gaia-x
Hub Germany

The Governance of Dataspaces

Actors, structures and phases of dataspace governance

Whitepaper 1/2024
February 2024

Dr. Abel Reiberg, Dr. Crispin Niebel,
Dr. Anna-Raphaela Schmitz

Gefördert durch:

Bundesministerium
für Wirtschaft
und Klimaschutz

aufgrund eines Beschlusses
des Deutschen Bundestages

acatech
DEUTSCHE AKADEMIE DER
TECHNIKWISSENSCHAFTEN

About the series

White Papers of the Gaia-X Hub Germany are intended for discourse and the exchange of ideas. They reflect the opinions of the authors and not necessarily those of the Gaia-X Association or any other institution of the Gaia-X ecosystem.

Editors

acatech - National Academy of Science and Engineering
Dr Abel Reiberg, Scientific Officer
Dr Crispin Niebel, Scientific Officer
Dr Anna-Raphaela Schmitz, Project Manager

Authors

acatech - German Academy of Science and Engineering
Dr Abel Reiberg, Scientific Officer
Dr Crispin Niebel, Scientific Officer
Dr Anna-Raphaela Schmitz, Project Manager

Publisher

Gaia-X Hub Germany c/o acatech - National Academy of Science and Engineering
Karolinenplatz 4
80333 Munich

Recommended citation

Reiberg, A., Niebel, C., Schmitz A.-R. (2024) Governance of data spaces. Gaia-X Hub Germany. White Paper 1/2024.

Table of contents

Summary

Governance, i.e. the coordination of the relevant actors, is a crucial part of setting up data spaces. Because data spaces are federated, they allow many actors to participate and operate. On the one hand, federated data spaces improve data sovereignty, competition and innovation, compared to centralised infrastructures. On the other hand, they require adequate governance arrangements to ensure efficient, effective, and fair interaction between stakeholders. This White Paper aims to introduce the main questions of data space governance, and to stimulate discussion on both theoretical solutions and their practical application. To this end, the paper briefly discusses basic questions such as actors, structures, phases and tasks. Each topic will be illustrated through two examples of advanced data space projects. These are the Mobility Data Space (MDS), a data space for mobility data, and Catena-X, a data space for the entire automotive supply chain.

After the introduction, the second section begins by establishing the definitions of “data space” and “governance”, which underpin this White Paper. Data space governance will be defined as coordinating stakeholders in an integrated data infrastructure. Section 3 describes the actors involved in data space governance. It shows that a variety of organisations may be involved – both as participants and as operators of a dataspace – provided that the requirements at the level of the dataspace initiative and at the level of the individual data space are met. Section 4 discusses governance structures. The examples of Gaia-X, MDS and Catena-X illustrate the complexity of governance arrangements and the breadth of participation possibilities. Section 5 presents specific legal requirements to consider, such as the Data Governance Act.

Section 6 describes the tasks of data space governance involved in the different phases of establishing a data space. It shows how the stakeholder networks and their interaction forms are expected to evolve during the design, establishment, and operation of data spaces. Section 7 deals with the tasks of data space governance. It highlights the importance of designing the participants group and interaction structures between them, with a special attention to network effects. The paper concludes that new data space projects can draw important lessons from projects already in development. However, advantages and disadvantages of specific governance solutions must be weighed.

1. Introduction

Several measures, in particular governmental ones, currently support the development of data spaces. These include regulations, such as the Data Governance Act (DGA), and funding measures such as those for European Data Spaces and Gaia-X.

The intensive development of data spaces offers important opportunities. It strengthens the European data economy while respecting values such as security, transparency and sovereignty. At the same time, developing data spaces faces many challenges: Decentralised infrastructures such as data spaces involve many different actors within various forms of coordination, more often than in the case of centralised infrastructures. This creates high demands on the needed governance arrangements.

To address these challenges, it is central to understand what data space governance entails and what the available design options are. This White Paper is one of several documents produced by the Gaia-X Hub Germany to address these issues and shed light on the emerging field of data space governance.

To this end, the next section by clarifying the meaning of the term “data space governance”. Sections 3 to 7 then consider the key issues in data space governance, including the involved actors, structures, phases, requirements, and tasks. Each topic is explained with reference to two specific projects in the context of Gaia-X: the Mobility Data Space (MDS) and Catena-X. The aim is to illustrate aspects of data space governance along with concrete implementation options.

2. Definition: Data space governance

To clarify the scope of data space governance, it makes sense to look at the two components of the term – “governance” and “data space” – before clarifying the concept of “data space governance.”

2.1. Data space

The current literature offers many different definitions and descriptions for the concept of data space. Comparing the different approaches, however, reveals many commonalities. Such commonalities suggest that a data space can be defined as a “federated, open infrastructure for sovereign data exchange based on common agreements, rules and standards” (Reiberg et al., 2022).

As this definition suggests, implementing data spaces heavily relies on governance. Firstly, because governance instruments such as “common agreements, rules and standards” are an essential part of data spaces. Secondly, inherent characteristics of data spaces such as openness and sovereignty can be understood as aspects of “good governance.”

2.2. Governance

Governance itself is a difficult concept to pin down because of the variety of definitional approaches. Until around the end of the 1980s, the term was seldom used, even in the English-speaking world, and mostly to describe state action, i.e. the actions of governments (Mayntz, 1998). From the 1990s, the term gained popularity in German-speaking countries and was extended to include forms of collective decision-making beyond the state (ibid.). In simple terms, governance means the steering or coordination of social actors [1](Benz, 2004, p. 25).

The increased use of the term depends on two main factors:

First, it highlights that societal coordination and rulemaking often involves, or even requires, action by both state and private actors – particularly to maintain a functioning market[2]. In regulatory contexts, for example, auditors, standardisers, and other private actors play an important role alongside government. In addition, private companies must interact more and more with state and other private organisations to function and to achieve other social goals. For example, when companies make commitments to non-governmental organisations to meet certain public welfare objectives.

Second, the term governance indicates that these kinds of societal coordination and rulemaking are rarely restricted to hierarchical structures. Instead, they usually include competitive or cooperative forms of interaction. This puts "governance" in contrast to terms such as "government" or "management".

In summary, governance research has significantly contributed to identifying and analysing the diverse actors and the forms of coordination and control that exist in many areas of society. This makes the concept particularly suited to analysing interaction and coordination in the data economy in general, and in the context of data spaces in particular.

2.3. Data space governance

Combining the two approaches above yields a working definition of data space governance. It means coordinating the actors involved in or (potentially) affected by what happens in a data space (i.e., a federated infrastructure for data exchange).

