

MDDM

Master Degree Program in
Data-Driven Marketing

DEVELOPMENT OF A DATA GOVERNANCE FRAMEWORK FOR CORPORATE DATA

Henrique Manuel Silva Matos

Dissertation

presented as partial requirement for obtaining the Master Degree Program in Data-Driven Marketing

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

DEVELOPMENT OF A DATA GOVERNANCE FRAMEWORK FOR CORPORATE DATA

By

Henrique Manuel Silva Matos

Master Thesis presented as partial requirement for obtaining the Master's degree in Data-Driven Marketing, with a specialization in Marketing Intelligence

Supervisor: Nadine Côte-Real

06/2023

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 acknowledge the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Henrique Matos

Lisbon, 27th June 2023

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my NOVA IMS thesis supervisor Nadine Côte-Real for all the guidance, feedback, and patience given throughout the process of creating this dissertation. The shared knowledge will always resonate in me.

I am extremely grateful to Casa Santos Lima, specially to Dr. José Luís Santos Lima Oliveira da Silva and Dr. Luís Olazabal de Almada, for believing in my potential and funding this journey that was this master's program.

I would be remiss in not mentioning all my coworkers that give me strength and practical knowledge to endure my best self to deliver this work.

Special thanks to my family, for the unrestricted and everlasting encouragement and support, each one at their own way.

Lastly, I would like to mention my friends that accompanied me unconditionally through every step of this journey, making it easy for me to finish this work.

ABSTRACT

With the amount of data exponentially growing, especially derived from the constant technological evolution, data has become not only essential for companies that want to retrieve value from it and create competitive advantages but also a necessity for ensuring compliance with current legislation and sustaining data quality attributes. Data Governance emerges to give data its importance in any corporate scenario by creating a new structure that guarantees data meets business needs, protecting it from any internal/ external threat, and ensuring stakeholders' commitment to maintain this system. It is applicable through the definition, approval, and communication of data strategies, policies, standards, and procedures necessary to enable better decision-making and ensure data security, privacy, and quality. Data Governance frameworks' importance arises due to their summarization capabilities, turning a Data Governance program into an easy-to-read structure. It was found that existing Data Governance frameworks have room for improvement when it comes to critical areas tackled, implementation and monitoring activities definition, and implementation roadmap creation. To fulfill these limitations, this study consists of a critical analysis of four existing frameworks according to Data Governance Critical areas defined by assessing the most relevant attributes of Data Governance and practical insights from already existing studies in this area. The result is a newly adapted framework that tackles all critical areas by building on each existing frameworks' strengths to maintain its assertiveness, simplicity, and pertinence; and its respective implementation roadmap to clarify its implementation and monitorization by increasing its practicality.

KEYWORDS

Data Governance; Data Governance Framework; Framework Comparison; Framework Creation; Implementation Roadmap

Sustainable Development Goals (SDG):



INDEX

1. Introduction.....	1
1.1. Research Context.....	1
1.2. Research Gap and Objectives	2
2. Literature review	3
2.1. Data Governance	3
2.1.1. Data Governance Definition	3
2.1.2. Data Governance Objectives	4
2.1.3. Data Governance Challenges	4
2.2. Data Governance Frameworks	5
2.2.1. Frameworks Choice	5
2.2.2. Non-Invasive DG Framework.....	5
2.2.3. DGI DG Framework.....	7
2.2.4. SAS DG Framework.....	8
2.2.5. DAMA-DMBOK Framework	10
3. Methodology	12
4. Empirical Study	14
4.1. Research Scope.....	14
4.2. Comparison Criteria Definition.....	14
4.2.1. Common Structure of Data Governance Activities	14
4.2.2. Data Governance Decision Domains	15
4.2.3. Data Governance Areas.....	15
4.2.4. Data Governance Critical Success Factors.....	16
4.2.5. Comparison Criteria Definition.....	17
4.3. Framework Comparison	18
4.4. Creation of an Improved Framework	20
4.4.1. Data Requirements.....	20
4.4.2. Data Strategy	20
4.4.3. Data Policies	22
4.4.4. Roles and Responsibilities	22
4.4.5. Employee Data Competencies	24
4.4.6. Processes and Procedures.....	24
4.4.7. Communication	25
4.4.8. Tools and Technology.....	25

5. Results and discussion	27
5.1. New Data Governance Framework	27
5.2. Implementation Roadmap of the New Framework	28
5.3. Discussion	30
6. Conclusions, Limitations and future works	31
Bibliographical References	33
Annexes	37
Annex 1 – Data Governance Critical Success Factors Exploration	37
Annex 2 – Frameworks Study on Each Comparison Criterion	40

LIST OF FIGURES

Figure 1 - Non-Invasive DG Principles (Seiner, 2014).....	6
Figure 2 - Non-Invasive Operating Model of Roles & Responsibilities (Seiner, 2014).....	6
Figure 3 - Complete Non-Invasive DG Framework (Seiner, 2014)	7
Figure 4 - DGI DG Framework (Thomas, n.d.)	8
Figure 5 - SAS DG Framework (SAS Institute Inc, 2018)	9
Figure 6 - DM Functions (Mosley et al., 2009)	10
Figure 7 - DSR Process Model - Adapted from (Peffer et al., 2007)	13
Figure 8 - Comparison Criteria (Own Construction).....	17
Figure 9 - DG three-part mission (Thomas, n.d.).....	21
Figure 10 - Data Issue Escalation Path (Mosley et al., 2009)	23
Figure 11 - New adapted DG Framework (Own Construction).....	27

LIST OF TABLES

Table 1 - DG Challenges (Benfeldt et al., 2020)	4
Table 2 - SAS DG Framework Components (SAS Institute Inc, 2018)	9
Table 3 - DAMA-DMBOK framework's environmental elements (Mosley et al., 2009)	10
Table 4 - The DAMA-DMBOK Functional Framework, Version 3 (Mosley et al., 2009)	11
Table 5 – DG activities' phases (Kuan Cheong et al., 2007)	14
Table 6 - Framework for data decision domains (Khatri & Brown, 2010)	15
Table 7 - DG Areas (Alhassan et al., 2016)	16
Table 8 – Criteria Definition Assessment (Own Construction)	17
Table 9 - Scale Description for framework comparison (Own Construction)	18
Table 10 - Framework Comparison Assessment Table (Own Construction)	19
Table 11 - New DG Framework's Implementation Roadmap based on (Mosley et al., 2009)	29
Table 12 - DG Process Summary – Adapted from (Mosley et al., 2009)	48

LIST OF ABBREVIATIONS AND ACRONYMS

CDO	Chief Data Officer
CIO	Chief information Officer
CSF	Critical Success Factor
DMBOK	Data Management Book of Knowledge
DG	Data Governance
DGI	Data Governance Institute
DGO	Data Governance Office
DM	Data Management
DQ	Data Quality
DSR	Data Science Research
GDPR	General Data Protection Regulation
HR	Human Resources
IT	Information Technology
KPI	Key Performance Indicator
SDLC	Software Development Cycle

1. INTRODUCTION

1.1. RESEARCH CONTEXT

Due to the global digitalization trend and most recently leveraged by the COVID-19 pandemic, it is estimated that the amount of data created, captured, copied, and consumed worldwide in 2021 will double by 2025 (Statista Research Department, 2022). This data overload is getting more challenging to manage, and firms tend not to be able to extract relevant insights from dispersed databases or inconsistent data silos (DELL Technologies, 2021). If not well-managed, data creation can go out of control and lack quality and usability (Oliveira et al., 2005), highlighting the relevance of studies on solutions that can help them overcome this challenge.

Data is vital for companies since it is the foundation of a good management system at any corporate level, enabling value and competitive advantages creation (Holsapple, 2008). Since data is not perishable, it is considered one of the most valuable assets of an organization (Panian, 2010). More than that, companies need to ensure their data is secure, complies with current legislation, and is relevant and accessible at a corporate level. (Abraham et al., 2019).

Data Governance (DG) appears as the planning and formalization of rules and processes over data classification, production, and application to guarantee its value to all stakeholders (Seiner, 2014). Its study and application intend to solidify data identity as corporate assets and assure its quality dimensions being intrinsically related with Data Quality (DQ) (Mosley et al., 2009). DG application enables better decision-making by structuring a cohesive data structure through data strategies, policies, standards, and procedures, ensuring data security, privacy, and quality (Al-Ruithe et al., 2016; Mosley et al., 2009; Thomas, n.d.).

The increasing importance and demand of DG solutions led academics to believe that there is no single approach to its implementation, and it will always depend on company specific data culture, stakeholders' resistance to changes, and corporate objectives, being a complex task when it comes to implement it in a real-case scenario (Weber et al., 2009). To reduce this complexity, DG is usually represented in a form of a framework that facilitates the understanding of a DG program by organizing concepts in a simplified manner which has been proven to be an efficient and successful way to deploy DG (Kuan Cheong et al., 2007; Thomas, n.d.).

Since enterprises are too diverse in all aspects, DG frameworks tend to be very broad to be adapted to each companies' reality, which led to a lack in details due to the necessity of a more subjective framework; or tend to be very specific to an exact industry and lack flexibility of implementation. Either way, existing DG frameworks tend to serve as a platform for all companies to build their own take on DG frameworks adapted to their business (Begg & Caira, 2012).

1.2. RESEARCH GAP AND OBJECTIVES

Through the research carried out, it was explored several DG frameworks and their advantages and possible points of improvement related to their scope and implementation potential. It was found in the frameworks' official documents that no matter how complete the framework is, there is space for improvement (Mosley et al., 2009; SAS Institute Inc, 2018; Seiner, 2019; Thomas, n.d.). For instance, DG Institute has a very complete framework in terms of tackled areas but does not offer the level of detail needed in terms of communication plans and responsibilities that the Non-Invasive DG framework offers, and so on (This point will be further explored in "[4.3. Frameworks Comparison](#)"). Moreover, it was found that very few frameworks explore the creation of a clear way of applying a DG framework to a company in a practical way (implementation roadmap) which can turn into a problem to companies that want to put DG concepts to practice (Black et al., 2023; Mosley et al., 2009). Furthermore, existing DG frameworks have room for improvement to make them tackle more important areas, more specifically, critical success factors (CSF) but maintaining their adaptability and scalability (Begg & Caira, 2012; Bento et al., 2022).

The aim of this study is to tackle this opportunity of improvement and build a "new"/adapted framework, flexible enough to be applied to a company new to DG and scalable to all types of organizations, and its respective path of implementation.

The question to be answered is: **"How to apply DG to a company through the creation of a "new"/ adapted framework based on existing literature?"**.

In order to answer this question and to tackle the main objective of creating the framework and its path of implementation, it is necessary to accomplish the following objectives:

- Gather information about DG and specially DG frameworks to emphasize the importance of the topics while contributing to DG literacy.
- Identify the critical areas for DG through literature.
- Compare existing DG frameworks to understand the best approaches to each critical area.
- Formulate an optimized perspective on each critical area based on existing frameworks approaches.

2. LITERATURE REVIEW

To start, it is necessary to do a literature review on the topics of interest: in this case, DG, and DG frameworks. The objective is to consolidate existing knowledge on these topics and to have a deeper comprehension of what already exists and how it can shape the research's path, so a **systematic literature review** will be done since it is the one with these underlined characteristics (Whitehead & Maude, 2016). The objectives of this literature review are to **(1)** gather valuable literacy to demonstrate the importance of corporate-level DG and **(2)** identify relevant existing DG frameworks and, if possible, their practical path of application to a company that wants to be data-driven.

2.1. DATA GOVERNANCE

2.1.1. Data Governance Definition

DG is not a new term in literacy. Its practice foundation is rooted in corporate governance and information Technology (IT) governance, both very mature disciplines in terms of literacy, where corporate governance is directed to how companies are administered, and IT governance is very centered on how software applications can sustain corporate's strategies and objectives (Khatri & Brown, 2010; Panian, 2010). Consequently, neither are data centered.

As seen before, due to the data-driven trend and exponential growth of data creation (Barham, n.d.), it was necessary to start giving data its importance and create a new structure that would guarantee data meets business needs, protect it from any internal/ external threat and ensure the commitment of all stakeholders with the maintenance of this system. Consequently, DG was born as an implied necessity to ensure that data was available across the company and centered on the business process's necessities (Panian, 2010). This way, it is necessary to build a new team responsible for data in a business-driven way, differentiated from IT that is software application centered and will function as support for the data team (Mosley et al., 2009; Seiner, 2014).

Like Data, DG has many definitions across literacy depending on the study area and the completeness necessary for the given context (Al-Ruithe et al., 2018; Khatri & Brown, 2010; Mosley et al., 2009). DG is "the exercise of authority and control (planning, monitoring, and enforcement) over the management of data assets" (Mosley et al., 2009). This definition implies that DG considers data as one of the most valuable assets of companies even though it is not embodied on the balance sheet; DG is a process driven by the executive power that needs to steer the plan for it to be effective; it requires the definition of accountabilities and power delegation to ensure an accurate implementation and monitoring; and is a complex process that implies a corporate effort to succeed, because if workers do not comply with the program, errors will be made and DQ may be compromised (Panian, 2010).

A DG program does not have a subtle start because it is a very rigorous process that needs a significant shift in the mentality of all stakeholders and a recognition of data as a valuable asset. In fact, Alhassan et. al. (2019) defend that a DG Program should start with a narrowed scope, restricted to data critical to the organization, and then build up when it makes sense to the company according to its objectives (Alhassan et al., 2019a). DG does not have an endpoint too because it is an ongoing process of improvement and acknowledgment that is set to have permanent adjustments in company's structure (Ladley, 2020; Mosley et al., 2009).

2.1.2. Data Governance Objectives

After reviewing several articles, it was possible to conclude that there are many DG objectives spread through diverse areas depending on its application purpose (Al-Ruithe et al., 2016; DGI, n.d.-a; Panian, 2010). For this project, DG objectives are divided into two categories: the data drivers which are directly involved with corporate data and the construction of the DG plan, and the business drivers that will be fulfilled with the application of DG.

Data-Driven objectives for DG:

- Definition, approval, and communication of data strategies, policies, standards, procedures, and metrics (Mosley et al., 2009);
- Enable better decision-making (Thomas, n.d.);
- Establishing data stewardship by defining decision-making roles/ responsibilities (Karkošková, 2022; Ladley, 2020);
- Management and resolution of data-related issues (Mosley et al., 2009);
- Ensure data security, privacy, and quality (Al-Ruithe et al., 2016).

Business-Driven objectives for DG:

- Ensure compliance with regulatory requirements like General Data Protection Regulations (GDPR) to avoid legal penalties (CMS, n.d.).
- Lower financial costs by reducing data redundancy and increasing operational efficiency and growing revenue by deep-understanding customer data (Panian, 2010).

2.1.3. Data Governance Challenges

Aiming the completion of the previously listed objectives, it is key to understand what type of struggles existing companies have felt when they were introducing a DG program so that these issues may be taken into consideration when it comes to build a cohesive DG strategy for the future (Benfeldt et al., 2020).

Table 1 - DG Challenges (Benfeldt et al., 2020)

Data Governance Challenges	
Perceiving Value	Produce a DG program that clearly transmits the true value-creation potential of enterprise-level data and DG, and foment its adoption by every data stakeholder.
Enabling Collaboration	Establish effective communication between departments so that the necessary data reaches each stakeholder.
Fostering Capabilities	Consider and develop each department's professional experience and data maturity level to create an enterprise-level DG program.
Data Overview	Have all enterprise-data easily accessible and under control.
Local Practices	Standardize data principles/ policies that need to be enforced to avoid department's differences on the way of work.
Political Ambience	Ensure that the executive level funds and enforces the correct adoption and implementation of the DG program.

2.2. DATA GOVERNANCE FRAMEWORKS

Being a complex process, DG is sometimes difficult to explain to all stakeholders. For that, a DG program is usually presented as a framework. A DG framework is defined as a logical structure for classifying, organizing, and communicating complex activities involved in making decisions about and acting on enterprise data (Thomas, n.d.). In other words, a DG framework simplifies a DG program created to facilitate its understanding through simplified concepts and better-organized ideas, and it was proven that its adoption would make data management (DM) a more efficient and successful process (Kuan Cheong et al., 2007). This way, the usage of DG frameworks is demonstrated as a relevant way to implement DG, complying with the main objectives of this project.

