information

MDPI

Article

Architecture Building Blocks for Data Governance in Data Spaces

Marta Zorrilla * and Juan Yebenes

Departamento de Ingeniería Informática y Electrónica, Facultad de Ciencias, Universidad de Cantabria, Santander, 39005 Cantabria, Spain; juan.yebenes@unican.es
* Correspondence: marta.zorrilla@unican.es

Abstract

The growing importance of data as a driver of the digital economy is promoting the creation of data spaces for the secure and controlled exchange of data between organizations. Data governance is emerging as an essential pillar to ensure efficient, ethical and transparent access and use of data in these ecosystems. The article reviews the state of the art to identify the specific requirements that data governance must address in data spaces and proposes a reference enterprise architecture to facilitate the design, development and implementation of a data governance system for a data space scenario. The proposed framework has already been formally defined and validated in the context of Industry 4.0, and is now adapted to the particular characteristics and needs of data spaces. This architecture focuses on key aspects of data governance in data spaces, such as new requirements, principles, organization, roles and responsibilities, and data quality, security and metadata management, as well as the data lifecycle in the data space. This research contributes to guiding data space government bodies to formalize data strategies and high-level governance principles in concrete architectural components that establish the capacities to be implemented within the data ecosystem. To support practical adoption, this work also provides clarifying examples of different blocks of architecture.

Keywords: data governance; data spaces; reference architecture; digital economy

Academic Editor:
Haridimos Kondylakis

Received: 29 July 2025
Revised: 7 October 2025
Accepted: 16 October 2025
Published: 22 October 2025

Citation: Zorrilla, M.; Yebenes, J. Architecture Building Blocks for Data Governance in Data Spaces. *Information* **2025**, *16*, 927. https://doi.org/10.3390/info16110927

1. Introduction

Data has emerged as the pivotal force within the digital economy, functioning as the most valuable strategic resource for business transformation and competitiveness [1]. In an era marked by continual technological advancement, the capacity to manage, analyze, and govern extensive volumes of data drives innovation and empowers organizations to adapt to an increasingly data-driven marketplace. Acknowledging this reality, Europe is taking the lead in establishing data spaces (DSs) that facilitate the secure and controlled exchange of data in public and private entities, thereby promoting cross-sector collaboration and generating added value [2].

The concept of data space (DS) is relatively new, and various definitions can be found within the literature. A. Reiberg et al. [2] define a DS as "a federated, open infrastructure for sovereign data sharing based on common policies, rules, and standards". The Data Spaces Support Centre [3] describes it as "an interoperable framework, based on common governance principles, standards, practices and enabling services, that enables trusted data transactions between participants". The International Data Spaces Association (IDSA) [4] characterizes a DS as "a decentralized infrastructure for trustworthy data sharing and

exchange in data ecosystems based on commonly agreed principles". Finally, the Spanish Data Office [5] defines it as "an ecosystem where participants voluntarily share data in an environment of sovereignty, trust, and security, established through integrated governance, legal, organizational, regulatory, and technological mechanisms". In summary, a DS is primarily characterized by members sharing and exchanging data within an environment of mutual trust, which occurs under the principles of sovereignty and decentralization, within an interoperable framework, guided by commonly agreed principles, standards, policies, and rules. This leads to the need for governance mechanisms to ensure the rights of participants and the fluent exchange of data [2] as well as to regulate the activities of participants within the ecosystem [6].

R. Braun et al. [7] note that DS governance involves "managing how the data space is accessed, controlled and used; assessing and controlling the generation of value from the data space and the redistribution of value among actors; reflecting and participating in decision making about the data space (who can make those decisions) and influencing how data and virtual space is accessed, controlled, used and benefited from".

On the other hand, the IDSA-RAM architecture [8] points out that DS governance "defines the roles, functions, and processes of DS from a governance and compliance perspective. In doing so, it defines the requirements that the business ecosystem must meet to achieve secure and reliable enterprise interoperability" and addresses this issue in a federated manner by distributing decision rights for governance and management activities among the different roles in the DS ecosystem [9]. Governance also "encompasses relationships with stakeholders and how their interests and concerns are articulated and taken into account" [10]. Finally, the Spanish Data Office [5] points out that the governance of a DS should allow for policies on access and use of data resources to be enshrined in the DS, for the use of data to be subject to external audit, for conflict resolution mechanisms to be put in place, and for transparency on the level of quality of data provided by participants to be ensured.

The complexity of these ecosystems, as pointed out by Fritzenkötter et al. [11], may require different levels of governance in a DS as part of a governance continuum. At a high level, a governance structure is needed for defining strategies and managing stakeholder interests. For other, more specific aspects, concrete governance structures may be needed, such as the establishment of standards and rules, as well as the management of asset interoperability. Between these levels, an oversight function must be defined to deliver transparency and accountability, which would include establishing roles and processes that allow recourse to the governing body if any stakeholder feels that those responsible for the DS have not fairly respected their needs.

Data governance (DG) within the DS governance, on the other hand, aims to enable and establish the necessary capabilities to exercise consensual and communicated decision-making, authority, and control over the management of data assets, and to define who has decision-making rights and responsibilities for data-related processes within DS [12]. It therefore focuses on how data that is shared and used within the DS is managed. According to [10], DG turns out to be a fundamental discipline for the proper functioning of DS. The DG system is shaped by the DS activities, objectives, standards, and purpose [13], ensuring that data is managed as a strategic asset. To be effective, the DG system must coordinate people, processes, and technologies [14] and include all DS members [15]. It also needs to define the scope of decision-making for data governance, identifying which data activities and aspects should be governed, as well as clarifying the roles, rights, authority, and responsibilities of those involved [16].

Although the objectives pursued by DG in general remain common to other environments, in the case of DS, it is also necessary to take into account the changes that this new

model entails, in terms of data, participants, and processes, as well as legal, regulatory, administrative, or economic aspects. In addition, the technological platform to use to make DG viable in these environments must also be considered [17].

Even though we have found references that address DG in the context of DS, the documentation analyzed does not provide guidance on how to develop the necessary structure for its implementation. To address this gap, we propose the TRENTI DG framework [18–20], whose aim is to solve the "translation problem": how to transform data strategies and high-level governance principles into tangible architectural components that can be designed, deployed, and maintained within a DS. The TRENTI framework was conceived precisely to address this gap. TRENTI is not merely a catalog of governance principles or procedures; it is a reference architecture that connects governance requirements with the enterprise architecture design through a modular and adaptable structure. The core contribution of this paper is the adaptation of TRENTI's Architecture Building Blocks (ABBs) to support the implementation of DG in DSs, based on newly identified requirements specific to this context. In particular, the following ABBs have been adapted: principles, organization, roles and responsibilities, data quality, security and metadata management, and the data lifecycle. For ease of understanding, several ABBs related to DS are instantiated. These ABBs have been chosen based on considerations and documentation from the organizations involved in European DSs initiatives.

The paper is structured as follows: After this introduction, Section 2 describes the method followed in this research. Section 3 presents the state of the art regarding data governance and data management in DS. Section 4 contextualizes data governance and identifies the requirements that the governance framework must address, which are conditioned by the specific characteristics of DS. Section 5 describes the reference architecture to support the implementation of data governance in DS. Section 6 discusses challenges and barriers that hinder the deployment of DS, as identified in the recent literature, and how the TRENTI DG framework can contribute to overcoming them. Finally, Section 7 draws the conclusions and contributions of this work and suggests future lines of research.

2. Method

This research adopts a common methodological approach in software engineering, aimed at designing and specifying a reference enterprise architecture for data governance in data spaces. The process began with the formalization of the problem, identifying the lack of structured guidance for implementing DG systems in DS environments. A thorough literature review was conducted, encompassing both academic sources and European initiatives (e.g., IDSA, GAIA-X, BDVA), to extract relevant concepts and requirements. Based on this analysis, we propose a solution derived from a framework originally developed for DG in Industry 4.0 [18], which has now been adapted and particularized to meet the specific needs and characteristics of DS. This framework includes a modular architecture composed of ABBs, each addressing distinct governance requirements. The ABBs are described and formally instantiated using different artifacts such as catalogs, matrices, and diagrams to facilitate reuse and alignment with data strategies boosted by the DS governance body.

3. Related Works

In this section, we review the existing literature and initiatives related to DS and DG to understand how the development of DG is approached in DSs and identify those characteristics that may affect or condition the development and deployment of DG, as well as the requirements to be fulfilled.

3.1. DS Features

In terms of the key factors that characterize a data space, data sharing is the most prominent [2]. Data sharing means enabling access and facilitating the exchange of data [5]. However, for this to happen, it is necessary to build trust between participants in the DS [21,22]. This means laying the groundwork to regulate the exchange or sharing of data and defining the conditions under which it takes place, through the formulation of appropriate agreements or contracts [23] so that an environment of trust is created between the members.

Owners' sovereignty over their data is another fundamental aspect of DS [24]. Sovereignty refers to an owner's ability to make decisions about their data [25] and to exercise control over that data and its use, even when shared with third parties [4]. To implement data sovereignty, data must be accompanied by metadata that unambiguously defines the restrictions on data use at each step of the information value chain [8]. According to IDSA, this can be ensured through (i) decentralization of data, which remains with the data owner and is not integrated into a common dataset, (ii) a precisely graded certification concept according to the profile of each participant (as much security as necessary), and (iii) security of infrastructures through new technological solutions and established security policies and functions.

In DS, the terms 'federated' and 'interoperable' go hand in hand. Interoperability is a fundamental aspect and must exist at different levels in DS [26] to enable its constituents to access or exchange data efficiently. There must be technical interoperability in terms of the physical connection between DSs, the APIs and protocols that enable communication between systems, and the data formats that must be supported homogeneously. There must also be conceptual or semantic interoperability so that both the data provider and the data consumer have the same interpretation of the meaning of the shared data, and organizational interoperability [27], which involves a common DG. This entails identifying the stakeholders involved in the management, provision and use of the data and the relationships between them and ensuring that the participants in the data exchange or sharing are aligned on the principles, guidelines, policies and rules governing the management of the shared data, especially with regard to quality, security and its life cycle, as well as the management of the metadata necessary for the operation of the DS. In relation to the latter, tools such as data catalogs, data glossaries, and labelling systems, among others, help to organize and facilitate access to information in a consistent manner. It is also important to highlight that, in order to facilitate interoperability, it is necessary to establish common standards at all levels, which must be assumed by all members of the DS, in addition to compliance with the current legislation.

3.2. Data Governance in Data Spaces

A common feature of DSs is the existence of governance mechanisms that protect the rights of participants and ensure the seamless exchange of data. These governance mechanisms are based on international standards, national and supranational laws, or rules and standards agreed upon by DS participants. DG must also be included as a crucial aspect of DS governance, acknowledging it as a sub-function of corporate/organizational governance [28]. However, to date, data space governance is scarcely researched [29]. For this reason, we have analyzed how this discipline is addressed in the context of DS. First, the contributions of European initiatives are considered. Next, contributions from the academic arena are analyzed.

Data governance plays a central role in the reference architecture proposed by the IDSA-RAM model [9], serving as a facilitator for the proper integration of an organization into a collaborative ecosystem such as a DS. The IDSA-RAM framework specifies

decision rights and processes for the definition, creation, processing and use of data. It takes a federated approach to the DG and introduces new roles to which decision rights for data governance and management activities are assigned. These roles include data owner, data provider, data consumer, broker, service provider, clearing house and app store. Responsibilities for these roles are assigned through an RACI matrix covering data management, metadata management, and the data lifecycle activities. The model also allows us to define the required standards, control, and compliance rules for data exchange between the different participants, specifying its components and mechanisms [30].