As a combination of terms, data space governance offers an alternative to similar terms, such as "data space management" or "control of data spaces".

The term "data space management", in fact, suggests coordination within organisational boundaries and in relation to economic processes. Coordination processes in data spaces, instead, usually involve different organisations. These may be public, private organisations, or

[1] As Benz (2004: 25) points out, this coordination generally helps managing interdependencies and uses specific regulatory systems.

[2] For example, work on the new institutional economics (see for example: Williamson, 1999) has shown that successful interaction between private actors often requires government regulation. Work on global governance (see for example Beisheim et al., 2011) has shown the need for private actors participation for successful state action.

more often, both. Furthermore, the coordination within data spaces relates not only to economic values, but also to other social ones.

Both “management” and “control” also suggest a hierarchical form of coordination. Using these terms would then overlook the many actors involved in data space design. These may not necessarily relate hierarchically to each other and often coordinate through cooperation or competition rather than through a top-down approach.

The concept of data space governance therefore allows to highlight the complexity of data spaces and to develop the necessary understanding for their design. This way, the complex reality must be unfolded stepwise, for example by differentiating and illuminating certain aspects of the governance of data spaces (in this case, actors, structures, phases etc.), as in the following sections.

3. Players in data space governance

The federated structure of data spaces offers many advantages, especially compared to centralised activities in the data economy. For example, the decentralised structure makes it easier to integrate existing companies, because the roles in the value creation process are open to other players, instead of being occupied by a single player. This in turn creates opportunities for competition and cooperation, ultimately promoting innovation and value creation.

However, a decentralised structure inevitably requires many different stakeholders to be involved, possibly in different forms and to a different degree. Therefore, diversity and complexity characterize data space governance.

To make this concept more tangible, the following sections distinguish between two groups of actors in a data space[3]: the participants or subjects in the data exchange, and the operators or federators.

3.1. Participants

The participants of a data space are the organisations and individuals who offer or request data or services in the data space. In the case of Gaia-X, participants are consumers and providers of the offerings of a Gaia-X data space. Offerings include data, data-related services, or infrastructure services. The following table provides some examples of offerings in the context of Gaia-X.

[3] This categorisation is particularly useful at the data space level. A distinction can also be made between the federation level and the orchestrator level (see section four), for which actor groups may be distinguished differently. For the sake of clarity, this section focuses only on the data space level.

Service Offering	Example
Cloud service	Infrastructure as a Service, Platform as a Service, Software as a Service
Data set	Data sharing in batch, stream and event driven
Software Licence	Perpetual or renewable licences for a product without an associated online service
Interconnection & Networking Service	Services beyond a regular Internet connection or with special features, such as bandwidth, latency, availability, or security-related settings

Table 1: Service offerings in Gaia-X (own presentation based on (Gaia-X AISBL, 2022a, p. 54))

Many organisations can take the role of participants. They can be for-profit or non-profit, small or large, governmental and non-governmental organisations. In principle, individuals could also participate in a data space.

Most current data space initiatives focus primarily on private companies that offer or request data or data-related services for commercial or business purposes. Other players are less addressed. The two examples of Catena-X and Mobility Data Space illustrate this: Catena-X mainly targets companies from the automotive manufacturing value chain, starting with the extraction of raw materials and the manufacturing of components, through vehicle production and sales, until dismantling and recycling. Companies from other sectors (e.g. software development), research institutions, and public authorities are also involved, albeit to a lesser extent.

The Mobility Data Space (MDS) focuses on companies directly or indirectly active in the transport sector, including road, rail, and waterway transport companies. This includes companies from the production sector (such as car manufacturers), the services sector (such as logistics companies) and the digital economy (such as software development). In addition, representatives from other sectors (such as insurance companies) as well as research and public administration are also involved in the MDS.

To participate in a data space, an organisation must meet certain conditions, in the form of technical and legal requirements. These are usually composed of federation-wide and data space-specific requirements.

For example, Gaia-X participants must provide descriptions of themselves and their services. Each offering in a Gaia-X data space thus carries a self-description, specifying at least the provider from which it originates and where to find the terms of use. Participants in MDS and Catena-X are required to provide more specific self-descriptions.

Such requirements are usually defined in the participation contract of each data space. These contracts regulate the rights and obligations of the participants, including prices for basic services.

The participation contracts generally mandate, among other things, the use of compatible software for data exchange. In the case of Gaia-X, for example, this could be the EDC connector. Both in Catena-X and in MDS, it is possible to install the connector independently or utilise external providers in the form of a *connector as a service*.

A second group of actors is also important to the requirements for participants: Federators of data spaces.

3.2. Federators

The federators provide the essential services for the operation of a data space[4]. In Gaia-X, these include federation services for identity management, secure data transfer, and labelling. These services enable participants in a data space to offer and request data and data-related services.

The data space design determines how many federators are needed and the extent of their responsibilities. For example, it is in principle possible for many federators to jointly provide almost all basic services. In this case, the operation of the data space would be largely decentralised. In current projects, however, a single federator provides a significant portion of the basic services. Therefore, the data space is operated centrally, at least to a certain degree. There can be various reasons for some amount centralisation. For example, differentiating between critical and non-critical services, with higher security requirements put on operators of the former and lower security requirements for operators of the latter.

Critical services of the Catena-X data space, such as identity management, are offered exclusively by Cofinity-X, the operator currently mandated for this purpose. Providing other services, such as listing offers in the Catena-X data space, is open to other operators.

Many projects establish an organisation specifically to mandate and supervise key operators as well as to coordinate the players in the data space more generally. This organisation thus acts as an "orchestrator", as it ensures the harmonious interaction of the players[5]. In the case of the MDS, the company DRM Datenraum Mobilität GmbH acts as such an orchestrator; in the case of Catena-X, the role is played by the association Catena-X Automotive Network e.V..

The governance of the data space also depends on how the orchestrator is constituted and how it relates with the federators. In the case of the MDS and Catena-X, articles of association or statutes govern the orchestrator's organisation, while cooperation and service agreements determine its relationship with federators. These stipulations influence the scope of interests and range of action for the federators.

For example, the orchestrator is often subject to neutrality obligations. This can mean, for example, that it is not meant to turn profits. In the case of Catena-X, the orchestrator is a non-

[4] Section 7 provides a more detailed explanation of the tasks of the federators.