2.2.1. Frameworks Choice

As seen before, over time several frameworks have been created trying to fulfill specific needs of a certain business area or being general enough to be adaptable to different areas (Begg & Cairn, 2012). The choice of the frameworks to be analyzed for this project fell in several factors such as their recognition among DG literacy (pertinence and relevance); their practicality and flexibility of implementation (falling into the objectives of this study); and finally, the level of integrity of areas addressed by the framework to enable the construction of a more complete and assertive analysis. To start, it was pondered the most relevant DG frameworks considered by Atlan (2022), an active metadata platform for modern data teams that works with the implementation of this type of programs worldwide. Examining each one according to the factors mentioned earlier, the chosen frameworks are:

Non-Invasive DG framework was chosen due to its flexibility in using existing corporate structures, people, and culture to shape DG into this ongoing system, as well as for its very detailed perspective on DG roles and responsibilities (Seiner, 2019).

DG Institute (DGI) framework was selected due to its recognition among literature (atlan, 2022; Bennett, 2023; Bento et al., 2022; Kimachia, 2022) and its completeness when it comes to areas considered (Thomas, n.d.).

SAS Institute DG framework was chosen due to its applicability when it comes to solving enterprises' real problems and due to SAS reputation in the business intelligence and software industries (SAS Institute Inc, 2018).

DAMA-DMBOK (Data Management Book of Knowledge) framework was also chosen due to its recognition among literature (atlan, 2022; Bennett, 2023; Kimachia, 2022) that results from the compiled knowledge of different experts on data over time and distancing itself from other frameworks because it has a more theoretical approach to DG has the central function of DM (Mosley et al., 2009).

2.2.2. Non-Invasive DG Framework

Taking into consideration the complexity of the application of a DG program, Seiner (2014), president of KIK Consulting & Educational Services and specialist in DG, has built a framework with the objectives of not being a threat to the existing enterprise culture, and tackling one of the most menacing challenges of DG: Corporate acceptance/ practice of the DG program. The Non-Invasive DG

approach states that companies already have the necessary people/ tools to put DG in a formal way in practice. The best practices of DG are compared to existing ones to identify current strengths and effectively overcome present weaknesses. In this approach, it is also crucial to ensure the administration understands and supports the program for it to be successfully implemented (Seiner, 2014). These principles are compiled in the figure below, where DG policy is explained through different principles that are measured by DQ dimensions:

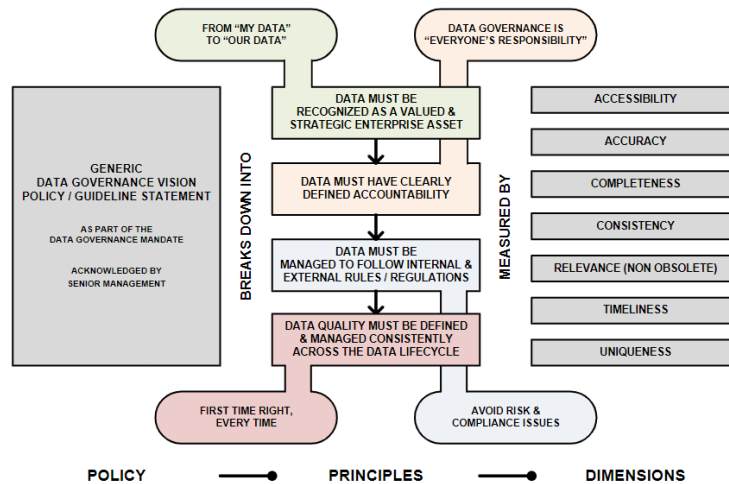


Figure 1 - Non-Invasive DG Principles (Seiner, 2014)

One of the major focuses of Non-Invasive DG is the roles and responsibilities definition. This model states that there are four layers of DG responsibilities (**Executive, Strategic, Tactical, and Operational**) and one extra layer of support (**DG team; DG partners**). Seiner (2014) consider that most organizations already have their executive, operational, and support levels in their company, they are just not recognized as so (Seiner, 2014).

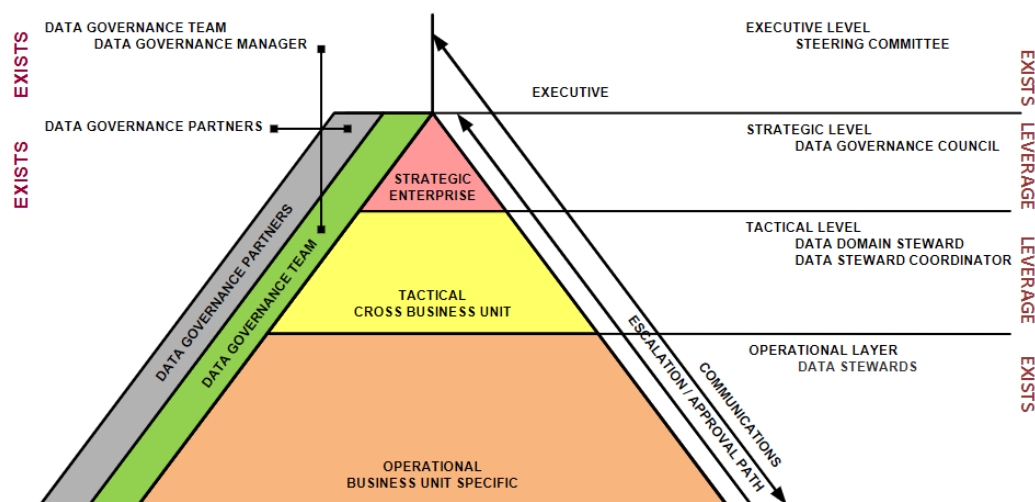


Figure 2 - Non-Invasive Operating Model of Roles & Responsibilities (Seiner, 2014)

The Non-Invasive DG framework is essentially an assessment table composed of levels that are people and authority-based (responsibility levels previously defined) and components that are core moving parts of governance (focus on the responsibilities and metrics). As a result, the framework

describes the different roles and pivotal characteristics of each level of accountabilities according to some previously defined CSFs. The diagram works as a guiding tool on how each component-level relation is conceived, developed, and implemented across the organization, demonstrating valuable-meaning at a corporate level (Seiner, 2019). It is possible to conclude that this model is very centered in people since they are the drivers of DG throughout the organization. Without people's commitment, a DG program is set to fail, which is what Non-Invasive DG Framework is trying to avoid. However, it is not very detailed regarding policy and rules creation around data, and it barely touches privacy concerns on data.

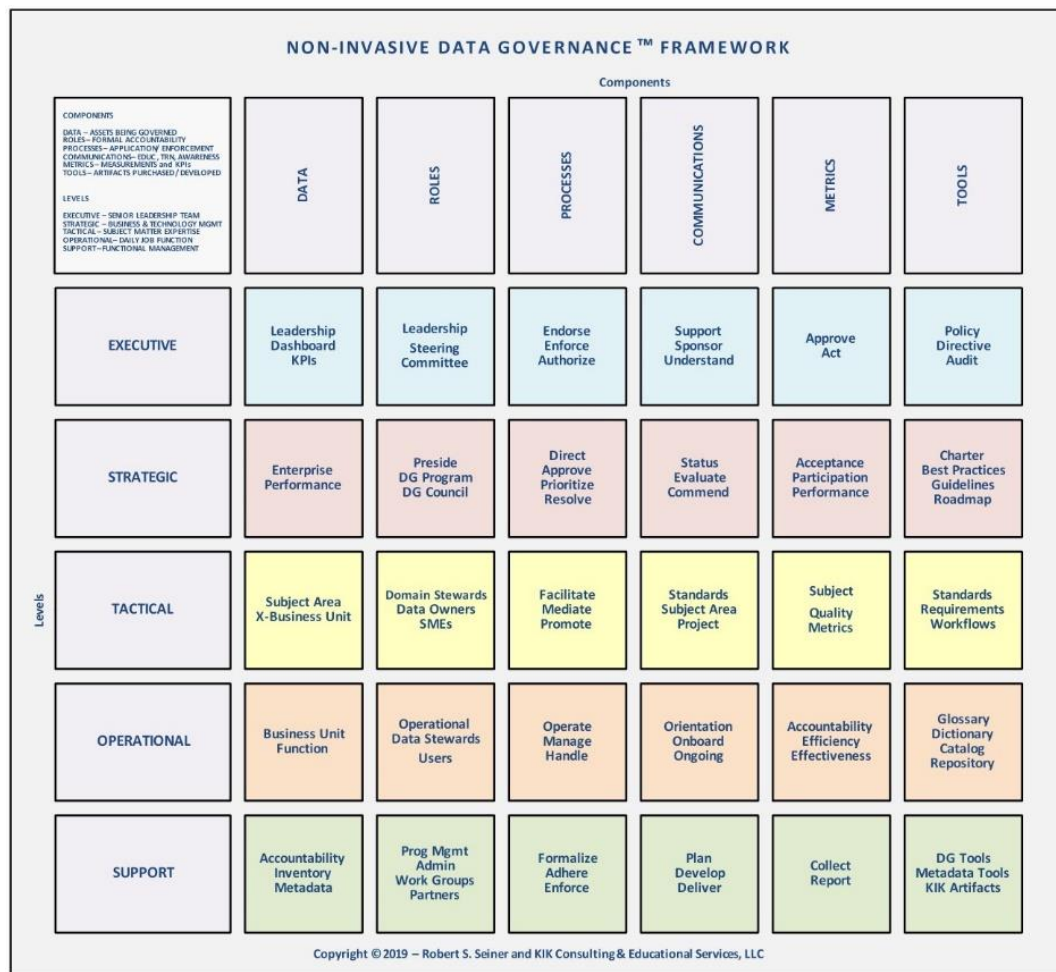


Figure 3 - Complete Non-Invasive DG Framework (Seiner, 2014)

2.2.3. DGI DG Framework

DGI presents a complete DG framework, compiled, and explained by Thomas (n.d.), that was created to tackle all stakeholders' needs in terms of data. The result is a very results-oriented framework that can meet all focus areas needs in a very systematic and objective manner. It comprises ten universal components divided into three broad areas: **Rules and Rules of Engagement, People and Organizational Bodies**, and **Processes**. Following these components will offer a path for DG planning and implementation (DGI, n.d.-a; Thomas, n.d.).

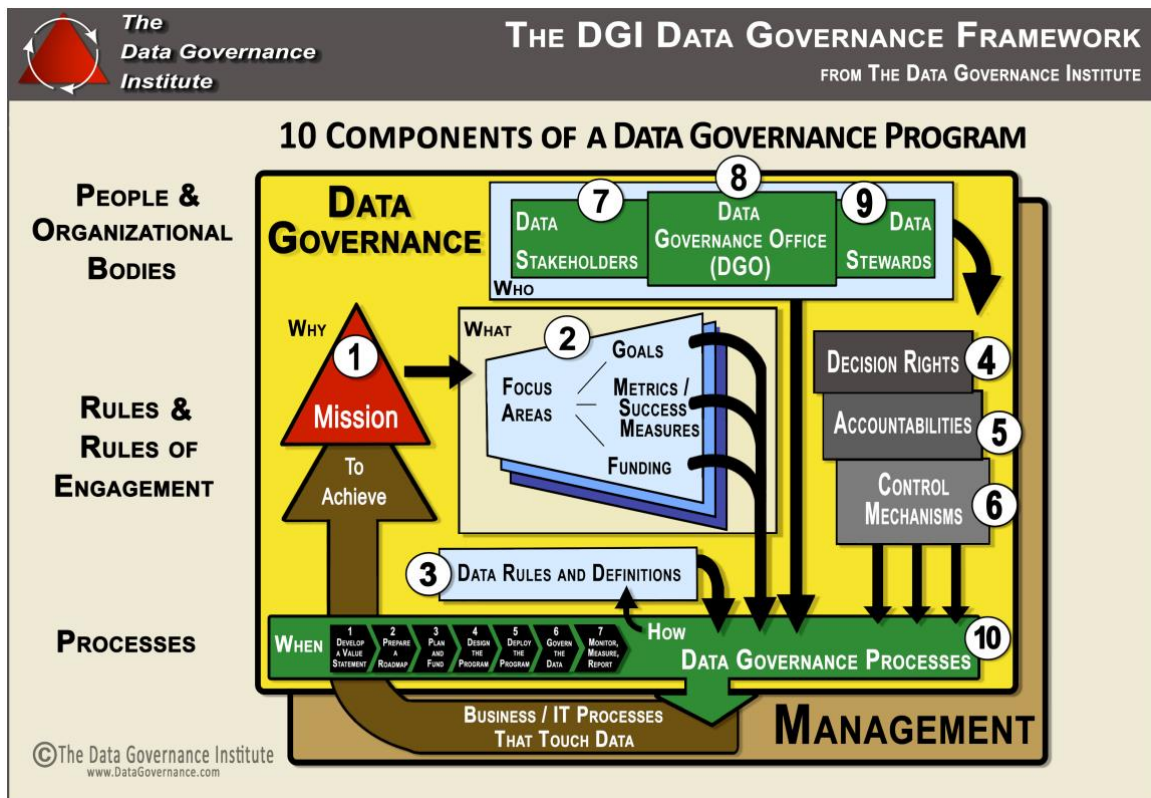


Figure 4 - DGI DG Framework (Thomas, n.d.)

In **Rules and Rules of Engagement**, there is very detailed planning on what exactly is necessary to govern and the definition of how the company wants to do it. This is stated through the “Mission and Vision” definition; “Goal, Governance Metrics and Success Measures, and Funding Strategies”; “Data Rules and Definitions”; “Decision Rights”; “Accountabilities”; and “Controls”. In **People and Organizational Bodies**, it is tackled how to put into practice the standards, rules, and policies earlier defined from the perspective of the workforce (Who is going to do what). This way, this part addresses “Data Stakeholders”; “A Data Governance Office”; and “Data Stewards”. Finally, in **Processes**, it is explained how the people in the second area will practice these rules and policies and maintain them from a process perspective through the “Proactive, Reactive, and Ongoing Data Governance Process” (Thomas, n.d.).

Despite being very complete and presenting many solutions that cover a big number of enterprise areas, this framework does not describe what type of technological solutions the company can have to automate DG activities since it is more technologically focused on data security in the “Controls” section. It also does not enter in much detail when it comes to communication plans amongst all stakeholders, being a point of improvement when it comes to the usage of this specific framework (Thomas, n.d.).

2.2.4. SAS DG Framework

As a data-driven software seller, SAS Institute Inc. created a DG framework as a tool to avoid data dysfunctions at a corporate level that aims to help stakeholders to raise their analysis capability by reducing data related problems and leverage overall reliability of data. To create a DG plan, a company must reunite two foundations: the **business** perspective strategies and inputs and the

technology required to put them into account in a holist-pragmatic approach. **Holistic** to such a degree that all aspects of data should be considered and interconnected to create a corporate-level view of data, and **pragmatic** in a way that the DG plan is implemented in phases to be as easy to assimilate and practical enough to bring “quick wins” to the company (SAS Institute Inc, 2018).