For GAIA-X, DG is also one of its fundamental pillars. Its goal is to ensure that data is handled ethically, securely, and in compliance with applicable regulations while respecting the sovereignty and privacy of data owners. To achieve this, GAIA-X relies on a set of key components outlined in the Gaia-X European Association for Data and Cloud AISBL (2024) report [31]. These include policies that define how data can be accessed and used within the DS, along with specific restrictions based on data types (e.g., sensitive, personal, or business data), as well as mechanisms to revoke permissions if conditions change. Another essential component is the use of standardized digital contracts, which outline the terms and conditions of data exchange, including clauses on privacy, permissible usage periods, and other constraints. Furthermore, all participants and services within the DS must be certified to ensure compliance with governance rules in the GAIA-X framework, including technical and legal verifications. A key feature is the federated catalog, which enables the registration of services and data, including metadata that specifies access policies, usage requirements, and licensing terms. Finally, identity and permission management ensure that DS participants have verified digital identities, establishing their legitimacy within the ecosystem. Access control systems ensure that only authorized entities can access the data. However, GAIA-X does not provide a specific architecture for implementing data governance.

The Big Data Value Association (BDVA) addresses data governance in data spaces through initiatives focused on interoperability, security, and respecting data sovereignty [32]. Its approach aligns with the European Data Strategy and involves close collaboration with initiatives like GAIA-X and IDSA to define standards and create common governance frameworks that ensure trustworthy and equitable data exchanges across data spaces. In contrast, FIWARE does not propose a specific solution for data governance but provides tools and solutions for interoperability and data sovereignty, where governance is implicitly addressed. These four organizations (IDSA, GAIA-X, FIWARE BDVA) have joined forces in the Data Spaces Business Alliance (DSBA) "with the common goal of accelerating business transformation in the data economy". In this regard, DSBA does not define either a specific architecture for data governance. However, the principles, standards, and tools it proposes constitute a robust framework to support the development of data governance in such environments.

Among the academic references, Torre-Bastida et al. [17] analyze the technological perspective of DG in data ecosystems (DSs). They suggest that, while traditional DG frameworks must evolve as current models are insufficient for complex DS scenarios, key elements of DG remain, such as defining clear objectives and addressing DG implementation from two perspectives: organizational (involving data, participants, processes, and regulatory domains) and technological (platforms enabling DG). However, DSs introduce new DG considerations, such as data sovereignty, quality, trust, and security. They also identify three levels of DG—micro (intraorganizational), meso (interorganizational), and macro (legislative)—and highlight challenges like ensuring data use control after sharing, supporting the principle of sovereignty, and the need for data catalogs for metadata management.

K. Suzuki and D. Yokozeki [23] propose several requirements for DG implementation, including managing dispersed data; ensuring interoperability across DS; mechanisms for discovering participants and data; governance of encryption-based data storage and data transmission; encrypted data processing; and processing platform/location governance to allow data owners to control over where and how their data are stored, processed, and distributed.

S. Cuno et al. [33] highlight the requirement for an appropriate organization that controls the data and pays special attention to the needs of the members of the DS. They also emphasize the need for clear roles and responsibilities for decision-making about data quality, access, and lifecycle management in Urban DS. They propose roles such as data committee, governance officer, data owner, data steward, and technology steward.

E. Curry [34] calls for research on decentralized DG models that support collaboration while considering the ethical, legal, and privacy interests of stakeholders. DG in ecosystems must acknowledge ownership, sovereignty, and data regulation while supporting sustainable economic models.

Datos.gob.es [35] underscores that DSs require collaboration among multiple organizations to establish a common vision on key strategic and operational aspects, legislation, standards, and governance models.

Data Spaces Support Centre [36] outlines the role of governance authorities in establishing rules and standards for secure, interoperable, and observable data transactions, operationalizing governance across the DS. This requires a central governing body for the entire DS.

Finally, M. Minghini et al. [10] point out that "Data governance within a common European data space should set out a framework outlining clear roles, duties, standards and responsibilities, to ensure that data is appropriately protected, while also supporting data sharing and openness to data mobility".

In short, DG in DSs is essential for enabling secure, interoperable, and collaborative data ecosystems. Frameworks such as IDSA-RAM and GAIA-X emphasize federated models, standardized roles, and ethical data handling. Academic research complements these efforts by highlighting the need for clear objectives, decentralized governance, and addressing challenges such as data quality, transparency, and security. Also, they point out that collaborative approaches and common standards are crucial to maintaining consistency across organizations while supporting data mobility and economic value creation. All of this requires a common DG framework outlining clear roles, duties, responsibilities and standards, as well as policies and processes to ensure that data are properly governed and managed in the DS. However, to our knowledge, there is no framework to guide the construction of an enterprise architecture for implementing DG in DSs.

4. Requirements for Data Governance in Data Spaces

A governance system must fulfil a variety of requirements, which involve both evolving current capabilities and incorporating new ones. These requirements can be organizational, process-oriented, or technology-related. Therefore, as a preliminary step in developing a Reference Enterprise Architecture for DG, it is crucial to establish the requirements of the DG system within a DS. This will highlight the basic needs that the architecture must address.

The requirements can be categorized into general ones, that means applicable to any DG system, and those specific to DG systems in DSs. Among the general requirements, it is essential to establish principles, as outlined by [4], that guide the conduct, behavior, and philosophy of the DS and its members regarding the use, management, and governance of the exchanged or shared data [37].

While each member participating in the DS may have its own data strategy, the DS itself should define a set of objectives related to data exchange and sharing, along with an overarching data strategy specific to the DS. Governance bodies for the DS need to be established, along with defined roles and governed entities. Decision rights, authority, and responsibilities over data assets need to be assigned to the defined roles [8,33].

A DG model must be established and should be based on the functions of Evaluate, Direct, and Monitor [38]. Also, data-related activities subject to governance must be identified [39]. These activities will form the data lifecycle (DLC), which, in the context of the DS, could be expanded to include Collect, Store, Process, Transmit, Exchange/Distribute, Use, and Destroy [40]. Furthermore, it is essential to determine the specific aspects of data to be governed for these DLC activities [41] (see Appendix A). Taking a data lifecycle approach to data governance is helpful because it allows for a thorough analysis of how data should be overseen at various stages, as well as how to create value from data use and reuse in a safe and equitable manner [42].

Additionally, standards, policies and rules must be established for DLC activities to manage data access and usage for participants, aligned with the various roles they may perform within the ecosystem [10], specifically addressing aspects of data sharing and usage, such as quality, security, and metadata. This includes transforming data to provide structure and format based on its intended use.

The organizational model for DG within the DS must be defined, whether it is centralized, federated, or otherwise [12]. Additionally, the extension of the DG system across the entire information value chain for the exchanged or shared data must be considered, which will involve signing agreements between DS participants [43]. Appropriate organizational structures, functions, and processes must be established to develop the strategies and policies defined for DG.

Emphasis should be placed on organizational and semantic/conceptual interoperability for data exchange, sharing, and exploitation. This includes developing a common language and reference model to manage the diversity of data and their relationships. Examples of this include creating business glossaries, data dictionaries, and data catalogs, along with implementing policies and processes for their management. Such initiatives enhance comprehension of the context, relevance, and associations among shared data. Semantic validation must be performed "ex ante," or before sharing the data, and "ex post," or at the time of receiving the data. The goal is to ensure that the data meets the promised or expected quality standards.

Another critical element of a DG system within DS frameworks is establishing guidelines for collecting the metadata needed to track data usage, as well as for defining security and regulatory classifications at both the data element and data set levels. Processes must also be defined to analyze the impact of any changes made to data elements. Furthermore, it is important to explore federated metadata management models within the DS, while also automating processes for metadata search, capture, interpretation, and enrichment. The integration of advanced techniques, including machine learning and artificial intelligence, into metadata cataloging and classification can significantly enhance these processes.

To ensure high data quality, it is imperative to establish comprehensive policies, standards, and processes. This involves identifying the criteria for measuring data quality, which may include accuracy, reliability, integrity, timeliness, etc., and defining corresponding quality rules. Moreover, mechanisms must be in place to assess and report on the quality level of the data contributed by participants [44]. Additionally, it is essential to implement processes for automatic, continuous, and real-time evaluation and validation of data quality at each stage of DLC. These processes should be capable of issuing real-time alerts and, when necessary, automatically correcting any identified issues.

Sensitive data must be identified, and appropriate security classifications should be established, taking into account usage conditions, relevant regulations, standards, and legislation. Agreements between parties should explicitly define policies that outline permissible actions with the data, designate who is authorized to perform those actions, and specify the terms and conditions based on the data classification. Such policies are designed to maintain the confidentiality, integrity, and high availability of shared data at all times. Additionally, robust mechanisms must be implemented to ensure the validation and authentication of both data providers and consumers.

Policies, processes, and a system of measures and Key Performance Indicators (KPIs) must be defined to monitor the performance of data usage within the DS. These measures should ensure that data-related strategies are being properly executed and that data usage and management are aligned with established agreements and policies. The monitoring process should encompass three categories of metrics: those assessing the maturity level of processes, those measuring the performance of processes within data management (DM) frameworks, and those evaluating data governance (DG) elements, such as policies, principles, and organizational structures.

In addition to the general requirements that are common to any decentralized governance system, the DG system within DSs must be capable of managing dispersed and distributed data while facilitating data sharing and exchange among participants. Therefore, it must adhere to the principles of trust, data sovereignty, transparency, interoperability, and legal and ethical compliance [45]. From this perspective, a DG authority for the data space must be established.

Furthermore, DSs must define agreements or contracts that regulate data sharing and exchange, specifying the conditions under which these activities will take place [23]. These agreements should outline the responsibilities and acceptable uses of shared data. Sharing and usage agreements must also describe applicable policies for various data lifecycle phases, including those related to the final destruction of data. These contracts and policies should be attached as metadata to the shared data and must be software-readable (termed self-aware contracts) to enable digital management. The data owner and/or the governance body of the DS should have the ability to audit the actual usage of their data and verify compliance with the agreed-upon terms.

Additionally, shared data assets and their usage restrictions, the employed data structures, vocabularies, taxonomies, and the technical means of access must be sufficiently described and made accessible to participants in the DS. Mechanisms should be placed to facilitate the discovery and availability of this information. This description should utilize standards and/or commonly agreed-upon criteria.

In the next section, we describe a reference enterprise architecture for DG that provides ABBs that satisfy these requirements with explanatory examples based on considerations and documentation from the organizations involved in European DSs initiatives.

5. Reference Enterprise Architecture to Deploy DG in DS

The TRENTI DG framework [18–20] consists of a reference architecture for representing the DG system, a method outlining the steps for architecture deployment, a list of adopted standards, and a maturity model. Together, these components enable the creation of DG systems aligned with the organization's data strategy. Although the framework was developed for Industry 4.0, its core is general and can be used in any sector. Likewise, the international standards on which it is based are those commonly used in organizations such as TOGAF® Standard [37], Reusable Asset Specification [46], ISO 38505 [38], ISO 8000 [47], etc., so its implementation is not a disruptive process.

The reference architecture consists of a series of Architecture Building Blocks (ABBs) (see Figure 1). Each of them specifies a functionality and capabilities that the architecture must implement. It translates abstract principles into actionable routines and structures [48]. Table 1 provides a brief description of each ABB in the reference model, as well as the adaptations made to each ABB for the DS context. Table 2 gathers the tasks that arise from the requirements for the DG in DSs described in Section 4 and the ABBs of the architecture that implement them. These elements are described across three architectural layers: business, information systems, and technology. In this paper, the artifacts are defined at the business level, leaving the selection and implementation of specific technical mechanisms to the DS governing body or whoever it delegates. This approach allows each DS to choose the most suitable tools and solutions according to its maturity level, regulatory constraints, and interoperability requirements. For example, a DS may opt to begin with a basic identity verification solution and later migrate to a more advanced one as its needs evolve. Appendix B presents a table listing the ABBs, the artifacts offered by Trenti for their development, and the tools that help develop these artifacts.

