0.4 to 0.7. Subsequently, an evaluation was conducted to assess the potential impact of excluding the RR6 indicator on both the composite reliability and AVE measures. The findings are presented in Table 3, which compares these quality criteria for the various constructs to determine the effect of removing the RR6 indicator

Table 3 – Quality criteria for Construct Reliability and Convergent Validity

Constructs	Before deleting the RR6 item			After deleting the RR6 item		
	Cronbach's Alpha	Composite Reliability	AVE	Cronbach's Alpha	Composite Reliability	AVE
DD	0.857	0.897	0.637	0.857	0.897	0.637
DC	0.917	0.932	0.633	0.917	0.932	0.633
TT	0.883	0.911	0.631	0.883	0.911	0.631
PP	0.909	0.930	0.688	0.909	0.930	0.688
RR	0.903	0.926	0.679	0.913	0.935	0.743
DQ	0.942	0.943	0.714	0.942	0.943	0.714
BV	0.915	0.932	0.662	0.915	0.932	0.662
RC	0.918	0.938	0.753	0.918	0.938	0.753

The analysis demonstrates significant improvements in both composite reliability and AVE measures upon the removal of the RR6 indicator. Consequently, the RR6 was excluded from the construct. Additionally, it is worth noting that the remaining loadings surpassed the threshold of 0.7, confirming the instrument's high indicator reliability.

When assessing construct reliability, researchers traditionally employ Cronbach's alpha, which estimates the reliability of a construct based on the intercorrelations among its indicators. However, Cronbach's alpha exhibits limitations, often leading to an underestimation of reliability values. To address this issue, Composite reliability has emerged as an alternative metric, providing a less conservative assessment, yielding higher reliability estimates. Both Cronbach's alpha and Composite reliability range from 0 to 1, with a threshold of 0.7 commonly used to indicate an acceptable level of reliability for a given construct. Moreover, Composite reliability values surpassing 0.95 indicate redundancy among indicators, suggesting the measurement of the same dimension (Hair et al., 2021).

As presented in Table 3, with the elimination of the RR6 indicator, all constructs exhibit Composite reliability and Cronbach's alpha values surpassing 0.7, while remaining below 0.95. These outcomes underscore the reliability of the constructs, as this range signifies a favourable level of internal consistency among the constituent indicators, implying an interconnectedness between the indicators in each construct.

The next step involves the assessment of convergent validity by examining the degree to which the different indicators designed to measure the same construct produce similar or correlated results. Researchers commonly use the AVE metric for assessing convergent validity in reflective constructs. This metric, ranging from 0 to 1, quantifies the variance explained by indicators within the construct, with higher values (closer to 1) signifying a stronger convergent validity. An established criterion sets an AVE threshold of 0.5 or above as suitable for confirming convergent validity (Hair et al., 2021). As illustrated in Table 3, all constructs demonstrate AVE values surpassing the threshold of 0.5. This suggests that the indicators are effectively converging to measure the intended concept, demonstrating that the construct is well-defined and cohesive.

The final step involves the evaluation of discriminant validity to determine the degree of distinctiveness among the various constructs. In other words, discriminant validity aims to demonstrate that the indicators of different constructs are not highly correlated with each other, indicating that they are measuring separate concepts. Traditionally, scholars have employed two methods to assess discriminant validity. The examination of cross-loadings serves as the first method, focusing on the relationships between indicators and constructs. When an indicator shows a significant loading on a construct other than the one it is intended to measure, a cross-loading occurs, indicating construct ambiguity (Hair et al., 2021).

A loading analysis was executed, revealing cross-loadings among the constructs of Data Competencies, Data Quality, and Processes and Procedures (Appendix E). These instances occurred since certain indicators exhibited higher loadings than some indicators designed to measure the specified constructs. In response, a decision was made to eliminate four indicators—namely, PP6, DQ4, DC3, and DC7—with the aim of mitigating the impact of cross-loadings. Following the removal of these indicators, a subsequent analysis of the loadings confirmed the absence of additional cross-loadings. Therefore, it was concluded that the prerequisites of the first method for discriminant validity analysis were met.

The Fornell-Larcker criterion is the second method for assessing discriminant validity, comparing the AVE square root for each construct with the correlations with other constructs. For discriminant validity to be established, the AVE square root for a specific construct must be greater than its correlations with other constructs. This logic underscores the notion that a construct shares more variance with its own indicators than with the indicators of other constructs, thereby signifying its distinctiveness (Hair et al., 2021). The results, as presented in Table 4, align with the principles of the Fornell-Larcker criterion.

Table 4 – Fornell-Larcker criterion with Square roots of AVE (in bold on diagonal)

	BV	DC	DD	DQ	PP	RC	RR	TT
BV	**0.813**							
DC	0.754	**0.818**						
DD	0.733	0.716	**0.798**					
DQ	0.771	0.794	0.737	**0.864**				
PP	0.736	0.764	0.729	0.832	**0.852**			
RC	0.808	0.780	0.726	0.796	0.814	**0.868**		
RR	0.697	0.729	0.685	0.737	0.823	0.750	**0.862**	
TT	0.633	0.673	0.670	0.654	0.672	0.728	0.628	**0.794**

However, recent research has highlighted the limitations of the prevailing methods for identifying discriminant validity issues. Specifically, the efficacy of cross-loadings diminishes in scenarios involving perfectly correlated constructs, while the Fornell-Larcker criterion proves inadequate, particularly when indicator loadings differ only slightly. These shortcomings consequently compromise the applicability of these two methods in empirical research. As a result, the heterotrait-monotrait ratio (HTMT) emerges as an alternative method, providing a more accurate estimation of construct correlation. Following established guidelines, an HTMT ratio below 0.9 indicates an acceptable discriminant validity, conforming that the constructs are distinct from each other (Hair et al., 2021).

In accordance with the specified guidelines, Table 5 displays the results of the HTMT criterion analysis. Notably, the calculated HTMT coefficients exhibit values lower than the predetermined threshold of 0.9. This adherence to the HTMT criteria supports the confirmation of discriminant validity, as the data complies with all three methods.

Table 5 – Heterotrait-monotrait ratio Matrix

	BV	DC	DD	DQ	PP	RC	RR	TT
BV								
DC	0.810							
DD	0.811	0.802						
DQ	0.808	0.855	0.814					
PP	0.782	0.832	0.820	0.896				
RC	0.887	0.854	0.814	0.850	0.887			
RR	0.748	0.798	0.764	0.789	0.897	0.816		
TT	0.707	0.749	0.762	0.709	0.742	0.807	0.699	

In conclusion, the outcomes of the measurement model underscore its alignment with recognised thresholds in the existing literature across the multiple measures. This alignment serves as strong evidence supporting the reliability of both indicators and constructs, validating convergent and discriminant validity within the model. Meeting these prerequisites ensures that the constructs not only display statistical distinctiveness but also maintain a reliable measurement through their respective indicators, making them suitable for testing in the structural model.

5.3. Structural Model Evaluation

Upon confirming construct reliability and validity, the next phase involves assessing the structural model's outcomes, focusing on two types of analyses. The first type examines the direct effects of Data Governance dimensions on organisational outcomes, testing the relationships among constructs and validating hypotheses. The second type explores the effects of contextual factors by analysing the influence of control variables on the dependent constructs.

To guide the process of hypothesis testing, this study adopts Hair et al.'s (2021) five-step approach: (1) collinearity assessment, (2) determination of path coefficients in the structural model, (3) calculation of the coefficient of determination (R^2 value), (4) computation of effect size (f^2), and (5) examination of predictive relevance using Q^2. Concerning the impacts of contextual factors, the analysis was conducted using the approaches employed for the evaluation of: (1) the reflective measurement model, and (2) the structural model.

5.3.1. Structural Model and Hypothesis Testing

This section focus on the analysis of the proposed hypotheses within the structural model, examining the relationship between data governance and organisational impact. To do so, the first step involves assessing the presence of collinearity within the structural pathways.

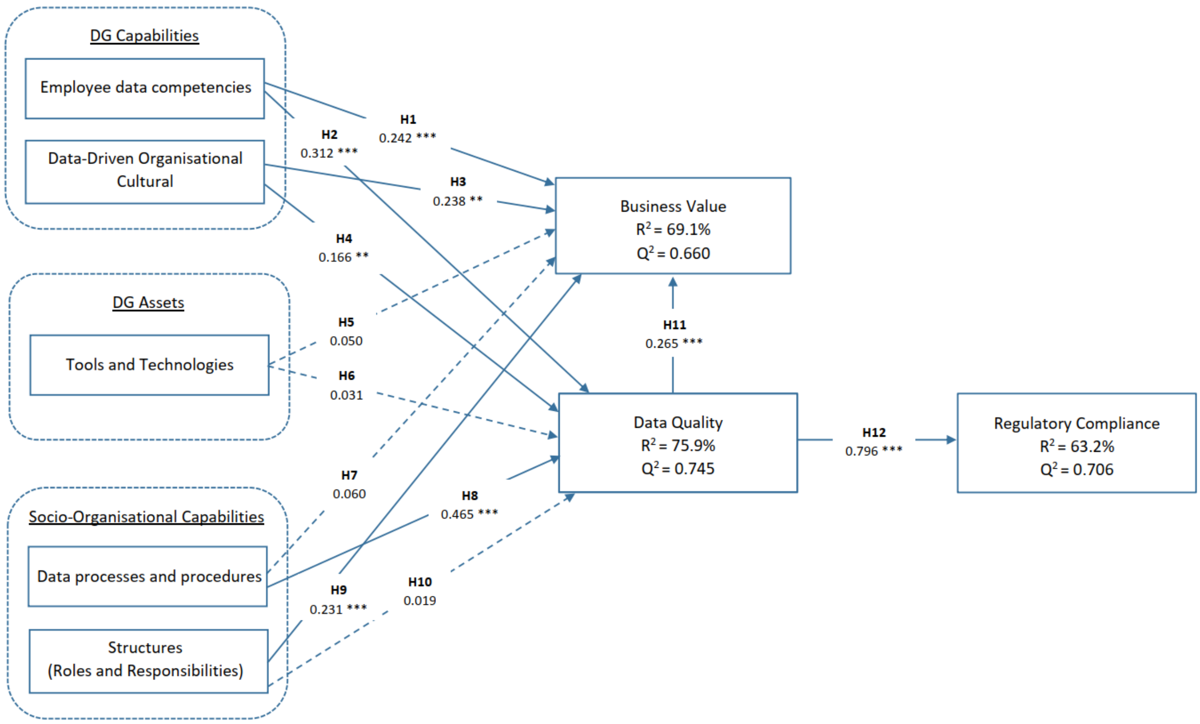
To accomplish this, the Variance Inflation Factor (VIF) is used as a metric to quantify the increase in standard error caused by collinearity. A VIF value of 1 signifies the absence of collinearity, while values from 1 to 5 are generally considered acceptable, indicating a moderate level of collinearity. Values exceeding 5 indicate a notable degree of collinearity which, in turn, implies a larger standard error due to the presence of collinearity with another construct. Consequently, this indicates reduced precision in estimating the effect of a variable on the outcome (Hair et al., 2021). Given these implications, Table 6 has been constructed to analyse the VIF values pertaining to each relationship between constructs.

