A Maturity Model for Data Governance in Decentralized Business Operations: Architecture and Assessment Archetypes

Vítor Hugo Machado Ribeiro
University of Coimbra, CISUC, DEI
Coimbra, Portugal *vhribeiro@dei.uc.pt*

João Barata
University of Coimbra, CISUC, DEI
Coimbra, Portugal *barata@dei.uc.pt*

Paulo Rupino da Cunha
University of Coimbra, CISUC, DEI
Coimbra, Portugal *rupino@dei.uc.pt*

Abstract

Organizations increasingly participate in inter-organizational partnerships that exploit business opportunities supported by shared data assets. Hence, data governance is required to establish collaborative operations between the partners, ensure accountability for shared data assets, define data ownership, identify data provenance, and comply with data-related regulations. This paper presents (1) the structure of a data governance maturity model for inter-organizational operations and (2) a set of maturity assessment archetypes for data governance. These results emerge from a research partnership with a major European technology and service provider involved in data collaboration ecosystems for digital and green logistics. Our contribution extends the state-of-the-art on distributed data governance, specifically for increasingly common business ecosystems built on shared data processing, and provides practical tools for organizations to conduct a data governance maturity assessment tailored to their role in such collaborative operations.

Keywords: Inter-Organizational Data Governance, Maturity Assessment Archetypes, Maturity Model.

1. Introduction and Motivation

Nowadays, data is one of the most valuable assets, requiring the definition and implementation of a data governance framework to provide "*the structure and guidelines necessary for the responsible use and management of data assets*" [24]. Additionally, data governance also allows organizations to develop innovative data-based products, foster data quality practices, and address data-related regulations [1].

However, business partnerships increase the complexity of data governance. Therefore, intra-organizational data governance boundaries are spanning [13], requiring organizations to implement inter-organizational data governance mechanisms that enable collaboration between the partners, establish data security requirements, define data ownership [26], and consider the data-related regulations affecting the partners in distinct locations [1].

Data governance maturity assessment can provide organizations with the means to conduct an initial diagnosis of their practices and develop a plan to improve and implement data governance [16]. However, existing data governance maturity models focus on a single organization scenario and are unsuitable for inter-organizational setups [21]. Our literature review has not found maturity models focusing on decentralized operations context, considering the multiple partners and mechanisms that are necessary to address inter-organizational operations. Furthermore, adopting a maturity model shared among different partners (e.g., data producers and data consumers) is not common practice, requiring the identification of adoption archetypes (defined as patterns of action).

We performed our research in collaboration with a European technology and service provider (TSP) involved in a 91.9M€ investment program to "*create an ecosystem of 28 products and services for the Green and Digital Transition in the transport and logistics sectors*". The company is developing new solutions based on artificial intelligence, using and producing significant volumes of data shared in a business ecosystem, including smart seaports and shipping companies. We defined two Research Objectives (RO):

- RO1: *Design the architecture of an inter-organizational data governance maturity model (IO-DGMM).*
- RO2: *Identity data governance maturity assessment archetypes.*

The remainder of this paper is structured as follows. Section 2 provides common ground on data governance, inter-organizational data governance, data governance maturity models, data governance, and maturity assessment archetypes. Subsequently, Section 3 describes the research methodology. Section 4 presents the architecture proposed for the IO-DGMM, followed by the maturity assessment archetypes. Section 5 shows a demonstration at the TSP company, and section 6 offers a discussion. Finally, Section 7 reports on the main conclusions, focusing on limitations and the next stages of our research.

2. Theoretical Background

2.1. Data Governance

Data governance focuses on defining, deploying, and exercising control and authority over data [5]. Thus, data governance is part of the organization-wide program that aims to capitalize on data as a key asset, improve data quality, and perform data-related risk management [1], [5]. Many organizations are developing frameworks for data governance implementation, including data standards, policies, processes, strategies, and organizational structures supporting data-related operations [5]. Furthermore, data governance involves managing contractual agreements, addressing the identified issues, and monitoring organizational performance [1]. Lastly, it considers the data itself as well as the systems (e.g., data collection tools and data analysis suites) that interact with it [14].

Data governance covers multiple dimensions, considering quality, security, life cycle, architecture, metadata, as well as storage and infrastructure [1]. More specifically:

- Data quality: defines a data quality strategy, appoints data quality roles and responsibilities, collects quality metrics, and manages data quality issues [5].
- Data security: defines data security processes and standards, appoints roles and responsibilities, and complies with data security regulations [17].
- Data life cycle: details processes (e.g., data collection, data analysis), analyzes data flow, and assesses the accomplishment of data-related regulations [5], [17].
- Data architecture: identifies enterprise data requirements, defines architectural policies, and develops an enterprise data model [5].
- Metadata: defines metadata standards, appoints metadata management roles and responsibilities, and follows a metadata management process [5], [17].
- Data storage: establishes processes to manage data storage, data storage policies, and planning of data storage requirements [2].

Data governance depends on the organizational context. On the one hand, at the intra-organizational level, there is a need to ensure the alignment between the business and strategic objectives, promote internal data management practices, and improve data quality [1]. On the other hand, inter-organizational operations involve multiple business partners, requiring mechanisms to ensure data integration, secure data access, and common data exchange processes, subsequently detailed in the next subsection.

2.2. Inter-Organizational Data Governance

Specific data governance mechanisms are required when progressing from intra to inter-organizational operations [13]. There is the need to handle multiple data sources that may lead to increased inconsistency in data quality [1], align the distinct business partner's strategies [13], improve inter-organizational collaborative processes [19], deal with the diverse regulations affecting the multiple entities [18], and promote data exchange [26].