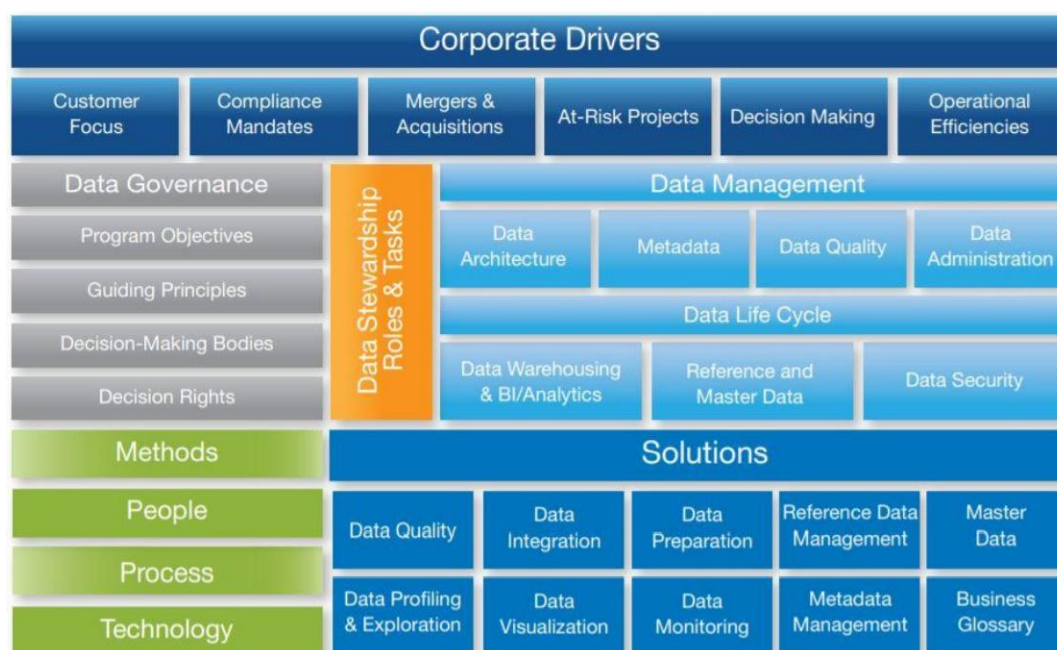


Figure 5 - SAS DG Framework (SAS Institute Inc, 2018)

SAS DG Framework is divided into three areas that correspond to the aspects of governance: **Strategic** (linkage between DG activities and achieving good performance in key business goals), **Organizing Framework** (development and monitoring of rules and policies that serve as guidelines for the DG activities), and **Tactical** (strategical execution of DG rules and policies previously defined) (SAS Institute Inc, 2018). These areas are subdivided into sections that are explained in the table below:

Table 2 - SAS DG Framework Components (SAS Institute Inc, 2018)

Governance Aspects	Framework Section	Role in the framework
Strategic	Business Drivers	Definition of DG Scope; Definition of corporate performance indicators; Measure relevance and pertinence of the DG plan; Steer the program.
Organizing Framework	Data Governance	Definition of objectives, guiding principles, decision-making bodies, and decision rights
	Methods	Development of DG plan (People, Processes, and Technology)
Tactical	Data Management	Execution of rules and policies defined in Data Governance
	Data Stewardship	Definition of the role and organization of data stewardship
	Solutions	Definition of applications to employ and automate Data Governance

This approach is very problem-solving-centered since it is dedicated to practical solutions to DG-related problems, resulting in “quick wins” for the companies that apply it successfully. Another vital part to highlight is the consideration of both people and technology as pivotal foundations of the implementation of DG in any company. On the other hand, because of its completeness, it can also be a hard framework to implement, requiring a more mature level in terms of existing processes and procedures and data requirements (Bento et al., 2022).

2.2.5. DAMA-DMBOK Framework

DAMA-DMBOK is a book that compiles knowledge from experienced practitioners about DM for over thirty years and is considered a reference in this area. In this book an overall governance framework describes all the functions that DM is comprised of. DG appears here has the central element of DM functions that is in charge for high-level planning, supervising, and controlling of DG (Mosley et al., 2009). Still, it is also important to take into consideration nine more vital functions that explore a part of DG in detail:

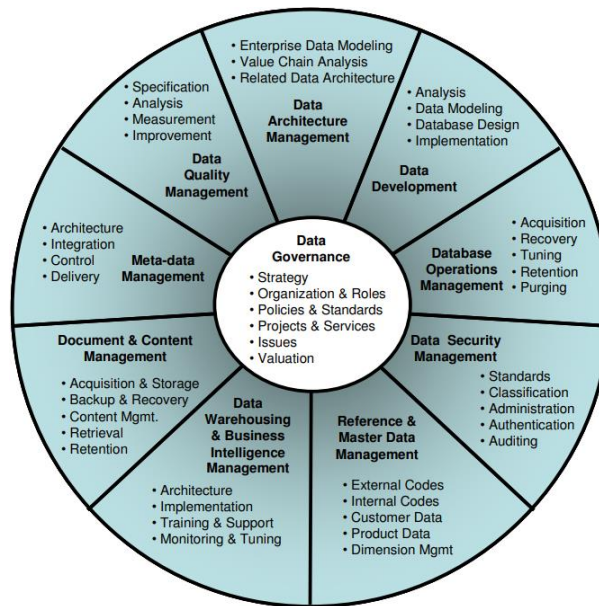


Figure 6 - DM Functions (Mosley et al., 2009)

To understand, plan, and structure this ten important knowledge areas, DAMA-DMBOK recommends the characterization of each one according to seven environmental elements that will guide the whole program and will culminate in the creation of the framework. These seven elements are divided into four **basic elements** (crucial structural points of each DQ area that represent their identity), and three **support elements** (represent processes/features that will ensure the correct implementation of what is previously defined) (Mosley et al., 2009):

Table 3 - DAMA-DMBOK framework's environmental elements (Mosley et al., 2009)

Type of Element	Environmental Elements	Definition
Basic	Goals and Principles	Core business principles, goals and vision that funnel each function's performance
	Activities	What each function does at the activities, sub-activities, and task levels
	Primary Deliverables	Information created and stored from the performed activities
	Roles and Responsibilities	Definition of all roles that will perform the defined activities and their specific responsibilities
Supporting	Practices and Techniques	Establishment of common practices and methods to help the best-execution of activities
	Technology	Definition of supporting technology to help goal achievement by enhancing the execution and monitoring of the defined activities
	Organization and Culture	creation of guidelines on any topic that is considered relevant for employees' understanding and acceptance of what is expected to be their role in the company

From the intersection of DM functions and environmental elements results the DAMA-DMBOK's functional DM framework. In this context, the part of interest is the matching between DG and the environmental elements that will work has a deconstructed DG framework (Mosley et al., 2009). From the four frameworks chosen this is the most different one in terms of structure because it may not be directly perceived has a DG framework on its own, despite being referenced multiple times in literature as one (atlan, 2022; Bennett, 2023; Kimachia, 2022).

Table 4 - The DAMA-DMBOK Functional Framework, Version 3 (Mosley et al., 2009)

Data Management Functions	Goals and Principles	Activities	Primary Deliverables	Roles and Responsibilities	Practices and Techniques	Technology	Organization and Culture
Data Governance							
Data Architecture Management							
Data Development							
Data Operations Management							
Data Security Management							
Reference and Master Data Management							
Data Warehousing and Business Intelligence Management							
Document and Content Management							
Meta-Data Management							
Data Quality Management							

3. METHODOLOGY

From the developed systematic literature review, it was possible to conclude that all frameworks explored have improvement points that other frameworks' advantages can tackle (SAS Institute Inc, 2018; Seiner, 2019; Thomas, n.d.). Thus, it will be necessary to assess these similarities and differences deeper and create a new/ adapted framework. Following this and considering the initial objectives, this research will try to answer the following questions:

- How to apply DG to a company through the creation of a “new”/ adapted framework based on existing literature?
- What is the roadmap for implementing this new DG framework?

Considering these questions, the following research will culminate in the creation of a DG framework and its respective roadmap of implementation. For this type of result, the best type of research to apply is Data Science Research (DSR), which is a type of research focused on problem-solving by enhancing existing knowledge and ultimately creating innovative artifacts (Brocke et al., 2020). For this research, the systematic literature review that was developed based on textbooks, peer-reviewed academic journal articles, and reports from technology consultants, will serve as a base for this next step and will culminate into a new artifact, a DG framework for corporate data.

A DSR research is a six-step process that starts with the **identification of the problem** that is affecting a real-world environment that sustains its importance and motivation (Brocke et al., 2020). In this case, as seen before, there are two main problems:

- The exponential data creation that is affecting companies' data quality and, consequently, its decision-making capability (Abraham et al., 2019; Bento et al., 2022).
- Existing DG frameworks have room for improvements and lack the definition of a clear roadmap of implementation (Begg & Caira, 2012; Bento et al., 2022).

This step is followed by the definition of **research objectives for the solution**, in this case, the DG framework (Brocke et al., 2020). To better understand what is really expected from the new/ adapted framework and from what was seen in other research studies on this area (CCI, 2016; Karkošková, 2022), it is necessary to establish the scope of the project (what is intended to achieve with the creation of the artifacts and the unit of analysis) and define what are the critical factors that will serve as comparison criteria to explore existing frameworks.

The third step will be the actual **design and development** of the framework and the roadmap of implementation, which will result from the critical comparison of the existing frameworks considering the areas of interest defined in the previous step (Brocke et al., 2020).

After developing the new framework, it is necessary to test it by **demonstrating** how the solution presented solves the problems posed initially (Brocke et al., 2020). For this project, this demonstration will have to be theoretical by explaining how the initial goals are met with the presented solution.

The complete DSR would be followed by the **evaluation** phase, where the proposed artifact would be used and tested in a real-world scenario by observing its effectiveness while implementing

it (Brocke et al., 2020). Due to the theoretical nature of this research, and because it is not part of its focus, the new framework will not be tested in a real case scenario, being not only one of this project's limitations but also highly encouraging other researchers to continue this work.

The last phase is the **communication**, where the artifact is presented to all stakeholders involved, and it is decided if it will be implemented or not (Brocke et al., 2020). For this academic research, this phase will culminate with the delivery and presentation of the dissertation.

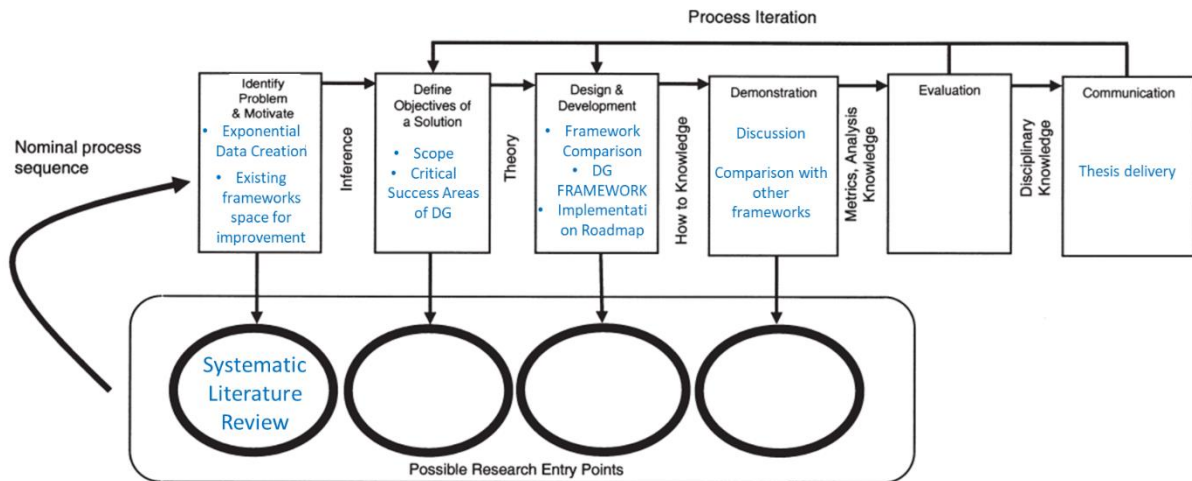


Figure 7 - DSR Process Model - Adapted from (Peffer et al., 2007)

4. EMPIRICAL STUDY

4.1. RESEARCH SCOPE

This project does not intend to find the best DG framework because has seen before that would be impossible due to fact that they are intrinsically dependent on each company's culture, objectives, and maturity level (Begg & Caira, 2012). Contrarywise, the present project intends to act on existing frameworks to improve their weaknesses and leverage their strengths (Bento et al., 2022) by prioritizing simplicity and flexibility, but at the same time maximizing the completeness in terms of DG matters, and using a unit of analysis of four frameworks previously identified:

- DGI Framework (DGI, n.d.-a).
- Non-Invasive DG Framework (Seiner, 2019).
- SAS DG Framework (SAS Institute Inc, 2018).
- DAMA-DMBOK Framework (Mosley et al., 2009).

4.2. COMPARISON CRITERIA DEFINITION

To proceed with the frameworks' comparison, it is first necessary to establish what are the factors that need to be compared. What is extremely necessary and important for a DG framework to tackle? For this, it is necessary to understand the common structure of DG activities, DG Decision Domains, DG areas of governance, and DG CSFs. This understanding will permit the choice of the factors that will be used to compare frameworks and build an improved one.

4.2.1. Common Structure of Data Governance Activities

As seen, a DG framework intends to logically structure and communicate complex activities involved with data creation, use, and communication (Thomas, n.d.). However, to structure this "complex activities" it is necessary to follow some specific actions in a kind of management process composed of three steps: 1) **Define**; 2) **Implement**; 3) **Monitor** (Alhassan et al., 2016).

Table 5 – DG activities' phases (Kuan Cheong et al., 2007)

Phase	Definition
Define	To start a DG program it is required that all DG aspects are defined.
Implement	How to put it in practice the defined aspects of the DG program.
Monitor	Creation of management policies and procedures that will incentivize compliance and assertiveness of the ongoing DG program that is constantly changing with the environment.

Even if not considered specific areas of comparison, these actions will have to be tackled by all of them to permit the structure of the improved framework and its implementation path that will directly be influenced by this logical process.

As seen in literature, there is a lack of research around implementation and monitoring phases for DG related activities (Alhassan et al., 2019), which is already an indicator that these areas have room for improvement and as such will be taken in consideration when comparing frameworks and presenting the new artifacts.

4.2.2. Data Governance Decision Domains

Khatri & Brown (2010) defined five decision domains that need to be assessed in order to have a complete spectrum of needs and responsibilities related to data at a corporate level, culminating with the presentation of a framework that explore the type of questions each DG decision domain needs to answer and potential roles that need to be addressed to fulfill those specific data needs (Khatri & Brown, 2010):

Table 6 - Framework for data decision domains (Khatri & Brown, 2010)

Data Governance Domains	Domain Decisions	Potential Roles or Locus of Accountability
Data Principles • Clarifying the role of data as an asset	• What are the uses of data for the business? • What are the mechanisms for communicating business uses of data on an ongoing basis? • What are the desirable behaviors for employing data as assets? • How are opportunities for sharing and reuse of data identified? • How does the regulatory environment influence the business uses of data?	• Data owner/trustee • Data custodian • Data steward • Data producer/supplier • Data consumer • Enterprise Data Committee/Council
Data Quality • Establishing the requirements of intended use of data	• What are the standards for data quality with respect to accuracy, timeliness, completeness and credibility? • What is the program for establishing and communicating data quality? • How will data quality as well as the associated program be evaluated?	• Data owner • Subject matter expert • Data quality manager • Data quality analyst
Metadata • Establishing the semantics or "content" of data so that it is interpretable by the users	• What is the program for documenting the semantics of data? • How will data be consistently defined and modeled so that it is interpretable? • What is the plan to keep different types of metadata up-to-date?	• Enterprise data architect • Enterprise data modeler • Data modeling engineer • Data architect • Enterprise Architecture Committee
Data Access • Specifying access requirements of data	• What is the business value of data? • How will risk assessment be conducted on an ongoing basis? • How will assessment results be integrated with the overall compliance monitoring efforts? • What are data access standards and procedures? • What is the program for periodic monitoring and audit for compliance? • How is security awareness and education disseminated? • What is the program for backup and recovery?	• Data owner • Data beneficiary • Chief information security officer • Data security officer • Technical security analyst • Enterprise Architecture Development Committee
Data Lifecycle • Determining the definition, production, retention and retirement of data	• How is data inventoried? • What is the program for data definition, production, retention, and retirement for different types of data? • How do the compliance issues related to legislation affect data retention and archiving?	• Enterprise data architect • Information chain manager

4.2.3. Data Governance Areas

Adding to the decision domains, it is important to consider what areas are DG comprised of. For this, Alhassan et al. (2016) defined eight DG areas that represent what can be explored in a DG plan for it to be meaningful and assertive (Alhassan et al., 2016):

Table 7 - DG Areas (Alhassan et al., 2016)

DG Areas	Definition
Data Roles and Responsibilities	What are the tasks, roles, and authority needed to be performed?
Data Policies	Establishment of data policies that steer the DG program (primary deliverable).
Data Processes and Procedures	Creation of processes necessary to ensure the correct execution of the DG program.
Data Standards	Establishment of data standards that serve as a basis to measure DQ.
Data Strategy	Determines how to enhance company's competitive position through data while ensuring its security, privacy, integrity, and governance.
Data Technologies	Definition of DG supporting systems (software/ hardware) that will support the DG program and leverage its potential (e.g., process automation).
Data Guidelines	Definition of guidelines for each data decision domain.
Data Requirements	Registration of stakeholders' data needs and mandatory data requirements that will be the basis of the DG program.