Table 6 – Collinearity Statistics VIF Values

Structural Paths	VIF
Data Competencies → Business Value	3.432
Data Competencies → Data Quality	3.036
Data-Driven Culture → Business Value	2.787
Data-Driven Culture → Data Quality	2.676
Data Quality → Business Value	4.259
Data Quality → Regulatory Compliance	1.000
Data Processes and Procedures → Business Value	3.994
Data Processes and Procedures → Data Quality	4.149
Roles and Responsibilities → Business Value	3.434
Roles and Responsibilities → Data Quality	3.432
Tools and Technology → Business Value	2.223
Tools and Technology → Data Quality	2.220

According to Table 6, the path "Data Processes and Procedures → Business Value" exhibits the highest VIF value with 4.259. Therefore, since all VIF values are uniformly below the threshold value of 5, the results suggest that collinearity does not reach critical levels and is not an issue for estimating the structural model paths.

To evaluate the hypotheses proposed in Chapter 3, the significance levels associated with path coefficients were examined. Adhering to Hair et al.'s (2021) methodological guidelines, this analysis employs a bootstrapping technique, involving 5000 iterations of resampling, with each iteration composed by the number of observations (i.e., 196 cases). Following the completion of the bootstrapping process, Figure 10 summarises the outcomes of the structural model, providing insights into key metrics, including path coefficients ($\hat{\beta}$), R^2, and Q^2.



Note: --- p≥0.1 (non-significant); *p<0.10; **p<0.05; ***p<0.01

Figure 10 – Structural Model Results

The path coefficients, ranging between -1 and +1, indicate the strength and direction of hypothesised connections. A coefficient nearing +1 suggests a strong positive relationship, while a negative coefficient indicates an inverse relationship. To establish statistical significance, researchers commonly rely on p-values. These values represent the probability of erroneously rejecting a true null hypothesis—obtaining the observed result if there were no true relationship. When the computed p-value falls below the designated significance level, defined as 0.05, it implies the statistical significance of the corresponding path coefficient (Hair et al., 2021). However, in accordance with other authors (Baptista & Oliveira, 2015; Kwon et al., 2014; Fink & Neumann, 2009), this study adopts a less conservative approach, setting the threshold for statistical significance at 0.1.

Based on the analysis of path coefficients and corresponding p-values, certain dimensions hold significance in the scope of this research. Specifically, the dimensions of Data Competencies ($\hat{\beta}$=0.242;

p=0.008), Data-Driven Culture ($\hat{\beta}$=0.238; p=0.026), and Roles and Responsibilities ($\hat{\beta}$=0.231; p=0.007) exhibit notable positive influence, establishing their statistical significance in explaining Business Value. This confirms the validation of hypotheses **H1**, **H3**, and **H9**. Conversely, the dimensions of Tools and Technologies ($\hat{\beta}$=0.050; p=0.624) and Processes and Procedures ($\hat{\beta}$=0.060; p=0.557) reveal no substantial statistical significance, thereby failing to provide support for hypotheses **H5** and **H7**.

Furthermore, an examination of the impact of Data Competencies ($\hat{\beta}$=0.312; p=0.000), Data-Driven Culture ($\hat{\beta}$=0.166; p=0.030), and Processes and Procedures ($\hat{\beta}$=0.465; p=0.000) on Data Quality reveals their considerable statistical significance in a positive manner. Consequently, hypotheses **H2**, **H4**, and **H8** receive validation. However, no substantial associations are detected between Tools and Technologies ($\hat{\beta}$=0.031; p=0.630) as well as Roles and Responsibilities ($\hat{\beta}$=0.019; p=0.801) and Data Quality, refuting hypotheses **H6** and **H10**.

In addition to these findings, it is noteworthy that Data Quality significantly contributes to explaining value on two distinct levels: Business Value ($\hat{\beta}$=0.265; p=0.004) and Regulatory Compliance ($\hat{\beta}$= 0.796; p=0.000). This confirms the hypotheses **H11** and **H12**, signifying a substantial and affirmative influence.

The third step leverages the R^2 metric to assess the model's prediction quality. This metric quantifies the degree to which the combined influence of connected independent constructs accounts for the variance in dependent constructs, serving as an overall indicator of the model's explanatory power. While no universal guideline exists, scholars often consider R^2 values of 75%, 50%, and 25% as signifying substantial, moderate, and weak explanatory abilities. However, the Adjusted R^2 takes a more balanced approach by factoring in model complexity, penalising the inclusion of non-significant variables to the model's explanatory capability. This approach mitigates the risks of overfitting and provides a more impartial basis for comparing different models (Hair et al., 2021). Hence, in line with this study's goals, opting for the Adjusted R^2 metric was deemed appropriate.

Based on the Adjusted R^2 outcomes, the structural model accounts for 69.1% of the variability in Business Value, 75.9% in Data Quality, and 63.2% in Regulatory Compliance. This underscores the substantive influence of the Data Governance Dimensions on Data Quality, alongside their moderate impacts on Business Value and Regulatory Compliance.

After assessing the R^2 values associated with each dependent construct, it is also valuable to analyse the effect size f^2. This analytical measure quantifies the degree to which the exclusion of a specific independent construct affects the dependent constructs. Practical benchmarks for interpreting f^2 indicate that values of 0.02, 0.15, and 0.35 correspond to the independent variable's weak, moderate, and strong effects. Effect size values below 0.02 suggest minimal discernible effects (Hair et al., 2021).

The f^2 values corresponding to each structural path are outlined in Table 7. The analysis indicates that the constructs Roles and Responsibilities and Tools and Technologies have limited impact on Data Quality. Similarly, the relationship between Processes and Procedures and Tools and Technologies regarding their influence on Business Value is not substantial. This aligns with expectations, as the path coefficients are small, and the associated p-values are high, indicating a weak relationship without statistical significance. In contrast, the remaining paths demonstrate f^2 coefficients exceeding 0.02, suggesting that the independent constructs significantly contribute to explaining the variability observed in the dependent constructs. Noteworthy instances include the influence of Process and

Procedures on Data Quality, the impact of Roles and Responsibilities on Business Value, and the effect of Data Quality on Regulatory Compliance. These instances showcase moderate to strong effects.

To gain a deeper understanding of the model's predictive capabilities and complement the insights provided by the R^2 values, the final step of the analysis focuses on assessing the Stone-Geisser's Q^2 value. This metric is an indicator of predictive relevance, specifically assessing the model's ability to accurately forecast outcomes for data points not part of the initial model estimation process. Q^2 values greater than 0 for a given dependent construct indicate that the path model exhibits predictive relevance. This underscores the model's competence in making accurate predictions for data beyond its original training scope (Hair et al., 2021).

Presented in Figure 10, the Q^2 values associated with the three dependent constructs are considerably above zero. More precisely, Data Quality stands out with the highest Q^2 value (0.745), succeeded by Regulatory Compliance (0.706), and ultimately, Business Value (0.660). These results unequivocally validate the predictive relevance of the model regarding the dependent variables.

An outline of the findings from the Structural model can be observed in Table 7. To summarise, among the twelve formulated hypotheses, eight received empirical support based on the collected data, while four displayed no statistically significant association with the explanation of business outcomes. Notably, one unexpected observation emerged from the hypotheses that lacked support: the absence of a significant impact of Tools and Technologies on both Business Value and Data Quality.

Table 7 – Summary of Significance Testing Results for Structural Model

Structural Paths	**Path Coefficients**	***p*-value**	**Effect size (f^2)**	**95% Confidence Interval**	**Conclusion**
DC → BV	0.242	0.008	0.057	[0.114; 0.368]	H1 supported
DC → DQ	0.312	0.000	0.135	[0.181; 0.424]	H2 supported
DD → BV	0.238	0.026	0.068	[0.092; 0.363]	H3 supported
DD → DQ	0.166	0.030	0.043	[0.054; 0.306]	H4 supported
TT → BV	0.050	0.624	0.018	[-0.125; 0.159]	H5 not supported
TT → DQ	0.031	0.618	0.009	[-0.071; 0.130]	H6 not supported
PP → BV	0.060	0.557	0.008	[-0.094; 0.243]	H7 not supported
PP → DQ	0.465	0.000	0.247	[0.304; 0.583]	H8 supported
RR → BV	0.231	0.007	0.210	[0.026; 0.244]	H9 supported
RR → DQ	0.019	0.825	0.008	[-0.104; 0.142]	H10 not supported
DQ → BV	0.265	0.004	0.055	[0.117; 0.425]	H11 supported
DQ → RC	0.796	0.000	1.795	[0.732; 0.850]	H12 supported

5.3.2. Contextual Factors

Understanding the relationship between data governance and its organisational impacts is a complex challenge. While the core constructs shed light on the direct connections between data governance dimensions and their subsequent impacts on organisations, it is essential to recognise that the context in which these relationships operate plays a pivotal role. Contextual factors act as influential variables that can either amplify or attenuate the observed connections. As such, this section explores these contextual factors, which include firm size, industry, country development, and DG adoption, and their influence on the structural model.