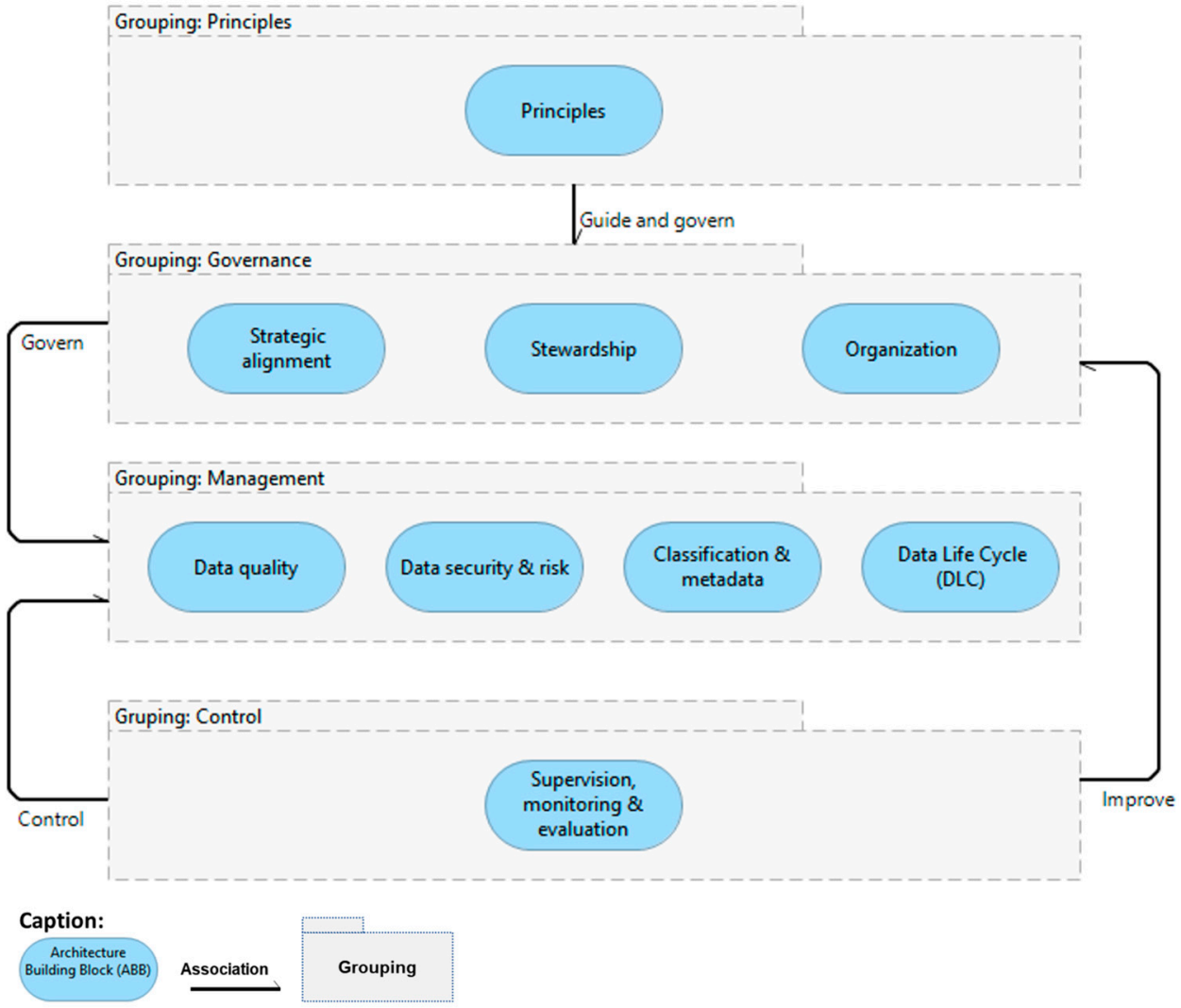

Figure 1. ABB reference model represented using Achimate® notation [49].

Table 1. Brief description of the ABBs.

ABB	Description
PRINCIPLES GROUP	
Principle	Specifies the enterprise architecture for the definition, cataloging, and management of a set of principles that guide DG. This ABB has been extended with principles specific to DSs, such as Sharing, Trust, Sovereignty, and Interoperability.
GOVERNANCE GROUP	
DG strategic alignment	Defines the enterprise architecture required to identify the information needs based on the DS objectives. This ABB has been extended to translate information needs into specific data requests for data providers within DS. Additionally, it defines the goals of DG in DS, ensuring they align with the DS overall strategy.
Stewardship	This ABB consists of three building blocks: (i) Roles and Responsibilities ABB, which establishes the enterprise architecture required for the definition of the DS-specific roles (data owner, data provider, data user, data consumer, etc.), their responsibilities, and decision-making rights. (ii) Policies and Standards ABB, which facilitates the transition from principles to the development of concrete policies and standards. (iii) Governance Model ABB, which defines a governance model centered on the key functions: Evaluate, Direct, and Monitor (EDM).
Organization	Specifies the enterprise architecture that helps define DS governing bodies, the bodies that are governed, and the organizational model of the DS governance.
MANAGEMENT GROUP	
Metadata management	Determines the enterprise architecture that implements DG system requirements to address data semantics, data representation, and the registration of data element descriptions, along with the management of metadata, which must include metadata about the quality of the data exchanged in the DS.
Data quality	Specifies the enterprise architecture for planning, development, and execution of data quality policies.
Data security, privacy and risk	Stipulates the enterprise architecture required to support the planning, development, and implementation of security and data protection policies, considering DS requirements, applicable legislation, and contractual obligations as defined in the data contracts signed between DS users.
DLC activities	Develops the enterprise architecture required to identify data-related activities and the specific aspects of data that are subject to governance in the DS. This ABB is further decomposed into several sub-blocks, each corresponding to a specific phase of the data lifecycle (Collect, Store, Process, Transmit, Exchange/Distribute, Use, and Destroy) where roles and responsibilities for these activities will be established and KPIs defined to measure their performance.
CONTROL GROUP	
Supervision, monitoring and evaluation	Specifies the enterprise architecture that enables the definition and management of KPIs, processes, and resources needed to monitor, measure, analyze, and report on the effectiveness and performance of the DG system.

Table 2. Tasks required to implement a DG system for DSs.

Task	Responsible ABB
Define principles guiding the conduct, behavior, and philosophy of the DS and its members regarding data usage, management, and governance.	ABB. Principles
Establish a DG plan for the DS, defining goals, objectives, and strategies aligned with DS objectives.	ABB. Strategic Alignment
Create a DG authority for the DS, identifying relevant actors and governance bodies. In DSs, this authority is part of the DS governing body.	ABB. Organization
Develop a federated model for governing dispersed and distributed data.	ABB. Organization
Ensure DS members commit to complying with the established DG system.	ABB. Organization
Define roles, responsibilities, and decision rights, assigning specific activities and ensuring alignment with frameworks like IDSA-RAM.	ABB. Roles, Responsibilities, and Decision Rights
Formulate agreements/contracts to regulate data sharing and exchange, considering standards and applicable legislation.	ABB. Policies and Standards. ABB Exchange/Distribute
Enable the establishment and management of software-readable agreements/contracts.	ABB. Policies and Standards. ABB Exchange/Distribute
Attach metadata to shared data to explicitly define usage conditions and ensure compliance, including metadata on data quality.	ABB. Metadata
Implement mechanisms allowing the data owner or DS governance body to audit the actual usage of shared data as well as the ethical handling of data.	ABB. Roles, Responsibilities, and Decision Rights; ABB. Supervision, monitoring & evaluation
Define parameters and processes for evaluating the quality of shared data. Quality validation must be performed "ex ante," or before sharing the data, and "ex post," or at the time of receiving the data. The goal is to ensure that the data meets the promised or expected quality standards.	ABB. Quality Management
Provide directories and catalogs to enable the discovery and access to information about DS participants and available data assets.	ABB. Metadata
Facilitate semantic or conceptual interoperability by selecting standards and describing vocabularies, taxonomies and ontologies.	ABB. Metadata
Establish security classifications for sensitive data and define policies for use in contracts, ensuring confidentiality, integrity, availability, and authentication of participants.	ABB. Data security and risks
Define KPIs to measure the fulfillment of the data governance program	ABB. Supervision, monitoring & evaluation

To formally specify, describe and manage reusable ABBs, we created ABB-profile (see Figure 2), an extension of the "Default Profile" of the Reusable Asset Specification (RAS), Version 2.2 [46]. standard and the ISO-42010:2011 [50] Systems and software engineering—Architecture Description standard. ABB-profile is made up of the following classes:

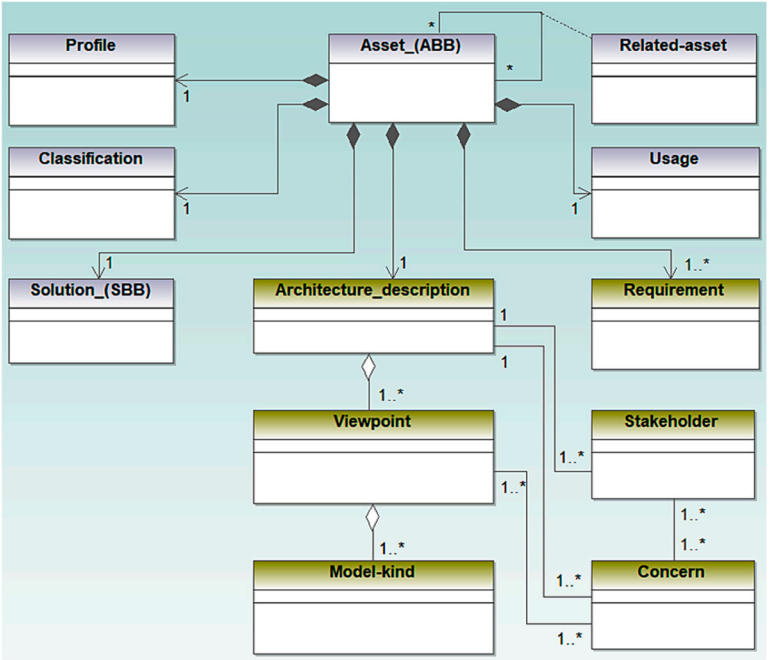

Figure 2. ABB_profile represented using UML® notation [51]. The elements that correspond to the RAS standard are shaded in gray. Elements that correspond to the ISO-42010 standard are shaded green.

Asset_(ABB). It defines the ABB. It contains three required attributes: Name, Id and Description. It is composed of the following elements of the RAS standard default profile:

- **Profile**. It describes the reusable asset type and provides information about its lineage. In this case, as stated above, it is an extension of the "Default Profile" of RAS.
- **Classification**. This class contains a set of descriptors to classify the ABB. Classification allows the ABB to be managed and located in a repository.
- **Solution**. It describes the Solutions Building Blocks (SBBs), which will be instantiated to implement the Governance System defined by the architecture. SBBs define the specific products and components that contribute to implementing the functionality and capabilities defined in the ABB.
- **Usage**. Depending on the level of detail of the ABB, this class describes the activities to be performed for implementing or using the asset, so that it can guide the development of SBBs.
- **Related-asset**. Exposes other related ABBs and the type of relationship (aggregation, similarity, dependency, parent, composition).
- **Requirement**. It describes each requirement that ABB implements.
- **Architecture-description**. This class describes and communicates different parts of the architecture defined by ABB, according to the ISO/IEC/IEEE 42010:2011 standard [50]. It is made up of the following classes:
  - **Stakeholder**. It contains information about an individual, team, organization or categories thereof, who has an interest in the DG system and for which the Viewpoints are built.
  - **Concern**. Describes aspects of the system that are of concern or interest to some stakeholder.
  - **Viewpoint**. It establishes the conventions for the construction, interpretation and use of ABB architecture views to frame specific DG system concerns.
  - **Model type**. This class establishes the modelling conventions for each type of model, related to a viewpoint.

In order to demonstrate the application of the model for the creation and management of ABBs, below we will use the above ABB_profile to describe some of the ABBs that are part of a DG system in a DS. These ABBs have been chosen based on considerations and documentation from the organizations involved in European DS initiatives. For the graphical descriptions, we will use the ArchiMate® Enterprise Architecture modeling language v3.1 standard.