[5] Orchestration is described here based on (Abbott et al., 2015). It is understood as a form of governance through which an individual actor brings together (or "enlists") other relevant actors, on a voluntary basis, to pursue a common goal. In comparison, coordination less direct and usually without harsh sanctions. Orchestration ensures that actors deploy their specific skills and work together harmoniously – akin to an orchestra performing together.

commercial association, for which profit is not a main purpose. In the case of MDS, the orchestrator is a limited liability company, and profit-making was excluded as a purpose in the articles of association.

It should be noted that data spaces are based on the effective cooperation between their participants. In this context, federators serve to provide the basis for this interaction. Participants then use these foundations to offer or use data or data-related services.

Cooperation between stakeholders necessarily relies on the stakeholders themselves having appropriate incentives and opportunities to act. Defining appropriate governance structures amounts to providing stakeholders with such opportunities and incentives.

4. Data space governance structures

Several aspects of a data space count as governance structures, depending on how broad a scope one considers. They generally correspond to rule systems, which permanently determine what interests and options for action are available to the data space actors. The following section uses practical examples to present some of the most important legal and technical control systems of data spaces. These will be characterized at three levels: that of the data space initiative, that of the data space and that of the orchestrators.

4.1. Level of data space initiatives (here Gaia-X)

A basic principle of data spaces is that they are organised as federations, i.e. they have a minimum level of interoperability. Several initiatives currently develop foundations for interoperability, in the form of agreements, standards, and software stacks, driving forward the development of data space federations. Such initiatives include particularly the International Data Spaces Association (IDSA) and Gaia-X. The IDSA serves the further development and application of the IDS reference architecture model, the development of which began in 2015 as a research project funded by the German Federal Ministry of Education and Research (BMBF). 16 Fraunhofer Institutes worked at the architecture's implementation, which is currently spearheaded by the IDSA. Gaia-X strives to further develop and apply the Gaia-X reference framework, with the associated governance arrangement. Gaia-X was launched in 2019 as an initiative of the German and French governments with numerous industrial partners. The Gaia-X Association currently drives the initiative.

IDS and Gaia-X complement each other, with IDS focusing on standards for data transfer and Gaia-X setting standards in trust protection (Boris Otto, 2023, p. 19). Both initiatives have also created comprehensive governance structures that determine how to bundle and balance interests as they further develop the reference frameworks. The following is an example of the Gaia-X structures.

An example of an important Gaia-X structure is the Gaia-X Association. This was created as an international non-profit association under Belgian law (*association internationale sans but lucratif* – AISBL). Its tasks include, on the one hand, representing the Gaia-X initiative externally – especially vis-à-vis other data space initiatives – and, on the other hand,

coordinating the internal workings of the Gaia-X initiative. This includes, in particular, organizing the committees that drive the development of the Gaia-X framework. Three committees are active (Policy Rules Committee, Data Spaces Business Committee, Technical Committee), each encompassing several working groups. The results of the committees' work form part of the Gaia-X framework and can be found in key documents, such as the Architecture Document (Gaia-X AISBL, 2022a).

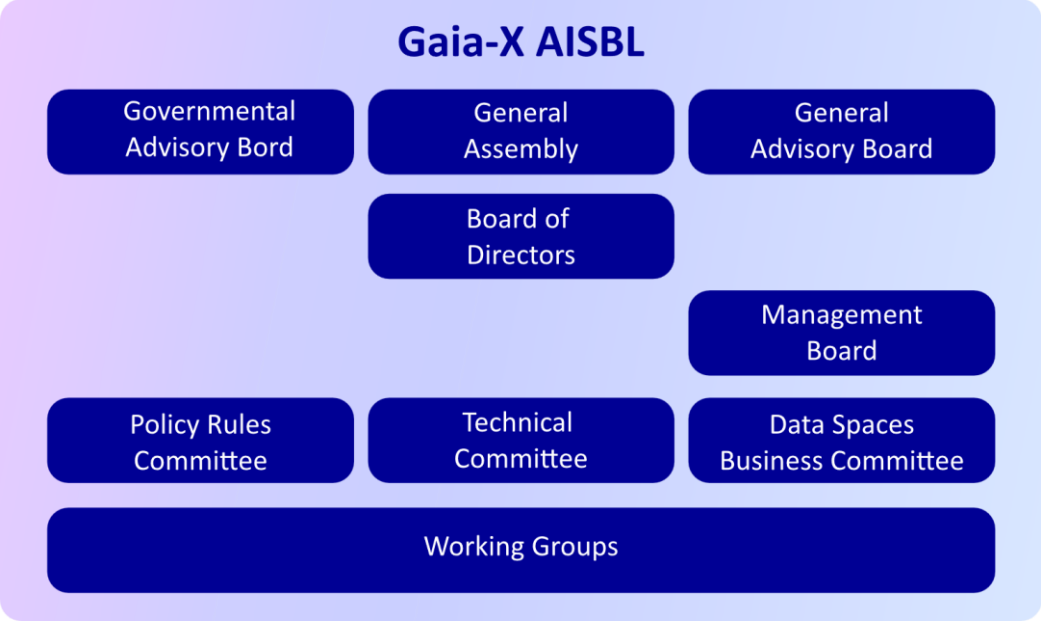


Figure 1 Structure of the Gaia-X AISBL. Source: Own representation based on (Gaia-X AISBL, 2024).

The national Gaia-X Hubs are another important Gaia-X structure, tasked with coordination and knowledge transfer at the national level. The Hubs introduce new members to the Gaia-X community, network ongoing projects on the application of the Gaia-X framework, and take care of both documenting and disseminating both basic findings and practical experiences for all Gaia-X community stakeholders. These activities take place in so-called domains, which reflect specific areas of society or industry, such as the public sector, healthcare, or mobility.

Alongside committees, working groups, and national Hubs, the technical frameworks themselves also serve guiding and coordinating functions. One example of this is the Gaia-X trust framework. This defines the basis for creating labels that confirm that individual participants or offers within a Gaia-X data space fulfil certain requirements.

Such requirements may include basic ones for participation in Gaia-X, as well as further requirements for individual data spaces. An example of general labels are the three basic Gaia-X labels (Level 1 to Level 3). Participants receive an appropriate label when they fulfil certain general criteria in terms of transparency, data protection, data security, portability and location. An example of data space-specific labels are the Catena-X labels, granted to offers and providers that comply with specific requirements of the Catena-X data space. For example, a provider that offers services for the traceability of individual components from production to recycling for supply chains in the automotive industry and fulfils the corresponding requirements for data sovereignty and interoperability from the Catena-X domain “Traceability”.

