



# Challenges to Implementing Effective Data Governance: A Literature Review

Carlos Alberto Bassi<sup>1</sup> <sup>a</sup> and Solange Nice Alves-Souza<sup>2</sup> <sup>b</sup>

<sup>1</sup>Information System Post-graduation Program (PPgSI), Escola de Artes, Ciências e Humanidades,  
Universidade de São Paulo, São Paulo, Brazil

<sup>2</sup>Departamento de Engenharia de Computação e Sistemas Digitais, Escola Politécnica,  
Universidade de São Paulo, São Paulo, Brazil

**Keywords:** Data Governance, Use Cases, Data Governance Projects, Data Governance Challenges.

**Abstract:** Implementing an efficient data governance program implies overcoming a series of identified challenges. Reviewing the scope of publications of case studies (CS) of Data Governance (DG) projects allowed identifying diverse types of challenges directly related to the data and/or the organization while others suffer from external influences. The market segment and the country of the operation can also influence DG projects. Investigating these challenges to find ways to face them can contribute to the successful implementation of a DG program. From 201 papers initially collected in the process of the literature review, 44 publications presented CS that implement DG projects. As a result, we identify the most impactful challenges for implementation of DG projects that should be prioritized.

## 1 INTRODUCTION

Organizations face challenges and problems in implementing a comprehensive and efficient Data Governance (DG) program. In many cases, there is a lack of knowledge on the part of the professionals involved in conducting DG implementation projects regarding which activities are necessary, who should be responsible for carrying out these activities, what the relationship and dependence are between these activities, as well as the impacts generated by not performing such activities properly.


DG is the comprehensive management of usability, availability, data privacy and security (DPS) and data quality (DQ) inside and outside the organization (Abraham et al., 2019). It includes establishing policies, standards, processes, and structures to ensure the correct use and effective protection of data (Al-Dossari and Sumaili, 2021). DG requires policy specification dynamics that can deal with problems related to the collection, storage, processing, sharing and use, reuse, and disposal of data throughout its life cycle (Filgueiras and Lui, 2022).


According to a McKinsey report, companies spend an average of 30% of their time on tasks with-

out added value due to poor DQ and availability (Zhang et al., 2022). Several studies have found, among other reasons, that many of the bad decisions are mainly due to the poor quality of the information generated from dirty, erroneous and incomplete data. This has led important companies worldwide to lose many thousands of dollars by managing information of low quality in their organization (Castillo et al., 2017).

Many businesses are currently adapting digital strategies and new business models. Studies have shown that organizations of all sizes recognize the need for DG (Lis and Otto, 2020). Increasing importance has been given to incorporating DG as a means of encouraging the strategic use of data, thus promoting data-driven innovation (Lis et al., 2022).

Implementing DG is a complex project that requires long-term commitment and continuous engagement and, as such, organizations usually need to formulate a series of actions towards these goals (Zhang et al., 2022). Mobilizing an organization to adopt DG has proven to be a challenge in practice. Taking stock of data inventory remains tedious, the potential for value creation seems abstract, and the importance of investing in DG is understood only if the company has already suffered major regulatory pressure or data breaches (Benfeldt et al., 2020).

<sup>a</sup>  <https://orcid.org/0009-0001-2524-7478>

<sup>b</sup>  <https://orcid.org/0000-0002-6112-3536>

There is a general lack for a clear understanding of what DG is and how it is currently implemented in companies (Krumay and Rueckel, 2020). A comprehensive review of the scientific and practice-oriented literature shows a lack of understanding the activities required for introducing a DG program (Alhassan et al., 2019b).

Knowing in advance the main challenges and their impacts can allow organizations to plan and allocate the necessary resources and efforts to implement DG, and consequently, reap the benefits of good governance.

This research consists of a literature review to identify the main challenges faced when implementing a DG project, investigating the impacts generated by these challenges and the possible relationship of the challenges with the market segment or country in which the organization operates.

## 2 RESEARCH METHODOLOGY

The challenges that organizations face and their impacts on the implementation of DG were identified by adopting a Scoping Review (SR) of publications that present case studies (CS) on DG projects. As part of the SR, the challenges faced, and their impacts are analyzed and consolidated.

The literature SR was performed according to the procedures defined by (Kitchenham, 2004), which employs the following steps: Planning, Research Execution and Publication Analysis.

Unlike a Systematic Literature Review (SLR), which includes clinical cases, a SR follows the systematization of an SLR, but all the analyses and conclusions refer to information extracted solely from articles selected according to the criteria and questions established for the research.

**Planning / Execution** - At Planning stage, the direction of the research was defined, as well as the research questions, parameters and search bases being formulated, besides the inclusion and exclusion criteria for publications (Turnbull et al., 2023). The research questions addressed were:

1. What are the challenges faced by an organization when implementing a DG project?
2. How do these challenges impact data and organizations?

The initial exploratory research allowed identifying adequate parameters for selecting publications: *Search databases*: Scopus, Web of Science and IEEE Xplore databases; *Period of publications*: 2016 to

2022, to analyze the latest case studies; *Language*: publications in English only and, *Title of publication and Keywords*: should contain the expressions “Data Governance” and “Case Study” or, “Data Governance” and “Project Management”.

At the Execution stage 201 publications were selected. Initially, 40 duplicated publications were excluded. Later, after reading the title, abstract, conclusion and, when necessary, a superficial reading of its content, we rejected publications that did not allow answering the 2 research questions formulated or that did not describe CS in the implementation of DG projects. As a result, 53 publications were selected for full reading and a more detailed analysis.

**Analysis** - At this stage, 9 publications obtained from exploratory research or citations that met the specified research parameters were added and the following were performed for each selected publications: classification of the CS described and collection of relevant information to support addressing the 2 research questions. From these, 44 publications regarded CS on organizations that implemented DG projects of the most varied scopes.

Table 1 presents the evaluation of the publications where the following topics were used for classification: *Data Governance* – addresses aspects of DG, *Case Study* – presents/analyzes a DG implementation CS, *Systematic Review* – performs a systematic review of the literature, *Framework* – presents/analyzes a DG framework, *Maturity Model* – analyzes the level of maturity in DG, *Metrics* – displays/analyzes metrics, *Data Quality* – addresses aspects of DQ, *Governance* – addresses aspects of corporate governance, *Analysis* – addresses aspects of data analysis and *Security* - addresses aspects of data/information security.