5.1. ABB DS_Principles

As noted in the previous sections of this document, all participants in a DS must adhere to common principles. Although the literature reviewed barely details specific DG principles in the DS, we have found mentions of different types of principles related to the DS, such as DS governance principles, DS design principles or principles of the Common European Data Spaces. The concept "Principle" in our architecture should be understood as a statement of intent that defines a fundamental norm or idea that guides thinking and behavior in DS. Next, we will explain how an enterprise architecture would collect and represent the principles related to DS, data and DG, without addressing the technological aspects for its implementation. Later, we will instantiate some of these principles specific to a DS.

To represent and manage DG principles in the DS, we have created the ABB DS_Principles, which has the structure reflected in Figure 3 and whose description is presented in Table 3.

Table 3. ABB DS_Principles.

Attribute	Description
Name	ABB DS_Principles
Description	This ABB makes it possible for the principles to be defined, cataloged and managed. The DS Architecture Principles establish, in turn, high-level requirements that condition the architecture process, affecting the design, development, maintenance and use of the DG architecture.
Requirement	The ABB contributes to the implementation of the following requirements: "Principles must be established to guide the conduct, behavior and philosophy of the DS regarding the sharing, use, governance and management of data." And "DG principles must be aligned with and support DS principles".
Classification	DG Principles; DS Principles; motivational elements; elements for decision-making.
Usage	It is used to record and manage the principles that guide the DG system, as well as a reference for decision-making; to justify other system requirements and to demonstrate the coherence between principles and the objectives and goals they support.
Related-asset	Name: Catalog management service. Description: Service for catalog management. Allows you to register, modify and delete principles in the catalog.
Related-asset	Name: Principles development process. Description: Defines the sequence of activities to be carried out for the development and approval of the principles.
Related-asset	Name: Principles catalog. Description: Defines the data entity where the established principles are recorded.

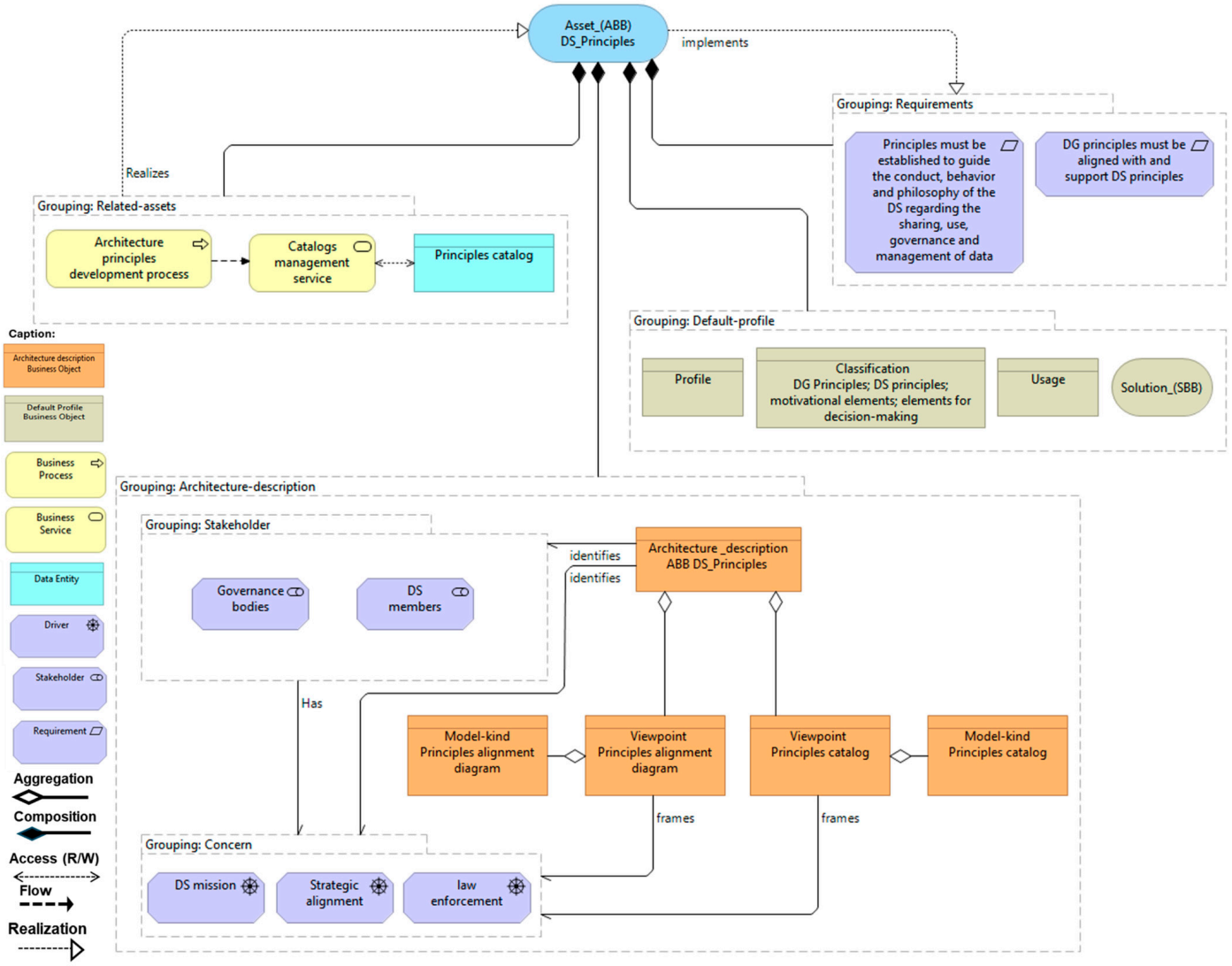

Figure 3. ABB DS_Principles represented using Achimate® notation [49].

For space reasons, below we will only develop the grouping related-assets of this ABB and the viewpoint principles catalog of the grouping architecture-description.

5.1.1. ABB DS_Principles. Grouping Related-Assets

This group (see Figure 4) implements the modules of the enterprise architecture necessary to establish the principles in the DS. To make this possible, a process must be established in the DS that executes the necessary activities to develop and approve the principles that will govern the DS, as well as to identify those responsible for each activity in the process. In addition, a data entity called the principles catalog was created to record the defined principles resulting from the application of the process. It also implements a business service that is responsible for managing the principles catalog. With all these components, DS members can develop the enterprise structure that allows them to define and manage the principles that guide DG in DS.

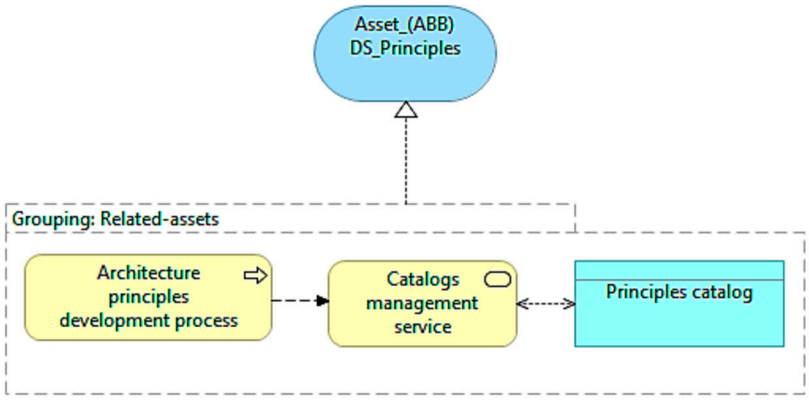

Figure 4. ABB DS_Principles. Related assets illustrated using Achimate® notation [49].

5.1.2. ABB DS_Principles. Grouping Architecture-Description

As shown in Figure 5, the architecture-description class in ABB DS_Principles is composed of the viewpoints: principles catalog and diagram of alignment of the principles with the goals and objectives of the DS.

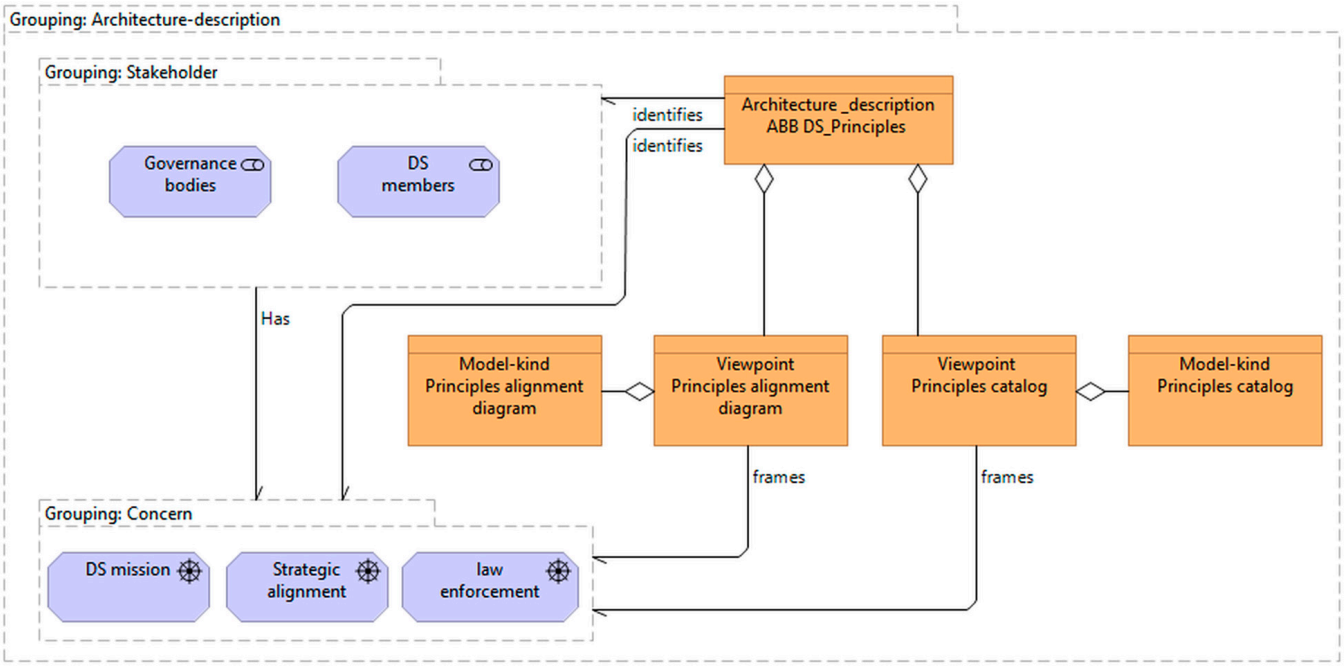

Figure 5. ABB DS_Principles. Architecture description depicted using Achimate® notation [49].

Principles catalog viewpoint

The principles catalog viewpoint is described in Table 4 and presents to the governing bodies and members of the DS, stakeholders of the DG system, how the principles should be recorded so that they can be managed. This contributes to satisfying their interests and concerns regarding the existence of common principles that guide data governance in the DS, aligned with current legislation and the mission and data strategy of the DS.

Table 4. Principles catalog viewpoint description.

Attribute	Description
Viewpoint Name	Principles Catalog
Description	This catalog captures the DG Principles for the DS according to an agreed structure (see model type in Table 5) that allows for their dissemination and management. A principle represents a statement of intent that defines a fundamental norm or idea that guides thinking and behavior in DS.
Type	Catalog
Stakeholders	Governing bodies and members of the DS
Concerns	DS mission, Strategic alignment, law enforcement

Table 5. Principle catalog model type description.

Attribute	Description
Id	Unique and exclusive identifier of the principle
Name	name given to the principle
Category	The following categories of principles apply: DS principles, DG principles
Owner	Responsible for defining and managing the principle
Statement	Statement that sets out the principle in an unequivocal, concise and clear manner
Rationale	Reasons justifying the principle, highlighting how it contributes to DS or DG objectives and strategies, the benefits it brings to the business and the relationships with other principles, including levels of priority or importance of some over others
Implication	Consequences of assuming or not assuming the principle. Exposition of the necessary requirements to comply with the principle, in terms of resources, activities and costs

The model type describes the principle catalog structure (see Table 5).