4.2.4. Data Governance Critical Success Factors

To better understand the relevance of the dimensions previously mentioned it is important to explore DG CSFs because they will take these areas and bring what are the most important factors for DG in general. According to Alhassen et al. (2019), CSF are areas of activity that should receive continuous and careful consideration from management (Alhassen et al., 2019a), in this case planning, and control activities over data, that are part of most DG definitions (Barham, n.d.; Panian, 2010; Seiner, 2019). Alhassen et al. (2019) consider the existence of seven CSFs of DG that will be presented in order of relevance (from the most relevant to the least relevant) and further explored in [Annex1](#) (Alhassen et al., 2019):

- **CSF#1: Employee data competencies.**
- **CSF#2: Clear data processes and procedures.**
- **CSF#3: Flexible data tools and technologies.**
- **CSF#4: Standardized easy-to-follow data policies.**
- **CSF#5: Established data roles and responsibilities.**
- **CSF#6: Clear inclusive data requirements.**
- **CSF#7: Focused and tangible data strategies.**

Adding to this seven CSFs, this project will also consider an eighth one that was pointed out by Bento et al. (2022), which is “**Communication**”. Communication is tied to all other CSF, all decision domains, and all parts of the DG process since all objectives, processes, or even data needs to be transmitted effectively for a correct execution of what is defined at a top management level (Bento et al., 2022). Since it is a driver so important that its absence will directly affect the implementation of a DG plan, it will also be included here as a CSF.

4.2.5. Comparison Criteria Definition

Considering the points previously explored, it is possible to conclude that the most complete criteria would fall into the CSF of DG. This is because all DG Decision Domains are tackled by one or more of the CSF presented, as well as the DG Areas:

Table 8 – Criteria Definition Assessment (Own Construction)

CSF	DG Decision Domains Tackles	DG Areas Tackled
CSF #1: Employee data competencies	All (all decision domains are executed by people that need to be capable to be involved in data activities)	Focus on Data Roles and Responsibilities due to emphasis on specific employees' responsibilities but is related to all DG areas since all areas need to be modeled to their enterprise's maturity level.
CSF #2: Clear data processes and procedures	All (all decision domains need to have processes and procedures defined)	Data Processes and Procedures.
CSF #3: Flexible data tools and technologies	All (all decision domains will be influenced by the software or hardware choices for the DG plan)	Data Technologies
CSF #4: Standardized easy-to-follow data policies	All (policies should be defined for all DG domains, but on its nature, are more related to Data Principals decision domain)	Data Policies, Data Guidelines
CSF #5: Established data roles and responsibilities	All (all decision domains have roles and responsibilities associated to their field of application)	Data Roles and Responsibilities
CSF #6: Clear inclusive data requirements	Mostly Data Principles but affects all	Data Requirements
CSF #7: Focused and tangible data strategies	Mostly Data Principles but affects all	Data Strategy
"CSF#8": Communication	Mostly Data Lifecycle and Data Principles but affects all	Mostly Data Requirements, Data Roles and Responsibilities, but is present in all of them

To have a more straightforward comparison criteria, and according to the correspondence made in the table above, the comparison criteria chosen to compare the frameworks will follow, **when possible**, the names of DG areas, which are:

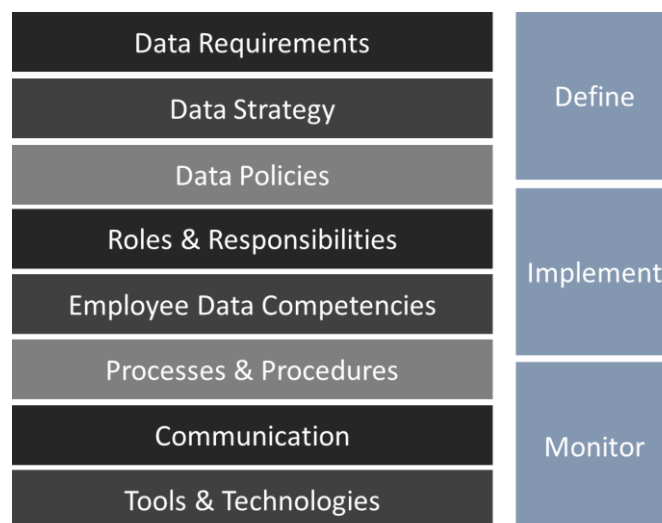


Figure 8 - Comparison Criteria (Own Construction)

Moreover, and because it was identified in the literature that there is a lack of DG focus on Implementation and Monitoring activities (Alhassan et al., 2016), it will be important to maintain the DG common “activities” structure as a parallel criterion on its own. This way the comparison criteria chosen above will ultimately be analyzed into small Define-Implement-Monitor processes to make sure every point is tackled and to facilitate the creation of the framework’s path of implementation.

4.3. FRAMEWORK COMPARISON

To move forward and understand what approach to consider for the creation of the new framework it is important to summarize each framework’s perspectives on each comparison criterion. For each comparison criterion, and accordingly to what was studied in [Annex2](#) it is necessary to schematically understand if the thematic in study is tackled, briefly mentioned, implied but not referenced, or not referenced at all (scale built by the author to fulfill the comparison full potential).

Table 9 - Scale Description for framework comparison (Own Construction)

Conceptual Scale	Description
Tackled	The topic is defined and explained. Definition of concepts and strategies to deploy them. Most attractive perspectives for DG development.
Briefly Mentioned	The topic is mentioned but not further developed. It may only be referenced but is not given enough structure to understand how to clearly deploy it.
Implied	The topic is not directly mentioned, but by definition it is implied by other strategies or decisions made in other topics.
Not Referenced	The topic is not directly or indirectly present. Least attractive perspectives in terms of pertinence.

The following table summarizes the presence of each comparison criterion on each framework on each phase of activities:

Table 10 - Framework Comparison Assessment Table (Own Construction)

Comparison Criteria	Structural Phase	Non-Invasive DG Framework	DGI Framework	SAS DG Framework	DAMA-DMBOK Framework
Data Requirements	Define	Tackled	Briefly Mentioned	Implied	Tackled
	Implement	Tackled	Implied	Implied	Briefly Mentioned
	Monitor	Implied	Implied	Implied	Briefly Mentioned
Data Strategy	Define	Briefly Mentioned	Tackled	Tackled	Tackled
	Implement	Briefly Mentioned	Tackled	Not Referenced	Tackled
	Monitor	Briefly Mentioned	Tackled	Not Referenced	Briefly Mentioned
Data Policies	Define	Briefly Mentioned	Tackled	Implied	Tackled
	Implement	Not Referenced	Implied	Implied	Tackled
	Monitor	Not Referenced	Implied	Not Referenced	Briefly Mentioned
Roles & Responsibilities	Define	Tackled	Tackled	Tackled	Tackled
	Implement	Tackled	Tackled	Implied	Tackled
	Monitor	Briefly Mentioned	Tackled	Implied	Briefly Mentioned
Employee Data Competencies	Define	Implied	Implied	Briefly Mentioned	Implied
	Implement	Not Referenced	Not Referenced	Implied	Not Referenced
	Monitor	Not Referenced	Not Referenced	Not Referenced	Not Referenced
Processes & Procedures	Define	Tackled	Tackled	Tackled	Tackled
	Implement	Tackled	Briefly Mentioned	Briefly Mentioned	Tackled
	Monitor	Briefly Mentioned	Briefly Mentioned	Implied	Tackled
Communication	Define	Tackled	Briefly Mentioned	Tackled	Tackled
	Implement	Briefly Mentioned	Implied	Implied	Tackled
	Monitor	Briefly Mentioned	Not Referenced	Implied	Briefly Mentioned
Tools & Technologies	Define	Tackled	Tackled	Tackled	Briefly Mentioned
	Implement	Tackled	Implied	Tackled	Implied
	Monitor	Briefly Mentioned	Tackled	Tackled	Implied

It is possible to denote that there are improvement points in every framework, and these are mostly towards the “Implement” and “Monitor” phases overall, and in the comparison criterion “Employee Data Competencies”, confirming this research’s pertinence. The framework that tackles the most comparison criteria in most activities’ phases is the DAMA-DMBOK approach. The second most complete framework (using the same system as before) is DGI approach, followed by Non-Invasive approach, and finally SAS approach. This does not necessarily mean that any framework is better than the other, just that it tackles more areas considered critical for DG. Because all of this, it is not possible to immediately choose only one framework to improve with other frameworks’ insights because it may be fallacious. To move forward, first is necessary to understand which approach this “new” framework will follow in each critical area and only after this it will be pertinent to build a framework that better represents what was defined.

4.4. CREATION OF AN IMPROVED FRAMEWORK

In this section, knowledge built in the previous sections will be considered to choose the better approach possible when it comes to each critical area of DG.

4.4.1. Data Requirements

Data requirements represent both an internal and external perspective of data. Internally, they assume the form of **data needs** that must reunite every department's needs to make sense at a corporate level. Externally, they appear as **regulatory requirements** and are extremely important to ensure data compliance with current legislation (Khatri & Brown, 2010). It is also important to start building the data life cycle, as data requirements are the first step to initiate this process (Mosley et al., 2009).

Definition: Considering the perspective of Non-Invasive DG, DGI, and SAS, each department's data manager must reunite, document, and present their business units' data needs and regulatory requirements to be approved by the DG Council as part of the scope of the DG program. It is important to understand what data needs to be governed considering enterprise's mission, goals, and operational execution (SAS Institute Inc, 2018; Seiner, 2019; Thomas, n.d.).

Implementation: The implementation of activities related with data requirements will be addressed here as all activities that will directly involve them. Following the definition part of data requirements, and to ensure consistency in this process a new role can be created to be responsible for the reunion and documentation of all business units' data requirements. In SAS perspective it would be called a data requirements manager (SAS Institute Inc, 2018) but it can also fall into Data Stewardships' responsibilities like in Non-Invasive DG (Seiner, 2019) depending on each company's size and willingness. What is important here is to guarantee the elimination of redundant information and facilitate internal communication by having a central figure that people trust.

Monitoring: To guarantee the continuity of the assertiveness of data requirements and due to the implied necessity of maintaining data requirements updated, it is extremely important to define and document strategies to monitor all activities involving data requirements. These activities must involve getting feedback from employees that work with data in a daily basis to maintain the assertiveness of data needs.

4.4.2. Data Strategy

In this point it is assessed the strategic planning of the DG program through the definition of how it will govern the data defined in data requirements and what are the objectives of the company for its data and its usage. To start, it is important to consider data an important asset of the company by all data stakeholders and most importantly from administrative bodies that will enforce the correct implementation of the program (Panian, 2010; Seiner, 2019).

Definition: From what was seen in the studied frameworks, the approach taken when it comes to what is part of the data strategy is not consensual and differs in level of complexity. The approach chosen will be a mix between DGI and DAMA-DMBOK approaches (DGI, n.d.-b; Mosley et al., 2009). First, it is important to define the strategic mission and vision. For this, it will be considered what was defined by the DGI. **DG Mission** is defined for stakeholders to have a clear understanding of what are

the DG program premises. It is a three-part mission that proactively defines data related rules (define), provides ongoing services to data stakeholders to facilitate the program's execution (implement), and reactively solves any problem that may arise (monitor). This mission will affect every other area of the framework as it works as a guiding rationale (Thomas, n.d.).



Figure 9 - DG three-part mission (Thomas, n.d.)

The **DG Vision** is a guiding statement that clearly declares what is the main enterprise objectives for the DG plan and aims to motivate employees to its correct execution (Thomas, n.d.).

Afterwards, it is extremely important to define **DG Goals, Objectives, and Metrics**. It is important to understand which problems or improvement points DG could fulfill and how to measure their execution. Goals are broad statements that represent what the company wants to achieve in a more strategic way, and must be materialized into SMART objectives (Specific, Measurable, Actionable, Relevant, and Timely) defined from short to long term and must always be revised to maintain pertinence over time (Bjerke & Renger, 2017; Mosley et al., 2009; Thomas, n.d.).

Last but not the least, it is important to define the **DG Implementation Roadmap**, which is a documented process with all the steps of the DG program in order of execution and is extremely important to give guidance and decrease doubts about the program's implementation (Mosley et al., 2009; Thomas, n.d.).

When defining a Data Strategy, it is also important to define **risk management strategies** that involve data privacy and security. This can be preventive (trying to block unwanted events); detective (to understand as quick as possible the existence of any unwanted event), and finally corrective (correct as quick as possible the unwanted event) which afterwards need to be translated into policies and procedures to permit their execution (Thomas, n.d.).

Implementation: Implementation activities must be established to all deliverables defined for data strategy. In the DG Mission, the definition of “ongoing services” is considered an implementation strategy; and the establishment of metrics and the implementation roadmap, due to their practical nature are very implementation oriented.

Monitoring: To complete the cycle, the establishment of a monitoring strategy is key to ensure DG keeps its relevance over time. This is addressed in the “issue resolution” part of the mission and can be controlled through the analysis of metrics timely.

4.4.3. Data Policies

Data Policies are a set of clear rules that permits the execution of DG operationally (Office for National Statistics, n.d.) and, promotes DQ execution (Mosley et al., 2009). They are intrinsically connected with all DG decision domains as they are key to define, document and sustain metadata; and define rules for data accessibility in terms of which data stakeholders can access, edit, or share (Khatri & Brown, 2010; Mosley et al., 2009). Here it is important to have in consideration a semantic “formula” to create a data policy that may represent a guide to their practical creation:

“If we do A, then we should expect B, with a result of C, otherwise, we should expect D, with a result of E.” (Thomas, n.d.)

Definition: To define data policies (applied to all DG decision domains) and considering DGI approach, first it is important to gather already existing data rules/ definition to understand gaps or overlaps in these rules that can be enhanced by aligning them with the data strategy that was previously defined. In the end, it is important to formalize these rules by document them and after approval from the DG council, communicate them to business units (Thomas, n.d.).

Implementation: The implementation of data policies is directly the exercise of DG since they must have an impact in every other DG decision domain, and even other critical areas of DG due to their legislative nature. As a preventive implementation strategy, all data policies should be documented, updated, and accessible by all respective data stakeholders and enforced not only by hierarchical power, but most importantly by everyone involved (Mosley et al., 2009).

Monitoring: The monitoring of data policies is one of the most important parts of DG because they are the ones who ensure DQ, metadata quality, and steer the data life cycle (Khatri & Brown, 2010; Mosley et al., 2009). The definition of strategies to ensure the correct implementation of this policies will boost employees’ confidence on data and thus will also work as a motivation to follow the DG program proactively (Alhassan et al., 2018). This could be in the form of data audits, and most importantly by automatizing some data processes to ensure consistency among datasets. All strategies defined must be documented and reviewed timely.

4.4.4. Roles and Responsibilities

The roles and responsibilities definition are probably the most important part of a DG plan since it is the most referenced thematic in all studied frameworks. Despite all of them having a very complete definition on this topic, none of them present an equal perspective which difficult their comparison. To move forward, it will be considered the roles that coincide between frameworks reaching the indispensable DG roles necessary to deploy the program.

Definition: In this point it is necessary to tackle decision rights (who gets to make decisions and in what situations) and accountabilities (who does what and when inside and between departments) that will form a hierarchical structure (Thomas, n.d.). This structure is formed of roles with specific responsibilities and comes mostly from Non-Invasive DG perspective:

- **Steering Committee/ DG Board:** Panel of executives that have the most hierarchical power and responsible to approve strategic objectives (e.g., DG Strategy) and for its endorsement (SAS Institute Inc, 2018; Seiner, 2019; Thomas, n.d.).