Therefore, "*trusted frameworks*" are required to support secure data-sharing operations [14]. Further, there are three distinct profiles of organizations:

- Data producers: create data assets that may be provided to other entities [6], [15].
- Data consumers: exploit data to develop products and services [6], [15].
- Data prosumers: fit both profiles, simultaneously sourcing and producing data and participating in various data exchange operations [15], [36].

The several inter-organizational data governance dimensions require specific dynamics when considering multiple collaboration partners, as described by [28]. As an example, a digital business ecosystem involves multiple partners with distinct roles, responsibilities, and contributions to the partnership, considering the producers (e.g., provide resources to the ecosystem), drivers (e.g., define a common vision for the ecosystem), and guardians of interest (e.g., assess the ecosystem's members for reliability) [34]. Table 1 details the inter-organizational data governance dimensions and corresponding mechanisms.

Table 1. Inter-Organizational Data Governance Mechanisms.

Dimension	Example of Mechanisms
Inter-Organizational Data Security	Establish secure remote data access mechanisms [27] and ensure non-repudiation for data-related operations [14].
Inter-Organizational Data Architecture	Deploy mechanisms that foster data integration from multiple partners [4] and develop an architectural model for the inter-organizational setup [7].
Inter-Organizational Metadata	Define inter-organizational metadata standards [17] and develop processes for decentralized metadata curation [23].
Inter-Organizational Data Ownership and Stewardship	Develop data ownership models [18] and appoint inter-organizational data stewards [33].
Inter-Organizational Data Life Cycle	Develop data exchange standards [27] and monitor data usage across multiple partners [18].
Inter-Organizational Data Privacy	Test shared data assets for personal identifying information [18] and ensure compliance with data privacy-related regulations affecting the various partners [1].

Organizations should focus on inter-organizational data security, inter-organizational data architecture, inter-organizational metadata, inter-organizational data ownership and stewardship, inter-organizational data life cycle, and inter-organizational data privacy [28]. The available data governance frameworks focus on a single enterprise scenario, lacking the mechanisms to address inter-organizational setups (e.g., defining data ownership) [13].

2.3. Data Governance Maturity Models

A maturity model provides a "*sequence of discrete maturity levels for a class of processes in one or more business domains, and represents an anticipated, desired, or typical evolutionary path for these processes*" [3], [32]. For example, distinguishing between 1-initial up to 5-optimized operations provides an overview of the organization's status and helps define a target for improvement [11, 29]. Moreover, maturity models can be used to compare the organization's capabilities with best domain practices and benchmarks.

Lower maturity levels are often associated with unpredictable processes, ad-hoc procedures to manage data, lack of data assets tracking, and reactive behavior in handling issues and problems [10]. At higher maturity levels, organizational processes are frequently automated, and enterprise-wide rigorous, documented, and improved procedures for data management focus on continuous improvement. The costs associated with data-related operations are also more controlled and optimized at higher maturity levels [10].

Our literature review on data governance maturity models found contributions focusing on specific contexts, such as cloud management [2] and telecommunications operations [33]. We also found more generic proposals, such as the IBM Data Governance Maturity Model [10]. Most of the available contributions come from the grey literature.

The existing data governance maturity models suit intra-organizational operations and dimensions (e.g., assessing a specific department). However, they are unsuitable for inter-organizational operations (e.g., data ecosystems) [21]. Therefore, there is a need to develop a new data governance maturity model that addresses inter-organizational setups, such as the case of data ecosystems [21].

2.4. Data Governance Archetypes and Maturity Assessment Archetypes

According to Merriam-Webster [37], archetypes are "*the original pattern or model of which all things of the same type are representations or copies*." They have been extensively used in the Information Systems field, based on a taxonomical analysis and morphological characteristics of a domain [22]. Archetypes are a "*mechanism for organizing, summarizing, and generalizing information*" [31], being a means towards achieving standardization [35].

The literature on data governance reveals examples of archetypes. For example, archetypes for data governance sophistication level based on the regulatory environment and the complexity of data assets [25]. The work of Hunke et al. [9] focuses on identifying archetypes for analytics-based services. The existing research focuses on the deployment and operationalization of data governance in organizations and, thus, are unsuitable for the context of a maturity assessment. Therefore, there is a need to develop a set of archetypes for inter-organizational data governance maturity assessment that considers the multiple abstraction levels at which the model may be deployed. For example, a diagnosis focusing on a data producer or an assessment focusing on the whole collaborative ecosystem.

3. Research Methodology

For the design and development of an architecture for the IO-DGMM and the corresponding archetypes, we have adapted the proposal by Becker, Knacksted, and Pöppelbuß [3]. The authors introduce an approach for developing IT management maturity models to handle technological changes, new regulations, and benchmarks [3]. Data governance is often considered a subset of IT governance [20]; thus, we found this approach suitable for developing our maturity model. Furthermore, we extended Becker, Knacksted, and Pöppelbuß [3]'s approach with steps to identify and develop data governance maturity assessment archetypes. Figure 1 describes the methodology [3].

Figure 1 summarizes the steps of the design and development of the IO-DGMM, which started by defining the existing problem ("Problem Definition" in Fig. 1.), including the target domain and group [3]. We conducted a literature review on data governance maturity models. For this purpose, we used Scopus and Web of Science (WoS), two of the most preeminent bibliographic databases [8]. We obtained 485 matches in Scopus and 17 in WoS Core Collection by applying the keywords "data governance" AND "maturity model" (green circle in Fig. 1). The analysis of the several maturity models revealed that the current contributions do not include mechanisms for inter-organizational data governance.