The year of publication was considered to ensure that the challenges to be identified represented the most current moment. Figure 1 presents the distribution of the CS by year. 75% of the publications occurred in the last 04 years indicating the current and growing interest in the implementation of DG. A lower number of articles in 2020 may reflect the peak of the COVID-19 pandemic.

The market segment could impose or bring greater complexity to DG implementation. Some publications contained CS in more than one market segment, as occurred in the CS by (Lis and Otto, 2020) that evaluated an engineering and construction industry, a telecommunication services company, and a digital service provider. Figure 2 presents the distribution of CS by market segment. The Government and Healthcare segments stand out in the implementation of DG

Table 1: Evaluation of the publications selected.

	Publication	DG	CS	SR	FR	MM	MT	DQ	GV	AN	SC	EV
01	(Li et al., 2019)	X	X									A
02	(Xiang, 2021)		X						X			R
03	(Varlina, 2017)		X						X			R
04	(Fylan and Fylan, 2021)	X	X									A
05	(Joshi et al., 2021)	X	X									A
06	(Qin et al., 2020)	X	X									A
07	(Gul and Ahsan, 2019)		X							X		R
08	(Park et al., 2019)	X	X									A
09	(Jiang et al., 2022)	X	X									A
10	(Kawtrakul et al., 2021)	X	X									A
11	(Bhardwaj and Singh, 2017)			X					X	X		R
12	(Kalkman et al., 2019)	X	X									A
13	(Jiya, 2021)	X	X									A
14	(Leiber, 2022)	X	X									A
15	(Yulfitri, 2016)	X	X									A
16	(Earley, 2016)	X					X					A
17	(Kurniawan et al., 2019)	X	X			X						A
18	(Pratiwi and Ruldeviyani, 2021)	X	X									A
19	(Wibowo and Sandikapura, 2019)	X	X									A
20	(Saltz et al., 2018)									X		R
21	(Weng and Hirata, 2022)								X			R
22	(Prado et al., 2021)	X	X	X								A
23	(Garcia et al., 2022)	X	X									A
24	(Permana and Suroso, 2018)	X	X			X						A
25	(Meyer et al., 2020)	X	X									A
26	(Carretero et al., 2016)	X	X									A
27	(Robinson et al., 2021)	X	X									A
28	(Fusi et al., 2018)	X	X									A
29	(Jones et al., 2020)	X	X									A
30	(Khan and Johnson, 2020)	X	X						X			A
31	(Lis and Otto, 2020)	X	X									A
32	(Benfeldt et al., 2020)	X	X									A
33	(Cerrillo-Martínez and Casadesús-De-mingo, 2021)	X	X									A
34	(Murtagh et al., 2018)	X	X									A
35	(Ryynänen and Harisalo, 2018)	X	X						X			R
36	(Zhang et al., 2022)	X	X		X							A
37	(Putro et al., 2016)	X	X									A
38	(Krumay and Rueckel, 2020)	X	X	X								A
39	(Mukhrizal et al., 2019)	X	X									A
40	(Linåker and Runeson, 2022)		X						X			R
41	(Alhassan et al., 2019b)	X	X									A
42	(Timotijevic et al., 2022)	X	X									A
43	(Wildenauer and Basl, 2021)	X			X							A
44	(Alhassan et al., 2019a)	X	X									A
45	(Whittard et al., 2022)	X	X	X								A
46	(Dutta et al., 2022)		X			X			X			R
47	(Lis et al., 2022)	X	X									A
48	(Carvalho et al., 2021)		X								X	R
49	(Keller et al., 2021)		X						X			R
50	(Bajjens et al., 2022)		X		X					X		R
51	(Vaghjiani et al., 2017)		X							X		R
52	(Saputra et al., 2018)	X	X		X	X						A
53	(Aisyah and Ruldeviyani, 2018)	X	X		X							A
54	(Castillo et al., 2017) *	X	X		X							A
55	(Riggins and Klamm, 2017) *	X								X		R
56	(Al-Dossari and Sumaili, 2021) *	X	X			X						A
57	(Juddoo et al., 2018) *	X						X				R
58	(Zygmuntowski et al., 2021) *	X	X									A
59	(Filgueiras and Lui, 2022) *	X	X									A
60	(Jim and Chang, 2018) **	X	X									A
61	(Kim and Cho, 2018) **	X	X		X							A
62	(Alhassan et al., 2018) **	X		X								A

**Caption**

**Columns:** ( DG ) - Data Governance, ( CS ) - Case Study, ( SR ) - Systematic Review, ( FR ) - DG Framework, ( MM ) - DG Maturity Model, ( MT ) - Metrics, ( DQ ) - Data Quality, ( GV ) - Corporate Governance, ( AN ) - Analytic, ( SC ) - Security, ( EV ) - Evaluation

**Evaluation of publication:** ( A ) - Publication accepted, ( R ) - Publication rejected

**Source:** ( \* ) - Exploratory Research, ( \*\* ) - Citations

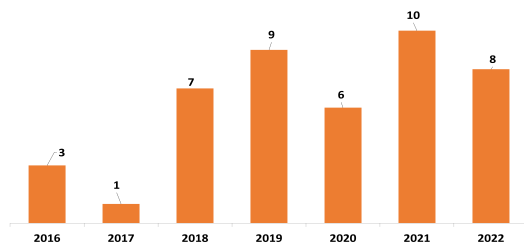


Figure 1: Year of publication of the case studies.

with approximately 40% of the cases. These market segments are highly regulated, data are shared among various organizations and considered to be of public interest, and there is a great concern on the part of citizens and patients regarding the security and privacy of their personal data. Industry and Agricultural segments are also affected by regulations and in recent years citizens have become more concerned with the origin of what they have been consuming.

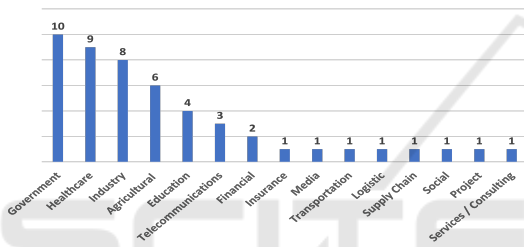


Figure 2: Case studies market segment.