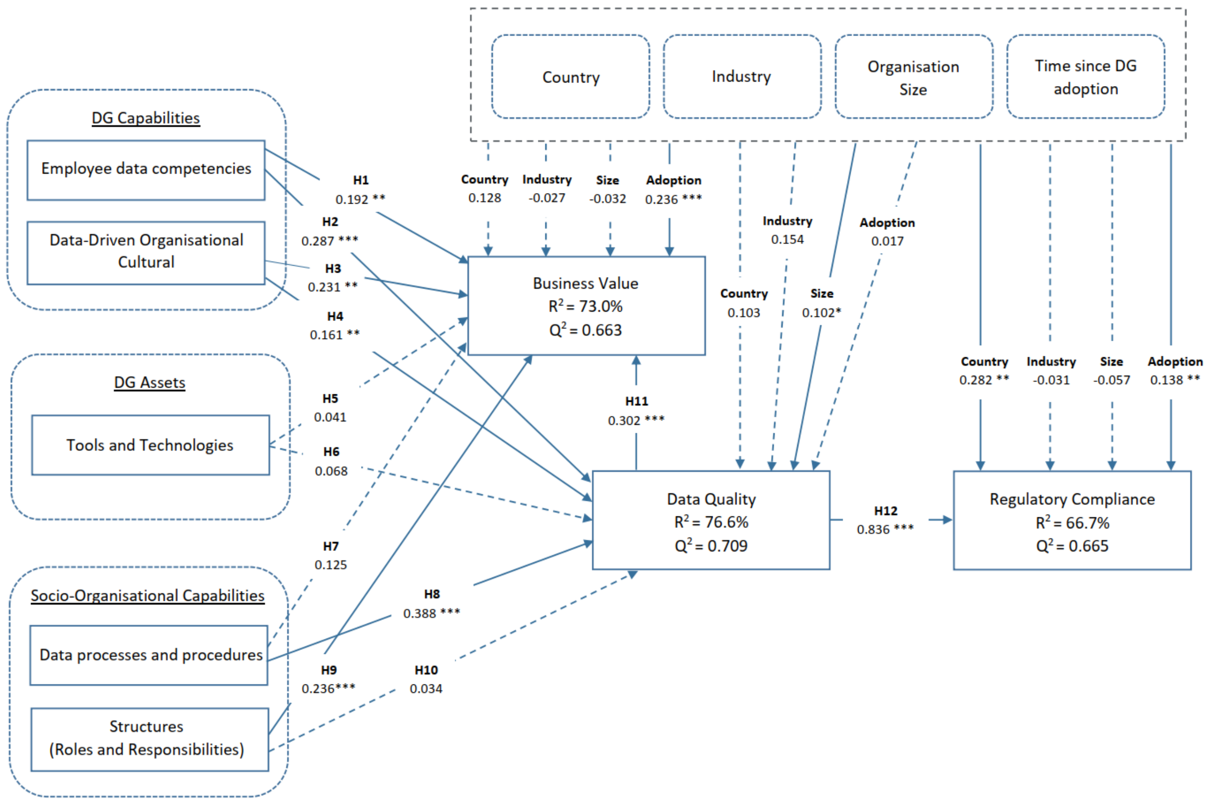
The initial phase involves evaluating how the introduction of contextual variables affects the remaining constructs. Following the statistical assessment of the reflective measurement model outlined in Section 5.2, an examination of the indicator reliability, construct reliability, convergent validity, and discriminant validity was conducted. Consistent with the guidelines in the preceding section, metrics for all four dimensions were observed to adhere to predefined thresholds, with the results being available upon request. This conclusion implies that exists compelling evidence of indicator reliability, construct reliability, convergent validity, and discriminant validity, even when considering contextual factors.

The second phase focuses on assessing the structural model. Following the procedures outlined in the previous subsection, the first step involves employing the VIF measure to examine the presence of collinearity. As with the evaluation of the reflective model, the outcomes are available upon request to the author.

The results indicate that all VIF values are below 5, signifying the absence of collinearity concerns within the structural paths. The dual confirmation of both the assessment of the reflective measurement model and VIF values underscores the robustness of the model, even in the presence of contextual factors. This result substantiates that the inclusion of these contextual factors does not compromise the reliability and validity between the constructs. Furthermore, the VIF values, which closely approximate 1 for the structural paths, indicate low levels of collinearity, thereby reinforcing the overall model's stability.

The next step involves the exploration of the impacts arising from data governance, particularly when considered alongside contextual factors. To do so, this analysis entails an examination of path coefficients and their associated *p*-values, encompassing both the direction of relationships and their statistical significance.

These outcomes are depicted in Figure 11, demonstrating the effects of incorporating contextual variables on the structural model.



Note: --- p≥0.1 (non-significant); *p<0.10; **p<0.05; ***p<0.01

Figure 11 – Structure Model Results with Contextual Factors

Together with data governance dimensions, the empirical results reveal the roles played by contextual factors in influencing organisational impacts. With the exception of the Industry variable, all other contextual factors exhibit discernible effects. In terms of Business Value, the analysis reveals a statistically significant relationship with the temporal dimension, regarding, the adoption of data governance ($\hat{\beta}$=0.236; *p*=0.000). In Data Quality, the model demonstrates a statistically significant relationship with organisational size ($\hat{\beta}$=0.102; *p*=0.045). Lastly, for Regulatory Compliance, this construct is positively influenced by two contextual factors: country development ($\hat{\beta}$=0.282; *p*=0.025) and data governance adoption ($\hat{\beta}$=0.138; *p*=0.015).

In addition, the introduction of contextual factors, along with the data governance dimensions, has increased the adjusted R^2 metric, contributing to an improved explanatory capacity of the model concerning observed variances in organisational impacts. When comparing the structural model with and without the inclusion of contextual factors, the impact on Business Value and Regulatory Compliance, exhibits a pronounced increase of 3.9% and 3.5% when contextual factors are considered. Conversely, the effect on Data Quality, while still discernible, appears reduced, with an increase of 0.7% when contextual factors are considered. This analysis reveals that contextual factors, while not pivotal in shaping the organisational impacts of data governance, exhibit some implications for Business Value and Regulatory Compliance. When it comes to Data Quality, while contextual factors do exert some influence, their role is notably less prominent in comparison.

The results presented in this section are summarised in Table 8.

Table 8 – Summary of Contextual Factors Findings

Structural Paths	**Path Coefficients**	***p*-value**	**95% Confidence Interval**	**Conclusion**
Country → BV	0.128	0.312	[-0.074; 0.337]	Not Statically Significant
Country → DQ	0.103	0.380	[-0.087; 0.302]	Not Statically Significant
Country → RC	0.282	0.025	[0.082; 0.492]	Statically Significant
Industry → BV	-0.027	0.854	[-0.285; 0.200]	Not Statically Significant
Industry → DQ	0.154	0.437	[-0.124; 0.457]	Not Statically Significant
Industry → RC	-0.031	0.833	[-0.313; 0.154]	Not Statically Significant
Firm Size → BV	-0.032	0.564	[-0.117; 0.069]	Not Statically Significant
Firm Size → DQ	0.102	0.045	[0.010; 0.164]	Statically Significant
Firm Size → RC	-0.057	0.337	[-0.154; 0.055]	Not Statically Significant
DG Adoption → BV	0.236	0.000	[0.151; 0.326]	Statically Significant
DG Adoption → DQ	0.017	0.843	[-0.069; 0.099]	Not Statically Significant
DG Adoption → RC	0.138	0.015	[0.041; 0.225]	Statically Significant

6. DISCUSSION

6.1. MAIN FINDINGS: DATA GOVERNANCE IMPACTS

Given the ambiguous understanding of Data Governance's scope and limited awareness of its organisational implications, it is essential to comprehend its diverse dimensions and complete value-chain. To address this, the proposed theoretical model uniquely combines Schryen's (2013) conceptualisation of IS business value with the organisational outcomes presented by Martijn et al. (2015) and Melville et al. (2004). Through this approach, the study assesses the value creation process through Data Governance, from its adoption to the consequent impacts on organisations.

The results strongly support Data Governance's pivotal role in fostering effective value creation, as evidenced by its significant impact on crucial organisational outcomes. The proposed model demonstrates that distinct Data Governance dimensions—Data-Driven Culture, Data Competencies, Roles and Responsibilities, and Processes and Practices—positively and significantly contribute to Data Quality (75.9%), Business Value (67.9%), and Regulatory Compliance (63.2%). Notably, the impact of Tools and Technologies on these outcomes was deemed statistically insignificant, challenging prior research.

In this section, the results of each Data Governance dimension are described and examined in accordance with both the prevailing literature and practical contexts.

6.1.1. Data Competencies

The results establish a significant positive relationship between the adoption of Employee Data Competencies and the creation of Business Value within organisations. The positive path coefficient indicates that an improvement in employee data competencies corresponds to an observable increase in business value. This indicates that organisations, by cultivating data skills through integrated training within Data Governance programs, experience improved operational efficiency and enhanced strategic planning, impacting overall business value.

The analysis also indicated a significant positive relationship between Employee Data Competencies and Data Quality. The statistical link emphasises that employees with adept data skills significantly contribute to maintaining accurate, reliable, and consistent data assets. This finding is particularly vital in the context of data-driven decision-making, demonstrating a virtuous cycle where the continuous refinement of employee skills translates into better data quality, culminating in high-quality reports, which is the foundation upon meaningful insights are built.

The direct connection between Employee Data Competencies and Business Value, as well as Data Quality is theoretically grounded in the Knowledge-Based View (KBV) theory. Previous studies (Côrte-Real et al., 2017; Nieves & Haller, 2014) employing KBV, explain the knowledge's role in driving performance effects. The studies indicate that an organisation's unique knowledge and intellectual assets profoundly influence its success, wherein knowledgeable employees play a crucial role in adapting to changing environments and contributing to successful implementations .

In addition, scholars, through a questionnaire-based approach (Ongena, G., 2023), and interviews in conjunction with a systematic literature review (Klee et al., 2021), further establish the direct and positive impact of data competencies on internal performance. The latter study even deems data competencies as a prerequisite for business value. This substantiates the imperative nature of data competencies in realising value and ensuring the efficacy of data analytics investments. Consequently, this study, alongside existing literature, lend support to the integration of data-related skill development as a core component of Organisational Data Governance strategies.

6.1.2. Data-Driven Culture

The findings reveal a significant positive connection between adopting a Data-Driven Culture and Business Value in the context of Data Governance. This supports the hypothesis that organisations embracing a culture of data-driven decision-making are more likely to experience enhanced business value outcomes. Aligned with the theoretical model, this suggests that organisations embracing data-driven practices excel in informed decision-making, process optimisation, and identifying growth opportunities, thereby elevating overall business value.

Furthermore, the analysis shows a significant direct relationship between the adoption of a Data-Driven Culture and improvements in Data Quality. This validates the initial hypothesis and emphasises the crucial role of organisational culture in a Data Governance program, shaping the effectiveness of data quality initiatives. As Organisations shift their culture towards valuing data-driven decision-making, the attention to data accuracy, completeness, and reliability intensifies, leading to enhanced data quality outcomes.

Previous research (Chaudhuri et al., 2021) stresses the necessity for organisational leaders to foster a data-driven culture. Their findings indicate that such culture enhances employee insights and raises stakeholders' attention to data quality, leading to improved business outcomes and overall organisational performance. Similarly, Yu et al. (2021) examined the role of data-driven culture in optimising the benefits from big data analytics, emphasising that it strengthens an organisation's capacity to integrate data into various business processes, fully unlocking the value of big data. These impacts are consistent with existing literature, reinforcing the importance of incorporating a Data-Driven Culture into the organisational Data Governance program.