The DG system includes three groups of principles: (i) DS principles, (ii) DG principles and (iii) data principles. All of them are collected in Figure 6, where the DG principles are general principles established from the ISO/IEC 38505-1 standard [52].

The catalog of principles will then be instantiated, using the model type exposed in Table 5, to define only the principles that are specific to DS, such as sharing, trust, sovereignty and interoperability (see Tables 6–9 below).

Table 6. Principle of sharing.

Attribute	Content
Id	DSDGP001
Name	Sharing
Category	DS principles
Owner	Governance bodies
Statement	DG must contribute to enabling DS participants to provide and receive data with a high degree of control, or to shape and contribute to this exchange.
Rationale	Data sharing can be considered as the most prominent principle, given that it is also the main objective of DS. Data sharing is understood in the sense of enabling access and facilitating the exchange of data, although it must be taken into account that data sharing does not necessarily imply the transfer of data from one entity to another.
Implication	Without this principle, the DS will not be able to operate.

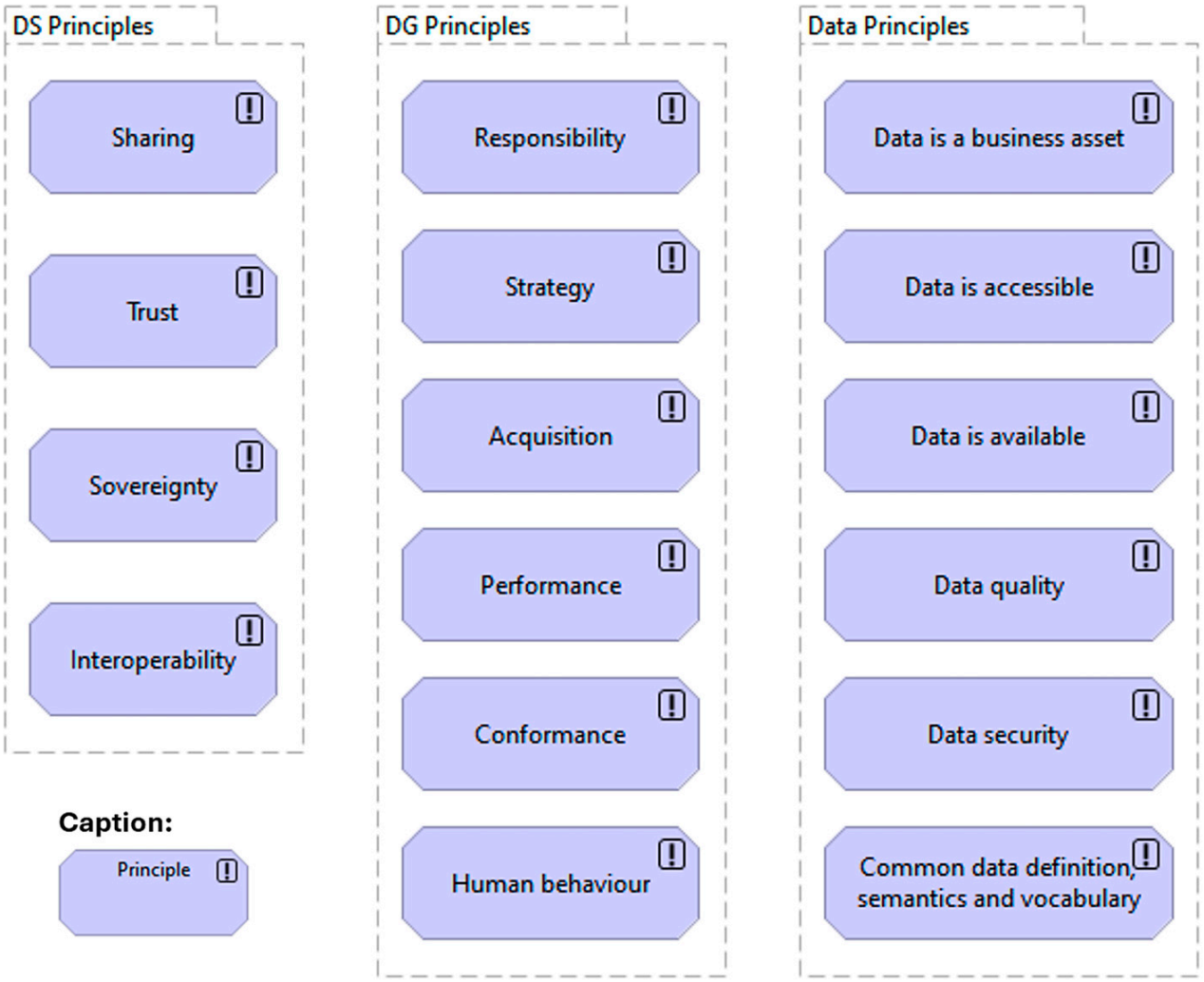

Figure 6. Principles of the DG system in DS represented using Achimate® notation [49].

Table 7. Principle of trust.

Attribute	Content
Id	DSDGP002
Name	Trust
Category	DS principles
Owner	Governance bodies
Statement	DG must establish the bases that regulate the exchange or sharing of data and the conditions under which it will occur, by formulating the corresponding agreements or contracts, so that an environment of trust is created between the parties.
Rationale	Data exchange/sharing is based on mutual trust between the parties involved.
Implication	Trust in the DS does not exist by itself but is achieved through the application of other principles, such as sovereignty and interoperability.

Table 8. Principle of sovereignty.

Attribute	Content
Id	DSDGP003
Name	Data Sovereignty
Category	DS principles
Owner	Governance bodies
Statement	The data owner should have the ability to exercise self-determination regarding the use of his or her data and, therefore, to make decisions and exercise control over that data and its use at all times.
Rationale	Participants should be able to decide whether to participate in the exchange according to their preferences regarding content, scope, purpose, duration and participants.
Implication	Without sovereignty, there will be no trust. The greater the sovereign control over data, the more participants in the data space can trust it. Sovereignty thus becomes an important factor in fostering trust among participants in the DS and also aligns with emerging legal and regulatory trends in various regions, particularly within the EU.

Table 9. Principle of interoperability.

Attribute	Content
Id	DGP004
Name	Data Interoperability
Category	DS principles
Owner	Governance bodies
Statement	DG should contribute to achieving organizational and conceptual or semantic interoperability in the DS.
Rationale	Interoperability is a fundamental aspect and must exist at different levels in the DS to enable its members to access or exchange data efficiently. DS members must be aligned with the DG that guides the management of data. Furthermore, both the data provider and the consumer must have the same interpretation of the meaning of the shared data.
Implication	Organizational interoperability involves identifying the stakeholders involved in the management, provision and use of data, as well as the relationships between them, and ensuring that participants in the exchange or sharing of data are aligned regarding the principles, guidelines, policies and rules that govern the management of the data being shared. Conceptual interoperability involves establishing a common language and reference model to address the wide variety of data and the relationships between them, such as the development of business glossaries, dictionaries and data catalogs, as well as the policies and processes for their management.

5.2. ABB DS_Data Exchange

We have also applied the Asset_ABB model to the exchange/distribute stage of the DLC in a DS (Collect, Store, Process, Transmit, Exchange/Distribute, Use, and Destroy [40]) and created the ABB DS_Data Exchange whose structure can be seen in Figure 7 and whose description is presented in Table 10.

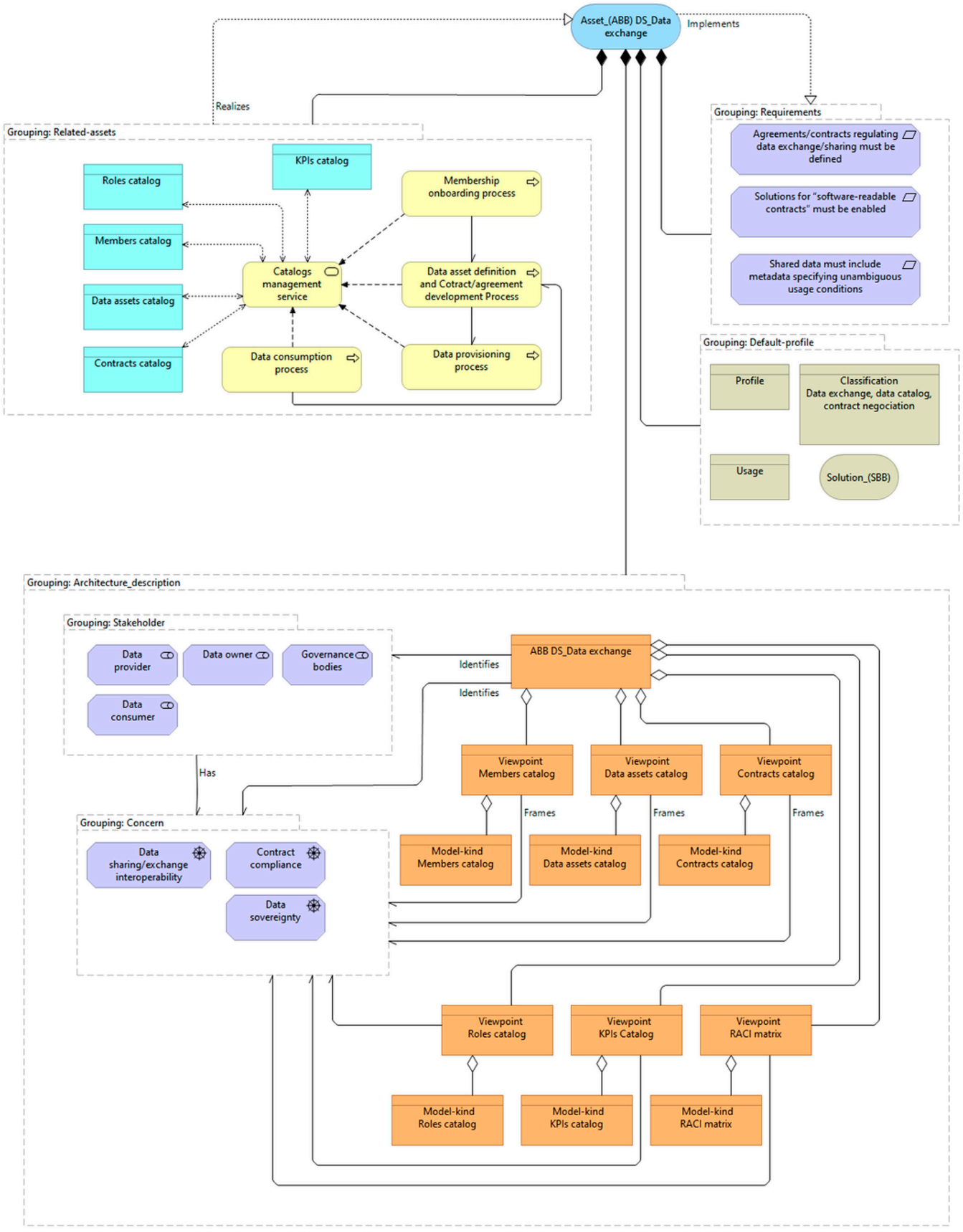

Figure 7. ABB DS_Data exchange illustrated using Achimate® notation [49].

Table 10. ABB DS_Data Exchange description.