### 3 ANALYSIS OF CHALLENGES

Challenges are related to everything that impacts organizations when pursuing their strategic objectives, e.g., decision-making, obtaining value, the provision of products/services or even regarding their image. These cause direct impacts when not treated or managed, by absence, non-conduction, existence, non-existence, precariousness, problems, or maintenance.

The challenges found in the CS were grouped into 11 categories, and it evaluated the impacts generated in relation to the points: *Governance* – DG, *Quality* – DQ, *Use* – data use, *Collaboration* – data collaboration, *Security* – DPS and *Organization* – about the organization.

Table 2 presents the challenges referenced in the publications related for each category and the Figure 3 presents the occurrences in the CS.

**Data** - definitions of data emphasize its role in representing facts about the world. Businesses use data to understand their customers, create new products and services, and improve operational efficiency by cutting costs and controlling risks (Data Management

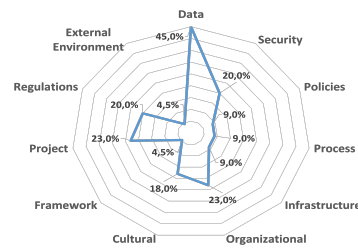


Figure 3: Occurrence in the case studies.

Association, 2017).

Figure 4 presents the number of occurrences of the challenges *Data* regarding the impacts generated. The existence of data silos, the lack of identification of responsibility for data and a clear view of master data impact the implementation of DG. Low DQ will have a significant impact as it affects decision-making.

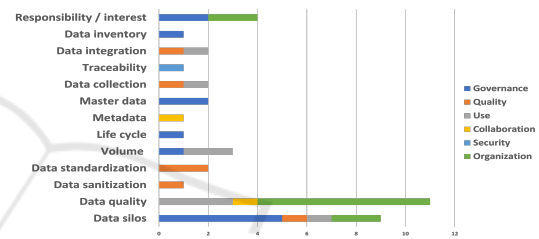


Figure 4: Impacts of challenges Data.

**Data Silos:** data stored in different places, with replication in different individual repositories that are not interconnected, or integrated, and are often inaccessible to many in the organization.

**Data Quality:** data need to be accurate, complete, consistent, reliable, current, and relevant so that they can serve a certain purpose and are essential to ensure assertive business decision-making and valuable information. Ensuring that the data were not intentionally or unintentionally manipulated, as well as its latency, the time elapsed between the generation or collection of the data and its availability for effective use, impact decision-making.

**Data Sanitization:** process to identify, correct and remove errors, replications and inconsistencies in the data processed for improving its quality.

**Data Standardization:** process that seeks to transform data to ensure their compatibility, consistency, use and integration between different applications using rules and transformations in their format, nomenclature, abbreviations, and units of measurement.

**Volume:** number of data collected, stored, and processed that must be managed in a certain period. This volume has a direct impact on the organization management and on the costs involved in processing, storage, security, and distribution.

Table 2: Reference of the category of challenges in the publications.

Publi- cation *	DT	SC	PL	PC	IF	OZ	CT	FM	PJ	RG	EE
01					X						
04	X	X									
05	X										
06	X	X									
08	X										
09	X										
10	X										
12		X								X	
13											
14											
15											
17											
18				X							
19	X										
22						X		X	X		
23		X									
24	X						X				
25	X	X				X	X		X	X	
26											
27	X										
28						X	X		X		
29											
30											
31	X	X		X	X	X			X	X	X
32	X				X	X	X		X	X	
33	X									X	
34		X								X	
36	X					X			X		
37											
38					X	X	X		X		
39											
41									X		
42	X	X								X	
44											
45	X						X		X		X
47				X		X	X			X	
52											
53	X		X			X			X		
54	X										
56			X								
58								X		X	
59	X		X							X	
60							X				
61	X	X	X	X		X			X		
Number of occu- rences	20	9	4	4	4	10	8	2	10	9	2
<b>Caption</b>											
(DT) - Data: data											
(SC) - Security: data security											
(PL) - Policies: specification or application of policies for implementing DG											
(PC) - Process: work processes necessary for implementing DG											
(IF) - Infrastructure: existing/necessary infrastructure to process the data											
(OZ) - Organizational: structure of the organization or the organization itself											
(CT) - Cultural: cultural or behavioral aspects in relation to data or its governance											
(FM) - Framework: adoption of frameworks for implementing DG											
(PJ) - Project: conducting the DG implementation project											
(RG) - Regulations: laws or regulations existing in the country/segment of operation											
(EE) - External Environment: environment in which the organization is inserted that escape the context of previous challenges											
(*) - The publication number is the same as the publication number shown in Table 1											

*Life Cycle*: describes the various stages that data go through since their creation, collection, validation, processing, analysis, storage, distribution up to their disposal. Knowledge of this cycle is relevant for better management, meeting the organization’s objectives and defining the most appropriate policies at each stage.

*Metadata*: relevant information that provides context and meaning to the data, helping to understand their origin, format, content, relationships, and usage.

They are used to manage, locate, classify, integrate, and share data throughout their life cycle.

*Master Data*: set of data identified as a unique and accurate reference of essential and critical information for an organization. Data are centrally managed, which helps in ensuring integrity, accuracy and consistency throughout systems and applications.

*Data Collection*: process of gathering data from various sources and in different formats for processing, treatment, analysis and sharing, to guarantee that the organization has accurate and reliable information to support its operations and decision-making.

*Traceability*: ability to follow the history of data from its origin to its current state, by recording and monitoring related events and activities. It is essential to identify the root cause of a problem and take the necessary actions to solve it, in addition to knowing what processing and treatments were performed.

*Data Integration*: process for combining data from multiple sources to create a single, unified, consistent view for access throughout the organization. Data integration can be accomplished with various techniques and the sources can be internal and/or external to the organization.

*Data Inventory*: detailed listing of all the organization data, including location, format, owner, usage, quality, and other relevant information. Inventory enables to understand their data assets and to manage them effectively.

*Responsibility / Interest*: obligation of a person or group of persons to ensure the quality, security, privacy, compliance, monitoring, and proper use of the data under their responsibility.

**Security** - set of practices and measures that aim to guarantee the protection of information against unauthorized access or disclosure, alteration, destruction, or theft, involving the implementation of technical, physical, and administrative security measures.