On the technical side, labels are realised together with the participants' self-descriptions as *verifiable credentials*. As these are machine-readable, it is also possible to take the labels into account when concluding and implementing smart contracts. For example, one can request that only providers or offers with a specific certification or based in a specific national territory can be considered for certain sectors (such as the exchange of personal data) or for particular kinds of transaction (for example those involving trade secrets).

The Gaia-X Trust Framework (Gaia-X AISBL, 2022b) determines who can define labels and who can assign labels, after reviewing the relevant criteria (Gaia-X AISBL, 2022b). For some labels (especially the basic labels for Gaia-X conformity) only so-called Gaia-X Digital Clearing Houses (GXDCH) define the label criteria and assign them. Participants of a specific Gaia-X data space can create further labels.

As these examples make clear, there are already many opportunities to participate in the governance structures at the data space initiative level. To make use of these opportunities, one should first complete the onboarding process at the relevant Gaia-X Hub. Being members of the AISBL and actively participating in its working can also be useful. There are also further opportunities for participation at the data space level.

4.2. Data space level

As the previous section explained, many basic rules for data spaces are defined at the federation level, i.e. by the respective data space initiative (such as Gaia-X). Nevertheless, individual data spaces have considerable leeway to specify further rules. Data spaces intended for specific areas of society or branches of industry, for example, may opt to make further specific agreements. This way the relevant infrastructure of the data ecosystem, i.e. the data space itself, can be designed to meet participant requirements or external regulations. Analysing the similarities and differences between MDS and Catena-X helps to elucidate the structures that enable the design, operation, and use of the data space.

In both projects, the structures roughly fall into three functional areas: A single organisation acts as an orchestrator and takes on most general organisation tasks. Technical operations and technical development, instead, are largely outsourced.

Both initiatives rely heavily on open-source software for technical development. The aim is simultaneously to benefit from the open-source community and to contribute to it.

The Tractus-X project, for example, was launched to further the software development within Catena-X (Eclipse Foundation, 2024). The project is driven within the Eclipse Foundation by several companies, which are members of both the Eclipse Foundation and the Catena-X Association. The aim of the project is to create reference implementations for the Catena-X framework.

Development activities will be opened even further in the future. Every member of Catena-X will be able to propose standards, which will then be reviewed by the Catena-X Association. After review, the standard will be used in operation.

The MDS, on the other hand, works with specific contractors for software development. These contractors are found through public tenders and the project values its independence from individual contractors.

In the area of technical operation, both projects rely on decentralisation, at different extents for different services. Individual services are subject to minimal requirements, such as to use a reference implementation. Such services can ultimately be offered by almost any of the interested parties and are therefore practically completely federated.

However, other services have stricter requirements, including in relation to technical standards and to the characteristics of the organisation in question. This most often affects sensitive services. Sometimes only a single provider offers these services, which therefore are largely centralised.

In both MDS and Catena-X, the orchestrator organisation plays a central role in the further development of governance, as it coordinates technical development and technical operation. In the case of the MDS, appropriate tenders (including service specifications) and subsequent contracts (including service contracts) are the main way to expend services. In the case of Catena-X, this is done through a certification process.

Behind this difference stands a different composition of the two orchestrators. The MDS is an organisation founded by the German Academy of Science and Engineering (acatech) as a neutral body. In the Catena-X Association, on the other hand, member companies have a stronger influence. The following section explains in more detail the organisation of the operating companies.

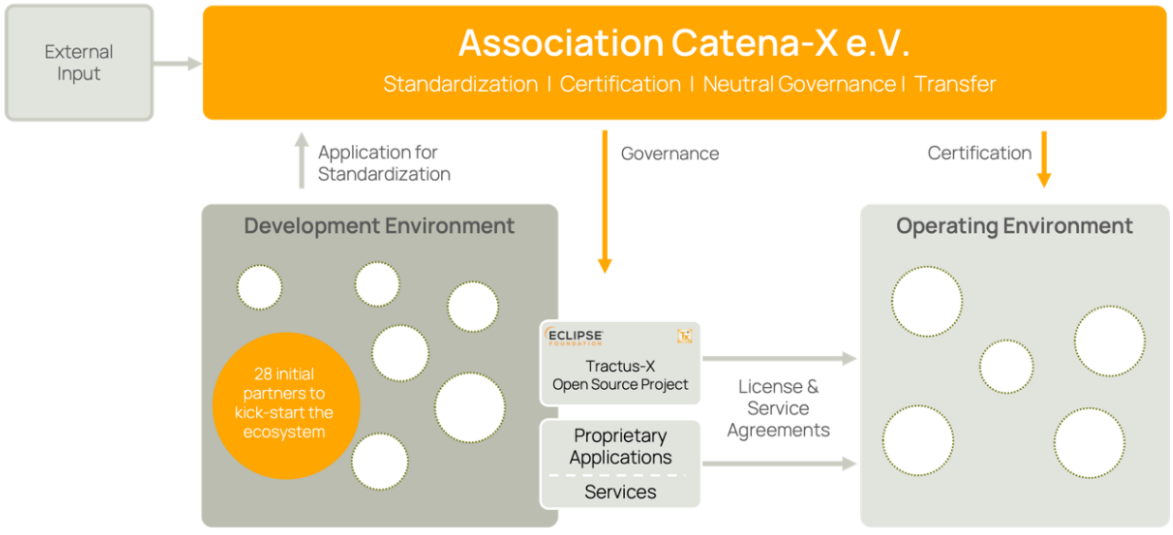


Fig. 2: Governance structures of the Catena-X data space. Source: (Catena-X Automotive Network e.V., 2022, p. 4)

All things considered, interested parties have several opportunities to participate in the rule-making process at the data space level. Companies can participate in the technical development of the data space – for example as part of the open-source community or as federators, in the technical operation of a service for the data space. Onboarding services for the data space, usually offered by orchestrators (in both MDS and Catena-X) or by