Attribute	Description
Name	ABB DS_Data Exchange
Description	Develops the enterprise architecture related to the business processes necessary for data exchange and sharing.
Requirement	Agreements/contracts regulating data exchange/sharing must be defined; Solutions for "software-readable contracts" must be enabled; Shared data must include metadata specifying unambiguous usage conditions.
Classification	Data exchange, data catalog, contract negotiation
Usage	It is used to design the enterprise structure related to data exchange and to register and manage DS members, contracts between members, and information about available data assets.
Related-asset	Name: Catalog management service. Description: Service for catalog management. Allows you to register, modify and delete elements in the catalog.
Related-asset	Name: Membership onboarding process. Description: Defines the sequence of activities to be carried out for the application, evaluation and certification/approval of the members in the DS.
Related-asset	Name: Data asset definition and contract/agreement development Process. Description: Defines the sequence of activities to be carried out for the creation and approval of a data asset to be shared, as well as the development of the conditions of quality, use, security, etc. under which it will be shared.
Related-asset	Name: Data provisioning process. Description: Defines the sequence of activities to be carried out for the internal data pipeline requirements definition, the operational implementation and test management for the technical implementation of the data exchange according to the specified standards
Related-asset	Name: Data consumption process. Description: Defines the sequence of activities to be carried out for defining consumer data needs, acceptance of the exchange contract and consumption and further use of data under consideration of the rules (policies) agreed with the data provider (partner).
Related-asset	Name: Members catalog. Description: Defines the data entity where the members of the DS are recorded.
Related-asset	Name: Data assets catalog. Description: Defines the data entity where the shared data assets are recorded.
Related-asset	Name: Contracts catalog. Description: Defines the data entity where the data sharing contracts are recorded.
Related-asset	Name: Roles catalog. Description: Defines the data entity where roles and responsibilities are recorded.
Related-asset	Name: KPIs catalog. Description: Defines the data entity where KPIs are recorded.

5.2.1. ABB DS_Data Exchange. Grouping Related-Assets

In this case, to implement the necessary enterprise architecture (see Figure 8) for data exchange and sharing, we need to incorporate a process for the application, evaluation, and certification/approval of members in the DS, before they can start sharing or consuming data. Throughout this process, a certification body belonging to the DS governing bodies supervises that the applicant meets all the requirements (regulatory, legislation, standards, organizational, and technical) to join as a member of the DS and to share or access data between the members. A process must also be implemented for the definition of data

assets to be shared and the usage contract associated with each of them. Finally, the data provision and consumption processes must be elaborated.

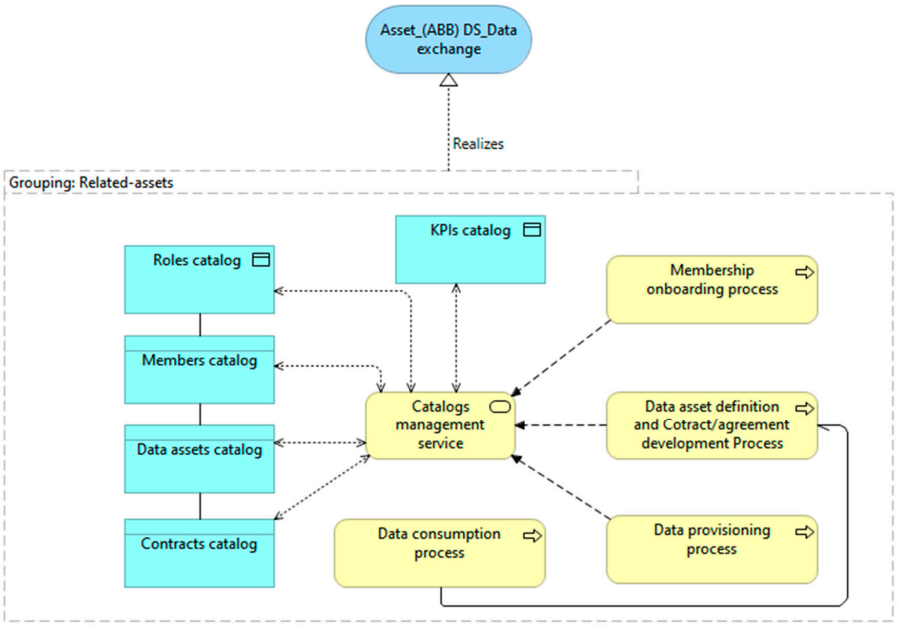

Figure 8. ABB DS_Data Exchange. Related-assets depicted using Achimate® notation [49].

5.2.2. ABB DS_Data Exchange. Grouping Architecture-Description

As shown in Figure 9, the architecture-description class in the ABB DS_Data exchange is composed of the viewpoints: members catalog, contracts catalog and data assets catalogs, roles catalog, KPIs catalog and RACI matrix. The corresponding viewpoints for the data asset catalog, KPI catalog, and RACI matrix are presented below, along with an example of each.

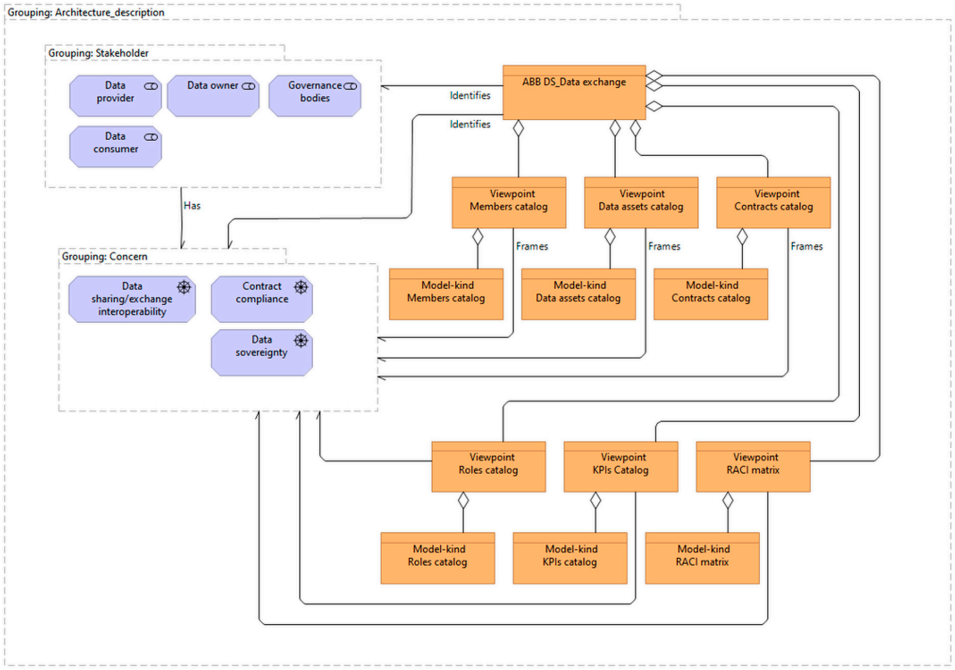

Figure 9. ABB DS_Data Exchange. Architecture description represented using Achimate® notation [49].

Data asset catalog viewpoint

Table 11 describes the data asset catalog viewpoint, which presents to stakeholders how information about data assets should be recorded so that they can be managed and located by data consumers. This contributes to satisfying their interest in ensuring interoperability and data sovereignty in the DS.

Table 11. Data asset catalog viewpoint description.

Attribute	Description
Viewpoint Name	Data Assets Catalog
Description	Defines the data entity that contains information about the different data assets of the DS. It allows data consumers to discover the data assets available in the DS, its format, structure and other metadata. The data asset catalog is a single source of reference for locating and understanding any data asset. It allows DS to organize all its data sets and make it easy for data consumers to locate them on demand. It is a fundamental tool for understanding data, as it provides information about the metadata associated with it.
Type	Catalog
Stakeholders	Data provider, Data consumer
Concerns	Interoperability, sovereignty

The data assets catalog is modeled after the structure shown in Table 12.

Table 12. Data assets catalog model type description.

Attribute	Description
Id	Unique code that identifies this data asset.
Name	Name assigned to the data asset.
Provider	The actor assigned the role of data provider for this data asset.
Description	Textual description of the data asset.
Logical-model	Link to the logical data model associated with the data asset (if applicable).
Physical-model	Link to the physical data model associated with the data asset (if applicable).
Lineage	Link to the information about the data lineage of the data asset (if applicable).
Data-type	Allowed values are: Not defined; Structured; Unstructured; Semi-structured.
Category	Categorization or taxonomy assigned to the data asset.
classification	Classification assigned to the data asset (e.g., Public, Internal, Confidential, Restricted)
Contract ref.	Link to the contract that determines the terms of use and other policies applicable to the data asset.
Store	Link to the storage, archiving or maintenance policies for the data asset (according to contract).
Use	Link to the policies applicable to the consumption or use of this data asset (according to contract).
retention	Link to the policies on the time and manner in which this information must be maintained (according to contract).
Security-policies	Link to the policies related to the security and privacy of the data asset.
Quality-policies	Link to the policies related to the quality of the data asset.

Table 13 shows an example of the instantiation of an entry in the data assets catalog that represents a hypothetical data asset on the mean times between failures of the hair dryer product line, which is shared in the DS.

Table 13. Data assets catalog example.

Attribute	Content
Id	DS_ACME_mtbf__line_3
Name	Mean Time Between Failures (MTBF)
Provider	ACME_ Hair dryer product-line
Description	Data set representing information regarding the mean times between failures for each product in product-line 3
Logical model	LMD_dbprol3
Physical model	FMD_dbprol3
Lineage	DL_Aigpl3TS
Data-type	Structured
Category	operating data
Classification	Internal
Contract ref.	DSC_240905A
Store	Storage audit policy; Operational data archiving policy;
Use	Analytics Sandbox
Retention	Operational Data Retention Policy
Security policies	Operational Data Backup Policy; Data Consumer Identity Verification Policy; Data sovereignty assurance policy
Quality policies	Threshold Check Rule; Correlation Check Rule

RACI matrix viewpoint

Table 14 describes the RACI matrix viewpoint, which presents to stakeholders the allocation of responsibilities for the defined roles. This contributes to satisfying their interest in ensuring that all the key activities have a designated person responsible for their execution.

Table 14. RACI matrix viewpoint description.

Attribute	Description
Viewpoint Name	RACI Matrix
Description	This matrix outlines the governance and data management responsibilities of each role in the DS.
Type	Matrix
Stakeholders	DG body, DS governance body
Concerns	Allocation of responsibilities

This artifact is modeled with the structure shown in Table 15.

Table 15. RACI matrix model type description.

Attribute	Description
Model type	It consists of the entities "Role" and "Activity," establishing the relationship "Role Performs Activity," which indicates that the role has a degree of responsibility in performing an activity. This relationship is represented by a matrix with roles in the columns and activities in the rows. The cells indicate the assignment of a role to an activity and the level of responsibility assigned. There are four types: R—Responsible: The person must perform the task and is responsible for its implementation. A—Accountable: The person has the faculty to authorize actions or tasks and their implementation and is accountable for them. C—Consulted: The opinion or advice of this person should be requested. I—Informed: The person must be informed promptly about the execution of the task.

Table 16 illustrates an example of the allocation of responsibilities for some data exchange activities carried out by this ABB.

Table 16. Responsibilities of some of the roles involved in data exchange.

Activity\Role	DOW	DPR	DCO	DUS	CBY
Membership onboarding	I,C	I,C	—	—	A,R
Data asset definition	A	R	—	—	—
Contractual conditions	A	I	—	—	R
Data provisioning	A	R	I	I	I,C
Data consumption	I	I	A	R	I

DOW: Data owner. DPR: Data provider. DCO: Data consumer. DUS: Data user. CBY: Certification body.

Key Performance Indicators Catalog (KPIs) viewpoint

Table 17 describes the KPI catalog viewpoint, which informs stakeholders of the effectiveness and efficiency of the data governance system, as well as its adherence to applicable DG rules and laws.

Table 17. KPI catalog viewpoint description.

Attribute	Description
Viewpoint Name	KPI Catalog.
Description	This catalog contains indicators that monitor the evolution of key aspects of the GD system.
Type	Catalog
Stakeholders	DG body, DS governance body
Concerns	Performance of the DG, compliance with policies, standards, and applicable legislation, compliance.

This artifact is modeled after the structure shown in Table 18.

Table 18. KPI catalog model type description.