Figure 5 presents the number of occurrences of the challenges *Security* regarding the impacts generated. Confidentiality and privacy, as well as the need to obtain consent, have a major impact on the use and processing of data considering the security aspect. DG will be impacted in countries where there is great concern or regulation about data privacy and confidentiality, such as the European Union and the United Kingdom.

*Consent*: process for obtaining an individual’s explicit permission for his/her personal data to be collected, processed, and stored by the organization or by third parties. This person must be informed how, when, why and by whom their personal data will be used and, whether he/she agree with how they will be



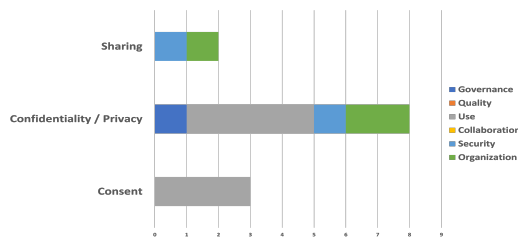


Figure 5: Impacts of challenges Security.

treated.

**Confidentiality / Privacy:** people’s right to decide what personal data and who may access these. Organizations are responsible for ensuring the privacy and confidentiality of the personal data they manage and must implement the necessary measures to protect from unauthorized access, theft, loss, or unauthorized sharing.

**Sharing:** process of transferring data between two or more people or organizations for the most diverse reasons, such as collaboration, service execution, decision making, analysis or research. Sharing must be performed in a safe, transparent, ethical manner and in accordance with applicable authorizations and laws.

**Policies** - set of principles, guidelines, rules, procedures, standards, and practices that establish the expected behavior and responsibilities that professionals in an organization must observe (Wies, 1994).

Figure 6 presents the number of occurrences of the challenges *Policies* regarding the impacts generated. The need to establish DG policies and DPS policies are impacting factors when DG is implemented. The Government segment has been heavily impacted by the need to specify DG, DPS and DQ management policies.

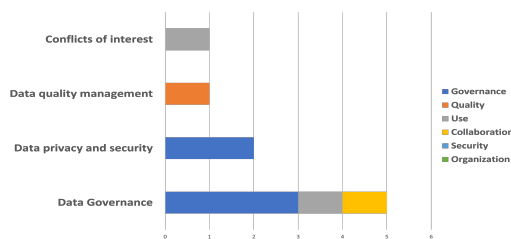


Figure 6: Impacts of challenges Policies.

**Data Governance:** the DG policy must deal with aspects such as identifying owners, defining responsibilities, managing the life cycle, standardizing collection processes, storage, distribution to ensure the efficient use of data.

**Data Privacy and Security:** the privacy policy defines the rules for how the collection, use, storage and shar-

ing of data should occur in compliance with the regulations and rights of data subjects. The security policy establishes the requirements to guarantee the confidentiality, integrity and availability of data and includes technical, organizational and risk management measures to prevent, detect and respond to threats and incidents.

**Data Quality Management:** set of guidelines, objectives, and standards to guarantee DQ at all stages of its life cycle, defining responsibilities, procedures, and criteria to continuously evaluate and guarantee this quality.

**Conflicts of Interest:** occurs when an individual or organization has conflicting interests regarding the data it uses or manages.

**Process** - set of structured and interrelated activities that transform inputs into outputs, aiming to achieve a specific objective. Processes can be sequential, parallel, or iterative, and can involve people, technology, information, and other resources (Association of Business Process Management Professionals International, 2019).

**Data Quality Control and Monitoring:** processes for verifying, validating, and continuously reviewing data to ensure their accuracy, integrity, consistency, compliance, and ensure that they are reliable in decision-making.

**Degree of Harmonization:** alignment and standardization between processes to ensure consistency and efficiency in conducting the organization’s business activities.

**Variations in Processes:** different ways of performing activities in a business process caused by the most varied factors such as personal, cultural, or regulatory.

**Infrastructure** - set of capabilities and technologies needed to store, manage, and protect data. This includes the hardware, software, networks, and services necessary to ensure data is securely and efficiently stored and managed.

**Compatibility Across Platforms and Standards:** ensuring interoperability, integration and data security in different software platforms, systems and devices following the same standards and protocols.

**Use of External Infrastructure:** use of infrastructure external to the organization.

**Alignment with Application Architecture:** ensuring that data architecture and applications are designed, developed, and implemented consistently with the technology and systems infrastructure.

**Fragmented Architecture with Legacy Systems:** data are stored in different Information Technology (IT) systems that may have been developed with different

technologies.

*Standardized Big Data Systems:* use of systems aimed at handling large volumes of data, ensuring that all data are collected, stored, processed, and analyzed according to the same principles and standards.

**Organizational** - it is about organizing companies in terms of activities and resources needed to fulfill their strategic objectives.

Figure 7 presents the number of occurrences of the challenges *Organizational* regarding the impacts generated. DG is impacted when there are localized practices, non-alignment/collaboration between units and, due to the non-existence of a unit responsible for conducting DG. Data collaboration is affected when there is a search for competitive advantages.

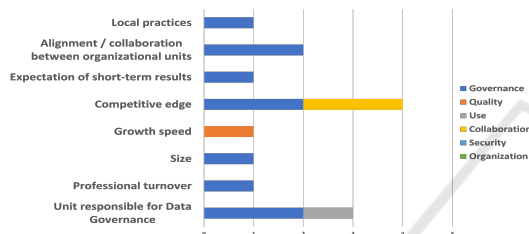


Figure 7: Impacts of challenges Organizational.

*Unit Responsible for Data Governance:* unit with attribution and responsibility for managing and monitoring DG. It must be multidisciplinary, that is, composed of professionals from different areas, such as IT, legal, compliance, risk, and business.

*Professional Turnover:* professionals who leave the organization and are replaced by new ones generate impacts on the retention of business knowledge and on the need for training in practices, policies, procedures, and processes.

*Size:* size and capacity of a company in terms of financial, organizational, human, and technological resources.

*Growth Speed:* rate at which an organization expands its revenue, size, profit, or number of employees. The pace of growth can be affected by factors such as competition, the market, investments in technology, human resources, and culture.

*Competitive Edge:* any characteristic, resource, or strategy that a company has and that differentiates it from its competitors, allowing it to obtain a competitive advantage that can be based on the quality of the product or service, technological innovation, operational efficiency, and customer service.