We chose a hybrid strategy to develop the IO-DGMM [3]. On the one hand, we have considered concepts and structures from other data governance maturity models (e.g., TM Forum's model [33] or the DMM/CMMI [11]). On the other hand, we developed new components for the IO-DGMM that were not yet available to integrate decentralized data governance maturity concepts (e.g., data sharing mechanisms [26]).

We conducted a literature review on inter-organizational data governance mechanisms to identify elements that could be used to develop the IO-DGMM. We used the keywords ("data governance") AND ("data platform" OR "decentralized" OR "inter-organizational" OR "digital platform" OR "collaborative network" OR "data value chain" OR "ecosystem" OR ("data ecosystem") OR ("platform ecosystem") OR ("digital ecosystem") OR ("digital business ecosystem")). We obtained 276 results in Scopus and 113 in WoS Core Collection. We removed 73 duplicate papers and identified 89 relevant ones by examining the title, keyword, and abstract. Analyzing the existing literature, we identified specific mechanisms for each data governance dimension, including metadata management, data privacy management, data security management, data quality management, data sharing procedures, and data strategy management (orange circle in Figure 1).

The "*Iterative Maturity Model Development*" included four sub-steps with (1) "*Select Design Level*," (2) "*Select Approach*," (3) "*Design the Model Section*," and (4) "*Test Result*" [3] (red rounded rectangles in Fig. 1), during which we developed the IO-DGMM iteratively. The development of data governance maturity assessment archetypes followed the development of the model. The "*Transfer and Evaluate Conception*" [3] stage started, on which we defined the means to transfer knowledge between the academic and industrial

communities, a set of maturity assessment archetypes, and the maturity model.

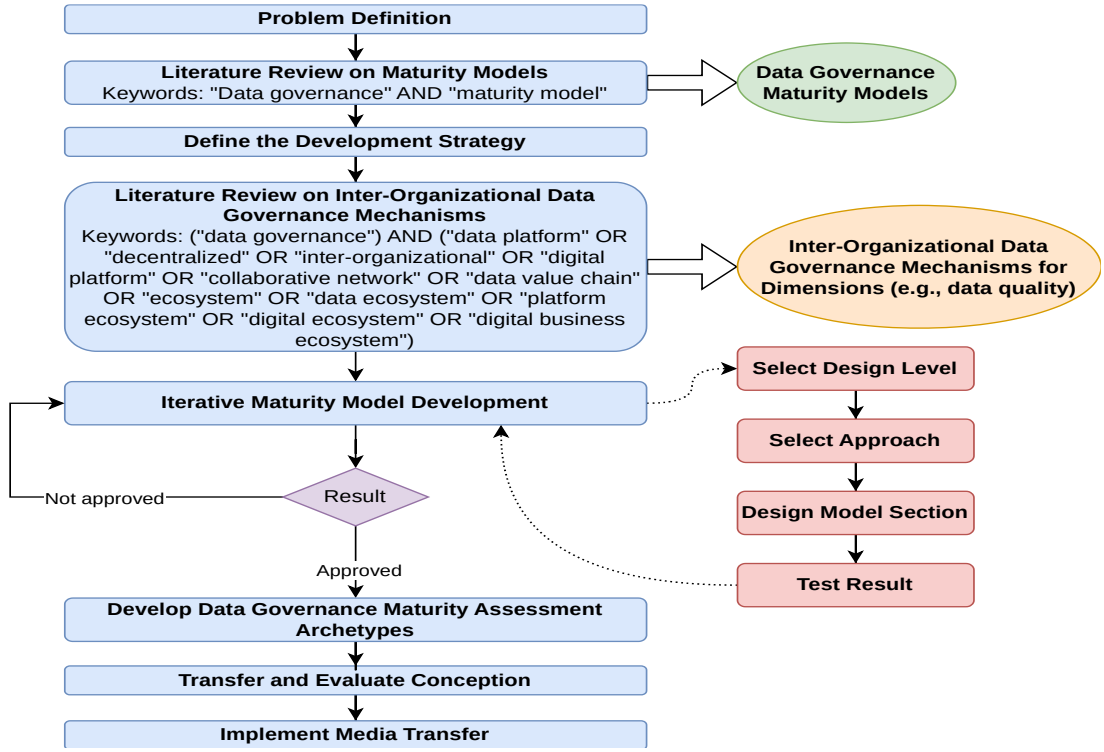

Fig. 1. IO-DGMM Design and Development Step (adapted from Becker, Knacksted, and Pöppelbuß [3]).

The "*Implement Media Transfer*" stage focused on making the IO-DGMM available to the targeted group and domain [3]. We piloted the maturity model in the TSP company, considering its participation in multiple data collaboration platforms. These platforms promote a collaborative ecosystem where members can access shared data resources and IT assets. Notwithstanding, each member is autonomous in developing its operations and strategy. The TSP develops products (e.g., platforms, hardware, software) that are operationalized by the end customers and the partners. At this stage, we focused on the smart green ports ecosystem, in which the TSP company is developing solutions that contribute to optimizing dry ports operations, route planning, and network management.

4. Data Governance Maturity Assessment Approach

4.1. Data Governance Maturity Model

In a previous stage of our research [28], we collected and analyzed multiple data governance maturity models in the available literature. This step provided us with the insights to compare, match, and derive a set of dimensions for our maturity model [28]. Figure 2 introduces the proposed IO-DGMM dimensions [5], [28], [33].