Attribute	Description
Model type	This catalog is made up of entities of the "Measure" type and includes the following attributes: • Id.: Unique identifier of the indicator. • Name: The name assigned to the indicator. • Description: A description of the indicator, its purpose, and its objective. It explains what the indicator is used for. • Category: If categories are established, this is the one assigned to the indicator. • Questions: Questions to be answered by the indicator. • Representation: How the indicator will be represented (e.g., graph, table, etc.). • Calculation: Describes the input data or variables involved, as well as the algorithm, formula, or operations applied to them to obtain the indicator value. • Aggregation: Explains how the indicator is presented (e.g., current value, accumulated by date, or average value). • Collection: Refers to how, when, and from what source the necessary data for calculating the indicator is collected. • Dissemination: It indicates how, through what means, and to whom the indicator is distributed, as well as when. • Retention and Disposal: Refers to the applicable storage, retention, and disposal policies of the indicator. • Interpretation: Explains how the indicator's value should be interpreted. • Related KPIs: References to related indicators. • Owner: The person responsible for defining and managing the indicator.

Table 19 provides an example of potential KPIs for evaluating the performance of activities conducted by this ABB.

Table 19. Some KPIs related to data exchange activities.

Activity	KPI Name	KPI Definition
Membership onboarding	Compliance Certification Accuracy (%)	Rate of correctly certificated participants
	Average Onboarding Time	Time required from application to full membership activation
Data asset definition	Metadata Completeness (%)	Rate of required metadata fields properly filled in
	Average Time to Register a Data Asset	Time from submission to asset availability in the catalog
Contractual conditions for data exchange	Policy Compliance Rate (%)	Rate of contracts that fully align with governance, legal, and data protection policies
	Contract Setup Time	Average duration to establish a contractual agreement for data use
Data provisioning	Data Availability (%)	Rate of data items available when required
	Data Quality Compliance	Compliance with the agreed-upon quality levels in the data contract
Data consumption	Consumer Satisfaction Score	User-reported quality of data consumption experience
	Usage Compliance Rate (%)	Rate of consumption issues that are not in compliance with contractual and policy rules

6. Discussion

DSs are still in an emergent stage [53]. They have emerged as a promising solution for sharing and reusing data across organizations [54]. The first publicly funded projects began in Germany in 2015, but it was not until 2020 that the Gaia-X Project was launched in the EU, focusing on a federated and sovereign data infrastructure. Other initiatives in Australia and India followed this trend [55,56]. However, widespread adoption requires not only qualified organizational and technical capabilities but also well-grounded frameworks to address existing barriers and foster the development of sustainable data-sharing ecosystems.

Copei et al. [57] identify several major challenges. One of the most significant is the reluctance of companies to share data due to concerns about exposing proprietary know-how and the potential loss of competitive advantage. Another critical issue is the difficulty in demonstrating the value of data sharing, along with the need for highly skilled IT personnel to deploy and maintain DS. At the technical level, barriers include identity management, lack of interoperability between different software stacks, and insufficient or overly complex access and usage control mechanisms. At the organizational level, reliance on a small open-source community is also considered a significant limitation. In this regard, Moritz et al. [54] add that cross-organizational coordination costs also represent an inherent obstacle in the development of data-sharing ecosystems, as they require consensus on both technical and organizational capabilities. This includes aligning data models, standards, and business processes across multiple actors, which can be complex and resource-intensive, as well as ensuring trust and active engagement within the ecosystem.

To address these challenges, Copei et al. [57] propose the creation of a robust governance framework that clearly defines roles, responsibilities, and decision-making structures to build trust. They also emphasize the importance of clearly communicating both the capabilities and limitations of data sovereignty, creating incentives and mechanisms for data valuation, and appointing a trusted data space operator with the financial resources, technical expertise, and legal authority necessary to oversee operations. Other researchers contribute with methodological frameworks and design principles. For instance, Moritz et al. [54] introduce a data service model supported by a set of design guidelines to clarify available choices and assist organizations in making informed decisions when building data services. Similarly, Jussen et al. [53] propose a four-phase guide—problematization, intéressement, enrolment, and mobilization—aimed at providing a structured pathway for establishing effective and trustworthy data-sharing ecosystems. In the same vein, Gieß et al. [55] offer a comprehensive solution space for data space designers to (re-)design data spaces more efficiently and gain a holistic understanding of the key elements to consider.

It is within this context that our work is situated. Our contribution aims to serve as a practical guide for DS governing bodies to build a solid governance system. The proposed framework enables the DS to define both an initial and a target architecture, which can be progressively achieved as the organization matures in the capabilities defined by each ABB. Furthermore, the framework is designed to support continuous improvement through a cycle of measurement, evaluation, and action. This approach allows the DS to monitor the maturity of their governance practices, identify gaps, and activate specific mechanisms to address them. By embedding this evaluative logic into the architecture, the DS can ensure alignment with strategic objectives, foster trust among participants, and build a resilient governance system capable of adapting to evolving technological and organizational challenges.

A limitation of this work is that the use case is fictional. After conducting interviews with various Spanish initiatives and projects for implementing DSs and analyzing the situation, we can confirm that these projects are still in the phase of implementing the technical solution and initiating stakeholder engagement. However, the proposed framework has already been successfully tested in industrial environments, which gives us confidence in its applicability to DSs, and it is expected to be put into practice in the near future in the Heleade DS project led by the University of Alicante.

7. Conclusions and Further Research

One of the goals of DS is to create value for its members, and this is achieved by exploiting the data that is shared. Therefore, data becomes a business asset that needs to be managed like any other asset, so DG emerges as a critical aspect for DS. Thus, after analyzing the characteristics of DS and the focus on DG in these environments, we have identified the requirements for DG in a DS and realized that there is a gap between the various proposals, guides and conceptual models for implementing a DS and its actual implementation in an organization.

Therefore, we have identified the DG requirements for their implementation in a DS and proposed a framework that allows organizations to define and implement the enterprise architecture (actors, roles, functions, processes, policies, standards, etc.) needed to operate in a DS, based on architectural building blocks that the DS can deploy and develop according to its specific needs. Additionally, we present a model, the ABB_profile, to formally specify, describe, and manage these ABBs, facilitating their reuse by members of the data space.

As future work, it would be beneficial to apply the proposed architecture to various use cases. This would provide deeper insights into the requirements and finer details of the business architecture. Additionally, steps should be taken to ensure that ABBs can be integrated into data space interoperability resources. This would facilitate their discovery, access, and adoption by data space members, promoting interoperability and reusability.

Author Contributions: Conceptualization, J.Y. and M.Z.; methodology, M.Z.; validation, J.Y. and M.Z.; investigation, J.Y.; writing—original draft preparation, J.Y. and M.Z.; writing—review and editing, J.Y. and M.Z.; supervision, M.Z.; project administration, M.Z.; funding acquisition, M.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by MCIN/AEI/10.13039/501100011033/FEDER "Una manera de hacer Europa" under grant PID2021-124502OB-C42 (PRESECREL) and "The APC was funded by Information Journal".

Data Availability Statement: This study is based on a fictional dataset created for illustrative and methodological purposes. As such, no real or proprietary data were used, and there are no data available for sharing.

Acknowledgments: During the preparation of this manuscript, the authors used DeepL and Copilot to improve the readability and language of the manuscript. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ABB	Architecture Building Block
BDVA	Big Data Value Association
DG	Data Governance
DLC	Data Lifecycle
DS	Data Space
EDM	Evaluate, Direct, and Monitor
IDSA	International Data Spaces Association
SBB	Solutions Building Block

Appendix A

Table A1 outlines the activities that should be managed in each phase of the DLC. Corresponding governance actions should be implemented for each activity, specifically in the areas of data quality assurance, security enforcement, and metadata administration.

Table A1. DLC activities description.

DLC Phase	Activities
Collect	Integration and interoperability of the data that is incorporated and how it should be extracted and processed to create useful data (clean, corrected, complete, and accurate) with the ability to be stored. It involves converting raw data into structured and formatted outputs tailored to their intended use, while also ensuring the collection and management of relevant metadata.
Store	Data storage, maintenance, and improvement (the data does not change intrinsically), or for archiving purposes.
Process	Extracting, preparing, and processing data based on its intended use. Data may be aggregated and/or combined with other data to enable exchange/distribution, analysis, archiving, or deletion.
Transmit	Validating data and contract compliance before departure and upon arrival at the destination, ensuring confidentiality and integrity during transmission.
Exchange/ Distribution	Membership onboarding, data asset definition and contract development, as well as data provisioning and data consumption.
Use	Consumption, use, and analysis of data in order to support decision-making.
Destroy	Permanent deletion of data.

Appendix B

Table A2 presents Trenti's ABBs, along with the corresponding artifacts and the tools that support their creation. The templates for the catalogs proposed by Trenti are available through a dedicated module developed for the Odoo ERP system, which can be accessed at: https://apps.odoo.com/apps/modules/16.0/data_gov#module-description (accessed on 28 July 2025).

Table A2. TRENTI's ABBs, artifacts and tools.

Grouping	ABB	Artifacts	Tools
Principles	Principles	• Catalog of principles. • Principles alignment diagram. • Catalog of requirements.	• Guide [1] for establishing DG principles and their alignment with business objectives. • Template for recording principles. • Guide for creating an alignment diagram. • Alignment diagram template. • Template for recording requirements.
Governance	Strategic alignment	• Company goals and objectives catalog. • Information needs catalog. • DG goals and objectives catalog. • Requirements catalog. • Business-DG alignment diagram. • Information needs-business goals/objectives alignment diagram.	• Guide for developing a DG plan. • Template for recording information needs. • Guide for creating alignment diagrams. • Template for developing diagrams. • Template for recording requirements.
	Organization	• DG Actor catalog. • Contract and SLA catalog. • Requirements catalog. • DG organization diagram. • DG functional diagram.	• Template for recording actors. • Template for recording contracts. • Template for developing diagrams. • Template for recording requirements.
	Stewardship—Governance model.	• Governance model diagram. • Requirements catalog.	• Template for developing procedures derived from the governance model. • Template for developing DG model diagram. • Template for recording requirements.
	Stewardship—Policies and standards.	• Policy catalog. • DG procedures catalog. • Requirements catalog. • Policy-principle alignment diagram.	• Template for recording policies. • Template for developing procedures. • Template for developing Policy-principle alignment diagram. • Template for recording requirements.
	Stewardship—Roles and Responsibilities	• DG Roles catalog. • RACI Matrix. • Requirements catalog.	• Template for recording roles. • Template for creating the RACI matrix. • Template for recording requirements.

Table A2. *Cont.*

Grouping	ABB	Artifacts	Tools
Management	Classification and metadata	• Business glossary catalog. • Data catalog. • Requirements catalog.	• Template for the business glossary and data catalog. • Template for recording requirements.
	Data quality	• Quality rules catalog. • Quality requirements catalog. • Quality characteristics catalog. • Data quality measures catalog. • Requirements catalog.	• Guide to data quality assessment. • Data quality plan template. • Templates for recording quality rules, quality requirements, and quality characteristics. • Template for recording requirements.
	Data security	• Data security classification catalog. • Requirements catalog. • CRUD security matrix. • CRUD security diagram.	• Guide to developing a data security plan. • Template for recording data security classifications. • Template for developing a CRUD matrix and diagram. • Template for recording requirements.
	Data Lifecycle (DLC)	• DLC activities catalog. • Requirements catalog.	• Template for recording DLC activities. • Template for recording requirements.
Control	Supervision, monitoring and evaluation	• KPI catalog. • Requirements catalog.	• Template for recording KPIs. • Template for recording requirements.

[1] Each guide/procedure will include a RACI matrix with the related roles and activities involved.