6.1.3. Tools and Technologies

Contrary to the initial hypothesis, the analysis found no statistically significant relationship between adopting Tools and Technologies and consequent Business Value. Similarly, the connection between integrating Tools and Technologies and Data Quality did not demonstrate statistical significance. These results challenge the conventional notion that the integration of advanced tools within a Data Governance program inherently enhances data quality and yields positive effects on business value. Therefore, the findings suggest that the direct impact of Technology on Business Value and Data quality might be more complex than initially assumed.

Multiple studies (Randhawa, 2019; Wagner et al., 2014) provide empirical evidence linking IT infrastructure to organisational outcomes. However, when explaining the causal relationship between technology and tangible value within the organisation, these studies also recognise the importance of aligning business and operational aspects with IT investments and their implementation.

Randhawa's (2019) qualitative research found that a robust IT infrastructure reduces data risks, enhances business data analysis, and improves operations. However, the key driver for realising these benefits lies in integrating robust IT infrastructure into the Organisational Data Governance Program. This integration is crucial for implementing consistent strategies across teams to assess the IT infrastructure's current state and address shortcomings. Through this, changes in data rules, application systems, and business processes become easily identifiable, ensuring accurate data and optimal IT performance.

From a quantitative perspective, Wagner et al. (2014) explored the role of Social capital in driving operational alignment and generating IT business value. The study identified that the exchange of knowledge and expertise between IT and business professionals, along with bridging the linguistic gap between technical and non-technical stakeholders, exerts a substantial impact on the business's grasp of IT. This, in turn, leads to improved business process performance.

The absence of statistical significance regarding IT organisational outcomes reinforces the complexity of Data Governance beyond a straightforward cause-and-effect relationship, whilst emphasising Data Governance's multifaceted nature. While tools and technologies undeniably play a role, this study asserts that technological investments alone are insufficient to drive tangible improvements in business outcomes and data quality. This aligns with the growing recognition that Data Governance involves more than just implementing technological solutions—it is about harmonising technology with existing processes, human expertise, and strategic objectives. Nevertheless, the lack of a direct link does not negate the potential influence of Tools and Technologies. The extent of alignment with organisational culture, processes, and data requirements impacts the effectiveness of technology adoption within a Data Governance Program.

6.1.4. Processes and Procedures

The results indicate a significant positive relationship between the adoption of Processes and Procedures and Data Quality. Organisations focusing on the structured implementation of DG processes and procedures witness tangible improvements in data quality metrics. Notable outcomes include increased data accuracy, completeness, consistency, and reliability. This substantiates the notion that Data Governance, when executed effectively, can act as a catalyst for enhancing data quality throughout the data lifecycle.

The study also identified an indirect effect between the adoption of Processes and Procedures in a Data Governance program and Business Value impacts. Despite the lack of a statistically significant direct connection, the presence of an indirect effect suggests that while DG processes and procedures might not directly influence business value, they can still contribute to value creation by enhancing data quality, which in turn positively affects business value. Improved data quality acts as a catalyst, enhancing the value generation process through accurate insights and informed decision-making. Earlier research has established the significance of defining and documenting processes and procedures as CSFs for Data Governance and overall work and operational contexts (Alhassan et al., 2019; Al-Ruithe et al., 2019). However, despite this acknowledgment, there has been limited exploration of the outcomes stemming from these efforts.

In a previous study by Lee & Strong (2003), the connection between understanding data production processes and organisational performance was examined. The research outcomes aligns with the present study, as the authors confirmed a relationship between knowledge and enhanced data quality. A more detailed analysis of data production processes showed that the knowledge possessed by distinct roles, from data collectors and managers to data consumers, exerts varying degrees of influence on data quality. These roles collaboratively engage within the data production process to identify and resolve issues, contributing to organisational data quality enhancement.

The indirect relationship between Processes and Procedures and Business Value may stem from a limited perspective on tangible organisational outcomes, overlooking intangible contributions. Data Governance Processes lead to tangible outcomes, such as data accuracy and reliability, establishing a solid foundation for intangible assets. These intangible outcomes, crucial for long-term success, are challenging to quantify, making it difficult to fully capture the impact of organisational processes on overall success.

Building upon Lee & Strong's (2003) findings, the integration of processes within an organisation proves to be a critical factor. The successful integration of processes is closely linked to the existence of roles that align data asset management with strategic goals/mission. Each role carries distinct responsibilities and accountabilities, fostering effective communication channels that transcend the silos of traditional departmental boundaries, promoting a culture of collaboration and cross-functional awareness. To study the intangible impacts of processes and their integration on organisations, these factors must be considered.

6.1.5. Roles and Responsibilities

The results demonstrated no significant connection between adopting Roles and Responsibilities in a Data Governance Program and a direct impact on Data Quality. This suggests that improving data quality involves multifaceted efforts that involve a confluence of factors, including data accuracy, completeness, consistency, and timeliness. While essential, merely designating roles and responsibilities might not encompass the full range of interventions required to address these diverse aspects. The lack of a direct connection implies that the impact of data stewardship on data quality requires a more intricate orchestration of processes, methodologies, and technological solutions.

Despite a non-significant relationship between Data Stewardship and Data Quality, the study identifies a statistically significant and direct connection between Roles and Responsibilities in a Data Governance program and the impacts on Business Value. This highlights the crucial role of data stewardship in driving tangible business outcomes within Data Governance programs.

Existing literature confirms Data Stewardship's pivotal role in Data Governance, guiding organisational processes related to data usage. Plotkin (2014) further elaborates on the central importance of Data Stewards in corporate initiatives, particularly their impact on data quality enhancement. The presence of individuals tasked with data stewardship responsibilities, combined with established procedures mandating their involvement in data-related decisions, ensures well-informed and collectively aligned decisions. The collaborative efforts of designated stewards, established processes, and a shared mission to manage data contribute to the improvement of data quality as a valuable organisational asset.

Accordingly, the findings reveal a discrepancy in relation to existing literature. Unlike Plotkin's conclusions of a direct link between data stewardship roles and data quality, this study shows a more intricate relationship. Instead, the primary connection appears to be between these roles and their impact on business value.

In line with Lee & Strong (2003) and Plotkin (2014), data stewardship roles do not operate in isolation, as they are part of data management processes. The formulation of roles is meaningful when they align with well-structured and efficient processes. Establishing roles without accompanying processes that facilitate data validation, reconciliation, and maintenance, may limit their direct influence on data quality outcomes. The impact of roles is contingent on how seamlessly they integrate with the larger data quality ecosystem. On the other hand, the findings show that the connection between roles and responsibilities and business value might be more direct, as it involves the translation of data efforts into concrete business outcomes.

6.1.6. Data Quality

The significant relationship between Data Quality and Business Value signifies a fundamental link between the quality of organisational data and its business value. This connection reveals that data, often considered an intangible asset, can become a valuable resource when enhanced, translating into tangible business value and competitive advantage. High-quality data acts as the lifeblood of informed decision-making, fostering a virtuous cycle where accurate, timely, and reliable information leads to strategic choices that optimise operational efficiency, enhance customer satisfaction, and drive innovation. As data quality improves, organisations are better equipped to navigate complex business landscapes, mitigate risks, and leverage emerging opportunities.

The analysis also demonstrated a significant direct connection between Data Quality and Regulatory Compliance, underscoring the critical role data quality plays in meeting industry-specific regulations and standards. As organisations invest in enhancing data quality, they proactively enhance their ability to fulfil regulatory obligations effectively. Specifically, higher data quality reduces the risk of errors/inconsistencies in compliance-related data, preventing regulatory violations, fines, and reputational damage. This research underscores that data quality initiatives are not just administrative tasks but strategic imperatives for sustaining a competitive advantage in regulated environments.

In addition, the analysis confirms an indirect relationship between the adoption of a Data Governance program and Regulatory Compliance through Data Quality. This underscores that compliance is not an isolated task but must be integrated into a broader Data Governance program that encompasses multiple dimensions, being natural outcome. Investing in Data Governance is not merely about improving data quality but also about fortifying the ability to meet regulatory obligations sustainably. Therefore, incorporating regulatory compliance in Data Governance provides a strategic advantage, as it shifts the approach from reactive, compliance-driven to proactive, data-driven approach.

Existing academic literature (Mahanti, 2019; Ji-fan Ren et al., 2017) has established the intricate relationship between Data Quality, Business Value, and Regulatory Compliance. These studies emphasise that data quality is not merely a technical concern but a strategic asset with significant implications for critical organisational outcomes. While prior research has explored these connections, this study goes a step further by empirically demonstrating how the elements incorporated within an organisation's Data Governance Program directly influence these outcomes. Additionally, it considers contextual factors when examining organisational results.

While contextual factors do exert some influence, their impact is relatively minimal when compared against the five dimensions analysed. Nevertheless, a noteworthy finding is the statistically significant direct connection between the recentness of Data Governance adoption and crucial organisational outcomes: Business Value and Regulatory Compliance. This underscores that Data Governance is an ongoing strategic journey, not a one-time project. As the program matures, it gradually becomes more adept at harnessing data as a strategic asset, contributing to the generation of greater business value. Additionally, it shows that regulatory compliance is integrated outcome of a Data Governance program, emphasising the need for sustained, well-structured, and evolving data governance initiatives.

6.2. THEORETICAL IMPLICATIONS

This section expands the theoretical contributions to the Data Governance literature. By addressing previously unexplored aspects in the existing body of knowledge, such as understanding DG's value creation process and operational mechanisms, it enhances the overall comprehension of Data Governance's function and its intricate relationship with organisational outcomes, providing valuable insights into the complexities of data-driven environments.

The Data Governance field has witnessed substantial growth in recent years, with numerous organisations recognising its potential benefits. However, some struggle to harness the full value of DG initiatives. Scholars have underscored the need for empirical assessments of the actual impact of DG investments and implementations, as well as the necessity to understand how to realize performance benefits from DG initiatives (Benfeldt et al., 2020; Abraham et al., 2019; Nielsen, 2017). Despite these calls for research, the DG value creation process remains relatively unexplored. This study addresses this gap by proposing and empirically validating a novel conceptual model. Drawing on strategic management theories, specifically the IS and IT conceptualisation of business value frameworks, this unique combination has not been previously applied to DG research.

This study contributes to DG literature by shedding light on the full DG value-chain, outlining key components and relationships involved in value creation through DG practices. By identifying five critical data management dimensions, the study empirically demonstrates that integrating these into an organisational DG program significantly enhances data quality, ultimately generating business value and ensuring regulatory compliance. This contribution carries profound implications for organisations seeking to harness the full potential of their data assets whilst navigating the intricate landscape of data management, quality assurance, and regulatory compliance.