- **Chief Data Officer (CDO):** Leader of the DG Council responsible for approving DG related issues with the steering committee and communicate with the DG Council (SAS Institute Inc, 2018; Seiner, 2019).
- **DG Council:** Enterprise-wide authority over data, normally composed of senior management that represents every business units. They are responsible for the creation and approval of the DG plan and its correct application (Mosley et al., 2009; SAS Institute Inc, 2018; Seiner, 2019).
- **Data Stewardship:** It is represented by the people that are responsible for ensuring that DG plan activities are being well-executed according to the policies and practices previously defined (Thomas, n.d.). A Data Steward is a business unit leader or other employee inside the business unit that is recognized as expert in data and are accountable for these responsibilities (Mosley et al., 2009). Typically, data stewardship is found inside the enterprise's already existing roles and must be given to a person that has some knowledge on DG related issues and express willingness to do the job (Mosley et al., 2009). Depending on the size of the company and its DG goals, Data Stewardship can be divided into three levels:
 - **Executive Data Stewards:** Senior managers who form the DG Council previously defined (Mosley et al., 2009).
 - **Coordinating Data Stewards:** lead and represent business data stewards, having meetings about the implementation of the DG plan (Mosley et al., 2009).
 - **Business Data Stewards:** operational level data stewards responsible for defining and controlling data daily (executers) (Mosley et al., 2009).

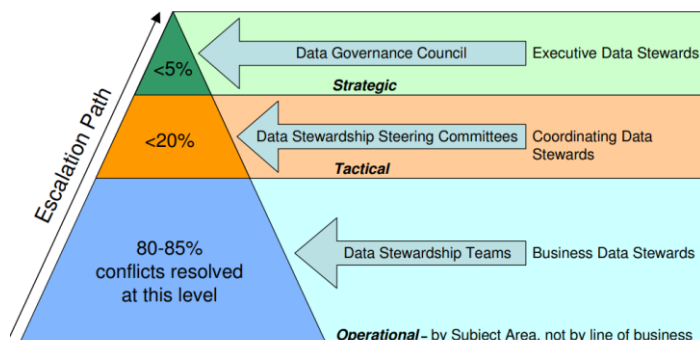


Figure 10 - Data Issue Escalation Path (Mosley et al., 2009)

- **Other important Roles:**
 - **Data Requirements Manager:** Responsible for documenting and updating data requirements with the approval of DG Council (SAS Institute Inc, 2018).
 - **IT Department:** Appear here to evidence their separate role from DG. This way, it is important to document what role will the IT Department play inside the DG plan (Thomas, n.d.).
 - **Other business constituents:** Other data stakeholders without any direct connection to DG activities but since they are in contact with data can have valuable feedback to monitor the DG plan (SAS Institute Inc, 2018).

Implementation: The implementation of data roles and responsibilities is intrinsically related to their definition as responsibilities are defined to each role as guidelines to deploy their jobs. Going

a little bit further, the definition of data processes may also be linked with the implementation of the roles and responsibilities as processes have a clear description of how to deploy each role.

Monitoring: To monitor the correct execution of each role and responsibility it is necessary to have a clear hierarchy of responsibilities that guarantee that each person has someone to report, not only to make sure that the work that person is developing follows the DG program but also for the manager to clarify any doubts the employee may have.

4.4.5. Employee Data Competencies

As seen before, data capabilities are both a CSF of DG and one of the biggest challenges to implement DG (Alhassan et al., 2019; Benfeldt et al., 2020). When it comes to the studied frameworks, this was the least explored area, appearing only in the form of contracting strategies that ensure the best people for each job (Mosley et al., 2009); consideration of data needs (here implying employees have a word to say in these regards); and training to deploy DG activities and facilitate communication (SAS Institute Inc, 2018).

Definition: This area is a direct continuation of the definition of “Roles and Responsibilities”, as it is important to address what is expected from people, to understand how to help them cope with their jobs. This way, it is important to define training strategies to maintain employees’ capable to understand the importance of data; how to deploy their jobs; and to realize that if they assure DQ they will have more trust in data that will enhance their trust in their roles inside the company (motivation). It is also recommended to define motivational activities that are related to the DG program so that employees feel part of it and mitigate the possibility of any misunderstanding of the DG plan itself (Sekhar et al., 2013).

Implementation: Due to its nature, the deployment of employees’ data competencies strategies must include the involvement of Human Resources (HR) to facilitate not only the definition of a relevant and complete strategy when it comes to employees’ training and motivation but also its implementation and documentation.

Monitoring: To keep the previously defined strategies relevant it is important to understand what employees really need in these regards since they are the ones affected by it (Sekhar et al., 2013). They should provide feedback on their personal needs and dissatisfaction with any part of their jobs (in this case related with DG activities) to their direct leader or to HR themselves that must reunite this feedback and update their strategies to meet employees’ needs.

4.4.6. Processes and Procedures

To execute a DG plan, it is essential to first establish **how** people will execute their jobs to achieve the objectives established. For that, it is necessary to define processes that guide people in their DG activities. As seen, all frameworks studied tackled the definition of processes and procedures highlighting their relevance in a DG framework.

Definition: Taking the DGI approach, processes must be defined in a proactive, reactive, and ongoing way, so that they can be standardized, documented, and repeatable to maximize efficiency in their execution (Thomas, n.d.). Following DAMA-DMBOK’s perspective, these processes need to be defined to realize a specific task and need to be documented and communicated to the respective

participants (Mosley et al., 2009). It is important to understand that processes need to be defined to every other DG critical area and decision domain (Seiner, 2019).

Implementation: Due to its practical nature, processes already define their execution on their own. This way, the implementation strategy for processes and procedures is mixed with their definition. However, it is crucial to take this in consideration and foment the creation of practical processes that can be easily implemented in a corporate scenario.

Monitoring: Contrasting with the implementation phase, monitoring strategies must be addressed and defined because they may not be directly referenced in some defined processes. However, they are extremely important to maintain each processes' relevance and correctness as internal and external environments keep changing. This monitoring is recommended to be made by everyone involved guaranteeing their correct completion (Mosley et al., 2009).

4.4.7. Communication

A very well-planned communication strategy can leverage performance and motivation to all stakeholders (Rajhans, 2009). At its core, a DG program has not only data, but also people that will deploy it, and these people need to know what they are doing to do it correctly. In terms of frameworks studied, communication is still very briefly explored in its entirety, being more an implied activity than explicitly stated which can lead some companies not taking them in consideration.

Definition: In this case, the perspectives chosen were the SAS and the DAMA-DMBOK in which communication is defined in a complete, effective, and concise manner with the objectives of monitoring and enforcing conformance by promoting the value of data (Mosley et al., 2009; SAS Institute Inc, 2018). The DAMA-DMBOK perspective suggests the definition of a reporting structure to facilitate communication between hierarchical levels that can also be seen in the Non-Invasive Operating Model of Roles and Responsibilities (Mosley et al., 2009; Seiner, 2019).

Implementation: The implementation of the communication strategy needs to be pre-defined as establishing who communicates what in what situations using what tools is important to push its implementation.

Monitoring: Due to the social nature of communication it may be difficult to monitor. However, the creation of monitoring strategies may facilitate the job of supervisors. For instance, the automation of informative emails helps internal communication, and the development of group dynamics foment interpersonal relationships among employees that will have better relations and thus have better communication (Rajhans, 2009).

4.4.8. Tools and Technology

In this point, supporting systems (software/ hardware) must be addressed to leverage the DG program full potential by automating processes, serving as databases to keep data secure and accessible, or data analysis and presentation tools (Seiner, 2019). Due to the exponential quantity of data that have been created (Statista Research Department, 2022), not touching in a theme like tools and technologies while defining a DG program may be fatal and that is expressed by its incorporation in every framework studied. Either way, all of them consider this point very important, and encourage

the usage of technology to facilitate daily processes in the software development cycle (SDLC) (Planning, Analysis, Design, Implement, Testing & Integration; Maintenance) (BWC, n.d.).

Definition: The definition of what types of tools and technologies to use while trying to deploy a DG program will depend on each companies' objectives and needs. Moreover, to understand if buying a new software/ hardware is the best suitable solution for the company it is recommended to go through a planning process with the help of IT department.

According to the Non-Invasive DG perspective, this process starts by identifying what is intended to be accomplished with the execution of the specific tool. After this, it is important to consider using already existing tools to take advantage of their full potential. If there is nothing internally that suits the specific needs, it must be considered the hypothesis of developing the new tool internally. Only if this step is not possible is that is recommended to the company to purchase a new tool externally (Seiner, 2019).

Implementation: The implementation of tools and technologies will have to be secured not only by DG teams but also by IT teams that will have the necessary knowledge to implement what the company needs businesswise (Mosley et al., 2009).

Monitoring: Following the same ideas from the implementation, it is important to define strategies to monitor the usage of the tools and technologies, and if the investment made on them is worthy or not. For this, it is still important to have both DG and IT departments' insights since one is more business related and the other as more knowledge with the technologies used.

5. RESULTS AND DISCUSSION

5.1. NEW DATA GOVERNANCE FRAMEWORK

To define the new DG framework structure, it was considered not only the most complete overview over a DG plan, but also the applicability and functionality of it. The objective of this project is to maintain the best parts of each framework studied but at the same time prioritizing simplicity to keep it easy to read and apply but maximizing completeness regarding DG matters.

As established before there are two main areas that this framework needs to tackle: Critical areas (main drivers of the framework); and Activities main phases (activity flow drivers). The corresponding approach followed was defined in the previous point.

When it comes to the structure itself, and from what was seen amongst the frameworks studied there are different types of layouts: Matrices (Non-Invasive and DAMA-DMBOK) which are good for detail but have a decreased practicality (forced into two dimensions - axes); organized scheme (SAS) which gives a broader view of the DG program but lacks interconnection; and processual scheme (DGI) that connects every frameworks component to each other. Following this framework's objectives, the chosen layout will be considered a mix between the SAS DG Framework and the DGI DG Framework to guarantee its practicality (both have already been practically used in real-case scenarios). Using the already tested and approved SAS DG framework as a basis and translating the approaches preciously defined it was possible to create the newly adapted framework (SAS Institute Inc, 2018):

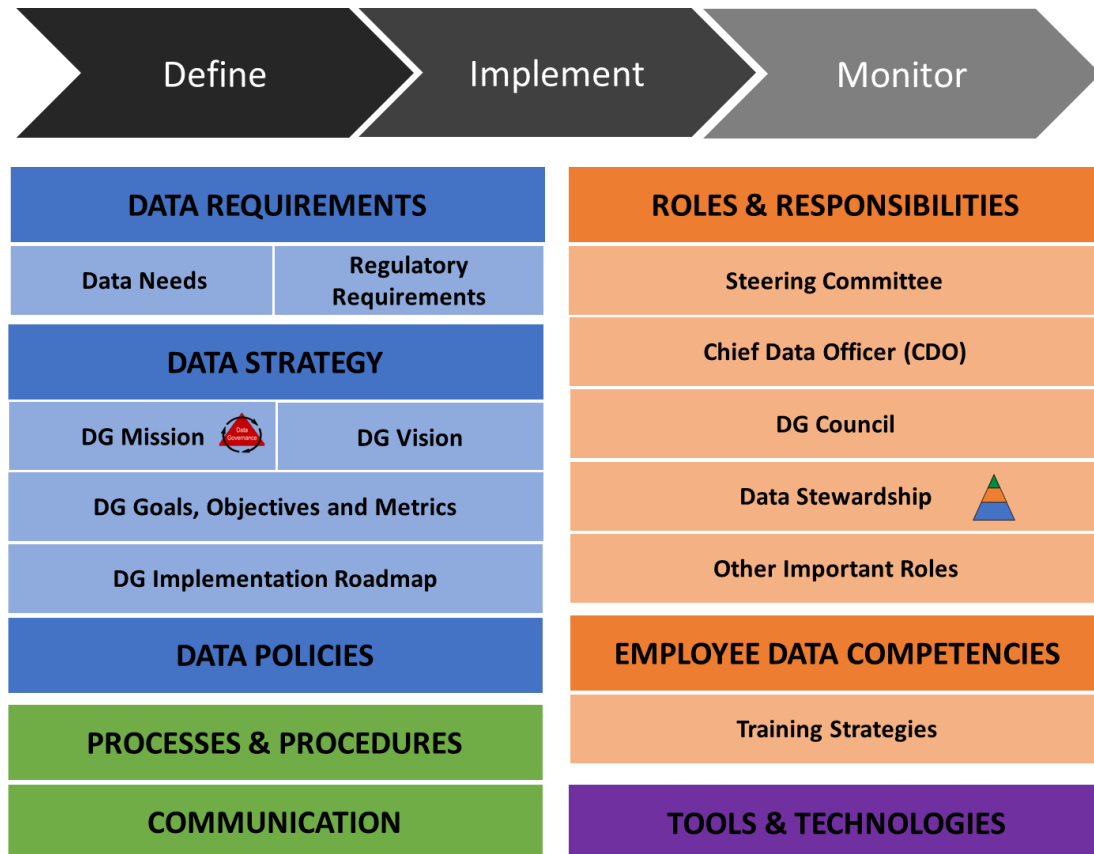


Figure 11 - New adapted DG Framework (Own Construction)

Continuing the logic behind the structure of the studied frameworks, the new one is also divided into similar areas. In this case, the approach used was from the DGI framework (Thomas, n.d.). In the new framework the division was made as follows:

- **Scope and Rules:** Here represented in blue are the **WHY** do the company needs the DG program (DG Mission and Vision) and the **WHAT** is being governed.
- **People & Organizational Bodies:** Here represented in orange is the **WHO** will deploy DG's activities.
- **Processes:** Here represented in green are the **HOW** and **WHEN** will the DG program translate its objectives into actionable processes and procedures.
- **Tools & Technology:** Here represented in purple are the supporting activities that leverage DG performance by tackling operational issues and automatizing tasks.

To emphasize the introduction of both definition, implementation, and monitoring strategies for each component, they were kept as framework's "components". They are not typical components that need to be defined on their own, but rather need to be addressed in all other components to ensure the correct application of the DG strategy.

5.2. IMPLEMENTATION ROADMAP OF THE NEW FRAMEWORK

To foment the correct creation of the DG framework in a real case scenario, it is important to define an implementation roadmap that serve as a guide and a reference to anyone who wants to apply it. In this case, the creation of the implementation roadmap will have as reference, structure wise and content wise, the one from DAMA-DMBOK's perspective because it was the only framework studied with one concretely defined ([Annex2](#)) (Mosley et al., 2009). The idea here, is adapting an already approved roadmap with the approaches previously defined:

Table 11 - New DG Framework's Implementation Roadmap based on (Mosley et al., 2009)

Activities	Deliverables	Responsible Roles	Approving Roles	Contributing Roles
Establish Strategic DG Roles & Responsibilities (P)	Steering Committee, DG Council	CDO, DG Council	Steering Committee	
Understand Strategic Enterprise Data Requirements (P)	Strategic Enterprise Data Needs; Regulatory Requirements	Data Requirements Manager	DG Council, CDO	Data Stewards, Data Users
Develop and Maintain the Data Strategy (P)	Data Strategy – Vision, Mission, Goals, Objectives, Metrics, Implementation Roadmap, Risk Management Strategies	CDO, DG Council	Steering Committee	Coordinating Data Stewards, Business Data Stewards
Identify and Appoint Data Stewards and other operational and support DG Roles & Responsibilities (P)	Business Data Stewards, Coordinating Data Stewards, ...	CDO, DG Council	Steering Committee	
Develop, Review and Approve Data Policies and Standards (P)	Data Policies, Data Standards (Metadata)	CDO, DG Council	Steering Committee	Coordinating Data Stewards, Business Data Stewards
Develop, Review and Approve Data Processes and Procedures (P)	DG Processes and Procedures	CDO, DG Council	Steering Committee	Coordinating Data Stewards, Business Data Stewards
Establish an Internal Communication Strategy (P)	Internal communication strategies	CDO, DG Council, HR Department	Steering Committee	Coordinating Data Stewards, Business Data Stewards
Establish Training Strategies to ensure Employee Data Competencies (P)	Training strategies	CDO, DG Council, HR Department	Steering Committee	Coordinating Data Stewards, Business Data Stewards
Define, Review and Approve what type of issues may be solved by technologies (P)	List of DG issues that may be solved by technologies	Coordinating Data Stewards	CDO, DG Council	Business Data Stewards, IT Department
Understand in what type of tools and technologies makes sense invest (P)	List of tools and technologies to adopt	DG Council, IT Department	CDO, Steering Committee	Coordinating Data Stewards, Business Data Stewards
Coordinate DG Activities (C)	DG Organization Schedules, Meetings, Agendas, Documents, Minutes	Coordinating Data Stewards	DG Council, CDO	Business Data Stewards
Manage and Resolve Data Related Issues (C)	Issue Log, Issue Resolutions	Coordinating Data Stewards, Operational Data Stewards, DG Council		Data Users
Monitor and Ensure Regulatory Compliance (C)	Compliance Reporting, Noncompliance Issues	Data Requirements Manager	DG Council	Coordinating Data Stewards, Business Data Stewards
Communicate, Monitor and Enforce Conformance with Data Policies, Standards, and Procedures (C)	Policy / Standards / Arch / Procedure Communication, Nonconformance Issues	Coordinating Data Stewards, Business Data Stewards	DG Council	Data Users
Implement and attend training sessions to enhance employees DG knowledge (O)	Training attendance registration	HR Department, Data Stewards, Data Users		

(P) – Planning Activities (C) – Control Activities (O) – Operational Activities

5.3. DISCUSSION

Through the research carried out it was possible to answer the proposed questions by tackling the existing problems and completing its proposed objectives. It was possible to create a new DG framework by identifying the critical areas of DG that a framework needs to tackle to be successfully applied to any company's reality, and the development of its respective implementation roadmap to fulfill existing literature gaps previously identified.