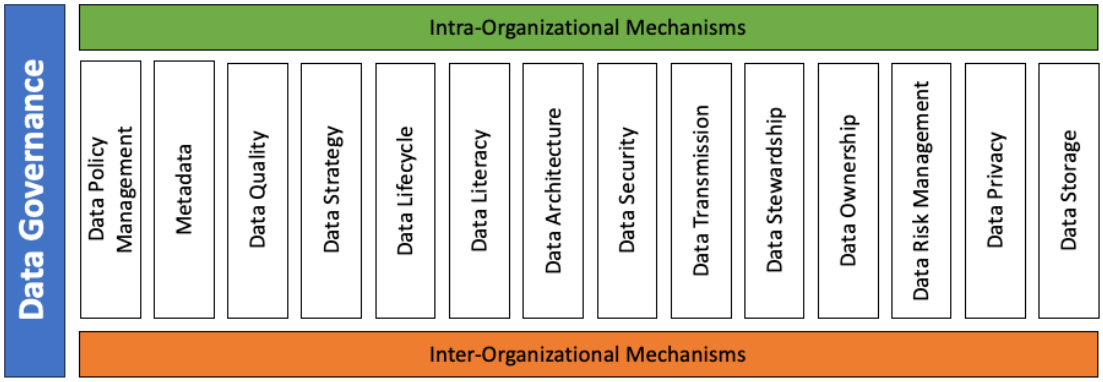

Fig. 2. IO-DGMM Dimensions (based on [5], [28], [33]).

Figure 2 depicts our model's data governance maturity dimensions. For each of the dimensions, we categorize intra-organizational and inter-organizational mechanisms. On the one hand, the intra-organizational mechanisms focus on internal cohesion, organizational practices, project development, and fulfilling the company's objectives. On the other hand, the inter-organizational mechanisms cover the partnership contribution estimation, accountability over shared operations, data-sharing operations and infrastructures, and collaborative processes. Table 2 presents an extract of the IO-DGMM, focusing on the intra-organizational mechanisms.

Table 2. IO-DGMM Excerpt – Intra-Organizational Mechanisms.

Dimension & Sub-Dimension						
Data Quality – Data Profiling						
Practice	**Level 0**	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**
Define, implement, monitor, and review a plan for data profiling in the organization.	There is no plan for data profiling in the organization.	The team defines an inconsistent and informal plan for data profiling in the organization for specific projects.	The team defines and implements a data profiling plan in the scope of its activities.	The organization defines and implements an organization-wide plan for data profiling in the organization.	The organization defines, implements, and monitors an organization-wide plan for data profiling in the organization.	The organization defines, implements, enforces, monitors, and reviews an organization-wide plan for data profiling in the organization.

The example in Table 2 focuses on the data quality dimension's data profiling sub-dimension. At level 0, data governance does not exist, or it is not performed. At level 1, the work is carried out based on reactive ad-hoc practices at the team level, and data assets are seen as a by-product of activities and products. At level 2, there are reactive high-level guidelines for conducting the activities at the team level, and the team is aware of the importance of data assets for its performance and improvement. At level 3, the work is developed based on proactive cross-organizational managed processes, and data is handled as a product that can be valuable for the organization for improvement and data-based product development. At level 4, the work is planned, monitored, and measured based on a set of cross-organizational defined goals; data is classified as a key organizational asset essential for organizational performance, and strategies are defined to identify possible problematic situations and exploit available opportunities. At level 5, the work is managed as part of an optimized strategy based on continuous improvement; data is classified as a critical asset for the organization's survival and a source of competitive advantages. Predictive systems are used to identify and prevent problematic occurrences.

Table 3. IO-DGMM Excerpt – Inter-Organizational Mechanisms.

Dimension & Sub-Dimension						
Data Strategy – Data Use Cases						
Practice	**Level 0**	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**
Identify, manage, communicate, monitor, and review business-relevant and profitable data for the partnership-	There is no identification of business-relevant and profitable data for the partnership.	The team identifies inconsistently and without coordination data assets that may be relevant and profitable for its specific projects.	The team identifies and manages business-relevant and profitable data for the partnership in the scope of their activities.	The organizations agree on, identify, manage, and communicate business-relevant and profitable data for the partnership.	The organizations agree on, identify, manage, communicate, and monitor the business-relevant and profitable data for the partnership.	The organizations continuously agree on, identify, manage, communicate, monitor, and review the business-relevant and profitable data for the partnership.

Table 3 presents an example of the inter-organizational mechanisms, focusing on the data strategy dimension's data use cases sub-dimension. At level 0, data governance does not exist or is not performed across the ecosystem. At level 1, activities are not managed, data is handled as a by-product of activities and artifacts across the ecosystem and its members, procedures, and decisions are carried out based on experience, and practices are locally performed without coordination by each partner. At level 2, activities are managed,

ecosystem members are becoming aware of the importance of data assets to capitalize on the collaboration, procedures and decisions are carried out based on experience, and practices are coordinated between the several partners. At level 3, standardized and coordinated processes are defined and managed between the ecosystem's members, data is used as a product collected from processes and services across the partners, and strategies are implemented to prevent identified situations and/or issues in the partnership. At level 4, standardized and coordinated processes are quantitively managed across the ecosystem, data is handled as an ecosystem resource, and strategies are performed to prevent identified situations and/or issues in the several partners and environmental conditions. At level 5, processes and activities are optimized across the ecosystem, data is a critical asset for the organization for improvement and product development across the several partners, and predictive models are used to anticipate events and/or issues and environmental conditions.

The IO-DGMM structure follows the same descriptive logic for each maturity level for the several practices of each dimension, aiming to facilitate the coherence of the analysis carried out by the users applying the model and the understanding by the practitioners targeted in the assessment. We complement the IO-DGMM with a questionnaire that allows the users to use the model and obtain the corresponding maturity assessment score.