References

1. Curry, E.; Scerri, S.; Tuikka, T. Data Spaces: Design, Deployment, and Future Directions. In *Data Spaces*; Springer International Publishing: Cham, Switzerland, 2022.
2. Reiberg, A.; Niebel, C.; Kraemer, P. What is a Data Space? In *White Paper 1/2022*; Gaia-X Hub Germany: Munich, Germany, 2022; pp. 1–20.
3. Data Spaces Support Centre. What Is a Data Space? 2024. Available online: https://dssc.eu/space/bv15e/766061351/Introduction+-+Key+Concepts+of+Data+Spaces#What-about-a-Definition? (accessed on 25 November 2024).
4. International Data Spaces Association (IDSA). *Design Principles for Data Spaces—Position Paper*; Open Dei: Rome, Italy, 2021.
5. Data Office. *Plan de Actuaciones Para el Despliegue de Espacios de Datos*; Data Office: Madrid, Spain, 2024.
6. Gaia-X European Association for Data and Cloud AISBL. *Gaia-X-Architecture Document—22.04 Release*; Gaia-X European Association for Data: Brussels, Belgium, 2022; p. 61.
7. Braun, R.; Laa, B.; Rohatsch, L. Governance challenges of urban dataspace—Transdisciplinary perspectives. In Proceedings of the Extended Abstracts of the Eu-SPRI 2022 Conference Challenging Science and Innovation Policy, Utrecht, The Netherlands, 1–3 June 2022; pp. 105–107.
8. International Data Spaces Association (IDSA). IDS-RAM 4. *IDSA Reference Architecture V4.0*. Available online: https://docs.internationaldataspaces.org/ids-knowledgebase/ids-ram-4/ (accessed on 11 September 2025).
9. Otto, B.; Steinbuß, S.; Teuscher, A.; Lohmann, S. *International Data Spaces Association. Reference Architecture Model. V3.0*; International Data Spaces Association: Dortmund, Germany, 2019; pp. 1–118.
10. Minghini, M.; Kotsev, A.; Posada, M.; Signorelli, S. *Scientific Insights into Data Sharing and Utilisation at Scale*; Publications Office of the European Union: Brussels, Belgium, 2023. [CrossRef]
11. Fritzenkötter, J.; Hohoff, L.; Pierri, P.; Verhulst, S.G.; Young, A.; Zacharzewski, A. *Governing the Environment-Related Data Space*; European Union: Brussels, Belgium, 2022.

12. DAMA International. *DAMA-DMBOK2: Data Management Body of Knowledge*, 2nd ed.; Technics Publications: Basking Ridge, NJ, USA, 2017; ISBN 978-1-63462-234-9.
13. Weber, K.; Otto, B.; Österle, H. One Size Does Not Fit All—A Contingency Approach to Data Governance. *J. Data Inf. Qual.* **2009**, *1*, 1–27. [CrossRef]
14. MDM Institute MDM Institute—Data Governance Definition. 2018. Available online: http://www.tcdii.com/whatIsDataGovernance.html (accessed on 23 October 2020).
15. Gray, C.; IBM Corporation. *An overview of data governance elevator Pitch (Final)*; IBM Corporation: Armonk, NY, USA, 2011.
16. Otto, B. Data governance (Otto). *Bus. Inf. Syst. Eng.* **2011**, *3*, 241–244. [CrossRef]
17. Torre-Bastida, A.I.; Gil, G.; Miñón, R.; Díaz-de-Arcaya, J. Technological perspective of data governance in data space ecosystems. In *Data Spaces: Design, Deployment and Future Directions*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 65–87.
18. Yebenes, J. *Marco para la Construcción de Sistemas de Gobernanza de Datos en Entornos de Industria 4.0*; Universidad de Cantabria: Santander, Spain, 2022.
19. Yebenes, J.; Zorrilla, M. A Data Governance Framework for Industry 4.0. *IEEE Lat. Am. Trans.* **2021**, *19*, 2130–2138. [CrossRef]
20. Yebenes, J.; Zorrilla, M. A reference framework for the implementation of Data Governance Systems for Industry 4.0. *Comput. Stand. Interfaces* **2021**, *81*, 103595. [CrossRef]
21. Dudoit, A. *European Common Data Spaces: A Structuring Initiative That Is Necessary and Adaptable to Canada*; CIRANO: Montreal, QC, Canada, 2023.
22. Alliance for Internet of Things Innovation (AIOTI). *Guidance for the Integration of IoT and Edge Computing in Data Spaces*; Alliance for Internet of Things Innovation: Brussels, Belgium, 2022.
23. Suzuki, K.; Yokozeki, D. Data Governance for Achieving Data Sharing in the IOWN Era. *NTT Tech. Rev.* **2023**, *21*, 49–54. [CrossRef]
24. Micheli, M.; Ponti, M.; Craglia, M.; Suman, A.B. Emerging models of data governance in the age of datafication. *BIG DATA Soc.* **2020**, *2020*, 1–15. [CrossRef]
25. Otto, B. A federated infrastructure for European data spaces. *Commun. ACM* **2022**, *65*, 44–45. [CrossRef]
26. Firdausy, D.R.; De Alencar Silva, P.; Van Sinderen, M.; Iacob, M.E. Towards a Reference Enterprise Architecture to enforce Digital Sovereignty in International Data Spaces. In Proceedings of the 2022 IEEE 24th Conference on Business Informatics (CBI), Amsterdam, The Netherlands, 15–17 June 2022; 1, pp. 117–125. [CrossRef]
27. Daclin, N.; Chen, D.; Vallespir, B. Developing enterprise collaboration: A methodology to implement and improve interoperability. *Enterp. Inf. Syst.* **2016**, *10*, 467–504. [CrossRef]
28. Bližnák, K.; Michal, M.; Pilková, A. A Systematic Review of Recent Literature on Data Governance (2017–2023). *IEEE Access* **2024**, *12*, 149875–149888. [CrossRef]
29. Hutterer, A.; Krumay, B. Scopes of Governance in Data Spaces Scopes of Governance in Data Spaces. In Proceedings of the Australasian Conference on Information Systems, Canberra, Australia, 4–6 December 2024.
30. Solmaz, G.; Cirillo, F.; Fürst, J.; Jacobs, T.; Bauer, M.; Kovacs, E.; Santana, J.R.; Sánchez, L. Enabling data spaces: Existing developments and challenges. In Proceedings of the 1st International Workshop on Data Economy, New York, NY, USA, 22 December 2022; pp. 42–48. [CrossRef]
31. Gaia-X European Association for Data and Cloud AISBL. GAIA-X Architecture Document—20.04. 2024. Available online: https://docs.gaia-x.eu/technical-committee/architecture-document/24.04/ (accessed on 2 November 2024).
32. Curry, E.; Tuikka, T.; Metzger, A.; Zillner, S.; Bertels, N.; Ducuing, C.; Dalle Carbonare, D.; Gusmeroli, S.; Scerri, S.; López de Vallejo, I.; et al. Data Sharing Spaces: The BDVA Perspective. In *Designing Data Spaces*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 365–382.
33. Cuno, S.; Bruns, L.; Tcholtchev, N.; Lämmel, P.; Schieferdecker, I. Data governance and sovereignty in urban data spaces based on standardized ICT reference architectures. *Data* **2019**, *4*, 16. [CrossRef]
34. Curry, E. *Real-Time Linked Dataspaces: Enabling Data Ecosystems for Intelligent Systems*; Springer: Berlin/Heidelberg, Germany, 2019.
35. Datos.gob.es. The Dataspaces Starter Kit. 2024. Available online: https://datos.gob.es/en/blog/dataspaces-starter-kit (accessed on 5 November 2024).
36. Data Spaces Support Centre. Data Spaces Blueprint v1.5—Home. 2024. Available online: https://dssc.eu/space/BBE/178422141/Data+Sharing+Governance (accessed on 5 November 2024).
37. The Open Group. *TOGAF® Standard Version 9.2*; The Open Group: San Francisco, CA, USA, 2018; pp. 1–532.
38. *ISO ISO/IEC 38505-2*; Implications of ISO/IEC 38505-1 for Data Management. International Standard Organization: Geneva, Switzerland, 2018; pp. 1–44.
39. Felici, M.; Koulouris, T.; Pearson, S. Accountability for Data Governance in Cloud Ecosystems. In Proceedings of the 2013 IEEE International Conference on Cloud Computing Technology and Science, Hangzhou, China, 2–5 December 2013; Volume 2, pp. 327–332. [CrossRef]

40. Liu, Y.; Gao, T.; Niu, D.; Zhang, H. Research on Collaborative Governance of Data Security in the Whole Life Cycle of Electric Power Manufacturing Data Space. In Proceedings of the 2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Hangzhou, China, 4–6 May 2022; pp. 119–126. [CrossRef]
41. Ballard, C.; Baldwin, J.; Baryudin, A.; Brunell, G.; Giardina, C.; Haber, M.; O'neill, E.A.; Shah, S. *IBM Information Governance Solutions*; IBM Corporation: Armonk, NY, USA, 2014.
42. Marcucci, S.; Alarcón, N.G.; Verhulst, S.G.; Wüllhorst, E. Informing the Global Data Future: Benchmarking Data Governance Frameworks. *Data Policy* **2023**, *5*, e30. [CrossRef]
43. Wallis, K.; Stodt, J.; Jastremskoj, E.; Reich, C. Agreements Between Enterprises Digitized By Smart Contracts in the Domain of Industry 4.0. *Comput. Sci. Inf. Technol. (CS IT)* **2020**, *10*, 23–32. [CrossRef]
44. UNE Normalización Española. *UNE 0087 Definición y Caracterización Espacios de Datos*; UNE Normalización Española: Madrid, Spain, 2025; pp. 1–67.
45. Tardieu, H. Role of Gaia-X in the European Data Space Ecosystem. In *Designing Data Spaces*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 41–59.
46. Object Management Group (OMG). Reusable Asset Specification, Version 2.2, no. November. 2005. pp. 1–121. Available online: https://www.omg.org/spec/RAS/2.2/PDF (accessed on 15 October 2025).
47. *ISO 8000-61*; Data Quality—Process Reference Model. International Standard Organization—ISO: Geneva, Switzerland, 2016; pp. 1–28.
48. Lnenicka, M.; Nikiforova, A.; Luterek, M.; Milic, P.; Rudmark, D.; Neumaier, S.; Kević, K.; Zuiderwijk, A.; Rodríguez Bolívar, M.P. Understanding the development of public data ecosystems: From a conceptual model to a six-generation model of the evolution of public data ecosystems. *Telemat. Inform.* **2024**, *94*, 102190. [CrossRef]
49. The Open Group. ArchiMate® 3.1 Specification. 2019, pp. 1–206. Available online: https://www.opengroup.org/sites/default/files/docs/downloads/n190p_5.pdf (accessed on 15 October 2025).
50. *ISO/IEC/IEEE 42010:2011*; Systems and Software Engineering—Architecture Description. International Standard Organization—ISO: Geneva, Switzerland, 2011; pp. 1–46.
51. Object Management Group (OMG). OMG® Unified Modeling Language® (OMG UML®) V2.5.1, no. December, p. 796, 2017. Available online: https://www.omg.org/spec/UML/2.5.1/PDF (accessed on 15 October 2025).
52. *ISO ISO/IEC 38505-1*; Application of ISO/IEC 38500 to the Governance of Data. International Standard Organization: Geneva, Switzerland, 2017; pp. 1–28.
53. Jussen, I.; Möller, F.; Schweihoff, J.; Gieß, A.; Giussani, G.; Otto, B. Issues in inter-organizational data sharing: Findings from practice and research challenges. *Data Knowl. Eng.* **2024**, *150*, 1022800. [CrossRef]
54. Guggenberger, T.M.; Schlueter Langdon, C.; Otto, B. Data spaces as meta-organisations. *Eur. J. Inf. Syst.* **2025**, *34*, 822–842. [CrossRef]
55. Gieß, A.; Schoormann, T.; Möller, F.; Gür, I. Discovering data spaces: A classification of design options. *Comput. Ind.* **2025**, *164*, 104212. [CrossRef]
56. Möller, F.; Jussen, I.; Springer, V.; Gieß, A.; Schweihoff, J.C.; Gelhaar, J.; Guggenberger, T.; Otto, B. Industrial data ecosystems and data spaces. *Electron. Mark.* **2024**, *34*, 41. [CrossRef]
57. Copei, S.; Rülicke, L. Best Practices to overcome challenges and barriers during the implementation of a data space for the energy domain: An experience report. *Data Br.* **2025**, *61*, 111838. [CrossRef] [PubMed]