Furthermore, the study introduces a theoretical framework with potential applicability beyond DG, offering a structured approach for organisations to assess the business value of various IT innovations at a process level. This framework can be applied to assess the value-chain of technology implementations, enhancing the theoretical foundations of DG and contributing to the broader field of IT innovation assessment and management.

In conclusion, this research advances the theoretical understanding of Data Governance, its value creation process, and its strategic importance in technology and organisational management. By exploring the DG value-chain, empirically validating the conceptual model, and establishing links between DG Dimensions, Data Quality, Business Value, and Regulatory Compliance, the study reinforces DG's pivotal role in contemporary organisations. Moreover, the introduced theoretical framework provides a valuable tool for assessing the business value of IT innovations, extending its applicability beyond the DG domain. These contributions enrich the scholarly research on Data Governance and its implications for organisational performance and competitiveness.

6.3. PRACTICAL IMPLICATIONS

For practitioners this study serves as a strategic roadmap for effectively harnessing the knowledge embedded in Data Governance systems and initiatives. It illustrates how organisations can develop capabilities that maintain and enhance their competitive edge. Specifically, this research provides compelling support to justify investments and initiatives related to Data Governance.

Despite the considerable investment required for DG technologies and practices, organisations operating in European and other contexts are increasingly recognising the potential value. The results indicate a growing awareness among executives, with most organisations reporting an operational DG program, highlighting the relevance of integrating DG into organisational strategies. By applying the guidelines derived from this research, practitioners can optimise their approach to DG, ensuring the alignment with strategic objectives and organisational needs.

Additionally, this study serves as a guide for practitioners and organisations, providing insights into the various stages of value creation through Data Governance. The core of DG value creation lies in establishing a thorough DG program, comprising five critical dimensions: Data Culture, Data Competencies, Technologies and Tools, Data Roles, and Data Processes. The integration of the DG program with critical organisational aspects ensures tangible benefits in terms of continuous data quality, business value, and alignment with regulatory standards. The findings provide practical tools for practitioners to formulate performance metrics, serving as quantifiable indicators that enable a consistent assessment of DG impact, and support a cycle of continuous improvement aligned with evolving organisational needs and industry standards

This research holds broad implications for various stakeholders in the Data Governance domain. For practitioners, the framework offers a strategic guide for measuring the DG's impact on knowledge conversion, process optimisation, and firm-level capabilities. By incorporating these guidelines, practitioners can enhance productivity, evaluate DG effectiveness, and align strategies with organisational objectives. For organisations contemplating DG adoption, this research provides a tangible perspective on the possibilities and benefits of DG implementation, offering insights into the transformative potential of DG in enhancing data assets and navigating the complexities of data management, quality assurance, and regulatory compliance.

DG Professionals, including vendors, also gain by understanding how organisations invest in and derive value from Data Governance. This empowers vendors to customize solutions, seamlessly integrating DG capabilities. This facilitates customers in attaining enhanced financial and strategic performance, boosting the vendors' market competitiveness.

7. CONCLUSIONS AND FUTURE WORK

7.1. CONCLUSIONS

This study aimed to explore the impact of data governance on business value, offering organisations a framework to evaluate the efficacy of their governance models and data practices. The research was developed within the context of an internship program at a service provider firm, adhering to the principles of natural science methodology.

Accordingly, a conceptual model was built, integrating elements from Schryen's (2013) IS business value conceptualisation, the organisational outcomes presented by Martijn et al. (2015) and Melville et al. (2004), and contextual factors. The model posited that a Data Governance Program, characterised by five dimensions (Data-Driven Culture, Data Competencies, Tools and Technologies, Data Processes, and Data Stewardship), has a measurable impact on organisational impacts, specifically, Data Quality, Business Value, and Regulatory Compliance.

To validate the conceptual model, a questionnaire was developed, encompassing all constructs within the model. Steps were taken to ensure the reliability and validity of the instrument, including input from experienced practitioners in the field of data governance. Over a two-month period, quantitative data was collected, achieving a total 196 responses. Checks were implemented to detect biases, such as non-response and CMB, revealing no notable influence on the results trustworthiness.

For data analysis, the PLS-SEM technique was employed due to its adaptability to studies with limited samples and complex models, as well as its flexibility in handling non-normal data distributions. The analysis commenced with an examination of the reflective measurement model. This encompassed evaluating indicator and construct reliability, convergent validity, and discriminant validity. Subsequently, the structural model underwent testing to validate the hypotheses formulated in the conceptual framework.

The findings confirmed several hypotheses, establishing that Data-Driven Culture and Data Competencies significantly impact both Data Quality and Business Value, Data Processes impact Data Quality, and Data Stewardship impacts Business Value. However, four hypotheses were not confirmed, indicating the non-significant impact of Tools and Technologies on Data Quality and Business Value, the lack of impact of Data Processes on Business Value, and the absence of influence from Data Stewardship on Data Quality.

Another outcome of the analysis was the identification of contextual factors influencing the organisational outcomes of a Data Governance Program. Specifically, organisational size, the country development in which an organisation operates, and the data governance program's maturity were significant contextual influencers.

In terms of implications, this research significantly advances the theoretical foundations of Data Governance by exploring the value creation process within DG and elucidating operational mechanisms, addressing existing research gaps in the literature. For practitioners, this study serves as a strategic roadmap, providing practical tools and guidelines to optimise DG initiatives, aligning them with organisational strategies and ensuring continuous improvement. The identified critical dimensions in the DG value-chain offer quantifiable indicators for practitioners to assess the impact of

DG initiatives and navigate the complexities of data management, quality assurance, and regulatory compliance.

7.2. LIMITATIONS

While the study contributes to understanding the relationships between data governance practices and organisational outcomes, it is crucial to recognise the limitations inherent in the research design, methodology, and data collection process.

The study focus primarily on the intended outcomes of data governance practices. However, the research design may not fully capture unintended consequences or side effects that emerge from the implementation of data governance initiatives. As such, the exploration of potential unintended consequences to provide a more comprehensive view of the broader organisational impacts.

The study focused on conventional data governance practices and did not extensively explore the implications of emerging technologies such as artificial intelligence on data governance outcomes. The rapidly evolving technological landscape necessitates ongoing examination of how cutting-edge technologies intersect with data governance.

Despite efforts to incorporate a diverse set of respondents, a majority of responses originated from European countries. Representation from other regions, notably Africa and South America, was limited. This geographical bias introduces a potential regional bias, impacting the generalisability of findings to a global context.

7.3. FUTURE WORK

Regarding future research initiatives, two potential areas have been identified to enhance our comprehension of the intricate connection between data governance practices and organisational outcomes.

The first area revolves around the absence of a direct link between the integration of Tools and Technologies in a Data Governance Program and organisational outcomes, specifically Data Quality and Business Value. This finding suggests that, while instrumental, technology does not operate in isolation and is not enough to ensure the success of a DG Program. To delve deeper into this complex interplay, future research should analyse the specific mechanisms through which tools and technologies contribute or hinder the achievement of organisational goals in the context of data governance. Understanding the specific pathways through which technologies contribute to data quality improvement and business value will provide organisations with actionable insights to optimise their technology investments. A similar analytical approach can be applied to further explore the connection between Data Stewardship and Data Quality, as well as Data Processes and Business Value.

The second area pertains to a deeper exploration of the contextual factors role in shaping the effectiveness of a DG Program. Therefore, expanding the geographical representation of participants is critical for achieving a more comprehensive understanding of global variations in data governance practices. Future research should strive for a more balanced representation across regions and continents, ensuring that insights are not overly influenced by specific cultural or regional aspects. By engaging organisations from diverse global contexts, researchers can uncover unique challenges and opportunities that transcend localised boundaries, contributing to a more universal body of knowledge.

Furthermore, exploring the interplay between organisational culture and data governance, and the dynamic adaptation of governance models over time, presents promising avenues for future research. The continuous evolution of technology and regulatory landscapes necessitates ongoing exploration to ensure the relevance and applicability of data governance practices in a rapidly changing environment.

BIBLIOGRAPHICAL REFERENCES

Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. In International Journal of Information Management (Vol. 49). https://doi.org/10.1016/j.ijinfomgt.2019.07.008

Alhassan, I., Sammon, D., & Daly, M. (2018). Data governance activities: a comparison between scientific and practice-oriented literature. In Journal of Enterprise Information Management (Vol. 31, Issue 2). https://doi.org/10.1108/JEIM-01-2017-0007

Alhassan, I., Sammon, D., & Daly, M. (2019). Critical Success Factors for Data Governance: A Theory Building Approach. Information Systems Management, 36(2). https://doi.org/10.1080/10580530.2019.1589670

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). https://doi.org/10.1007/s00779-017-1104-3

Anant, V., Donchak, L., Kaplan, J., & Soller, H. (2020). The consumer-data opportunity and the privacy imperative. McKinsey & Company, April.

Anderson, J. C., & Gerbing, D. W. (1988). Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. Psychological Bulletin, 103(3). https://doi.org/10.1037/0033-2909.103.3.411

Attard, J., & Brennan, R. (2018). Challenges in value-driven data governance. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11230 LNCS. https://doi.org/10.1007/978-3-030-02671-4_33

Aydiner, A. S., Tatoglu, E., Bayraktar, E., Zaim, S., & Delen, D. (2019). Business analytics and firm performance: The mediating role of business process performance. Journal of Business Research, 96. https://doi.org/10.1016/j.jbusres.2018.11.028

Bany Mohammad, A., Al-Okaily, M., Al-Majali, M., & Masa'deh, R. (2022). Business Intelligence and Analytics (BIA) Usage in the Banking Industry Sector: An Application of the TOE Framework. Journal of Open Innovation: Technology, Market, and Complexity, 8(4), 189. https://doi.org/10.3390/JOITMC8040189

Baptista, G., & Oliveira, T. (2015). Understanding mobile banking: The unified theory of acceptance and use of technology combined with cultural moderators. Computers in Human Behavior, 50. https://doi.org/10.1016/j.chb.2015.04.024

Barua, A., Kriebel, C. H., & Mukhopadhyay, T. (1995). Information technologies and business value: An analytic and empirical investigation. Information Systems Research, 6(1). https://doi.org/10.1287/isre.6.1.3

Barker, J. M. (2016). Data governance: The missing approach to improving data quality. ProQuest Dissertations and Theses.