*Expectation of Short-Term Results:* goals and objectives related to sales, profit, customer satisfaction, or market growth that an organization expects to achieve in a short period of time.

*Alignment / Collaboration Between Organizational Units:* ability of different units to work together towards a common goal, sharing information and resources efficiently and effectively to meet defined goals and objectives.

*Local Practices:* refers to the use of specific procedures and policies by a given unit, regardless of whether they are compatible with practices adopted globally by the organization. This may arise from cultural, regulatory or market differences, making it necessary to adapt policies and procedures.

**Cultural** - set of values, beliefs and behaviors that define how the organization conducts its business and how it treats its customers and partners.

Figure 8 presents the number of occurrences of the challenges *Cultural* regarding the impacts generated. The lack of perception of the value of data as an asset and the lack of understanding and training of those involved in the concepts, technologies and best practices will significantly impact the implementation of DG in organizations from different market segments and countries.

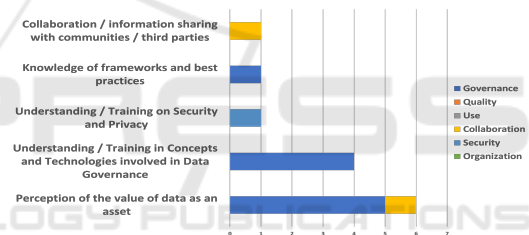


Figure 8: Impacts of challenges Cultural.

*Perception of the Value of Data as an Asset:* recognition that data are a valuable and strategic resource, capable of generating competitive advantage and of adding value to the business.

*Understanding / Training in Concepts and Technologies Involved in Data Governance:* mastering and understanding concepts, technologies, methodologies, and practices involving DG.

*Understanding / Training on Security and Privacy:* mastering and understanding the concepts, technologies, methodologies, and practices involving DPS.

*Knowledge of Frameworks and Best Practices:* mastering and understanding the concepts, methodologies, and practices of models for implementing DG.

*Collaboration / Information Sharing with Communities / Third Parties:* practices involved in the exchange of information and data between companies, government institutions or other communities.

**Framework** - it is developed at different levels of abstraction to provide a range or perspectives on how to

approach data management. These perspectives provide insight that can be used to clarify strategy, develop road maps, organize teams, and align functions (Data Management Association, 2017).

*Adapted to the Needs of the Organization:* use of a DG framework that can meet the specific needs of the organization, considering its business objectives, structure, culture, and size.

*Capacity to Promote Data Sharing:* the DG framework must be able to assist in data sharing, establishing rules, policies, requirements, and processes in a safe and effective way.

**Project** - a temporary endeavor undertaken to create a unique product, service, or result. The temporary nature of projects indicates the beginning and end of project work or a phase of project work (Project Management Institute, 2021).

Figure 9 presents the number of occurrences of the challenges *Project* regarding the impacts generated. The success of the DG implementation project is impacted by factors, such as the lack of engagement and commitment by all those involved, the non-existence or poor performance of a management committee, not carrying out a project approach for the entire organization, as well as a lack of understanding the activities involved in running a DG program.

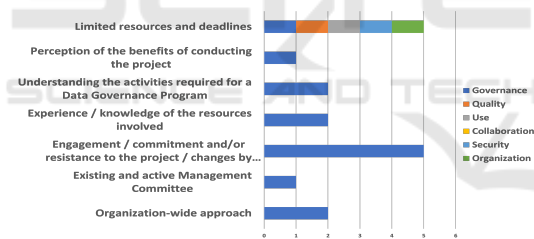


Figure 9: Impacts of challenges Project.

*Organization-Wide Approach:* adoption of an approach that seeks to involve all units and professionals to guarantee success in the implementation of a project.

*Existing and Active Management Committee:* importance of having an active management committee to ensure that DG policies and guidelines are effectively implemented, monitored, updated, and followed by all those involved.

*Engagement / Commitment and/or Resistance to the Project / Changes by Those Involved:* need to engage and obtain commitment from all those involved in addition to mitigating obstacles due to resistance to the project and organizational changes that may be generated by the implementation of a project.

*Experience / Knowledge of the Resources Involved:* technical and practical knowledge regarding DG con-

cepts, technologies and best practices by the human resources involved increases the probability of successful adoption of DG.

*Understanding the Activities Required for a Data Governance Program:* knowledge of the activities specified in the DG program and the responsibilities of each of those involved are critical to the success of the project.

*Perception of the Benefits of Conducting the Project:* depending on how the project is presented, the organization does not identify the benefits that can be obtained with the implementation of DG.

*Limited Resources and Deadlines:* the organization must be aware of the limited resources and deadlines it has so that it can plan and define priorities in the implementation of DG.

**Regulations** - they are rules established by government agencies or regulatory entities that aim to guarantee the safety, quality and efficiency of the activities and resources necessary for the fulfillment of the strategic objectives of the companies.

Figure 10 presents the number of occurrences of the challenges *Regulations* regarding the impacts generated. DPS regulations, as well as regulations that affect the use and processing of data in the public interest and transparency, impact DG. The Government and Healthcare segments are eventually impacted by existing regulations on the treatment and use of data. Countries that have been applying regulations in DPS, in the public interest and in data processing impact the implementation of DG.

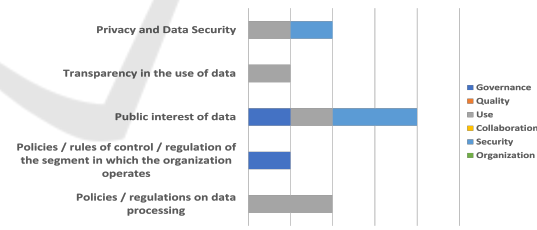


Figure 10: Impacts of challenges Regulations.

*Policies / Regulations on Data Processing:* establish rules, procedures, and responsibilities to ensure that data processing is performed in an appropriate, legal, ethical, and secure manner. They must cover aspects such as data collection, storage, sharing, access, protection, and disposal, in addition to defining roles and responsibilities in data processing.

*Policies / Rules Of Control / Regulation of the Segment in Which the Organization Operates:* regulations established by regulatory bodies and government authorities for the segment in which the organization operates.



*Public Interest of Data:* concept that certain data are of importance to society as a whole and, therefore, must be treated to guarantee their availability and accessibility to all interested parties, highlighting those data related to public health, national security, environment, and education.