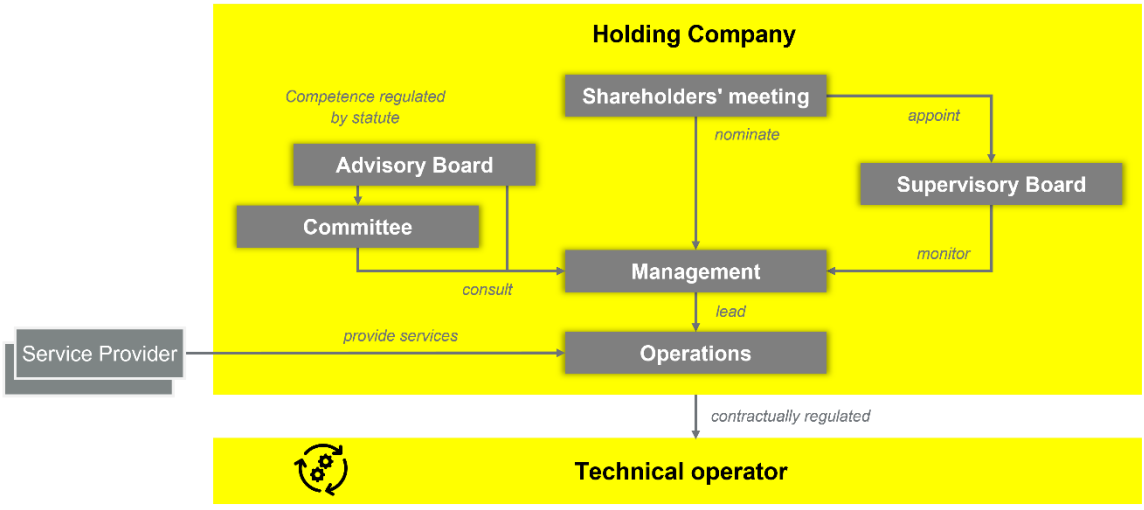
independent third parties (in the case of Catena-X), provide good first point of contact for interested parties and facilitate their entrance in the data space.

4.3. Orchestrator level

We have analysed the structures at the level of data spaces federation and at the level of individual data spaces. The next level to consider is that of internal structures of the data space orchestrators. Orchestrators play a prominent role in the governance of a data space. Therefore, they are found in many projects, including the MDS and Catena-X. These orchestrators are set up as independent organisations, with little focus on developing and representing particular interests. They rather aim to bundle and harmonise the interests of all relevant actors in a data space, to ensure effective and efficient operations, and to grow the data space. While the design of the orchestrators may vary, it necessarily reflects this goal.

For example, an orchestrator can be implemented as a corporation or as a cooperative. The limited liability company (GmbH in German) is a particularly frequent choice, for example with the MDS. In other cases, such as Catena-X, the orchestrator can be a registered association (German: eingetragener Verein, e.V.).

At a closer look, the examples of MDS and Catena-X reveal many similarities, despite implementing their orchestrators as different legal entities: Both cases present a management body (board of directors/management), an assembly (of shareholders or members), one or more supervisory or advisory bodies, and several committees.



Source: Mobility Data Space, 2025

Fig. 3: Structure of the Mobility Data Space operating company. Source: DRM Datenraum Mobilität GmbH.

In the case of Catena-X, the association board and the management, appointed by the board, have the central decision-making competences. In the case of the MDS, these tasks are assigned to the management of the GmbH. Both these management bodies are responsible, in particular, for representing the organisation externally and for day-to-day business. Thus,

they make central decisions on the data space design, particularly on who can participate and how to facilitate their interactions (see Section 7).

In the case of the MDS, the main shareholder (> 50% share, i.e. acatech) holds the management positions. Because acatech is not itself involved in the value creation processes in this sector, it acts as a neutral body. In the case of Catena-X, the board of the association is composed of representatives of the member companies, according to a specific requirements table. The requirement table is intended to ensure that there is an appropriate representation of all sectors (OEMs, suppliers and equipment manufacturers) and company types (SMEs and multinational companies). Working groups and specialised committees are appointed by the Executive Board to fulfil some of its key tasks.

In both cases, assemblies represent the broadest body and the basic authority that legitimise the decisions for the organisation. For MDS, this is the shareholder assembly, for Catena-X the members assembly. These groups meet at regular intervals to make key decisions and to ratify the decisions of the management bodies if necessary.

At Catena-X, the general assembly appoints the Executive Board, which in turn appoints the management. In the case of MDS, the shareholders assembly appoints the management directly.

In the case of Catena-X, all member organisations participate in the general assembly, but their voting rights depending on the type of organisation. Only companies in the automotive industry can be full members with voting rights. Other organisations or individuals can become associate members without voting rights.

In the case of the MDS the assembly includes all shareholders. In addition to acatech, these are companies from the mobility sector (e.g. Deutsche Bahn AG from the rail transport sector, Deutsche Post DHL Group from the logistics sector and Volkswagen AG from the automotive industry), from other sectors (e.g. HUK Coburg from the insurance sector) and from the German Federal States of Bavaria, Baden-Württemberg and North Rhine-Westphalia. The number of votes in the meeting corresponds to the shares held by each company.

An advisory board was created in both cases to support the management bodies. Half of the members of the Catena-X Advisory Board are appointed by the General Assembly and half by the Executive Board. The members of the MDS Advisory Board, on the other hand, are appointed by the Shareholders Assembly. In both cases, they are experts in the respective area of society, including members of organisations from science, politics and administration. Their tasks include in particular: providing expert advice to the management bodies and shaping the exchange with external organisations.

In the case of the MDS, a Supervisory Board also exists to advise but also monitor the management bodies. As in the case of the Advisory Board, also the members of the Supervisory Board are appointed by the Shareholders Assembly. In addition to supporting the work of the management, the Supervisory Board also scrutinises it in the interests of the shareholders.

The described similarities and differences in internal structures between MDS and Catena-X show that similar goals can be achieved in different ways. For example, both organisations aim to ensure a fair representation of the interests of all data space participants and to prevent specific interests from prevailing. To this end, the MDS strongly relies on the participation of acatech as a neutral body, whereas Catena-X relies on the diversity of member companies. At the same time, both organisations strive for efficient and factually sound decision-making by appointing a supporting Advisory Board – and possibly a Supervisory Board – alongside the permanent management.

5. Legal requirements of data space governance (here Data Governance Act)

The data space governance structures explained in the previous section provide all data space actors with important rules. Other legal requirements also need to be considered and will be discussed using a single example. The relevant regulations come from in a variety of legal areas. These include company law, antitrust law, intellectual property law and data protection law. Recently, regulations specifically targeting data spaces or aspects of dataspaces have also gained relevance. The Data Governance Act (Regulation 2022/868) is a particularly useful example, as it addresses the intermediary figure that plays a central role in all current data space projects.

In particular, the DGA contains provisions about so-called data intermediation service providers, a category to which many data space federators will likely belong (Reiberg et al., 2023, pp. 14–15). A data intermediation service is "[...] a service which aims to establish commercial relationships for the purposes of data sharing between an undetermined number of data subjects and data holders on the one hand and data users on the other, through technical, legal or other means [...]" (Art. 2 N. 11 DGA).