Batini, C., Cappiello, C., Francalanci, C., & Maurino, A. (2009). Methodologies for data quality assessment and improvement. ACM Computing Surveys, 41(3). https://doi.org/10.1145/1541880.1541883

Begg, C., & Caira, T. (2012). Exploring the SME Quandary : Data Governance in Practise in the Small to Medium-Sized Enterprise Sector. The Electronic Journal Information Systems Evaluation, 15(1).

Benfeldt, O., Persson, J. S., & Madsen, S. (2020). Data Governance as a Collective Action Problem. Information Systems Frontiers, 22(2). https://doi.org/10.1007/s10796-019-09923-z

Benitez, J., Henseler, J., Castillo, A., & Schuberth, F. (2020). How to perform and report an impactful analysis using partial least squares: Guidelines for confirmatory and explanatory IS research. Information and Management, 57(2). https://doi.org/10.1016/j.im.2019.05.003

Berntsson Svensson, R., & Taghavianfar, M. (2020). Toward Becoming a Data-Driven Organisation: Challenges and Benefits. Lecture Notes in Business Information Processing, 385 LNBIP. https://doi.org/10.1007/978-3-030-50316-1_1

Bhansali, N. (2014). *Data Governance: Creating Value from Information Assets*. CRC Press.

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. https://doi.org/10.1007/978-3-030-44999-5_36

Cai, L., & Zhu, Y. (2015). The challenges of data quality and data quality assessment in the big data era. Data Science Journal, 14. https://doi.org/10.5334/dsj-2015-002

Cave, A. (2017). Exploring Strategies for Implementing Data Governance Practices (Walden University). Retrieved from https://scholarworks.waldenu.edu/dissertations

Cao, G., & Duan, Y. (2014). A path model linking business analytics, data-driven culture, and competitive advantage. ECIS 2014 Proceedings - 22nd European Conference on Information Systems.

Chaudhuri, R., Chatterjee, S., Vrontis, D., & Thrassou, A. (2021). Adoption of robust business analytics for product innovation and organisational performance: the mediating role of organisational data-driven culture. Annals of Operations Research. https://doi.org/10.1007/s10479-021-04407-3

Chen, L., & Mohamed, S. (2008). Contribution of knowledge management activities to organisational business performance. Journal of Engineering, Design and Technology, 6(3). https://doi.org/10.1108/17260530810918289

Chen, Y., Wang, Y., Nevo, S., Jin, J., Wang, L., & Chow, W. S. (2014). IT capability and organisational performance: The roles of business process agility and environmental factors. European Journal of Information Systems, 23(3). https://doi.org/10.1057/ejis.2013.4

Cheong, L. K., & Chang, V. (2007). The need for data governance: A case study. ACIS 2007 Proceedings - 18th Australasian Conference on Information Systems.

Cohen, J. (1992a). A power primer. Psychological Bulletin, 112(1), 155–159.

Cohen, J. (1992b). Statistical Power Analysis. Current Directions in Psychological Science, 1(3). https://doi.org/10.1111/1467-8721.ep10768783

Cohen, J. (2013). Statistical Power Analysis for the Behavioral Sciences. In Statistical Power Analysis for the Behavioral Sciences. https://doi.org/10.4324/9780203771587

Côrte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of Big Data Analytics in European firms. Journal of Business Research, 70. https://doi.org/10.1016/j.jbusres.2016.08.011

DalleMule, L., & Davenport, T. H. (2017). What's your data strategy? In Harvard Business Review (Vol. 2017, Issue May-June).

Dai, W., Wardlaw, I., Cui, Y., Mehdi, K., Li, Y., & Long, J. (2016). Data profiling technology of data governance regarding big data: Review and rethinking. Advances in Intelligent Systems and Computing, 448. https://doi.org/10.1007/978-3-319-32467-8_39

DAMA International. (2017). DAMA-DMBOK DATA MANAGEMENT BODY OF KNOWLEDGE (2nd ed.). Technics Publications.

Damanpour, F. (1996). Organisational complexity and innovation: Developing and testing multiple contingency models. Management Science, 42(5). https://doi.org/10.1287/mnsc.42.5.693

Dommeyer, C. J., & Gross, B. L. (2003). What consumers know and what they do: An investigation of consumer knowledge, awareness, and use of privacy proteciton strategies. Journal of Interactive Marketing, 17(2). https://doi.org/10.1002/dir.10053

Elbashir, M. Z., Collier, P. A., Sutton, S. G., Davern, M. J., & Leech, S. A. (2013). Enhancing the business value of business intelligence: The role of shared knowledge and assimilation. Journal of Information Systems, 27(2). https://doi.org/10.2308/isys-50563

Fink, L., & Neumann, S. (2009). Exploring the perceived business value of the flexibility enabled by information technology infrastructure. Information and Management, 46(2). https://doi.org/10.1016/j.im.2008.11.007

Freeze, R. D., & Raschke, R. L. L. (2007). An assessment of formative and reflective constructs in is research. Proceedings of the 15th European Conference on Information Systems, ECIS 2007.

Gregor, S., Martin, M., Fernandez, W., Stern, S., & Vitale, M. (2006). The transformational dimension in the realisation of business value from information technology. Journal of Strategic Information Systems, 15(3). https://doi.org/10.1016/j.jsis.2006.04.001

Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realising value from big data. Journal of Strategic Information Systems, 26(3). https://doi.org/10.1016/j.jsis.2017.07.003

Hair, Hult G., Ringle, C. M., & Sarstedt, M. (2021). A primer on partial least squares structural equation modeling (PLS-SEM) Second Edition. In SAGE Publications Inc.

Hüner, K. M., Otto, B., & Österle, H. (2011). Collaborative management of business metadata. International Journal of Information Management, 31(4). https://doi.org/10.1016/j.ijinfomgt.2010.12.002

IBM Institute. (2007). The IBM Data Governance Council Maturity Model : Building a Roadmap for Effective Data Governance. Governance An International Journal Of Policy And Administration, October.

Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organising data for trustworthy Artificial Intelligence. Government Information Quarterly, 37(3). https://doi.org/10.1016/j.giq.2020.101493

Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., & Childe, S. J. (2017). Modelling quality dynamics, business value and firm performance in a big data analytics environment. International Journal of Production Research, 55(17). https://doi.org/10.1080/00207543.2016.1154209

Khatri, V., & Brown, C. V. (2010). Designing Data Governance. Communications of the ACM, 53(1), 148–152. Retrieved from https://dl.acm.org/doi/pdf/10.1145/1629175.1629210

Kim, H. Y., & Cho, J. S. (2018). Data governance framework for big data implementation with NPS Case Analysis in Korea. Journal of Business and Retail Management Research, 12(3). https://doi.org/10.24052/jbrmr/v12is03/art-04

Kim, S. S., & Kim, Y. J. (2017). The effect of compliance knowledge and compliance support systems on information security compliance behavior. Journal of Knowledge Management, 21(4). https://doi.org/10.1108/JKM-08-2016-0353

Klee, S., Janson, A., & Leimeister, J. M. (2021). How Data Analytics Competencies Can Foster Business Value– A Systematic Review and Way Forward. Information Systems Management, 38(3). https://doi.org/10.1080/10580530.2021.1894515

Kock, N., & Hadaya, P. (2018). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. Information Systems Journal, 28(1). https://doi.org/10.1111/isj.12131

Koltay, T. (2016). Data governance, data literacy and the management of data quality. IFLA Journal, 42(4). https://doi.org/10.1177/0340035216672238

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

Ladley, J. (2020). Data Governance: How to Design, Deploy, and Sustain an Effective Data Governance Program (2nd ed.). Academic Press.

Laudon, K. C., & Laudon, J. P. (2019). Management Information Systems Managing The Digital Firm sixteenth edition. In Pearson.

Lee, Y. W., & Strong, D. M. (2003). Knowing-Why About Data Processes and Data Quality. In Journal of Management Information Systems (Vol. 20, Issue 3). https://doi.org/10.1080/07421222.2003.11045775

Mahanti, R. (2019). Data Quality: Dimensions, Measurement, Strategy, Management, and Governance. Quality Press

Mahanti, R. (2021). Data Governance and Compliance. In Data Governance and Compliance. https://doi.org/10.1007/978-981-33-6877-4

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. Decision Support Systems, 15(4). https://doi.org/10.1016/0167-9236(94)00041-2

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. https://doi.org/10.1007/978-3-319-25013-7_37

Martins, C., Oliveira, T., & Popovič, A. (2014). Understanding the internet banking adoption: A unified theory of acceptance and use of technology and perceived risk application. International Journal of Information Management, 34(1). https://doi.org/10.1016/j.ijinfomgt.2013.06.002

Masa'Deh, R., Obeidat, Z., Maqableh, M., & Shah, M. (2021). The Impact Of Business Intelligence Systems on an Organisation's Effectiveness: The Role of Metadata Quality From a Developing Country'S View. International Journal of Hospitality and Tourism Administration, 22(1). https://doi.org/10.1080/15256480.2018.1547239

McGilvray, D. (2021). Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information (TM). In Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information (TM). https://doi.org/10.1016/B978-0-12-818015-0.09991-6

Melville, N., Kraemer, K., & Gurbaxani, V. (2004). Review: Information technology and organisational performance: An integrative model of it business value. In MIS Quarterly: Management Information Systems (Vol. 28, Issue 2). https://doi.org/10.2307/25148636

Mikalef, P., & Krogstie, J. (2018). Big data governance and dynamic capabilities: The moderating effect of environmental uncertainty. Proceedings of the 22nd Pacific Asia Conference on Information Systems - Opportunities and Challenges for the Digitised Society: Are We Ready?, PACIS 2018.

Mikalef, P., Krogstie, J., van de Wetering, R., Pappas, I. O., & Giannakos, M. N. (2018). Information governance in the big data era: Aligning organisational capabilities. Proceedings of the Annual Hawaii International Conference on System Sciences, 2018-January. https://doi.org/10.24251/hicss.2018.615

Moody, D., & Walsh, P. (1999). Measuring The Value Of Information: An Asset Valuation Approach. Seventh European Conference on Information Systems (ECIS'99).

Monks, R., & Minow, N. (2011). *Corporate governance* (5th ed.). Chichester, England: John Wiley and Sons Ltd.

Ndamase, Z. (2014). The impact of data governance on corporate performance : the case of a petroleum company. (Study). University of Cape Town ,Faculty of Commerce ,Department of Information Systems. Retrieved from http://hdl.handle.net/11427/13323

Nielsen, O. B., Persson, J. S., & Madsen, S. (2019). Why governing data is difficult: Findings from Danish local government. IFIP Advances in Information and Communication Technology, 533. https://doi.org/10.1007/978-3-030-04315-5_2

Nielsen, O. B. (2017). A Comprehensive Review of Data Governance Literature. Selected Papers of the IRIS, (8). Retrieved from http://aisel.aisnet.org/iris2017/3

Niemi, E. (2013). Working Paper : Designing a Data Governance Framework. In Proceedings of the IRIS Conference (Issue August 2013).