4.2. Data Governance Maturity Assessment Archetypes

The IO-DGMM can be deployed at multiple abstraction levels in intra-organizational and inter-organizational scopes. These possibilities were identified, analyzed, and discussed based on the context of the TSP company. The maturity assessment can be focused on the interaction/relationship between specific partners in a collaboration platform (e.g., a dry port company acts as a data producer by providing data for a TSP that uses data to improve its services). This assessment may occur at distinct levels in the organization, considering a specific department, involved processes, the provided service, or even the organization as a whole (e.g., a maturity assessment focusing on the port slot allocation management platform development at the TSP, that integrates data from the dry port's companies uses).

Table 4. Data governance maturity assessment archetypes.

Select the Inter-Organizational Level		
Category	**Inter-Organizational Level**	**Description**
Ecosystem Interaction	Consumer	The maturity assessment focuses on the interaction between two organizations, considering the data governance mechanisms associated with data consumers obtaining data from producers.
	Producer	The maturity assessment focuses on the interaction between two organizations, considering the data governance mechanisms associated with data producers providing data to consumers.
	Prosumer	The maturity assessment focuses on the interaction between two organizations, considering the data governance mechanisms associated with data prosumers.
Ecosystem Wide	Ecosystem Governance	The maturity assessment focuses on the collaborative partnership (e.g., data ecosystem), involving all the partners and their cohesion as a whole.
	Ecosystem Membership	The maturity assessment focuses on the collaborative partnership, emphasizing an organization that is a candidate to integrate the ecosystem.
Select the Intra-Organizational Level		
Category	**Intra-Organizational Level**	**Description**
Local	Process	The maturity assessment focuses on a specific process (or a set of processes) of the organization(s), considering process-level practices.
	Product/Service	The maturity assessment focuses on a specific product/service of the organization(s), considering the associated data governance practices.
	Department	The maturity assessment focuses on a specific department(s) or a team(s) in the organization(s).
	Organizational	The maturity assessment covers the entire organization(s).
Archetype = Inter-Organizational Level + Intra-Organizational Level		

In a distinct scenario, the IO-DGMM maturity model can be deployed to evaluate the cohesion and the inter-organizational collaboration between the several partners that integrate an ecosystem or to assess the possibility of integrating a new organization in the partnership. Therefore, data governance maturity assessment may happen at distinct abstraction levels in the organization. The possible archetypes for adopting the proposed maturity model were identified in collaboration with the case company (Table 4).

Table 4 depicts the data governance maturity assessment archetypes, that are generated by combining an inter-organizational and intra-organizational level. The maturity assessment archetype selection starts by defining the inter-organizational level, considering Ecosystem Wide or Ecosystem Interaction (first column). When choosing "Ecosystem Interaction," the maturity assessment may focus on the consumer, producer, or prosumer relationships within the ecosystem. Applying the "Ecosystem Wide," the assessment may target the ecosystem governance (e.g., the data platform collaboration-wide maturity assessment) or the ecosystem membership adherence (e.g., a third party that wishes to integrate the data collaboration platform). Then, there is the need to choose the intra-organizational level of maturity assessment, considering specific contexts (e.g., process, product/service, department) or the entire organization. For example, a prosumer-product archetype focuses on the consumer + producer data governance mechanisms in the scope of an inter-organizational product. The following section describes a demonstration of the data governance maturity assessment.

5. Demonstration

Considering the preliminary results of our research, we decided to focus on the assessment of metadata, data architecture, data strategy, and data transmission dimensions presented in Fig. 2. The goal was to obtain initial feedback on the application of the maturity model and to conduct necessary adjustments, before performing a full assessment. For this specific company, two archetypes were selected: (1) the prosumer-organizational archetype and (2) the ecosystem governance-organizational archetype.

Together with the TSP, we defined the dates for the presentations and interviews with the specific members of the organization selected as the target of the maturity assessment. An initial presentation focused on explaining the goals of the maturity assessment, the basic concepts of the maturity model, and the role of each party in the process.

Next, we conducted the preliminary stages of the data governance maturity assessment. We compiled and analyzed the retrieved information to classify each data governance dimension defined for the pilot survey. Fig. 3. reports the preliminary findings of our pilot survey on the TSP using the prosumer-organizational archetype.

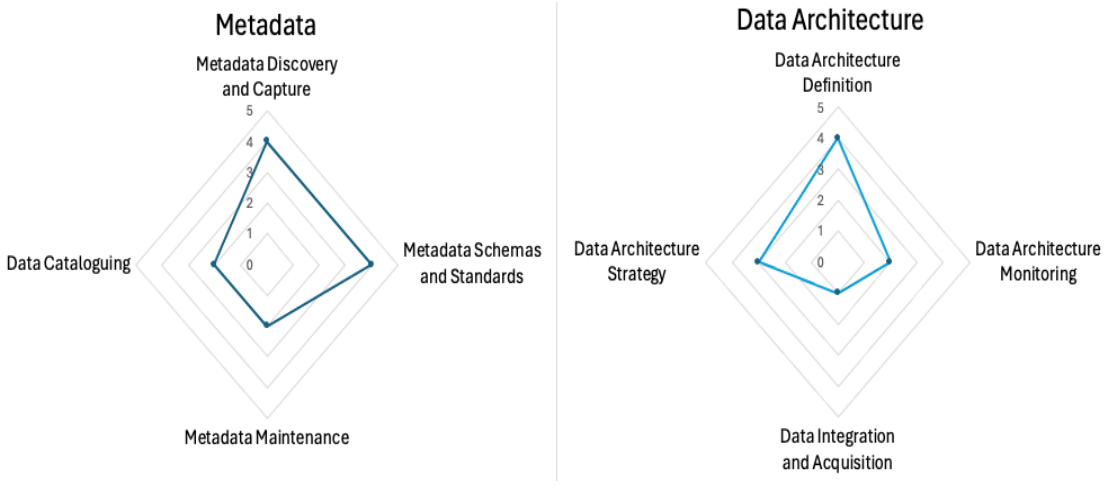

Fig. 3. Maturity assessment using the prosumer-organizational archetype at the TSP, applying the metadata and data architecture dimensions.