To start the development and implementation of the new DG framework, it is important that companies access data as a strategic asset that needs to be managed and maintained, at a corporate level (Panian, 2010). Everyone needs to understand the importance of data and need to be motivated to deploy the new DG program. This maintenance needs to be enforced by executive members that are key to steer the program in the most relevant and pertinent way, and fund it when necessary (Seiner, 2014). The creation of the DG framework is represented in the suggested implementation path, and it is flexible enough to be adapted by every company that wants to design and deploy their DG framework but needs to start by defining the strategic DG roles and responsibilities: the teams that will be responsible to define, approve, and communicate the DG program.

The DG framework translates the DG program, so when defining it, it is recommended that each area of the framework is considered in each activity phase. This process is important to maximize the pertinence of the framework, but in a flexible way, because each area can be emphasized depending on the focus of each project. The consideration of the activities phases on the framework adds value to this approach since it will encourage the assessment of implementation and monitoring activities, tackling the gap identified earlier (lack of representation on existing frameworks) (Alhassan et al., 2019).

The new DG framework builds on knowledge from each existing framework studied, maintaining their strengths in each critical area. Comparing the existing frameworks to the new one, it is possible to denote that the new one tackles more critical areas than the others, fulfilling the identified gap of existing frameworks lacking development in some DG CSFs (Begg & Caira, 2012; Bento et al., 2022). When it comes to disadvantages, the new framework was not tested in a real-case scenario while the studied frameworks are already reference models when it comes to DG application. To bridge this point, the new framework was always created based on the previous ones at every level, from structure, to approaches considered, to maintain its assertiveness when it needed to.

The creation of the implementation path on itself is already an advantage comparing the new framework to the studied ones (except for the DAMA-DBOK approach), but it also encourages not only real-case applications to test it, but also a deeper study on implementation and monitoring activities to strengthen the application of a DG program.

6. CONCLUSIONS, LIMITATIONS AND FUTURE WORKS

This research relevance starts with the gathering and summarization of pertinent information about DG critical areas which are areas that any DG program need to follow to succeed. The implementation of a DG framework is extremely important to clarify a DG program to all stakeholders, leveraging enterprise's decision-making power and ensuring data security, privacy, and quality.

It was found in literature that existing DG frameworks have room for improvement, namely when it comes to critical areas tackle; implementation and monitoring activities definition; and implementation roadmap creation. To understand how to deploy DG to a company through the usage of a new/ adapted framework, it was intrinsically necessary to analyze the existing DG frameworks to understand how they tackle the points identified. This analysis was based on four DG frameworks identified due to their recognition among DG literacy, practicality and flexibility, and level of integrity of areas addressed. Despite considered pertinent, the unit of analysis of four frameworks is a limitation of this project because other DG frameworks can have important insights to strengthen the creation of the enhanced DG framework that are missing. For future works, it is recommended that this analysis is extended to other existing frameworks to extract more value from literature.

Through the frameworks' comparison, it was possible to conclude that all have points of improvement in one or more critical areas; and only DAMA-DMBOK approach objectively defines an implementation roadmap, confirming this research's pertinence. The scale utilized for the framework comparison was made by the author, being another limitation of this project. This decision was taken to leverage the relevance of the specific comparison taken and with the intention to maintain the crossing between critical areas and Define-Implement-Monitor activities since it was not found a scale in literature to sustain this comparison's objectives.

After this comparison it was possible to formulate an improved approach on each critical area that ultimately resulted in the creation of the new DG framework that tackles all critical areas by building on existing frameworks' strengths to maintain its assertiveness, simplicity, and pertinence; and its respective implementation roadmap to clarify its implementation and monitorization by increasing its practicality. Due to the scope of this research, it was not tested in a real-case scenario, which limits this project's applicability, and being strongly recommended for future works to proceed this research to that next step.

As explained, it is not possible to state that any framework is always better than the others due to each company's reality and objective when it comes to DG. However, it is possible to state that compared to the other frameworks, the proposed artifact tackles more critical areas, fulfills a wider spectrum of activities phases, and presents an implementation roadmap that will help guiding its practical implementation. However, and due to the scope of this study, this project is limited to already defined implementation and monitoring activities. Being a under-developed area, it is strongly encouraged that further studies have a deeper understanding on how to objectively create implementation and monitoring activities applied to DG critical areas.

With the creation of the new DG framework and its respective implementation roadmap it was possible to answering the proposed questions and contribute to the DG literacy by understanding how to apply DG to a company using a DG framework created using DG critical areas. All defined objectives were also accomplished by gathering information on DG and DG frameworks, identifying the DG critical

areas, analyzing, and comparing four existing frameworks accordingly to these critical areas, and formulating an optimized perspective on each one that resulted in the creation of the proposed artifacts.

BIBLIOGRAPHICAL REFERENCES

- Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, 49, 424–438. <https://doi.org/10.1016/J.IJINFOMGT.2019.07.008>
- Alhassan, I., Sammon, D., & Daly, M. (2016). Data governance activities: an analysis of the literature. *Journal of Decision Systems*, 25, 64–75. <https://doi.org/10.1080/12460125.2016.1187397>
- 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, pp. 300–316). Emerald Group Publishing Ltd. <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), 98–110. <https://doi.org/10.1080/10580530.2019.1589670>
- Al-Ruithe, M., Benkhelifa, E., & Hameed, K. (2016). A Conceptual Framework for Designing Data Governance for Cloud Computing. *Procedia Computer Science*, 94, 160–167. <https://doi.org/10.1016/J.PROCS.2016.08.025>
- Al-Ruithe, M., Benkhelifa, E., & Hameed, K. (2018). A systematic literature review of data governance and cloud data governance. <https://doi.org/10.1007/s00779-017-1104-3>
- atlan. (2022, September 29). 6 Commonly Referenced Data Governance Frameworks in 2022. <https://atlan.com/data-governance-framework/>
- Barham, H. (n.d.). *Achieving Competitive Advantage Through Big Data: A Literature Review*. Retrieved December 6, 2022, from <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8125459>
- 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, 3–13. www.ejise.com
- Benfeldt, O., Persson, J. S., & Madsen, S. (2020). Data Governance as a Collective Action Problem. *Information Systems Frontiers*, 22(2), 299–313. <https://doi.org/10.1007/s10796-019-09923-z>
- Bennett, T. (2023, March 8). *Exploring 8 Popular Data Governance Frameworks*. Integrate.io. <https://www.integrate.io/blog/popular-data-governance-frameworks/>
- Bento, P. I. P. R., Neto, M., & Côte-Real, N. (2022). *How data governance frameworks can leverage data-driven decision making: a sustainable approach for data governance in organizations*.
- Bjerke, M. B., & Renger, R. (2017). Being smart about writing SMART objectives. *Evaluation and Program Planning*, 61, 125–127. <https://doi.org/10.1016/j.evalprogplan.2016.12.009>

- Black, S., Davern, M., Maynard, S. B., & Nasser, H. (2023). Data governance and the secondary use of data: The board influence. *Information and Organization*.
<https://doi.org/10.1016/j.infoandorg.2023.100447>
- Brocke, J. vom, Hevner, A., & Maedche, A. (2020). *Introduction to Design Science Research*. 1–13.
https://doi.org/10.1007/978-3-030-46781-4_1
- BWC. (n.d.). *Software Development Life Cycle (SDLC)*. Retrieved April 20, 2023, from
<https://bigwater.consulting/2019/04/08/software-development-life-cycle-sdlc/>
- CCI. (2016). *Building a Data-Driven Culture: A Video Learning Series*.
<https://www.careinnovations.org/resources/building-a-data-driven-culture/>
- CMS. (n.d.). *GDPR Enforcement Tracker - list of GDPR fines*. Retrieved November 18, 2022, from
<https://enforcementtracker.com/>
- DELL Technologies. (2021). *The Data Paradox RESEARCH FINDINGS*.
- DGI. (n.d.-a). *Defining Data Governance - The Data Governance Institute*. Retrieved November 20, 2022, from <https://datagovernance.com/defining-data-governance/>
- DGI. (n.d.-b). *DGI Data Governance Framework - The Data Governance Institute*. Retrieved October 29, 2022, from <https://datagovernance.com/the-dgi-data-governance-framework/>
- Holsapple, C. W. (2008). Decisions and Knowledge CHAPTER 2 Decisions and Knowledge. *Handbook on Decision Support Systems 1*, 21–54. https://doi.org/10.1007/978-3-540-48713-5_2
- Karkošková, S. (2022). *Data Governance Model To Enhance Data Quality In Financial Institutions*.
<https://doi.org/10.1080/10580530.2022.2042628>
- Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, 53(1), 148–152. <https://doi.org/10.1145/1629175.1629210>
- Kimachia, K. (2022, September 8). *An overview of data governance frameworks*. TechRepublic.
<https://www.techrepublic.com/article/data-governance-framework/>
- Kuan Cheong, L., Chang, V., & Kuan, L. (2007). Association for Information Systems AIS Electronic Library (AISeL) *The Need for Data Governance: A Case Study Recommended Citation “The Need for Data Governance: A Case Study” The Need for Data Governance: A Case Study*. 999.
<http://aisel.aisnet.org/acis2007/100>
- Ladley, J. (2020). *Data Governance: How to Design, Deploy, and Sustain an Effective Data ... - John Ladley - Google Livros*. https://books.google.pt/books?hl=pt-PT&lr=&id=AkW9DwAAQBAJ&oi=fnd&pg=PP1&dq=data+governance+for+non+data-driven+companies&ots=OO6PKLBHwJ&sig=bSEzG62SOhuKZwN22psKFkS6thU&redir_esc=y#v=onepage&q&f=false
- Mosley, M., Brackett, M., Earley, S., & Henderson, D. (2009). *The DAMA Guide to The Data Management Body of Knowledge (DAMA-DMBOK Guide)*. www.technicspub.com/books.htm

- Office for National Statistics. (n.d.). *Data Principles*. Retrieved April 20, 2023, from <https://www.ons.gov.uk/aboutus/transparencyandgovernance/datastrategy/dataprinciples>
- Oliveira, P., Henriques, P. R., & Henriques, P. (2005). *A Formal Definition of Data Quality Problems. Source code analysis View project Query and Answer system View project A FORMAL DEFINITION OF DATA QUALITY PROBLEMS*. <https://www.researchgate.net/publication/220918803>
- Panian, Z. (2010). Some Practical Experiences in Data Governance. *World Academy of Science, Engineering and Technology*, 62, 939–946.
- Peppers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 24(3), 45–77. <https://doi.org/10.2753/MIS0742-1222240302>
- Rajhans, K. (2009). Effective Organizational Communication: a Key to Employee Motivation and Performance. *InterScience Management Review*, 145–149. <https://doi.org/10.47893/imr.2009.1040>
- SAS Institute Inc. (2018). *The SAS ® Data Governance Framework: A Blueprint for Success*.
- Seiner, R. S. (2014). *Non-invasive Data Governance : The Path of Least Resistance and Greatest Success*. <https://viewer.ebscohost.com/EbscoViewerService/ebook?an=1005063&callbackUrl=https%3a%2f%2fdiscovery.ebsco.com&db=e020mww&format=EB&proflid=eds&lpid=&ppid=&lang=pt&location=https%3a%2f%2fdiscovery.ebsco.com%2fc%2ft73ky3%2fresults%3fq%3drobert%2520s.%2520seiner%26autocorrect%3dy&isPLink=False&requestContext=>
- Seiner, R. S. (2019). *The Non-Invasive Data Governance Framework – The Framework Structure – TDAN.com*. <https://tdan.com/the-non-invasive-data-governance-framework-the-framework-structure/24945>
- Sekhar, C., Patwardhan, M., & Singh, R. Kr. (2013). A literature review on motivation. *Global Business Perspectives*, 1(4), 471–487. <https://doi.org/10.1007/s40196-013-0028-1>
- Statista Research Department. (2022, September 8). *Total data volume worldwide 2010-2025 | Statista*. <https://www.statista.com/statistics/871513/worldwide-data-created/#statisticContainer>
- Thomas, G. (n.d.). *DGI Data Governance Framework - The Data Governance Institute*. Retrieved October 16, 2022, from <https://datagovernance.com/the-dgi-data-governance-framework/>
- 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>
- Whitehead, D., & Maude, P. (2016). *Nursing and Midwifery Research: methods and appraisal for evidence-based practice* (5th ed.). <https://books.google.pt/books?hl=pt-PT&lr=&id=INvWDwAAQBAJ&oi=fnd&pg=PA53&dq=Searching+and+reviewing+the+research+lit>

erature&ots=CTzCcoM_FO&sig=-
M0lAN0eRijfsXpt_Up9wowkwjM&redir_esc=y#v=onepage&q=Searching%20and%20reviewing
%20the%20research%20literature&f=false

ANNEXES

ANNEX 1 – DATA GOVERNANCE CRITICAL SUCCESS FACTORS EXPLORATION

- **CSF #1: Employee data competencies.**

Establishes an employee's capacity to cope with DG activities which are valuable due to their contribution and involvement with DG activities (even if just data users), and impacting the complete DG scope of defining, implementing, and monitoring (Alhassan et al., 2019a). Reinforcing the fundamental DG principal, employees need to understand the importance of data as an important asset of the company (Panian, 2010), and this can only happen if they know why, how, and what they are doing. This will always depend on each one's role and responsibilities. For example, a top-level manager responsible for defining data policies, processes or procedures need to have a high-level understanding of DG, while a data creator needs to understand how to do it with homogeneity and where to store it properly, and a data-user needs to understand what they are looking for and where to look for. To overcome employee data lack of competencies it is recommended to train them continuously with the tools necessary to perform their jobs properly. It is also suggested that this training is made internally (with business-specific related themes), and externally (with DG theory and practical tools) (Alhassan et al., 2019a).

- **CSF #2: Clear data processes and procedures.**

In order to have data's full potential, companies need to have clear data processes and procedure defined for employees to understand what to do in every critical situation, preserve data quality, and cope with legislation. These processes and procedures involve activities related to data flow and use, data quality, data access, and data recording and storage, but can include more areas. To assure this clarity in terms of processes and procedures, it is recommended that they are synthesized and computerized to assure their compliance and execution quality (Alhassan et al., 2019a).

- **CSF #3: Flexible data tools and technologies.**

This CSF is related to how companies store and present their data in the form of software or hardware. This means that to have succeed in terms of DG application, a company must define activities related to dealing with these types of tools once they will inevitably have to use them. In fact, the usage of data tools allows companies to automate processes like data validations that will provide better data quality, meaning it will fuel the best implementation of all DG areas. The usage of technologies to work on data raise some security/ privacy issues that must be foreseen in a way that the company is always prepared to any attack that would affect data integrity (Alhassan et al., 2019a).

To accomplish this CSF, it is key that companies have an appropriate IT infrastructure that can integrate data in a flexible, scalable, and secure way. This means DG must be totally aligned with IT Governance to build the best base for data storage and exploration as possible that will impact not only all DG steps, but also how each company sees data and how they can access it (Alhassan et al., 2019a).