Nieves, J., & Haller, S. (2014). Building dynamic capabilities through knowledge resources. Tourism Management, 40. https://doi.org/10.1016/j.tourman.2013.06.010

Ongena, G. (2023). Data literacy for improving governmental performance: A competence-based approach and multidimensional operationalisation. *Digital Business*, *Volume 3*(Issue 1). https://doi.org/10.1016/j.digbus.2022.100050

Otto, B. (2011a). Data governance. Business and Information Systems Engineering, 3(4), 241–244. https://doi.org/10.1007/s12599-011-0162-8

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). https://doi.org/10.17705/1cais.02903

Panian, Z. (2010). Some practical experiences in data governance. World Academy of Science, Engineering and Technology, 38.

Plotkin, D. (2014). Data stewardship : an actionable guide to effective data management and data governance (1st edition). Elsevier.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. In Journal of Applied Psychology (Vol. 88, Issue 5). https://doi.org/10.1037/0021-9010.88.5.879

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. Acta Psychologica, 104(1). https://doi.org/10.1016/S0001-6918(99)00050-5

Radcliffe, J. (2009). The Seven Building Blocks of MDM : A Framework for Success. Gartner, 22(October).

Randhawa, T. S. (2019). *Incorporating Data Governance Frameworks in the Financial Industry* [Doctoral Study, Walden University]. Walden Dissertations and Doctoral Studies.

Rockart, J. F. (1979). Chief executives define their own data needs. Harvard Business Review, 57(2).

SAS Institute Inc. (2018). The SAS ® Data Governance Framework: A Blueprint for Success [White Paper]. Retrieved from https://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/sas data-governance-framework-107325.pdf

Saunders, M., Lewis, P., & Thornhill, a. (2009). Research Methods for Business Students. In Business (Vol. 5th).

Schryen, G. (2011). Seeking the "Value" in IS Business Value Research - An Agenda for Investigating Synergies between Socio-Organisational Change, IS Capabilities Change, and IS Innovation. ECIS, 2011.

Schryen, G. (2013). Revisiting IS business value research: What we already know, what we still need to know, and how we can get there. In European Journal of Information Systems (Vol. 22, Issue 2). https://doi.org/10.1057/ejis.2012.45

Sheikh, K., & Mattingly, S. (1981). Investigating non-response bias in mail surveys From the MRC Epidemiology and Medical Care Unit, Harrow. Journal of Epidemiology and Community Health, 35.

Sidi, F., Shariat Panahy, P. H., Affendey, L. S., Jabar, M. A., Ibrahim, H., & Mustapha, A. (2012). Data quality: A survey of data quality dimensions. Proceedings - 2012 International Conference on Information Retrieval and Knowledge Management, CAMP'12. https://doi.org/10.1109/InfRKM.2012.6204995

Shao, X., & Oinas-Kukkonen, H. (2019). How Does GDPR (General Data Protection Regulation) Affect Persuasive System Design: Design Requirements and Cost Implications. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11433 LNCS. https://doi.org/10.1007/978-3-030-17287-9_14

Spruit, M., & Pietzka, K. (2015). MD3M: The master data management maturity model. Computers in Human Behavior, 51. https://doi.org/10.1016/j.chb.2014.09.030

Srinivasan, R., & Swink, M. (2018). An Investigation of Visibility and Flexibility as Complements to Supply Chain Analytics: An Organisational Information Processing Theory Perspective. Production and Operations Management, 27(10). https://doi.org/10.1111/poms.12746

Tallon, P. P., Ramirez, R. v, & Short, J. E. (2014). Toward a Theory of Information Governance. Journal of Management Information Systems, 30(3).

Thomas, G. (2006). The DGI data governance framework. The Data Governance Institute, Orlando, FL (USA)

Ullman, J. B., & Bentler, P. M. (2012). Structural Equation Modeling. In Handbook of Psychology, Second Edition. https://doi.org/https://doi.org/10.1002/9781118133880.hop202023

van Helvoirt, S., & Weigand, H. (2015). Operationalising data governance via multi-level metadata management. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9373. https://doi.org/10.1007/978-3-319-25013-7_13

Vasconcelos, J. B. D., & Barão, A. (2017). *Ciência dos Dados nas Organizações*. FCA.

Wang, R. Y. (1998). A Product Perspective on Total Data Quality Management. Communications of the ACM, 41(2). https://doi.org/10.1145/269012.269022

Wagner, H. T., Beimborn, D., & Weitzel, T. (2014). How social capital among information technology and business units drives operational alignment and it business value. Journal of Management Information Systems, 31(1). https://doi.org/10.2753/MIS0742-1222310110

Weber, K., Otto, B., & Österle, H. (2009). One size does not fit all -A contingency approach to data governance. Journal of Data and Information Quality, 1(1). https://doi.org/10.1145/1515693.1515696

Wende, K. (2007). A model for data governance - Organising accountabilities for data quality management. ACIS 2007 Proceedings - 18th Australasian Conference on Information Systems.

Yu, W., Wong, C. Y., Chavez, R., & Jacobs, M. A. (2021). Integrating big data analytics into supply chain finance: The roles of information processing and data-driven culture. International Journal of Production Economics, 236. https://doi.org/10.1016/j.ijpe.2021.108135

APPENDIX A – STATISTICAL POWER TABLE

Statistical power is a statistical test for avoiding Type II errors (false negative), the probability of correctly rejecting the null hypothesis when it is actually false. Dependent on the significance level (α), sample size (N), and effect size (ES), statistical power analysis determines the optimal N for reliably detecting an existing effect (Cohen, 1992b). Power tables, such as Cohen's (1992a) (Fig. 12), are commonly used, displaying various effect sizes and significance levels, along with the corresponding minimum sample sizes required to achieve a desired level of statistical power, often set at 80% (Kock & Hadaya, 2018). This threshold, widely applied across research fields, balances the risk of Type II errors against the necessary sample size for adequate power.

Exhibit 1.7 Sample Size Recommendation a in PLS-SEM for a Statistical Power of 80%

Maximum Number of Arrows Pointing at a Construct	*Significance Level*											
	1%				*5%*				*10%*			
	Minimum R^2				*Minimum R^2*				*Minimum R^2*			
	0.10	*0.25*	*0.50*	*0.75*	*0.10*	*0.25*	*0.50*	*0.75*	*0.10*	*0.25*	*0.50*	*0.75*
2	158	75	47	38	110	52	33	26	88	41	26	21
3	176	84	53	42	124	59	38	30	100	48	30	25
4	191	91	58	46	137	65	42	33	111	53	34	27
5	205	98	62	50	147	70	45	36	120	58	37	30
6	217	103	66	53	157	75	48	39	128	62	40	32
7	228	109	69	56	166	80	51	41	136	66	42	35
8	238	114	73	59	174	84	54	44	143	69	45	37
9	247	119	76	62	181	88	57	46	150	73	47	39
10	256	123	79	64	189	91	59	48	156	76	49	41

Figure 12 – Sample size Power Table. (Cogen, 1992a)

The significance level, the first determinant of minimum sample size, represents the probability of a Type I error (false positive), which occurs when the null hypothesis is wrongly rejected. Decreasing the significant level reduces the risk of Type I errors but increases the risk of Type II errors (Cohen, 2013). In scientific research, a significance level of 5% commonly represents a reasonable balance between Type I and Type II errors (Cohen, 1992b). This study follows the recommended practices, acknowledging the consequences of both Type I (erroneously concluding organisational impacts of data governance practices) and Type II (erroneously dismissing the impact of data governance practices on the organisation) errors.

The effect size, the second determinant of minimum sample size, quantifies the magnitude of the relationship between independent and dependent variables. Larger effect sizes signify stronger relationships, requiring smaller sample sizes to achieve a given level of power. Cohen's coefficient, calculated as $\Delta R^2/(1-R^2)$, is a widely used effect size measure, with R^2 measuring the proportion of variance in the dependent variable explained by the independent variables (Kock & Hadaya, 2018). Given the absence of prior research supporting specific effect size magnitudes, a conservative approach was taken, setting the minimum R^2 value at 0.10 to determine sample size.

The final variable for minimum sample size is the maximum number of arrows directed at a model's constructs. In the proposed research model, the dependent constructs— Data Quality, Business Value, and Regulatory Compliance—have 6, 5, 2 and 1 arrows pointing at them. Since the maximum value corresponds to 6 this was the value set for this final variable.

In summary, in the power table approach, the minimum sample depends on the significance level, the effect size and the maximum number of arrows pointing at a model's construct. For this study, the values for these three variables were defined at 5%, 0.10 and 6. Therefore, based on the power table, the sample collected by the questionnaire must consist of at least 157 valid responses.

APPENDIX B – DEMOGRAPHICS ITEMS

Variables	Items	Source(s)
Industry Sector	What is your Organisation's industry sector? Arts and Entertainment Banking/finance Business services Communications Defence E-Commerce Education Food and Beverage Government/Municipalities Health services Insurance Logistics Manufacturing Oil and Gas Real estate Retail/Distribution Technological development Tourism and Accommodations Transportation Utilities Other	(Fink & Neumann, 2009; Ji-fan Ren et al., 2017)
Position	What is your position within the Organisation? Top/executive manager Middle/first line manager Middle/knowledge worker Operational/Data worker Operational/Production worker Don't Know /Don't want to answer	(Laudon & Laudon, 2019)
Organisation Size	How many full-time employees work in your Organisation? <50 50-250 251-500 501-1000 1001-5000 >5000 Don't Know /Don't want to answer	(Aydiner et al., 2019)
Time since DG adoption	Which of the following statements best describes your Organisations' Data Governance program? The program has been operational for less than 1 year The program has been operational for 1 to 3 years The program has been operational for 3 to 5 years The program has been operational for over 5 years We do not have an operational program Don't Know /Don't want to answer	(Adapted from Côrte-Real et al., 2017)
DG Framework	What is the framework on which your Organisation's Data Governance program is based? DAMA-DMBOK Functional Framework DGI Data Governance Framework Eckerson Modern Data Governance McKinsey Data Governance PwC Data Governance Framework SAS Data Governance Framework Don't Know /Don't want to answer Other	(Al-Ruithe et al., 2019) Indicated by Practitioners