The results show a high maturity score for the "Metadata Discovery and Capture" (e.g., the existence of an optimized metadata collection process) and "Metadata Schemas and Standards" (e.g., periodically review metadata architecture) sub-dimensions (on the left of Fig. 3.). It scored low on the "Metadata Maintenance" (e.g., lack of a metadata versioning system) and "Data Cataloguing" (e.g., the use of local data catalogs only in specific projects) sub-dimensions. For "Data Architecture," the company scored high for "Data Architecture Definition" (e.g., the definition and periodic review of data architecture requirements) and "Data Architecture Strategy" (e.g., the periodic review of the data architecture strategy) sub-dimensions. It scored low for the "Data Integration and

Acquisition" (e.g., there is no data infrastructure mapping in the organization) and "Data Architecture Monitoring" (e.g., the company doesn't have KPIs for data architecture) sub-dimensions. Fig. 4. highlights the ecosystem governance-organizational archetype.

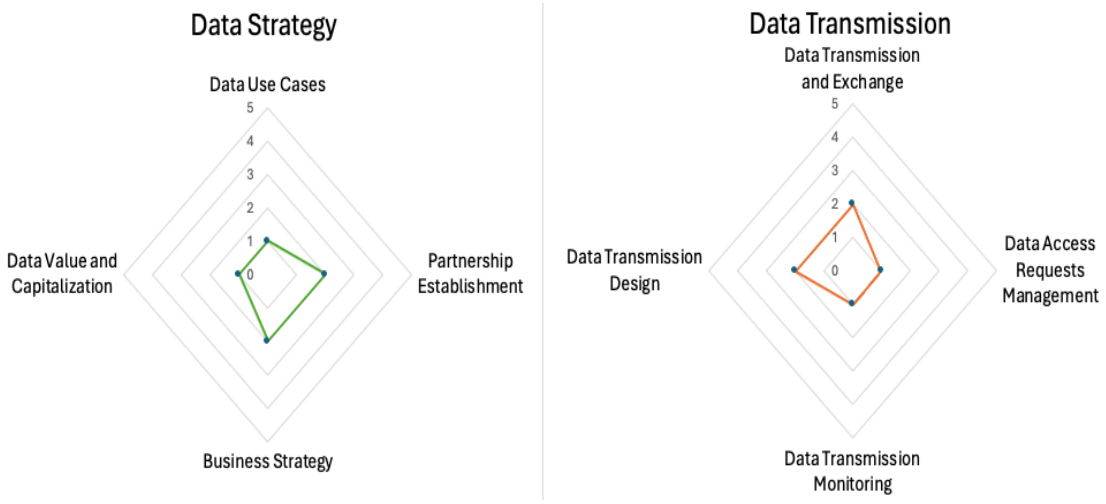

Fig. 4. Maturity assessment using the ecosystem governance-organizational archetype at the TSP, applying the data strategy and data transmission dimensions.

The company obtained a low maturity score for the several sub-dimensions of "Data Strategy," including "Data Use Cases" (e.g., the organizations inconsistently and without coordination identify data assets that may be relevant to the remaining partners), "Data Value and Capitalization" (e.g., the organizations determine the value of a specific data assets for the ecosystem using ad-hoc methods), "Business Strategy" (e.g., the organizations agree on, develop, and implement business cases for the ecosystem in the scope of their activities), and "Partnership Establishment" (e.g., the organizations agree on and identify relevant partners to share data in the ecosystem). The company also obtained a low score for most of the sub-dimensions on "Data Transmission," considering "Data Transmission Design" (e.g., the organizations agree on, define, and implement data sharing standards for the ecosystem that cover the scope of the teams' shared activities), "Data Access Requests Management" (e.g., the organizations use a basic formulary to request access or transmission of specific data assets in the ecosystem), "Data Transmission Monitoring" (e.g., the organizations define inconsistent and uncoordinated basic methods to monitor and log data transactions in the ecosystem), and "Data Transmission and Exchange" (e.g., the organizations agree on, define, and implement guidelines for data transmission for sharing data assets in the ecosystem).

The pilot survey allowed the development of a roadmap for improving the targeted dimensions at this stage of the maturity assessment, emphasizing inter-organizational governance-related operations and structures. Based on the assessment's maturity score, we developed a report containing level classification for the multiple data governance dimensions and specific recommendations for the organization to improve its current practices. For data quality, as an example, we suggested the organization to develop a list of trusted sources for obtaining data. Moreover, for data architecture, the team suggested the development of logical and physical models for the infrastructures.

6. Discussion

The IO-DGMM can be used to conduct the maturity assessment of organizations, considering their increased collaboration in data sharing operations, that require compliance with data-related regulations, collaborative processes, and inter-organizational coordination mechanisms. By applying the model, it is possible to assess the practices that cover partnership collaboration, contribution estimation, and joint data strategy, highlighting a maturity assessment score for the collaborative partnership and its multiple interactions. The maturity assessment archetypes allow the application of the IO-DGMM at distinct abstraction levels, with the possibility of focusing on an ecosystem-wide

diagnosis or specific interactions between two specific members.