- **CSF #4: Standardized easy-to-follow data policies.**

Data policies are extremely important for establishing a DG program. They are statements/rules that work as guidelines and rules needed for dealing with data in all its decision domains. Well-defining data policies will dictate the success of DG since they will leverage the execution and effectiveness application of data quality, data principles, metadata, data access, and data life cycle. The absence of data policies may cause employees trust in data, affecting decision-making based on data; and can also conduct to an unorganized data structure that will lack privacy control which is a serious problem in terms of data integrity and compliance with legislation leaving the possibility of legal problems, meaning this CSF is directly tied to CSF#6 (clear inclusive data requirements) once it is required to understand data requirements to create effective data policies (Alhassan et al., 2019a).

To achieve this CSF, it is key to define “strong, clear, simple, and easy-to-follow” data policies that should be documented in a certain template to maintain cohesion and facilitating its comprehension by all employees. More than that, and according to what was seen before it is recommended to formulate implementation methods, and here it is always useful to automate processes such as embedding data policies into a system that will facilitate DG control, and finally establish monitoring techniques to maintain data policies updated and relevant (Alhassan et al., 2019a).

- **CSF #5: Established data roles and responsibilities.**

In order to sustain a DG program, just like in every other program, it is necessary to define tasks, roles, and responsibilities. “Who will do what?” is the question that this CSF compiles to give structure to DG as well as fulfilling the “taking account of data” part of DG definition (Alhassan et al., 2019a; Mosley et al., 2009). This CSF comes to clarify the role of DG inside the organization being a critical point to achieve to build a DG framework. Due to its nature of clarifying the DG role in the company, the establishment of roles and responsibilities falls more into the decision domain of data principles (Alhassan et al., 2019a).

According to Alhassan et al. (2019), an essential DG role is the DG committee (Alhassan et al., 2019a) that steers the program by defining, implementing, and monitoring its effectiveness, as well as solving any problem that can arise throughout (Seiner, 2019).

- **CSF #6: Clear inclusive data requirements.**

Data requirements are the data needs of all stakeholders. These requirements are based on business owners detailed perspectives that are passed to IT department that will “translate” them into real applicable data needs to incorporate in the DG program. These needs may arise from internal and external regulators, being tied to legal compliance since they are the main source of external regulations which as seen before are critical to maintain any company. This CSF is directly related to CSF#1 (Employee data competencies) in a way that data requirements need to be communicated and understood by employees, which according to Alhassan et al. (2019) it is still very poorly explored in terms of literature (Alhassan et al., 2019a).

- **CSF #7: Focused and tangible data strategies.**

The last CSF calls back to one of the most important aspects of DG which is considering data as an important asset of any company (Panian, 2010). Since it is important to plan for DG to achieve not only this goal, but to consider short- and long-term objectives that relates to it. By planning what is intended to achieve with the DG program, companies are setting goals and steer the plan to accomplish them, making it a more effective and meaningful way to work DG. As expected, this CSF is directly tied with the data principles decision domain once it is basically its definition in a form of CSF (Alhassan et al., 2019a).

Due to the importance given to establishing data strategies, it is logical that the opposite treatment of poor planning and definition of goals negatively affects the program since its nature is diluted as data tend not to be considered a strategic asset of the company. For this, it is crucial to define the top management committee established in CSF#5 (Established data roles and responsibilities) that will enforce the execution of the program in favor of the completion of its goals. This way, it is recommended that top managers are trained to be more focused on the strategic elements of DG (tying this CSF with CSF#1: Employee data competencies) and to take account of DG program, transmitting this message to other employees (top to bottom) (Alhassan et al., 2019a).

- **“CSF#8”: Communication**

Adding to this seven CSFs, this project will also consider an eighth one that was pointed out by Bento et al. (2022), which is **“Communication”**. Communication is referenced in Alhassan et al. (2019) CSF#6 (Clear inclusive data requirements), has data requirements need to be communicated to employees to be successfully understood. Since it is a driver so important that its absence will directly affect the implementation of a DG plan, it will also be included here as a CSF (Alhassan et al., 2019a; Bento et al., 2022). Communication is tied to all other CSF, all decision domains, and all parts of the DG process since all objectives, processes, or even data needs to be transmitted effectively for a correct execution of what is defined at a top management level. The implementation of communication channels reduces uncertainty and raises stakeholders’ confidence not only on the process but ultimately on data (Bento et al., 2022).

ANNEX 2 – FRAMEWORKS STUDY ON EACH COMPARISON CRITERION

Framework 1 - Non-Invasive DG Framework (Seiner, 2019)

- 1. Data Requirements:** Data needs are addressed by the “**Data**” component of the framework. In this framework, data needs are separated by employees’ roles inside the DG structure (executive level, strategic level, tactical level, operational level, and support level) being a more personalized view to data needs as they are already documented in the framework, facilitating their implementation. It is also responsibility of the DG Council to manage all data needs with the objective of raising overall efficiency and effectivity. This way, the definition, implementation, and monitoring of data needs are planned inside the framework and are part of the responsibilities of the DG Council will enforce their compliance (Seiner, 2019).
- 2. Data Strategy:** In this framework data strategy is only defined in the form of metrics. Metrics play an essential role as they work as inner objectives of each role defined. In this case, metrics are defined per accountability level, having a more clarified way to implement them. The monitoring part of these metrics will be responsibility of each level’s managers (upper levels). In this case implementation and monitoring techniques are briefly defined, but the definition of short- and long-term objectives/ metrics is very narrowly explored (Seiner, 2019).
- 3. Data Policies:** This framework does not explore data policies on its own. Their definition is just briefly mentioned as part of the responsibilities of the administration of the program and its use and control is responsibility of Data Domain Stewards. However, nothing in specific can be said when it comes to the Define-Implement-Monitor process (Seiner, 2019).
- 4. Roles & Responsibilities:** As seen before, roles and responsibilities are one of the most important fields of this framework as it is considered “the foundational component of a successful DG program” (Seiner, 2019). Its structure is composed of four layers of DG responsibilities (Executive, strategic, tactical, and operational) and one extra layer of support (DG team; DG partners).

In a very concise way, the **executive level** is composed of the steering committee and is responsible for high-level decision-making. Most of the time, it is composed of the high-level executives of the business and is represented by the CDO. Ultimately, the steering committee approves the DG charter, strategy, and funding requests. It is also accountable for solving conflicts involving implementing the program among the department’s different opinions and ways of working (Seiner, 2014).

The **strategic level** is composed of the **DG Council**. They are responsible for approving data policies, methods, priorities, and tools; for assigning DG tactical level roles; and for reinforcing the application of the DG program at the tactical level (and, in some extreme cases, operational level too). For this, they need to be skilled in DG issues and understand the business structure and culture. It is generally formed by the heads of each business unit in order to reunite some enterprise-level knowledge when it comes to rules and policy decisions (Seiner, 2014).

The **tactical level** is composed of **data domain stewards** and **data steward coordinators**. These people need to be recognized among their business units since they need to promote objectives and understand/ impose the importance of the DG. They are responsible for identifying data stakeholders and assigning data stewards to resolve data issues. Since these people need to be somehow skilled in terms of data quality, they will also provide resolutions to enterprise-level problems. They will drive the communication between their business unit and the DG program. They are responsible for informing the different stakeholders of

decisions, action items, and scope of work, representing one of the most critical roles of the entire structure (Seiner, 2014).

The **operational level** is composed of the **operational data stewards**. Operational data stewards already exist in the organization. They are the people that work with data daily. Still, they have yet to make any decision on it, being more of a relationship with data and not necessarily a position. They are only responsible for how they use data and register data according to rules and policies defined by higher levels (Seiner, 2014).

The **support level** is represented by the **DG Partners** and the **DG Team**. The DG Partners is usually from Information Technology (IT); Regulatory and Compliance team; Information Security; Project Management Office; and Audit and Legal. As the name implies, DG Partners is responsible for supporting the DG program in their areas of specialization, e.g., the Audit and Legal team will help the data team build and update DG policies and rules around data compliance with current legislation. The DG team is composed of people already in the company who have a portion of their time allocated to work toward the definition, development, and deployment of DG and is responsible for defining, delivering, and sustaining DG program activities (Seiner, 2014).

It is possible to understand that this framework has a very detailed definition of the roles and responsibilities' structure that tackles all the Define-Implement-Monitor process. However, being so detailed and complete may bring some stiffness to the model when it comes to its implementation and monitoring in certain companies which may be a serious problem when it comes to its utilization in a real case scenario.

5. **Employee Data Competencies:** This point is not directly explored in this framework. However, it is recognized the need for skilled employees in certain roles' description like the DG Council that needs to be skilled in DG issues and understand the business structure and culture to properly execute their jobs, but without specifying any measure to ensure employees' competencies (Seiner, 2019).
6. **Processes & Procedures:** Processes are considered the second most important component of this framework. Again, their identification is applied to each employees' role on the roles and responsibilities model. This component is more related to allocate the right people to do the existing processes to obtain the expected result in a standardized and repeatable way rather than defining new DG processes which makes it a little bit different than the other frameworks. This way if any part of the process implies endorsing, enforcing, or authorizing anything related to data activities, it is part of the executive level's responsibilities to solve it. If it implies direction, approval, prioritization, or resolution of any data-related problem, it is part of the strategic level's responsibilities to solve it. On the other hand, if it is necessary to facilitate, mediate, or promote any data-related issue, it is part of the tactical level's responsibilities to solve it. If it is necessary to operate, manage, or handle any data-related issue, it is part of the operation level's responsibilities to solve it. And finally, if the data-related issue involves formalization, adherence, or enforcement, it is part of the support level's responsibilities to solve it (Seiner, 2019).
7. **Communication:** This framework considers Communication as an essential part of DG, being represented by its own component. It affects all other components from data definition to its delivery and must be defined and documented mirroring the roles previously defined. This way, the Non-Invasive DG framework motivates companies to define communication plans per level of responsibilities. The definition part only seems to lack on documenting these plans

that will negatively affect the implementation and monitoring phases, that are only briefly mentioned in some levels of responsibility (Seiner, 2019).

8. **Tools & Technology:** To finalize, tools are also a very important part of Non-Invasive DG framework corresponding to a specific component with its name. In this context, tools are used to improve data knowledge, to comply with legislation, and to record and store data. The framework acknowledges the growing market of technological tools to enhance DG through “big data, smart data, metadata, and all data used for analytics” (Seiner, 2019). Choosing the correct tools will have to have in consideration their practicality, simplicity, and usefulness for DG. This way it is recommended to first assess what is intended to accomplish with the acquisition of the tool, consider using already existing tools to take advantage of existing resources, if not, consider developing something internally to culminate the existing needs, and after that if not possible to accomplish what was intended consider buying a new tool. In these regards, this criterion seems to be very well explored by this framework as it tackles how to define the tools needed, how to implement them and how to monitor their effectiveness by attributing that task to each level responsible for it (depends on the type of toll) (Seiner, 2019).

Framework 2 - DGI Framework (DGI, n.d.-a; Thomas, n.d.)

1. **Data Requirements:** Data requirements are not very explored by the DGI framework. Internal requirements (data needs) are briefly explored in Component #7 (Data Stakeholders) as each stakeholder’s data needs are addressed but only to separate them according to their respective roles and responsibilities, but nothing is defined or documented specifically in terms of data needs. However, while defining goals in Component #2 (Goals, Governance Metrics / Success Measures, Funding Strategies) it is necessary to take in consideration the “4Ps” (Programs, Projects, Professional Disciplines, and People as individuals) showing a certain level of attention to these areas needs as its implied that they will serve as a basis for goals definition. When it comes to external requirements (legislation) it is possible to briefly understand that they are taken into consideration when it comes to Data Rules and Definitions (Component #3) has compliance requirements are part of these rules’ definition. This way, it is possible to conclude that this point is very narrowly explored and only into the “Define” part of the process (Thomas, n.d.).
2. **Data Strategy:** This area is very present in DGI framework as its first component is explicitly the definition of a mission and a vision for DG (Component #1: Mission and Vision) that is materialized by the definition of goals, metrics, and funding strategies in the following component (Component #2: Goals, Governance Metrics/ Success Measures, Funding Strategies). Planning ahead is a primary part of this framework as a mission and vision are developed to give all stakeholders a notion of the scope of the DG program, which is a motivational tool for its good implementation. This mission is built in a three-part mission starting with the definition of rules, implementation of these rules, and problem resolution. As seen, the definition of objectives is also very present in Component #2 with the definition of SMART objectives (Specific, Measurable, Actionable, Relevant, and Timely) to be achieved over time with the aim of increasing revenue, manage costs, and ensuring survival (risk assessment). These two components are very strategic-centered, and while the “Mission and Vision” as a very complete Define-Implement-Monitor process, the “Goals, Governance

Metrics/ Success Measures, Funding Strategies” tend to focus more on the “Define” area, lacking implementation and monitoring strategies (Thomas, n.d.).

3. **Data Policies:** Rules definition appears as one of the three-part mission components of DGI framework (Component #1: Mission and Vision), being represented as an essential part of it. As seen before, it is considered that these “rules” will serve as guidelines for DG activities being treated here as policies. The other two parts of the DG mission are related with the implementation of the defined rules as an “ongoing service” and finally the reactive issue resolution of ongoing problems that can be seen as the monitoring part for data policies. Furthermore, the framework is flexible enough for companies to decide if they want to tackle all this three-part mission or not. To concretize this policies definition, the DGI framework as the Component #3 (Data Rules and Definitions) that takes into consideration already existing rules/ definitions to create new ones based on current gaps or overlaps. Except for the mission definition, that is more strategic driven, it was not expressly found any implementation or monitoring technique to maintain these rules updated. However, the definition of these Data Policies is very important and prominent in this framework (Thomas, n.d.).

4. **Roles & Responsibilities:** This is the most prominent factor of DGI framework as it is present in five components: Component #4 (Decision Rights); Component #5 (Accountabilities); Component #7 (Data Stakeholders); Component #8 (Data Governance Office); and Component #9 (Data Stewards).

First, decision rights are tackled as they intended to understand who gets to make decisions and in what situations, playing a significant part on the roles’ definition process. This decision rights definition needs to be documented and facilitated as necessary that can be seen as a Monitor technique if reviewed regularly.

After defining who has the power to decide on DG related themes, the DGI framework defines accountabilities in Component #5 that will specify who does what and when inside and most importantly a cross-functional departments. To understand if the accountabilities defined are relevant and will have a positive impact on the company it is necessary to implement them, control the implementation, document it and finally prove that the necessity that led to the creation of the specific accountability is satisfied. Despite not specifying any technique when it comes to the last part of the process that is related with the monitoring it, this framework presents a very complete Define-Implement-Monitor process.

In Component #7 (Data Stakeholders) it is established that everyone that uses data are part of the DG process and it is suggested that a DG Board is created that is responsible for providing insights of the program, create data policies, and solve DG-related issues.

Component #8 (Data Governance Office) establishes the creation of a DGO that will be responsible for monitoring the DG process and report to data stakeholders, being the pivotal part of the program to communicate what is defined with stakeholders.

Finally, in Component #9, data stewards are defined as people who assure the correct performance of data-related activities such as ensuring legal compliance and the correct implementation of data policies. A Data Stewardship Council can also be formed of data stakeholders with the objective of creating recommendations about the program to be approved by the DG Board.

This way, the inherent role structure of DGI framework is composed of: DG Board; DGO; Data Stewardship Council; Data Stewards; Other Data Stakeholders (Thomas, n.d.).