The aim of the DGA is to promote such services by strengthening trust in them (Recital 2 DGA). To this end, the DGA sets up a series of provisions, two types of which are particularly relevant in the context of this work: Firstly, those defining which services many data space operators may offer. Secondly, those specifying how these services can be provided. Both will be briefly discussed here.

The DGA stipulates a functional separation between entities that facilitate the exchange of data and entities that use data. For example, providers of data intermediation services are largely prohibited from using those data for their own commercial purposes (Art. 12 lit. a DGA). The only exceptions are procedures used on behalf of the data provider and intended to facilitate the exchange of data, for example the pseudonymisation of data (Art. 12 lit. e DGA).

Furthermore, the DGA formulates requirements on how data space services can be provided. For example, providers must offer services fairly, transparently, and non-discriminatorily, as well as meeting certain security requirements (Art. 12 lit. f to o DGA). In addition, interoperability must be guaranteed in several respects.

The operating companies of MDS and Catena-X must also fulfil these requirements. Both projects endeavour to give participants a high degree of control over their data. Participants can, for example, determine for whom the data are available, as well as set parameters such as the duration and purpose of data use. Federators have no access to the shared data, unless expressly requested by the data provider (and, if applicable, by the data recipient). The MDS offers *trusted data services*, through which participants ensure a certain data quality or can convert data between specific formats. As a result, the MDS federators already fulfil the functional separation between intermediation and use of data required by the recently enacted DGA (as it is currently assessed).

This also applies to other requirements of the DGA. For example, both Catena-X and MDS designed their data space to keep access as open as possible. This does not only apply to the group of participants, but also to the more narrowly defined group of federators. In principle, Catena-X allows any interested organisation to offer secondary services for the data space. Only critical services, such as those necessary for security, are offered by a selected group of providers.

Security and interoperability requirements are also fulfilled by design. As the data spaces comply with the Gaia-X framework, they employ Gaia-X labels, which ensure a high level of security. Moreover, the common basis of all Gaia-X data spaces ensures a high degree of interoperability.

In conclusion, the operation of data spaces must observe several legal requirements. However, utilising the existing frameworks of the data space initiatives considerably streamlines the task.

6. Phases of data space governance

The actors, structures and legal requirements described above demonstrate the complexity of data space governance. Because of this complexity, developing a data space becomes a challenging and usually long-term endeavour, often spanning several years. To better analyse the governance along this articulate process, it makes sense to differentiate between development phases.

In the following section, we distinguish three phases: Initial phase, Building phase, and Operational phase. This framework will serve to highlight the different actors involved, the dominant forms of interaction, and the most important funding sources for each phase. The number of actors ranges from a medium to a high number, interactions will be on a spectrum from cooperative to competitive to hierarchical, and different proportions of public and private sources can provide funding.

6.1. Initial phase

The initial phase includes all planning work for the data space that is not aimed at a concrete organisational or technical solution.

Generally, the first step is evaluating the potential for the envisaged data space. This involves preliminarily narrowing down the possible participants in the data space and identifying the possible interactions between them.

For example, both the MDS and Catena-X needed to clarify which players have an interest in a data space in their respective sectors (mobility and automotive manufacturing) and which use cases could be developed for the exchange of data and data-related services.

To better define this first pool of participants, it is best to involve a relatively broad spectrum of stakeholders in the initial phase. This allows to collect many possible perspectives and requirements, before concretising them. Low-threshold forms of cooperation lend themselves to this goal, as they increase the number of perspectives that actually come into consideration.

In the case of MDS and Catena-X, existing exchange formats such as the so-called “Concerted Action on Mobility” and the “Industry 4.0 Platform” proved to be important forums. In these, representatives from industry, research and administration exchanged views on current challenges, such as the transformations in the automotive industry, and discuss possible solutions, including the realisation of open and transparent data spaces.

This preliminary work can be funded through private or public sources, as well as public-private partnerships. In general, data spaces fit the description of public goods. In fact, one can hardly assume the use of data space to be excludable, particularly because of requirements imposed by law or by data space initiatives. The DGA, for example, has specific non-discrimination clauses, and Gaia-X requires data spaces to be open. Therefore, a significant proportion of state funding will likely be necessary. State involvement also improves visibility and representation for the nascent data space. Particularly in the initial phase of development, seed funding is a pressing need, while refinancing is a relatively distant prospect. In addition, this phase requires mediating between a relatively large number of stakeholders, in part because several actors – at the local, national, European, and international level – currently seem willing to fund data space projects. For example, German Federal States participate in individual Gaia-X projects (such as the Mobility Data Space) and the German federal government made a strong commitment within the funding competition “Innovative and practical applications and data spaces in the Gaia-X digital ecosystem” as well as in EU programmes, such as the “Digital Europe Programme”.

Early government support helped both MDS and Catena-X, for example by funding their seminal exchange formats, like the “Concerted Action on Mobility” and the “Industry 4.0 Platform”.

6.2. Building phase

The building phase is the timespan from the creation of concrete organisational and technical solutions until to the start of regular operations. The aim of this phase is to transfer the previously defined data space participants and their interaction forms into concrete technical systems and organisational structures. This can take place, for example, through a reference

implementation and the establishment of an operating company. This phase also includes the testing of the systems and structures that are necessary before the start of regular operations.

The building phase generally involves a smaller network of stakeholders compared to the initial phase – at least in terms of active participants. The aim in fact shifts from dialogue between many different perspectives to effective collaboration between a manageable number of participants. This phase still relies on cooperative interactions, e.g. between stakeholders who develop complementary work packages. Alongside them, hierarchical interactions come into play, e.g. between the stakeholders undertaking the overall management and the ones taking on sub-tasks. In many projects, a consortium of organisations builds the data space. These consortia may encompass both private companies and public research organisations.

The building phase often requires both private and state funding. Private investors gain an early access to developing business models and can build up expertise for the realisation of a data space. Thus, they gain future advantages and business activities. Public funders can incentivise rapid implementation and have a formative influence on the nascent ecosystem.

The German government, for example, supported the MDS as part of the so-called MFund programme for the development of data-based business ideas. Three research projects addressed fundamental questions about the structure, operation, and use of the Mobility Data Space. Their results led to the design of technical and organisational solutions, before the operating company was founded by acatech. This company then operationalised the MDS through to the start of operations.

Catena-X received public funding as part of the “KoPa35”. This was an economic stimulus package of the German government aimed at supporting a transformation process in the automotive manufacturing and its supply chain. The Catena-X consortium was funded as part of the module on production modernisation. The consortium ultimately led to founding the Catena-X Automotive Network association, and developed the technical and organisational foundations until the start of operations.