*Transparency in the Use of Data:* the practice of ensuring that everyone involved understands how data are collected, processed, and shared.

*Privacy and Data Security:* regulation aimed at protecting individuals' personal and sensitive data, as well as ensuring that companies that collect, store, process and share these data do so in an ethical and legal manner.

**External Environment** - external factors that can influence the performance and activities of the company. These factors include political, economic, social, technological, and environmental aspects.

*Political and/or Institutional Instability:* frequent and unpredictable changes in the political or institutional environment of a country/region in the economic, social, political, or legal spheres that directly affect organizations.

*Political Support:* need for political support to deal with regulatory or legal issues, as well as to overcome institutional or bureaucratic obstacles that may arise in a country or market segment.

## 4 DISCUSSION AND TECHNOLOGY

The related challenges were those identified in the analysis of the data collected from the CS within the scope of this research. Other challenges may be faced by organizations depending on a plethora of factors, from the way in which the project was performed, the level of maturity in DG that the organization is at, the applicable laws and regulations, the financial and technological resources available and, the importance of data to the business.

**Most Impacting Challenges** - The analysis indicates that certain challenges are more relevant and more referenced in the research CS. These challenges generate different impacts, are present in different market segments and do not depend on the country.

Some of the most impacting challenges observed are the existence of data silos (Aisyah and Ruldeviyani, 2018), the need for DQ, the need to guarantee DPS, the perception of the value of data being an asset (Benfeldt et al., 2020), the limitation of resources and deadlines in conducting a DG implementation project and the data being a differential of the organization

before its competitors (Lis et al., 2022).

When conducting a DG implementation project, these challenges should be treated as a priority in view of the impacts they generate and the benefits that can be reaped when solved.

**Influence of the Segment** - The market segment in which the organization operates influences the challenges to be overcome when implementing a DG project. In general, these are highly regulated segments that require transparency in data processing, need to cover an extensive and dispersed geographic area, and involve different profiles of people with the most different socioeconomic levels, including Government and Healthcare.

Some of the challenges influenced by the segment of activity are the relevance of the DQ treated, the DPS guarantee, greater engagement (Whittard et al., 2022), and commitment of all those involved, establishment of clear DG policies, understanding and/or training in the concepts and technologies involved in DG (Krumay and Rueckel, 2020), the attribution of well-defined responsibilities (Cerrillo-Martínez and Casadesús-De-mingo, 2021) and the occurrence of political and/or institutional instabilities.

**Data / Organization / Environment** - The challenges faced when conducting DG implementation projects can be classified according to the characteristics of the data and/or organization or the organization's operating environment. Depending on the complexity, some challenges can even be classified in both situations.

*Challenges Related to Data/Organization* - are those inherent to the characteristics of the data or the organization itself, such as data silos, DQ, volume, life cycle, data inventory, confidentiality/privacy, alignment with application architecture, staff turnover, size, speed of growth, perceived value of data as an asset, understanding / training on concepts and technologies involved in DG and, limited resources and deadlines. These challenges are not influenced externally by the market segment and should be prioritized when implementing a DG project, as it depends only on the organization.

*Challenges Related to the Environment* - are those directly influenced by factors external to the organization, such as consent, sharing, use of external infrastructure, policies / regulations in data processing, policies / regulations for control / regulation of the organization's segment of activity, public interest in data and political and/or institutional instabilities. These challenges are more complex to be dealt with because the organization often does not have control

or power to act.

## 5 CONCLUSIONS

The research performed from a literature SR allowed identifying the most common challenges faced by organizations when implementing DG projects. Many of these challenges are independent of the environment in which the organization operates and are influenced by the characteristics of the data or by the organization itself.

Previous knowledge of these challenges helps organizations that need to implement effective governance of their data to better plan and take the necessary actions to mitigate the impacts that these challenges generate.

The results of the review presented here will support the specification of a guide with practical and coordinated actions that help organizations in the implementation of DG, overcoming the main challenges and, thus, evolving in their maturity and obtaining the desired benefits for an effective governance of their data. The follow-up of this work will be presented in another article as soon as the research is completed.

## REFERENCES

- Abraham, R., Schneider, J., and vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, pages 424–438.
- Aisyah, M. and Ruldeviyani, Y. (2018). Designing data governance structure based on data management body of knowledge (DMBOK) framework: A case study on indonesia deposit insurance corporation (IDIC). In *International Conference on Advanced Computer Science and Information Systems*, pages 307–312. IEEE.
- Al-Dossari, H. and Sumaili, A. A. (2021). A data governance maturity assessment: A case study of saudi arabia. *International Journal of Managing Public Sector Information and Communication Technologies*, 12:19–30.
- Alhassan, I., Sammon, D., and Daly, M. (2018). Data governance activities: a comparison between scientific and practice-oriented literature. *Journal of Enterprise Information Management*, 31:300–316.
- Alhassan, I., Sammon, D., and Daly, M. (2019a). Critical success factors for data governance: a telecommunications case study. *Journal of Decision Systems*, 28:41–61.
- Alhassan, I., Sammon, D., and Daly, M. (2019b). Critical success factors for data governance: A theory building approach. *Information Systems Management*, 36:98–110.
- Association of Business Process Management Professionals International (2019). *BPM CBOK - BPM Common Body of Knowledge*. Association of Business Process Management Professionals International, version 4.0 edition.
- Baijens, J., Huygh, T., and Helms, R. (2022). Establishing and theorising data analytics governance: a descriptive framework and a vsm-based view. *Journal of Business Analytics*, 5:101–122.
- Benfeldt, O., Persson, J. S., and Madsen, S. (2020). Data governance as a collective action problem. *Information Systems Frontiers*, 22:299–313.
- Bhardwaj, A. and Singh, W. (2017). Systematic review of big data analytics in governance. In *International Conference on Intelligent Sustainable Systems*, pages 501–506. International Conference on Advanced Computer Science and Information Systems.
- Carretero, A. G., Freitas, A., Correia, R., and Piattini, M. (2016). A case study on assessing the organizational maturity of data management, data quality management and data governance by means of MAMD. In *International Conference on Information Quality*, pages 72–81. MIT Information Quality Program.
- Carvalho, A. P., Canedo, E. D., Carvalho, F. P., and Carvalho, P. H. P. (2021). Anonimisation, impacts and challenges into big data: A case studies. In *Lecture Notes in Business Information Processing*, volume 417, pages 3–23. Springer Science and Business Media Deutschland GmbH.
- Castillo, L. F., Raymundo, C., and Mateos, F. D. (2017). Information architecture model for data governance initiatives in peruvian universities. In *9th International Conference on Management of Digital EcoSystems*, volume 2017-January, pages 104–107. Association for Computing Machinery, Inc.
- Cerrillo-Martínez, A. and Casadesús-De-mingo, A. (2021). Data governance for public transparency. *Profesional de la Información*, 30.
- Data Management Association (2017). *DAMA-DMBOK - Data Management Body of Knowledge*. Data Management Association, 2nd edition.
- Dutta, A., Roy, R., and Seetharaman, P. (2022). An assimilation maturity model for it governance and auditing. *Information and Management*, 59.
- Earley, S. (2016). Metrics-driven information governance. *IT Professional and Management*, 18:17–21.
- Filgueiras, F. and Lui, L. (2022). Designing data governance in brazil: an institutional analysis. *Policy Design and Practice*.
- Fusi, F., Manzella, D., Louafi, S., and Welch, E. (2018). Building global genomics initiatives and enabling data sharing: Insights from multiple case studies. *OMICS A Journal of Integrative Biology*, 22:237–247.
- Fylan, F. and Fylan, B. (2021). Co-creating social licence for sharing health and care data. *International Journal of Medical Informatics*, 149.
- Garcia, R. D., Ramachandran, G. S., Jurdak, R., and Ueyama, J. (2022). A blockchain-based data governance with privacy and provenance: A case study for