5. **Employee Data Competencies:** This point is the least explored in DGI framework as it only may be implied in the form of a “stakeholders care” service that is responsibility of DGO mentioned in Component #8 without any further clarification on how it works (Thomas, n.d.).
6. **Processes & Procedures:** Processes are a very important part of DGI framework as they have their own category and represented by Component #10 (Proactive, Reactive, and Ongoing Data Governance Processes). In this component all processes and procedures should be defined in a standardized, documented, and repeatable way to support data compliance. However, this process definition has very few guidelines when it comes to implementation or monitoring techniques associated with them (Thomas, n.d.).
7. **Communication:** Communication strategies are very briefly explored by DGI framework, being from total responsibility of the DGO defined in Component #8 to communicate what is defined in all terms DG to all stakeholders. This way, it is very dubious what to represent in a Define-Implement-Monitor process since that responsibility is drawn to the DGO (Thomas, n.d.).
8. **Tools & Technology:** This area is first slightly tackled in Component #5 (Accountabilities) where it is possible to understand that, if necessary, the DGI framework encourages the usage of a software to embed data accountabilities into daily processes in the so-called software development cycle (SDLC). However, the most prominent component related to this topic is Component #6 (Controls) where it is necessary to plan to anticipate any privacy/ security event in order to act faster if they happen, in the so-called organization’s risk management strategies. Accordingly, there are three types of organizational risk management strategies: preventive (trying to block unwanted events); detective (to understand as quick as possible the existence of any unwanted event), and finally corrective (correct as quick as possible the unwanted event). There is also four levels of controls: network/ operating system, database, application, and user processes. This way, these controlling strategies follow the Define and the Monitor part of the process but are not very clarifying when it comes to implement them. Adding to this fact, there is no specific mention when it comes to software or hardware usage (Thomas, n.d.).

Framework 3 - SAS DG Framework (SAS Institute Inc, 2018)

1. **Data Requirements:** There is no specific component of the SAS DG framework to tackle data requirements. However, they are present in other components as a need to their correct application and execution. In “Data Security”, regulatory requirements (external) are essential to implement a system that not only complies with them, but also to give confidence to all stakeholders when it comes to data reliability and credibility. In “Data Life Cycle”, data needs (internal) are taken into consideration because they will steer the data life cycle inside the company (from data production to utilization). Data requirements also appear in the form of a specific role in the DG process which is the “Data requirements manager”. Since this role is not developed in any other part, this indicates that data requirements only appear in the DG plan as data providers. Concluding, the definition, implementation, and monitoring of these data requirements are not under DG responsibility but are very important for its success (SAS Institute Inc, 2018).
2. **Data Strategy:** In SAS DG framework, Data Strategy does not appear in the form of its own component but are still very important parts of it since they are part of the objectives and “Guiding Principles”. They first appear here as a required element to ensure this framework’s main objective of avoiding data errors by creating a well-planned DG initiative and more

specifically defining a policy development process (SAS Institute Inc, 2018). When it comes to “Guiding Principles” data strategy will be one of the main topics to have in consideration when it comes to funding requests has, they will be one of the decision points to executive managers (SAS Institute Inc, 2018). Concluding, data strategy is defined in the form of guiding principles, but there is no mention to strategies regarding executing or monitoring those activities.

3. **Data Policies:** In this framework, policies are not isolated in one component. They rather take part in various components as their “definition phase”. For instance, in “Data Security” the first step is to set policies to be put in practice in order to assure that data is safe, or even in “Data quality”, where establishing data policies will ensure that data meets business expectations. This way, data policies are defined in each field separately, and implemented by each part responsible. Since they are separated, data policies may not be cohesive with each other, making it difficult to maintain data cohesion at a corporate level.
4. **Roles & Responsibilities:** This framework as several points where it defines data roles and responsibilities. First, a list of DG stakeholders appears being constituted of: Enterprise DGO; Steering committee; DG Council; Data steward team; Architecture team; Data requirements manager; Metadata manager; Data Quality manager; Security and access manager; and Business Constituents. The responsibilities of each stakeholder are encouraged to be defined using a RACI tool (Responsible; Accountable; Consulted; Informed). From the stakeholders mentioned, the framework explores more in depth the roles and responsibilities of the data steward team, the DG council, and data producers or consumers. The rest of the stakeholders are only responsible for providing feedback from their areas to the DG council and have the right to obtain frequent updates on the DG program (SAS Institute Inc, 2018).

The Data Steward may be divided into four roles: Chief data steward (responsible for managing the Data Steward team, and has the most decision-making power when it comes to stewardship); Business data stewards (Responsible for understanding the usage of data and to enforce data compliance with previously defined policies and objectives within their business area); Technical data stewards (responsible for “translating” business needs into data needs using their technological-oriented knowledge as well as monitoring data quality when it needs to be evaluated in a root-cause investigation); and Working Groups (low-level data stewards that are responsible for data creation and use according to previously defined policies and objectives) (SAS Institute Inc, 2018).

The DG council are responsible for defining high-level policies that will steer the entire DG program and to approve any strategic or high investment required change to it (SAS Institute Inc, 2018).

To conclude, data producers/ consumers are responsible for creating data through the usage of an application or just use data to enable decision-making in other business areas (SAS Institute Inc, 2018).

5. **Employee Data Competencies:** Employees’ competencies are narrowly explored by this framework as they appear inside the “Process” area, under the “Communication” program, where effective employees training is crucial to deploy communication inside the program (SAS Institute Inc, 2018).
6. **Processes & Procedures:** As seen in with the Data Policies, data processes and procedures are dispersed over almost every area of the framework. It first appears in the “Data Architecture”, where the definition of procedures is important to reduce work duplication and leverage employees understanding of their tasks which will led to work optimization. Then, in “Data

Quality” and “Data Security” the procedures definition is a constituent of data quality and data security definition, respectively. In “Data Administration” it is possible to understand that the responsibility for defining data processes and procedures is given to data administration itself. Finally, and most importantly, there is a component specific to “Process” where it only shows two main drivers of process definition which are Measurement and Communication. This way, as happened with data policies, data processes definition may not be cohesive between areas, making it difficult to maintain data cohesion at a corporate level (SAS Institute Inc, 2018).

7. **Communication:** Communication appears inside the Process area of the framework where its importance is exalted. In this regard, the communication strategy should be complete, effective, and concise. Its importance is also highlighted inside the second guiding principle in the phrase “Data governance policies and decisions will be clearly communicated and transparent” (SAS Institute Inc, 2018), however, the theme is not further explored. This way, it is possible to conclude that despite defining a communication strategy, there are not a concrete approach when it comes to their implementation or monitoring, despite being considered a key to the implementation of DG in a company (SAS Institute Inc, 2018).
8. **Tools & Technology:** The SAS framework complies a complete approach to tools and technologies definition and implementation. Inside the “Technology” area it is possible to understand that this framework addresses the technologies as a very important tool to sustain a DG program as the number of data is rising in all companies. It is suggested to include automation tools to perform standardized data activities with the objective of raising data quality and making it easier to monitor it. A little bit further, in the “Solutions” area, the framework offers several “tools” for the most diverse data-related issues that can possibly happen to any company. They are divided into: Data Quality; Data Integration; Data Preparation; Reference Data Management; Master Data; data Profiling & Exploration; Data Visualization; Data Monitoring; Metadata Monitoring; Metadata Management; and Business Glossary (SAS Institute Inc, 2018).

Framework 4 – DAMA-DMBOK Framework (Mosley et al., 2009)

1. **Data Requirements:** In “Primary Deliverables” the main focus is to plan and enforce what is ultimately required from (in this case) DG. Therefore, it is underlined that the data needs and requirements are assessed in here to understand concretely what the expected outcomes from DG activities are. Furthermore, in “Organization and Culture” element it is expressed the inclusion of “Expectations and Attitudes” that may be interpreted has the consideration of what stakeholders expect in terms of data reaching the concept of data needs which are also the first step to implement DG in a company (Table 12). Regulatory compliance is taken in consideration only in the DG application process, not appearing directly on the framework. Data needs are also stated has the first step to implement DG in a company and the development and maintenance of a data strategy the second step, emphasizing their importance has precursors to the creation of a meaningful DG program (Mosley et al., 2009).
2. **Data Strategy:** In this framework data strategy is defined in the form of objectives in the “Goals and Principles” element. In this element it is denoted the vision and mission desired for DG, as well as the strategic goals and specific objectives for the program. Data Strategy is also present in the “Organization and Culture” element with the definition of CSFs and management metrics that need to be updated as monitoring actions over the execution of DG activities occur. In terms of implementation activities, Data Strategy appears as the second step where

it is developed and maintained. In this viewpoint, the strategy is composed of the vision, mission, business goals, objectives, principles, components, metrics, and implementation roadmap (Mosley et al., 2009).

3. **Data Policies:** In this framework, policies are defined in the “Goals and Principles” among the definition of the Data Strategy. In this element guiding principles are defined to steer the program into the completion of the objectives previously defined. Due to the complexity of this task, the DAMA-DMBOK’s framework goes a little bit further in the “Practices & Techniques” and explore the creation of guidelines that represent best practices and common approaches to DG-related issues, as well as considering alternative techniques that stakeholders can use to execute their jobs. Data policies are also part of the implementation process of DG appearing in the sixth step (“Develop, Review and Approve Data Policies, Standards, and Procedures”) as a crucial point to establish guidance to the program (Mosley et al., 2009).

4. **Roles & Responsibilities:** As the name suggests, this comparison criteria is represented in the DAMA-DMBOK framework through “Roles and Responsibilities” element. Here all individual, organizational, and business-related roles are defined, as well as IT relation with DG in the program specific context. It is also one of the first steps to create a DG program here referenced has the establishment of DM professional roles (since DAMA-DMBOK primary focus is DM) followed by the definition and appointment of data stewards (also accenting their importance in this framework). When it comes to the roles directly related with DG, DAMA-DMBOK states that they are defined, communicated, and enforced by the Chief information Officer (CIO) and have separated them into three main “groups”(Mosley et al., 2009):

The **DG Suppliers**, here represented by the business executives, IT executives, data stewards, and regulatory bodies, and are the ones that are directly involved in the creation, execution, and maintenance of DG inputs (business goals, business strategies, IT objectives, IT strategies, data needs, data issues, and regulatory requirements) and are managed by the CIO and the DG Council (Mosley et al., 2009).

The **DG Participants** are the people that despite not being responsible for the definition of DG strategy, rules, or processes, are directly involved in DG activities (e.g., data creation, database updating). In other words, they are the main executors of DG activities, and are represented by executive data stewards, coordinating data stewards, business data stewards, and DM executive. Beyond their extreme importance in execution activities, this group is also a key participant in the maintenance of the DG program with the introduction of their feedback related to the ongoing implementation of DG (Mosley et al., 2009).

The **DG Consumers** are the ones that despite not directly being involved in DG activities, will be related to them through data usage or data creation. They are composed of data producers, knowledge workers, managers, and executives (here represented has the part of them that are not directly involved in the creation and sustenance of the DG program), and customers (Mosley et al., 2009).

To clearly identify the roles and responsibilities of their employees, DAMA-DMBOK created a table with all the steps to create the DG function inside their DM framework that has each step’s “**responsible roles**” (the ones responsible to plan and control the task), the “**approving roles**” (the ones with executive power and responsibility to validate the “responsible roles” proposals), and the “**contributing roles**” (the ones that execute or give feedback about the issue in hands) (Mosley et al., 2009):

Table 12 - DG Process Summary – Adapted from (Mosley et al., 2009)

Activities	Deliverables	Responsible Roles	Approving Roles	Contributing Roles
1.1.1 Understand Strategic Enterprise Data Needs (P)	Strategic Enterprise Data Needs	DM Executive	DG Council, CIO	Data Stewards, DM professionals
1.1.2 Develop and Maintain the Data Strategy (P)	Data Strategy – Vision, Mission, Business Case, Goals, Objectives, Principles, Components, Metrics, Implementation Roadmap	DM Executive	DG Council, CIO	Data Stewards, DM professionals
1.1.3 Establish DM Professional Roles and Organizations (P)	DM Services organizations and staff	CIO	DG Council	DM Executive
1.1.4 Establish DG and Stewardship Organizations (P)	DG Council, Data Stewardship Committee, Data Stewardship Teams	DM Executive, CIO, DG Council	Senior Management	Data Stewards, DM professionals
1.1.5 Identify and Appoint Data Stewards (P)	Business Data Stewards, Coordinating Data Stewards, Executive Data Stewards	DM Executive, Executive Data Stewards	DG Council	Coordinating Data Stewards, DM professionals
1.1.6 Develop, Review and Approve Data Policies, Standards, and Procedures (P)	Data Policies, Data Standards, DM Procedures	DM Executive	DG Council, CIO	Data Stewardship Committee, Data Stewardship Teams, DM professionals
1.1.7 Review and Approve Data Architecture (P)	Adopted Enterprise Data Model, Related Data Architecture	DG Council	DG Council, CIO	Enterprise Data Architect, Data Stewardship Committee, Data Stewards, Data Architects, DM Executive
1.1.8 Plan and Sponsor DM Projects and Services (P)	DM Projects, DM Services	DG Council	DG Council, CIO	DM Executive, DM professionals, Data Stewards
1.1.9 Estimate Data Asset Value and Associated Costs (P)	Data Asset Value Estimates, DM Cost Estimates	Data Stewards	DG Council	DM Executive, DM professionals, Data Stewards
1.2.1 Supervise Data Professional Organizations and Staff (C)	DM Services organization(s) and staff	DM Executive(s)	CIO	DM professionals
1.2.2 Coordinate DG Activities (C)	DG Organization Schedules, Meetings, Agendas, Documents, Minutes	DM Executive, Enterprise Data Architect, Data Architects	DG Council, Data Stewardship Committee, Data Stewardship Teams, CIO	DM professionals
1.2.3 Manage and Resolve Data Related Issues (C)	Issue Log, Issue Resolutions	Data Stewardship Teams, Data Stewardship Committee, DG Council		DM Executive, DM professionals
1.2.4 Monitor and Ensure Regulatory Compliance (C)	Compliance Reporting, Noncompliance Issues	DM professionals	DG Council	DM Executive, CIO
1.2.5 Communicate, Monitor and Enforce Conformance with Data Policies, Standards, Procedures, and Architecture (C)	Policy / Standards / Arch / Procedure Communication, Nonconformance Issues	DM professionals, Data Stewards	DG Council, Data Stewardship Committee	DM Executive
1.2.6 Oversee DM Projects and Services (C)		DM Executive	DG Council	DM professionals
1.2.7 Communicate and Promote the Value of Data and DM (C)	DM Website, DM Newsletter, Understanding and Recognition	DM Executive, DM professionals, Data Stewards, CIO	DG Council	Data Stewards

(P) – Planning Activities (C) – Control Activities

This table will be important to serve as a starting point for the construction of the implementation path of the new framework because despite being very complete in its scope, it is very connected with the DM area which was what the DAMA-DMBOK intended. For this project, since the focus is solely DG it must be revised accordingly.

5. **Employee Data Competencies:** This point is not directly tackled by this framework, however there are some planning when it comes to roles and responsibilities (“Roles and Responsibilities” element) and contracting strategies that a priori will allow the choice of the best person for each job, (“Organization & Culture” element). Furthermore, there is a creased willing of creating ways to facilitate all DG tasks since various processes are accompanied by documented best practices and techniques that are constantly reviewed. In this context it is possible to understand that this criterion, despite not necessarily being directly mentioned, is still managed (Mosley et al., 2009).
6. **Processes & Procedures:** DG-related processes and procedures are assessed in both “Activities” and “Practices and Techniques” elements, as well as being part of one of the activities of the process showed in the previous table (1.1.6 Develop, Review and Approve Data Policies, Standards, and Procedures (P)). In “Activities” element phases, tasks, steps, and processes are defined as well as user case scenarios to facilitate the comprehension of them by all employees. As seen before, in “Practices and Techniques” there is the definition of best practices, common approaches, and alternative techniques to data-related issues which are considered documented processes and procedures, pushing this framework to a very complete perspective on this criterion (Mosley et al., 2009).
7. **Communication:** Directly, it is possible to see some preoccupation when it comes to communication in the element “Organization and Culture” where it is recommended the creation of a reporting structure, as well as the existence of group dynamics where this point is leveraged at a corporate level. Furthermore, in the previous table it is possible to understand that communication is a very serious matter since it is part of various activities in the control part: “1.2.5 Communicate, Monitor and Enforce Conformance with Data Policies, Standards, Procedures, and Architecture (C)” and “1.2.7 Communicate and Promote the Value of Data and DM (C)” (Mosley et al., 2009).
8. **Tools & Technology:** In this framework, this point is tackled in the element “Technology” where it is defined criteria to choose the best technology tools to satisfy the data needs previously defined and fulfill the DG strategy that was underlined. However, this point is presented very briefly and there is no specific strategy when it comes to what types of software and hardware to use (Mosley et al., 2009).