The IO-DGMM incorporates intra/inter-organizational dimensions. When comparing with the available maturity models, the IO-DGMM incorporates a new data literacy dimension for both intra-organizational and inter-organizational scopes. Additionally, the IO-DGMM can establish the basis for the integration and standardization between the multiple partners in the ecosystem since they are working using a unified maturity model. Moreover, it allows organizations to develop a coordinated and agreed-upon data governance improvement roadmap, considering the assessment results. According to their roles in the ecosystem (e.g., a data producer), the organizations may prioritize the implementation of specific mechanisms tailored to their needs. Lastly, the maturity model can establish a dependency between the intra-organizational and inter-organizational scope dimensions (e.g., if the organization is at level 3 in intra-organizational data quality, it cannot achieve higher than that for inter-organizational data quality).

The business partners can use the IO-DGMM to assess the possibility of establishing agreements with candidate entities (e.g., a new partner data provider). More specifically, the model can be used to conduct an intra-organizational assessment of the candidate. Based on the maturity assessment's results, the organizations can decide to integrate (or not) the candidate as a partner or ecosystem member.

The research team also concluded that the maturity model can be used to score data producers regarding the quality of their assets, the provided services, and the shared data. For example, Singh et al. demonstrate that an increased reputation may allow data producers to attract additional data consumers [30].

The data governance maturity assessment archetypes enable the deployment of the IO-DGMM at distinct abstraction levels according to the organization's needs. By conducting an assessment at the inter-organizational scope, it is possible to adapt the model to assess interactions between specific partners (e.g., a data producer that provides data to a partner in an ecosystem) and the collaborative partnership dynamics (e.g., using the partnership-wide level collaboration archetype). The IO-DGMM intra-organizational scope allows organizations to perform an individual-level maturity assessment utilizing the product, department, or organizational archetypes. Therefore, based on the selected archetypes, the organization can conduct a tailored maturity assessment that is suitable and prioritized toward its objectives.

7. Conclusion

This paper describes the architecture and assessment archetypes for a data governance maturity model for inter-organizational operations. It is the result of a joint collaboration with a TSP company that is currently adopting it to develop a comprehensive data governance strategy adapted to the challenges of new digital solutions, with increasing requirements in data governance within their entire data collaboration platform.

Some limitations must be stated. First, the artifacts were created and demonstrated in a single case that was not representative of the entire industry. Future work should focus on deploying the developed artifacts in other cases (e.g., software and hardware development, healthcare). Second, the artifacts were created in a highly regulated context, such as the port management sector. Future iterations should focus on deploying the artifacts in other sectors to compare the maturity assessment results and extend the maturity model. Third, our work contributes toward data governance maturity assessment in inter-organizational setups. However, we currently do not have evidence of organizational improvements that may come from applying the data governance implementation roadmap based on the maturity assessment results. Future work may consider steps to evaluate the developed artifacts formally. Fourth, our maturity model provides organizations with a generic assessment and improvement path. Future research can focus on creating a tailoring system like the one incorporated in ISACA's Control Objectives for Information Technologies [12], aiming to support organizations in prioritizing the assessment and improvement of specific domain dimensions (e.g., achieving level 4 for data architecture and level 3 for data security), based on a set of pre-defined factors (e.g., industry sector, compliance requirements). Fifth, our research focused on the data ecosystems context. Future work

may consider other typologies of inter-organizational relationships, such as the case of digital business ecosystems [34]. Lastly, the current maturity model and the corresponding support documentation are supported on a spreadsheet. Currently, the team is developing a more sophisticated web platform for the purpose, allowing organizations to track their maturity status and continuously update the results, including evidence of compliance. This digital tool will be essential in audits and in creating a shared data space for the data collaboration platform, including suppliers, partners, and data end users.

Acknowledgments

This research was funded by Project "Agenda Mobilizadora Sines Nexus" ref. No. 7113, supported by the Recovery and Resilience Plan (PRR) and by the European Funds Next Generation EU, following Notice No. 02/C05-i01/2022, Component 5 - Capitalization and Business Innovation - Mobilizing Agendas for Business Innovation. It was also financed through national funds by FCT - Fundação para a Ciência e a Tecnologia, I.P., in the framework of the Project UIDB/00326/2020 and UIDP/00326/2020. The first author is funded by FCT - Foundation for Science and Technology, I.P., under the Ph.D. grant 2023.00740.BD.