- e-prescription. In *IEEE International Conference on Blockchain and Cryptocurrency*. Institute of Electrical and Electronics Engineers Inc.
- Gul, R. and Ahsan, A. (2019). Big data and analytics: Case study of good governance and government power. In *European Conference on Intangibles and Intellectual Capital*, pages 128–134.
- Jiang, R., Wang, W., Xie, Y., and Yin, X. (2022). Research and design of infrastructure monitoring platform of intelligent high speed railway. In *6th Information Technology and Mechatronics Engineering Conference*, pages 2096–2099. Institute of Electrical and Electronics Engineers Inc.
- Jim, C. K. and Chang, H. C. (2018). The current state of data governance in higher education. In *Association for Information Science and Technology*, volume 55, pages 198–206. John Wiley and Sons Inc.
- Jiya, T. (2021). Responsible data governance in projects: Applying a responsible research and innovation (RRI) framework. *Journal of Technology Management and Innovation*, 16:31–37.
- Jones, K., Daniels, H., Heys, S., Lacey, A., and Ford, D. V. (2020). Toward a risk-utility data governance framework for research using genomic and phenotypic data in safe havens: Multifaceted review. *Journal of Medical Internet Research*, 22.
- Joshi, D., Pratik, S., and Rao, M. P. (2021). Data governance in data mesh infrastructures: The saxo bank case study. In *21th International Conference on Electronic Business*, volume 21, pages 599–604. AIS Electronic Library.
- Juddoo, S., George, C., Duquenoy, P., and Windridge, D. (2018). Data governance in the health industry: Investigating data quality dimensions within a big data context. *Applied System Innovation*, 1:1–16.
- Kalkman, S., Mostert, M., Udo-Beauvisage, N., Delden, J. J. V., and Thiel, G. J. V. (2019). Responsible data sharing in a big data-driven translational research platform: lessons learned. *Medical Informatics and Decision Making*, 19.
- Kawtrakul, A., Chanlekha, H., Waiyamai, K., Kangkachit, T., D’Orazio, L., Kotzinos, D., Laurent, D., and Spyrtatos, N. (2021). Towards data-and-innovation driven sustainable and productive agriculture: BIO-AGRI-WATCH as a use case study. In *International Conference on Big Data*, pages 3407–3415. Institute of Electrical and Electronics Engineers Inc.
- Keller, J., Burkhardt, P., and Lasch, R. (2021). Informal governance in the digital transformation. *International Journal of Operations and Production Management*, 41:1060–1084.
- Khan, Z. T. and Johnson, P. A. (2020). Citizen and government co-production of data: Analyzing the challenges to government adoption of VGI. *The Canadian Geographer*, 64:374–387.
- Kim, H. Y. and Cho, J.-S. (2018). Data governance framework for big data implementation with nps case analysis in korea. *Journal of the Academy of Business and Retail Management Research*, 12:36–46.
- Kitchenham, B. (2004). Procedures for performing systematic reviews. Keele university. technical report tr/se-0401, Keele University, Department of Computer Science, Keele University, UK.
- Krumay, B. and Rueckel, D. (2020). Data governance and digitalization – a case study in a manufacturing company. In *Pacific Asia Conference on Information Systems Proceedings*, pages 6–22. AIS Electronic Library.
- Kurniawan, D., Ruldeviyani, Y., Adrian, M. R., Handayani, S., Pohan, M. R., and T, R. K. (2019). Data governance maturity assessment: A case study in it bureau of audit board. In *International Conference on Information Management and Technology*, pages 629–634.
- Leiber, T. (2022). Performance data governance and management of learning and teaching in higher education: the sqelt project. *Quality in Higher Education*, 28:6–10.
- Li, Q., Lan, L., Zeng, N., You, L., Yin, J., Zhou, X., and Meng, Q. (2019). A framework for big data governance to advance rhins: A case study of china. *IEEE Access*, 7:50330–50338.
- Linåker, J. and Runeson, P. (2022). Collaboration in open government data ecosystems: Open cross-sector sharing and co-development of data and software. *International Conference on Electronic Government*.
- Lis, D., Arbter, M., Spindler, M., and Otto, B. (2022). An investigation of antecedents for data governance adoption in the rail industry - findings from a case study at thales. *IEEE Transactions on Engineering Management*.
- Lis, D. and Otto, B. (2020). Data governance in data ecosystems - insights from organizations. In *Americas Conference on Information Systems Proceedings*. AIS Electronic Library.
- Meyer, A., Bannister-Tyrrell, M., Mackenzie, C., Stegeman, A., and Cameron, A. (2020). Barriers to the adoption of a fish health data integration initiative in the chilean salmonid production. *Computers and Electronics in Agriculture*, 179.
- Mukhrizal, F. F., Aziz, N. A., and Akhir, E. A. P. (2019). Data governance practice for outbound supply chain management (SCM): A case study of malaysian textile industry. *International Journal of Supply Chain Management*, 8:965–970.
- Murtagh, M. J., Blell, M. T., Butters, O. W., Cowley, L., Dove, E. S., Goodman, A., Griggs, R. L., Hall, A., Hollowell, N., Kumari, M., Mangino, M., Maughan, B., Mills, M. C., Minion, J. T., Murphy, T., Prior, G., Suderman, M., Ring, S. M., Rogers, N. T., Roberts, S. J., Straeten, C. V. D., Viney, W., Wiltshire, D., Wong, A., Walker, N., and Burton, P. R. (2018). Better governance, better access: practising responsible data sharing in the METADAC governance infrastructure. *Human Genomics*, 12.
- Park, J., Yamashita, T., Takada, A., Nojiri, C., Izukura, R., Nohara, Y., Hotta, T., Kang, D., and Nakashima, N. (2019). Effect of governance functionality for data standardization management of the medical information database network project. In *International Medi-*