6.3. Operational phase

The operational phase denotes the period in which development advanced sufficiently to allow the data space to operate and involve a significant number of participants. In this phase, the primary goal is to ensure operations without major interruptions or incidents.

Growing the number of players becomes a primary concern, as it increases the benefits of the data space as a whole (including for existing players). This usually requires careful planning and a growth strategy. In most cases, the strategy will first target few potential participants and applications, before gradually expanding the functionality of the data space. In the case of both Catena-X and the MDS, for example, the initial focus was on companies headquartered in Germany. Intensive work is currently underway to attract companies headquartered in EU member states and other countries. Catena-X initially recruited large companies, before strengthening its focus on small and medium-sized enterprises. MDS shifted its targeting

towards companies in the rail and shipping sectors, after having covered much of the private transport area.

Cooperative and hierarchical forms of interaction likely emerged in previous phases, particularly between and within data space operators. Alongside those, competitive forms of interaction increasingly become part of the data space ecosystem during the operating phase. In this phase, in fact, a functioning market for data and services should develop, both among the participants and among the federators, at least to some extent.

An exchange of data and data-based services should arise accordingly. The revenues it generates should increasingly finance the data space operations. The need for government funding therefore wanes during the operating phase, replaced by private investment, drawn by the effective value creation.

At Catena-X and the MDS, the pricing of data space services is gradually shifting from purely membership fees (Catena-X) and shareholder contributions (MDS) to use-based pricing.

7. General tasks of data space governance

As argued in Section 1, data spaces can be understood data exchange infrastructures. Specifically, because they enable or facilitate the exchange of data between data providers and data recipients.

From an economic perspective, data spaces can also be viewed as multi-sided platforms (Evans 2003): Providers of data and data-related services form at least one side. The consumers of data and data-related services represent at least one other side. The data space orchestrators and federators act as intermediaries, connecting the two sides. All parties involved benefit from positive network effects. For example, the benefits of data space for consumers generally grow with the number of participating providers. Vice versa, the more consumers there are, the greater benefits producers reap. Network effects also occur within each side. For example, providers of data-related services benefit from the presence of other providers in the data space, as this makes it easier for them to develop or provide products and services in cooperation and generate further profits.

Data spaces thus fulfil two central criteria for multi-sided platforms (Evans, 2003, p. 192). Therefore, the literature on multi-sided platforms helps to understand the goals and tasks of data spaces and data space governance.[6]

Like Gawer (Gawer, 2021, pp. 2–7) argues, multi-sided platform companies face fundamental decisions in at least two more areas than conventional companies do. First, they need to commit to the (at least two) sides that the platform serves. Secondly, they must decide on the interface, i.e. the interaction form between these sides. Both aspects are addressed below.

[6] As argued here, the concept of the multi-sided platform can be used to understand the tasks of a data space. On the one hand, data spaces have similar tasks to integrated companies such as Facebook and Google, which are often referred to as “platforms” for short. On the other hand, it becomes clear that these two cases fulfill their tasks in significantly different way and to different outcomes.

7.1. Organisation of the group of participants

In principle, the group of data space participants can be defined as desired. However, the definition always encounters de facto boundaries (more or less restrictive). Resources for participant recruitment and other tasks can be deployed in a targeted manner. This will drive the development of the data space in line with the relevant requirements.

It is important to include beneficial participants in the data space, for example those who offer particularly relevant data or services. Meanwhile, other participants may have to be excluded, for example if they could pose a security risk to other participants.

It is often necessary to gather a critical mass of certain participant groups for the data space. If this critical mass is not reached, the benefit for other participants will not offset the participation costs. The lack of network effects leads to long-term negative growth and possibly to the complete dissolution of the data space. Reaching the critical mass, on the other hand, consolidates growth in the long term, up to a high saturation degree. Therefore, even high initial investments in participant acquisition can generate positive long-term returns and be amortised once critical mass is reached.

Participant boundaries usually follow established economic sector categorisations. For example, there are projects to establish data spaces in the sectors of healthcare (such as the HEALTH-X dataLOFT and Team-X projects), education (such as the MERLOT project), as well as in the financial sector (see the EuroDaT project). The link to established categories also makes it easier to pick participants that have a history of mutual interactions. These are more likely to seek further interactions and therefore to need a suitable infrastructure. Furthermore, existing institutions that already represent or orchestrate the sector can sustain the establishment of data space.

The data space usually develops from a specific segment of the projected long-term member base. Once a critical mass is reached in this smaller group, more segments can be targeted and stepwise added. For some use cases, one could first recruit all companies within a specific part of a value chain, before moving to companies in other parts of the value chain. For other use cases, it may be better to start with few companies spanning the entire value chain, before recruiting further participants for each part.

The benefits of larger participant numbers can be assessed along many dimensions. As platforms demonstrate, one may want to lower entry barriers or make stronger integration efforts for a specific side of the multi-sided market, whose participation generates profits. These additional profits then subsidise the participation from another side. The demand for data-using membership, for example, often has low price elasticity. As a result, data-using participants will face relatively high membership costs, which will help keep the costs for data-providing participants comparatively low.

The ways to draw participant boundaries are diverse and too many to list here. For example, how prices are set (e.g. the amount due for membership), how technical systems are designed

(e.g. minimum requirements on client software), and how the organisation functions (e.g. the onboarding process). Limited resources demand a careful analysis of the measures to include or exclude participants: from costs and benefits (especially regarding the importance of the respective participants) to their effectiveness. Legal requirements and, where applicable, the requirements of the data space initiative (for example Gaia-X) also matter. For example, any boundary must respect the relevant requirements regarding non-discrimination of participants.

7.2. Designing the interaction

At its core, a data space aims to facilitate the interactions between participants. Consequently, the design of the offering that enables and brokers this interaction is a central task in the governance of data spaces. In particular, the governance must clarify which technical services support the exchange of data and services.

In many cases, the design phase includes defining not only the group of participants, but also the range of functions of the data space. However, due to the open design of data spaces, both the participant boundaries and the interaction forms considered will need to be continuously adjusted.

Specific use cases often underpin this process. For example, a use case from the MDS focuses on the exchange of information about hazardous situations. In this use case, data from vehicle sensors is collected and processed, according to each user's consent and specifications. This generates, e.g., information on extreme weather events, used to warn drivers according to the route they chose on their vehicle's navigation system. This results in an innovative product that contributes to road safety. One example from the Catena-X data space is the sharing of demand and capacity data across the automotive manufacturing value chain. The goal is to faster identify and counteract disruptions in the production and their extent. This optimises the production processes, avoiding or reducing costs for both producers and consumers. Employing such concrete use cases allows data spaces to pinpoint the nature of participant interactions and identify the support they require.

In general, opportunities to reduce transaction costs for participants play an important role in the planning. Such costs arise in many aspects of the interaction. For example, data exchange incurs costs when finding potential transaction partners, checking the partners' offers, negotiating the transaction, closing the negotiation in a legally compliant manner, materially transferring the data, transferring monetary or other considerations, monitoring the compliance with the agreements as well as with external requirements, and so on.

Accordingly, there is a variety of avenues to reduce transaction costs for participants. The governance of the data space needs to continuously clarify whose tasks are taken over by whom and in what way.

The framework of the data space initiative usually provides the essential foundations. The Gaia-X framework, for example, offers the reference implementations, which will be available

in the form of GXFS-DE and GXFS-FR federation services, as well as the basis for a range of other technical and organisational services.

As part of the GXFS-DE, the “Identity and Trust” work package created services based on the principle of *self-sovereign identities*, which guarantee participants the certainty that others are indeed who they claim to be. The “Sovereign Data Exchange” package developed services that automate contract conclusion and logging, allowing compliance control. This allows participants to exercise reliable and granular control over the data they provide. The work packages “Federated Catalogue”, “Portal and Integration”, and “Compliance” developed further services. These foundations developed alongside the framework of the data space initiative and are freely available as reference implementations. It is up to each data space to adapt and supplement them through in-house development. Catena-X and the MDS have largely completed this work. For example, both data spaces have started operating with several basic services. Both projects use freely available solutions that have been customised for their purposes – including the EDC connector for data transfer. They are also developing their own specific solutions. One example from Catena-X is the so-called “Golden Record”, a directory of company data. Catena-X uses this to ensure that data on business partners, locations and addresses from different sources are identified, linked and harmonised for all participants. One example from MDS is the data catalogue, in which participants search through all data offerings and view the description of each data offering (metadata). Another example is the planned MDS Participant Portal, an integrity and access management system that will be used to manage participants' access rights to the data space.

As a rule, the federators design and offer services following a comprehensive business model (including an analysis of the value proposition for the customers). Comprehensive legal requirements must also be observed here, such as the Data Governance Act mentioned above. Among other things, this regulation largely prohibits entities from using data, the exchange of which they enable or facilitate – as is the case for data space federators.

8. Conclusion

Current projects have already prepared complex governance arrangements to develop, set up and operate data spaces. The previous chapters addressed some of the key questions: Which actors can participate in data spaces? Which structures underly their interaction? Through what phases does a data space take shape? The tasks that need to be mastered when designing a data space were also addressed.

Overall, this work should make tangible the complexity of building data spaces. This complexity is both a curse and a blessing. On the one hand, constructing a data space can take many forms, so the data space can be expediently tailored to its area of application, for example a particular industry. On the other hand, one usually needs to develop specific solutions, which comes with costs.

Many specific projects can serve as role models. However, the examples mentioned in this White Paper demonstrate the diversity of avenues that can be taken to realise similar objectives, such as the neutrality of central players. Many conceivable solutions, each with its advantages and disadvantages may apply to each specific case. For example, certain participation structures may favour broader representation, but hinder effective decision-making. Therefore, realising a data space requires difficult decisions at every stage. In conclusion further intensive study of the design options is urgently required – both for scientific work on the concept of data spaces and for the practical implementation of this concept.

Bibliography

Abbott, K. W., Genschel, P., Snidal, D., & Zangl, B. (Eds.). (2015). *International Organizations as Orchestrators*. Cambridge University Press. https://doi.org/10.1017/CBO9781139979696

Beisheim, M., Börzel, T., Genschel, P., & Zangl, B. (2011). *Wozu Staat? Governance in Räumen begrenzter und konsolidierter Staatlichkeit: Vol. Baden-Baden*. Nomos.

Benz, A. (2004). Governance—Modebegriff oder nützliches sozialwissenschaftliches Konzept? In A. Benz (Ed.), *Governance—Regieren in komplexen Regelsystemen: Eine Einführung* (pp. 11–28). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90171-8_1

Boris Otto. (2023). *Synthese_19_09_23.pdf*. 1.

Catena-X Automotive Network e.V. (2022). *Catena-X Operating Model Whitepaper V2*. https://catena-x.net/fileadmin/_online_media_/CX_Operating_Modelv2.1_final.pdf

Eclipse Foundation. (2024). *Eclipse Tractus-X*. Eclipse Tractus-X. https://projects.eclipse.org/projects/automotive.tractusx/who

Evans, D. S. (2003). *Some Empirical Aspects of Multi-Sided Platform Industries* (SSRN Scholarly Paper 447981). https://doi.org/10.2139/ssrn.447981

Gaia-X AISBL. (2022a). *Gaia-X Architecture Document*. https://gaia-x.eu/wp-content/uploads/2022/06/Gaia-x-Architecture-Document-22.04-Release.pdf

Gaia-X AISBL (Ed.). (2022b). *Gaia-X Trust Framework 22.04.* https://gaia-x.eu/wp-content/uploads/2022/05/Gaia-X-Trust-Framework-22.04.pdf

Gaia-X AISBL. (2024). *Who We Are*. Gaia-X.Eu. https://gaia-x.eu/who-we-are/association/

Gawer, A. (2021). Digital platforms' boundaries: The interplay of firm scope, platform sides, and digital interfaces. *Long Range Planning*, *54*(5), 102045. https://doi.org/10.1016/j.lrp.2020.102045

Mayntz, R. (1998). New Challenges to Governance Theory. In *Über Governance: Institutionen und Prozesse politischer Regelung*. Campus-Verl.

Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European Data Governance and Amending Regulation (EU) 2018/1724 (Data Governance Act). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022R0868

Reiberg, A., Appelt, D., Kraemer, P., & Smoleń, A. (2023). *Data trusts, data intermediaries and Gaia-X* (White Paper 2/2023). Gaia-X Hub Germany. https://gaia-x-hub.de/en/publication-en/wp-data-trusts-gaia-x/

Reiberg, A., Niebel, C., & Kraemer, P. (2022). *What is a Data Space?* Gaia-X Hub Germany. https://gaia-x-hub.de/wp-content/uploads/2023/11/GX-White-Paper-Data-Space.pdf

Williamson, O. E. (1999). *The Mechanisms of Governance*. Oxford University Press, U.S.A.