References

1. Abraham, R., Schneider, J., vom Brocke, J.: Data governance: A conceptual framework, structured review, and research agenda. Int J Inf Manage. 49 (January), 424–438 (2019)
2. Al-Ruithe, M., Benkhelifa, E.: Cloud data governance maturity model. In: 2nd International Conference on Internet of Things and Cloud Computing, ICC 2017. (2017)
3. Becker, J., Knackstedt, R., Pöppelbuß, J.: Developing Maturity Models for IT Management. Business & Information Systems Engineering. 1 (3), 213–222 (2009)
4. Bruhn, J.: Identifying useful approaches to the governance of indigenous data. Int Indig Policy J. 5 (2), 1–32 (2014)
5. DAMA International: The DAMA Guide to The Data Management Body of Knowledge. Technics Publications, LLC Post. 406 (2009)
6. Gelhaar, J., Gürpinar, T., Henke, M., Otto, B.: Towards a taxonomy of incentive mechanisms for data sharing in data ecosystems. In: PACIS. p. 121. (2021)
7. Gelhaar, J., Otto, B.: Challenges in the Emergence of Data Ecosystems. PACIS. 175 (2020)
8. Harzing, A.W., Alakangas, S.: Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison. Scientometrics. 106 (2), 787–804 (2016)
9. Hunke, F., Heinz, D., Satzger, G.: Creating customer value from data: foundations and archetypes of analytics-based services. Electronic Markets. 32 (2), 503–521 (2022)
10. IBM: The IBM Data Governance Council Maturity Model : Building a roadmap for effective data governance. Governance An International Journal Of Policy And Administration. (October), 1–16 (2007)
11. ISACA: CMMI, https://cmmiinstitute.com/, Accessed: July 29, 2022, (2023)
12. ISACA: COBIT®2019 Framework: Governance and Management Objectives. ISACA (2018)
13. Jagals, M., Karger, E.: Inter-organisational data governance: A literature review. Twenty-Ninth European Conference on Information Systems. (June), 1–19 (2021)
14. Janssen, M., Brous, P., Estevez, E., Barbosa, L.S., Janowski, T.: Data governance: Organizing data for trustworthy Artificial Intelligence. Gov Inf Q. 37 (3), 101493 (2020)
15. Jussen, I., Schweihoff, J., Dahms, V., Möller, F., Otto, B.: Data Sharing Fundamentals: Definition and Characteristics Ilka. In: HICSS 2023 Proceedings. (2023)
16. Karkošková, S.: Data Governance Model To Enhance Data Quality In Financial Institutions. Information Systems Management. 40 (1), 90–110 (2023)
17. Khatri, V., Brown, C.V.: Designing data governance. Comm. ACM 53(1), 148–152

(2010)
18. Lee, S.U., Zhu, L., Jeffery, R.: Data governance decisions for platform ecosystems. Proceedings of the Annual Hawaii International Conference on System Sciences. 2019-Janua 6377–6386 (2019)
19. Lis, D., Otto, B.: Data governance in data ecosystems - Insights from organizations. 26th Americas Conference on Information Systems, AMCIS 2020. (2020)
20. Mäntymäki, M., Minkkinen, M., Birkstedt, T., Viljanen, M.: Defining organizational AI governance. AI and Ethics. (2022)
21. Marcelo, M.I., Barros Lima, G. de F., Farias Lóscio, B.: Investigations into Data Ecosystems: a systematic mapping study. Knowl Inf Syst. 61 (2), 589–630 (2019)
22. Möller, F., Bauhaus, H., Hoffmann, C., Niess, C., Otto, B., Isst, F.: Archetypes of Digital Business Models in Logistics Start-UPS. In: ECIS. (2019)
23. Oliveira, M.I.S., Lóscio, B.F.: Louvre: A framework for metadata curation in data ecosystem. In: ACM International Conference Proceeding Series. Association for Computing Machinery (2019)
24. Pansara, R.: Unraveling the Complexities of Data Governance with Strategies, Challenges, and Future Directions. Transactions on Latest Trends in IoT. 6 (6), 46–56 (2023)
25. Petzold, B., Roggendorf, M., Rowshankish, K., Sporleder, C.: Designing data governance that delivers value. Mc Kinsey Digital, June. (2020)
26. De Prieelle, F., De Reuver, M., Rezaei, J.: The Role of Ecosystem Data Governance in Adoption of Data Platforms by Internet-of-Things Data Providers: Case of Dutch Horticulture Industry. IEEE Trans Eng Manag. 69 (4), 940–950 (2022)
27. Rasouli, M.R., Trienekens, J.J.M., Kusters, R.J., Grefen, P.W.P.J.: Information governance requirements in dynamic business networking. Industrial Management and Data Systems. 116 (7), 1356–1379 (2016)
28. Ribeiro, V., Barata, J., Cunha, P.: Decentralizing Data Governance: A Case Study in TELCO Data Ecosystems. In: Americas Conference on Information Systems (AMCIS). (2023)
29. Röglinger, M., Pöppelbuß, J., Becker, J.: Maturity models in business process management. Business Process Management Journal. 18 (2), 328–346 (2012)
30. Singh, P.K., Singh, R., Nandi, S.K., Nandi, S.: Designing a blockchain based framework for IoT data trade. In: I4CS 2020, Bhubaneswar, India, January 12–14, 2020, Proceedings 20. pp. 295–308. (2020)
31. Souza, E., Lencastre, M., Melo, R.C.F., Ramires, L., Alves, K.: Analyzing Problem Frames together with Solution Patterns. In: WER. pp. 179–189. (2007)
32. Tarhan, A., Turetken, O., Reijers, H.A.: Business process maturity models: A systematic literature review. Inf Softw Technol. 75 122–134 (2016)
33. TM Forum: TM Forum Data Governance Guidebook, https://www.tmforum.org/resources/standard/gb1023-data-governance-guide-book-v3-0-0/, Accessed: July 11, 2022, (2022)
34. Tsai, C.H., Zdravkovic, J.: A survey of roles and responsibilities in digital business ecosystems. In: 13th IFIP WG 8.1 Working Conference on the Practice of Enterprise Modeling (PoEM 2020), Riga, Latvia (virtual), November 25-27, 2020. pp. 44–53. (2020)
35. van der Valk, H., Haße, H., Möller, F., Otto, B.: Archetypes of Digital Twins. Business and Information Systems Engineering. 64 (3), 375–391 (2022)
36. Vesselkov, A., Hämmäinen, H., Töyli, J.: Design and governance of mhealth data sharing. Communications of the Association for Information Systems. 45 (1), 299–321 (2019)
37. Archetype Definition & Meaning - Merriam-Webster, https://www.merriam-webster.com/dictionary/archetype, Accessed: April 19, 2024