- cal Informatics Association*, volume 264, pages 1562–1563. IOS Press.
- Permana, R. I. and Suroso, J. S. (2018). Data governance maturity assessment at PT. XYZ. case study: Data management division. In *International Conference on Information Management and Technology*, pages 15–20. Institute of Electrical and Electronics Engineers Inc.
- Prado, R., Prado, E. P. V., Grotta, A., and Barata, A. M. (2021). Benefits of the enterprise data governance in industry: A qualitative research. In *23rd International Conference on Enterprise Information Systems*, volume 2, pages 699–706. Science and Technology Publications.
- Pratiwi, M. and Ruldeviyani, Y. (2021). Data governance to improve data quality for statistical process control (SPC): A case study PT. XYZ. In *International Conference on Software Engineering and Information Management*, pages 99–105. Association for Computing Machinery.
- Project Management Institute (2021). *PMBOK - Project Management Body of Knowledge*. Project Management Institute, seventh edition.
- Putro, B. L., Surendro, K., and Herbert (2016). Leadership and culture of data governance for the achievement of higher education goals (case study: Indonesia university of education). In *International Seminar on Mathematics, Science, and Computer Science Education*, volume 1708. American Institute of Physics Inc.
- Qin, H., Liu, X., Qin, H., and Zhu, J. (2020). How to be an enabler of digital transformation for media organizations? In *International Conference on E-business, Management and Economics*, pages 153–156. Association for Computing Machinery.
- Riggins, F. J. and Klammer, B. K. (2017). Data governance case at krausemcmahon LLP in an era of self-service BI and big data. *Journal of Accounting Education*, 38:23–36.
- Robinson, C. J., Kong, T., Coates, R., Watson, I., Stokes, C., Pert, P., McConnell, A., and Chen, C. (2021). Caring for indigenous data to evaluate the benefits of indigenous environmental programs. *Environmental Management*, 68:160–169.
- Ryynänen, S. P. and Harisalo, R. (2018). A strategic and good governance perspective on handling patient complaints. *International Journal of Health Care Quality Assurance*, 31:923–934.
- Saltz, P. J., Goul, M., Armour, F., and Sharda, R. (2018). Key management and governance challenges when executing data science / analytics projects. In *Twenty-Fourth Americas Conference on Information Systems*, pages 1–4.
- Saputra, D. A., Handika, D., and Ruldeviyani, Y. (2018). Data governance maturity model (DMG2) assessment in organization transformation of digital telecommunication company: Case study of PT telekomunikasi indonesia. In *International Conference on Advanced Computer Science and Information Systems*, pages 325–330.
- Timotijevic, L., Carr, I., Cueva, J. D. L., Eftimov, T., Hodgkins, C. E., Seljak, B. K., Mikkelsen, B. E., Selnes, T., Veer, P. V., and Zimmermann, K. (2022). Responsible governance for a food and nutrition e-infrastructure: Case study of the determinants and intake data platform. *Frontiers in Nutrition*, 8.
- Turnbull, D., Chugh, R., and Luck, J. (2023). Systematic-narrative hybrid literature review: A strategy for integrating a concise methodology into a manuscript. *Social Sciences & Humanities Open*.
- Vaghjiani, K., Oppermann, I., Moon, J., Rogers, E., and Stacey, H. (2017). A case study of lean data driven innovation in government. In *Proceedings of the 5th International Conference on Innovation and Entrepreneurship*, pages 170–177. Academic Conferences and Publishing International Limited.
- Varlina, V. (2017). Big data utilization in the dynamic governance concept case study of Indonesian society's shifting consumption pattern. *Atlantis Press*, pages 26–29.
- Weng, Y. H. and Hirata, Y. (2022). Design-centered HRI governance for healthcare robots. *Journal of Healthcare Engineering*, 2022.
- Whittard, D., Ritchie, F., Musker, R., and Rose, M. (2022). Measuring the value of data governance in agricultural investments: A case study. *Experimental Agriculture*, 58.
- Wibowo, S. and Sandikapura, T. (2019). Improving data security, interoperability, and veracity using blockchain for one data governance, case study of local tax big data. *International Conference on ICT For Smart Society*.
- Wies, R. (1994). Policy definition and classification: Aspects, criteria, and examples. In *International Workshop on Distributed Systems: Operations & Management*.
- Wildenauer, A. and Basl, J. (2021). Unlocking the full potential of building information modelling by applying the principles of industry 4.0 and data governance such as COBIT. *Workshop on Intelligent Computing in Engineering*.
- Xiang, H. (2021). The role and application of big data technology in decision-making management of colleges and universities. In *5th International Conference on Data Science and Business Analytics*, pages 141–147. Institute of Electrical and Electronics Engineers Inc.
- Yulfitri, A. (2016). Modeling operational model of data governance in government case study: Government agency X in Jakarta. In *International Conference on Information Technology Systems and Innovation*. Institute of Electrical and Electronics Engineers Inc.
- Zhang, Q., Sun, X., and Zhang, M. (2022). Data matters: A strategic action framework for data governance. *Information and Management*, 59.
- Zygmuntowski, J. J., Zoboli, L., and Nemitz, P. F. (2021). Embedding European values in data governance: A case for public data commons. *Internet Policy Review*, 10.