APPENDIX C – QUESTIONNAIRE ITEMS

Constructs	Items	Code	Source(s)
Data-Driven Culture	Please indicate to what extent you agree with the following statements regarding data-driven practices in your Organisation:		(Wong et.al., 2021)
	We consider data a tangible asset	DD1	
	We base our decisions on data rather than on instinct	DD2	
	We are willing to override our own intuition when data contradict our viewpoints	DD3	
	We continuously assess and improve business processes based on the insights extracted from data	DD4	
	Employees are continuously coached to make decisions based on data	DD5	
Data Competencies	Please indicate to what extent you agree with the following statements regarding data competencies in your Organisation:		(Ongena, G., 2023)
	We have a broad understanding of different data sources, and we are able to select the most relevant ones	DC1	
	We have a high awareness of data quality criteria (e.g., complete, consistent, accurate, duplicate, up to date)	DC2	
	We can identify whether the data is trustworthy and locate alternate sources if required	DC3	
	We can detect and treat invalid records using technologies that support data cleaning	DC4	
	We are able to work with advanced statistics and select the most appropriate tool for each analytical task	DC5	
	We are able to create interactive charts and dashboards	DC6	
	We are able to effectively communicate our results and findings, tailoring the narrative to our audience	DC7	
	We can interpret data outputs and results confidently and critically	DC8	
Tools and Technologies	Please indicate to what extent you think data tools and technologies help your Organisation to:		(Wang et al., 2012)
	Enhance the responsiveness to data and information requests	TT1	
	Enhance the quality and amount of data available to the business units	TT2	(Masa'Deh et al., 2021)
	Enhance the data quality of business processes' outputs	TT3	
	Reengineer internal data processes	TT4	
	Provide thorough metadata to help understand the meaning of data	TT5	
	Provide all the information needed to understand the lineage of data	TT6	
Data Processes and Procedures	Please indicate to what extent you agree with the following statements regarding data processes and procedures in your Organisation:		(Xu, 2003)
	There are adequate processes for improving Data Quality, with appropriate DQ controls and resolution procedures	PP1	(Spruit & Pietzka, 2015)
	There is a defined process on how to access data and rules for which roles data access can be granted	PP2	(Lee & Strong, 2003)
	We know the sources and the processes by which the data used in our tasks is collected	PP3	
	We know the processes and team(s) involved in storing and maintaining the data used in our tasks	PP4	
	We know the processes and team(s) that require the data used in our tasks, and for which purpose the data is used	PP5	
	We know the standard procedures for correcting deficiencies/problems in the data when storing, collecting or using it	PP6	
Roles and Responsibilities	Please indicate to what extent you agree with the following statements regarding data stewardship practices in your Organisation:		(Spruit & Pietzka, 2015)
	Data stewards and owners are established for data domains	RR1	
	Data owners and stewards are announced and known throughout the Organisation	RR2	
	The data owners and stewards have defined responsibilities for treatment and management of data	RR3	
	The data owners define the usage purposes and quality standards of data and promote their adherence	RR4	
	Every data source that an employee might need is communicated and their access is managed upon request	RR5	
	The information in metadata repositories is maintained regularly and is kept up to date	RR6	

Constructs	Items	Code	Source(s)
Data Quality	The data sources used in your Organisation provide:		(Ji-fan Ren et al., 2017)
	A complete set of data	DQ1	
	The data needed for our tasks	DQ2	
	Up-to-date data	DQ3	(Lee & Strong, 2003)
	Well-formatted data	DQ4	
	Correct and reliable data	DQ5	
	Timely available data	DQ6	(Kwon et al., 2014)
	Easily accessible data when needed	DQ7	
	Consistent data across all the different sources	DQ8	
Business Value	Please indicate to what extent the data governance practices in place in your Organisation contribute to:		(Ji-fan Ren et al., 2017)
	Reducing operating costs	BV1	
	Avoiding the need to increase the workforce	BV2	
	Enhancing employee productivity	BV3	
	Enabling quicker response to change	BV4	
	Creating competitive advantage	BV5	
	Aligning analytics with business strategy	BV6	
	Providing higher quality products or services to customers	BV7	
	Supporting innovation	BV8	
Regulatory Compliance	Please indicate to what extent the data governance practices in place in your Organisation contribute to:		(Bany Mohammad et al., 2022)
	Providing all information with transparency and clarity	RC1	
	Reporting and publishing required data according to local and international standards	RC2	(Xu, 2003)
	Achieving better assessments in internal and external audits of data systems	RC3	
	Responding proactively to various legal and regulatory risks	RC4	(Kim & Kim, 2017)
	Understanding useful compliance information related to my role / task	RC5	

APPENDIX D – OUTER LOADINGS

Construct	Indicator	Outer Loading
Data-Driven Culture	DD1	0.765
	DD2	0.783
	DD3	0.754
	DD4	0.853
	DD5	0.830
Data Competencies	DC1	0.789
	DC2	0.791
	DC3	0.767
	DC4	0.782
	DC5	0.792
	DC6	0.827
	DC7	0.764
	DC8	0.847
Tools and Technologies	TT1	0.806
	TT2	0.835
	TT3	0.803
	TT4	0.765
	TT5	0.770
	TT6	0.785
Processes and Practices	PP1	0.853
	PP2	0.876
	PP3	0.824
	PP4	0.834
	PP5	0.828
	PP6	0.758
Roles and Responsibilities	RR1	0.846
	RR2	0.823
	RR3	0.908
	RR4	0.894
	RR5	0.798
	RR6	**0.647**
Data Quality	DQ1	0.841
	DQ2	0.864
	DQ3	0.892
	DQ4	0.739
	DQ5	0.886
	DQ6	0.847
	DQ7	0.821
	DQ8	0.859

Construct	Indicator	Outer Loading
Business Value	BV1	0.809
	BV2	0.747
	BV3	0.767
	BV4	0.891
	BV5	0.826
	BV6	0.817
	BV7	0.829
	BV8	0.833
Regulatory Compliance	RC1	0.867
	RC2	0.892
	RC3	0.894
	RC4	0.875
	RC5	0.807

APPENDIX E – CROSS LOADINGS AFTER INDICATORS REMOVAL

Constructs		BV	DC	DD	DQ	PP	RC	RR	TT
BV	BV1	**0.809**	0.553	0.548	0.571	0.536	0.620	0.504	0.564
	BV2	**0.746**	0.469	0.466	0.429	0.388	0.583	0.444	0.521
	BV3	**0.765**	0.480	0.505	0.462	0.407	0.626	0.454	0.476
	BV4	**0.891**	0.728	0.677	0.729	0.681	0.758	0.693	0.539
	BV5	**0.828**	0.632	0.603	0.710	0.716	0.640	0.650	0.437
	BV6	**0.818**	0.733	0.657	0.741	0.723	0.718	0.613	0.533
	BV7	**0.830**	0.619	0.665	0.643	0.626	0.684	0.545	0.552
	BV8	**0.833**	0.568	0.653	0.694	0.730	0.694	0.662	0.648
DC	DC1	0.638	**0.790**	0.675	0.728	0.694	0.665	0.590	0.574
	DC2	0.570	**0.810**	0.606	0.688	0.707	0.636	0.574	0.510
	DC4	0.515	**0.808**	0.467	0.555	0.553	0.598	0.539	0.470
	DC5	0.637	**0.824**	0.502	0.588	0.542	0.624	0.612	0.568
	DC6	0.665	**0.827**	0.562	0.622	0.544	0.620	0.610	0.532
	DC8	0.655	**0.845**	0.666	0.691	0.683	0.674	0.640	0.630
DD	DD1	0.539	0.533	**0.763**	0.559	0.526	0.564	0.463	0.434
	DD2	0.500	0.509	**0.783**	0.531	0.544	0.502	0.454	0.538
	DD3	0.589	0.541	**0.754**	0.538	0.480	0.516	0.508	0.480
	DD4	0.665	0.663	**0.854**	0.686	0.677	0.651	0.649	0.636
	DD5	0.614	0.591	**0.830**	0.609	0.659	0.646	0.629	0.568
DQ	DQ1	0.628	0.695	0.632	**0.836**	0.728	0.704	0.615	0.549
	DQ2	0.658	0.709	0.658	**0.877**	0.752	0.706	0.630	0.600
	DQ3	0.625	0.658	0.613	**0.898**	0.766	0.673	0.645	0.510
	DQ5	0.665	0.673	0.645	**0.871**	0.689	0.709	0.575	0.599
	DQ6	0.628	0.646	0.587	**0.868**	0.702	0.614	0.619	0.518
	DQ7	0.711	0.717	0.612	**0.835**	0.686	0.694	0.701	0.548
	DQ8	0.735	0.698	0.703	**0.859**	0.702	0.706	0.667	0.621
PP	PP1	0.681	0.783	0.616	0.743	**0.854**	0.760	0.759	0.592
	PP2	0.700	0.756	0.682	0.759	**0.884**	0.731	0.808	0.571
	PP3	0.569	0.534	0.567	0.626	**0.828**	0.656	0.645	0.557
	PP4	0.546	0.599	0.614	0.721	**0.858**	0.640	0.632	0.548
	PP5	0.625	0.550	0.620	0.681	**0.836**	0.672	0.645	0.594
RC	RC1	0.712	0.650	0.658	0.75	0.751	**0.868**	0.639	0.620
	RC2	0.713	0.698	0.623	0.669	0.676	**0.891**	0.661	0.661
	RC3	0.733	0.748	0.640	0.754	0.759	**0.894**	0.720	0.654
	RC4	0.744	0.745	0.601	0.632	0.647	**0.874**	0.648	0.588
	RC5	0.643	0.537	0.626	0.632	0.687	**0.807**	0.579	0.636

Constructs		BV	DC	DD	DQ	PP	RC	RR	TT
RR	RR1	0.598	0.562	0.621	0.618	0.718	0.649	**0.855**	0.575
	RR2	0.538	0.533	0.509	0.491	0.577	0.581	**0.797**	0.547
	RR3	0.635	0.678	0.672	0.709	0.791	0.677	**0.928**	0.571
	RR4	0.675	0.718	0.598	0.668	0.709	0.733	**0.894**	0.578
	RR5	0.548	0.632	0.541	0.669	0.734	0.582	**0.829**	0.438
TT	TT1	0.454	0.531	0.550	0.556	0.555	0.589	0.471	**0.807**
	TT2	0.528	0.614	0.615	0.650	0.670	0.646	0.536	**0.836**
	TT3	0.461	0.462	0.549	0.492	0.485	0.536	0.386	**0.802**
	TT4	0.521	0.467	0.504	0.513	0.560	0.538	0.547	**0.766**
	TT5	0.504	0.596	0.457	0.411	0.393	0.544	0.495	**0.768**
	TT6	0.546	0.532	0.503	0.465	0.502	0.607	0.546	**0.784**

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa