10-9-2023

Design Principles for Institutionalized Data Ecosystems – Results from a Series of Case Studies

Patrick Weber
Ferdinand-Steinbeis-Institut, Stuttgart und Heilbronn, Germany, patrick.weber@ferdinand-steinbeis-institut.de

Maximilian Werling
Ferdinand-Steinbeis-Institut, Stuttgart und Heilbronn, Germany, maximilian.werling@ferdinand-steinbeis-institut.de

Henning Baars
Universität Stuttgart, Chair of Information Systems, henning.baars@bwi.uni-stuttgart.de

Design Principles for Institutionalized Data Ecosystems – Results from a Series of Case Studies

Research Paper

Patrick Weber[1], Maximilian Werling[1] and Henning Baars[2]

[1] Ferdinand Steinbeis Institute, Stuttgart and Heilbronn, Germany
{patrick.weber,maximilian.werling}@ferdinand-steinbeis-institut.de
[2] University of Stuttgart, Chair of Information Systems I, Stuttgart, Germany
henning.baars@bwi.uni-stuttgart.de

Abstract. Sharing and collaborating on data across organizational boundaries is increasingly important for building a comprehensive data foundation for a variety of relevant analytical models and reports. We argue that a formalized set of rules and responsibilities - data governance - is needed to guide such data sharing activities and thus provide the foundation for an institutionalized data ecosystem. To this end, we propose a set of design principles. Based on three case studies from different application domains, we derive the design principles using Service-Dominant Logic as our theoretical lens. We distinguish between dynamic and static design principles. Our approach supports the delineation and specification of data governance structures for data ecosystems.

Keywords: Data Ecosystem, Data Governance, Service-Dominant Logic, Data Collaboration

1 Introduction

Data sharing across organizational boundaries has become increasingly important as it is recognized as a powerful opportunity for value creation, especially in the context of the Internet of Things (IoT) (Lehrer et al. 2018). This has led to the design and establishment of trusted and secure data infrastructures, so-called data spaces (Pretzsch et al. 2020; Braud et al. 2021). However, building and deploying infrastructures is not enough. In an increasing number of cases, it has become apparent that it is advisable to embed data spaces in holistic Information Systems (IS) that also consider non-technical aspects (Reiberg et al. 2022). Data governance is a central aspect of such IS, providing the shared institutional logics that govern data sharing (Baars et al. 2021; Gelhaar & Otto 2020; Vargo & Lusch 2016).

The topic of data spaces has attracted the attention of management and IS researchers and is discussed under the term *data ecosystems*. However, the solutions discussed in research are rarely found in practice: There, many data sharing projects do not get beyond the prototype phase due to the lack of a framework for binding collaboration (Weber et al. 2022). In this paper, we examine three different data ecosystems that have moved beyond this stage. The focus is on the rules and regulatory structures - data governance - that have been established in these cases. Following the research gap outlined above, this paper focuses on the following two Research Questions (RQ):

- RQ1: *What are data governance design principles for the creation of an institutionalized data ecosystem?*
- RQ2: *How can design principles be grouped for the purpose of providing a structure for an institutionalized data ecosystem?*

The next chapter presents a literature review that provides the conceptual background and current state of research on the topic. The next section presents our methodology and an overview of our case studies. We then propose our data governance design principles. We conclude with a discussion of the limitations of this study and our contributions to theory and practice.

2 Related Work and Background

To gain a better understanding of the current state of research on our topic, we conducted a systematic literature search (Levy & Ellis 2006; Vom Brocke et al. 2009). The search process was documented according to PRISMA 2020 (Page et al. 2021). We searched titles and abstracts for the keywords "*data*" AND "*ecosystem*" AND "*governance*" in the following databases: AIS eLibrary, ScienceDirect, and IEEE eXplorer. Since we use the Service-Dominant Logic (SDL) as our theoretical lens, we conducted a second search to identify publications that also use it. The following keywords were used: "*service-dominant logic*" AND "*governance*", searched in the title and abstract, respectively. The results of both searches were checked for title and abstract relevance.

A total of 172 accessible publications were identified in both searches. Of these, 72 were excluded because of obvious irrelevance (especially in the second search). A further 54 titles were excluded after evaluation of the abstracts and 3 as duplicates. A total of 41 publications were examined in more detail, of which 6 were finally eliminated in this step. The following sections provide an overview of the relevant publications.

2.1 Data Ecosystems: An SDL Perspective

Recent developments around the IoT and the increasing availability of smart objects in value-added processes provide new opportunities for companies to offer innovative digital services (Lehrer et al. 2018). Smart objects continuously generate status data, which forms the basis for improved processes and digital services (Porter & Heppelmann 2014). However, for many companies, such efforts are accompanied by various technical and organizational challenges that they can hardly overcome on their own (Gelhaar & Otto 2020; Weber et al. 2020). This leads to the need to collaborate across company boundaries, building partnerships that are often rooted in a cooperative value proposition that aligns the economic activities and resources of the actors (Adner 2017; Shipilov & Gawer 2020). Digital ecosystems that focus on the aspect of data sharing are referred to as data ecosystems. Organizations are connected to each other cooperatively and/or competitively through digital platform infrastructures, often referred to as data spaces, for their value creation activities (de Reuver et al. 2022; Oliveira & Lòscio

2018). Data spaces are a form of digital platform infrastructure, providing a trust-based space for the exchange or transaction of data (Reiberg et al. 2022).

To better understand these phenomena, we use the Service-Dominant Logic (SDL) as our theoretical lens. The SDL originated in marketing (Vargo & Lusch 2008; Vargo & Lusch 2004), but is increasingly used in IS research, especially in conceptualizing data ecosystems (Azkan et al. 2020; Kannan et al. 2022; Blaschke et al. 2019; Hein et al. 2018; Lehrer et al. 2018). Instead of product offerings, SDL focuses on the exchange of service, where one actor uses its capabilities to benefit another actor (Lusch & Nambisan 2015). Lusch and Nambisan provide a framework for describing collaborative service innovation based on three key concepts: the service ecosystem, the service platform, and value co-creation (Lusch & Nambisan 2015).

Service ecosystem**:** The service ecosystem refers to a set of "*[…] of mostly loosely coupled social and economic (resource-integrating) actors connected by shared institutional logics and mutual value creation through service exchange*" (Lusch & Nambisan 2015, p 162). As a conceptual pillar of our paper, we interpret a data ecosystem as a variant of a service ecosystem.

Service platform: The value creation activities are facilitated by the service platform, which is defined as "*[a] modular structure that consists of tangible and intangible components (resources) and facilitates the interaction of actors and resources […]*" (Lusch & Nambisan 2015, p 162). In our case, the service platform is the digital platform infrastructure that facilitates the exchange of data and information between actors in the ecosystem (de Reuver et al. 2022).

Value co-creation: Value co-creation is defined as "[t]he processes and activities that underlie resource integration and incorporate different actor roles in the service ecosystem" (Lusch & Nambisan 2015, p 162). In our understanding, value is created by actors who collect and share data, generate new insights by applying their specific know-how to the data, and use these insights to contribute to the collaborative value proposition of the data ecosystem (Adner 2017; Weber et al. 2020).

2.2 An SDL Perspective on Data Governance

The SDL emphasizes the importance of shared institutional logics within service ecosystems. Shared institutional logics facilitate both the exchange of service and the creation of added value. They also support the development of network effects: The more actors adhere to shared logics, the greater the potential benefits (Vargo & Lusch 2016). We argue that inter-organizational data governance is not only enabled by such logics, but that they form its backbone.

Data governance generally refers to the distribution of decision rights and responsibilities for data-related decisions (Weill & Ross 2004). Historically, the scope of data governance was mostly limited to a single organization (intra-organizational data governance). More recently, the literature has shifted towards the study of data governance in a decentralized ecosystem setting (inter-organizational data governance) (Abraham et al. 2019; Lis & Otto 2021). A relevant data governance needs to facilitate the creation of a collaborative environment in which all actors act in accordance with, and further benefit from, the shared values of the ecosystem (van den Broek & Fleur van Veenstra 2015).

Data governance challenges in service ecosystems: Data governance for an ecosystem setting must ensure that members behave in a manner consistent with collective goals, prevent or resolve conflicts between actors, and ensure that collective resources - in particular, shared data - are used effectively and transparently. Ownership and valuation of shared data, considering the interests of the ecosystem and its individual members, is a challenge that governance needs to address (Lis & Otto 2021).

Data governance challenges regarding service platforms: A key challenge is the distribution of decision rights over which actors use the platform infrastructures and how. This translates into the task of finding an appropriate mix of centralized and decentralized approaches. Key platform-oriented data governance issues include data security and quality, standardization, regulation of competition among members, and questions of revenue distribution (Scholz et al. 2022; Perscheid et al. 2020).

Data governance challenges regarding value co-creation: In the context of value co-creation, there are four main challenges that inter-organizational data governance must address: First, it is imperative to define the roles of each partner in the value creation process. Second, since all activities in the data value chain process require an investment and thus generate costs, but revenues are only generated in a data exploitation phase that is often not fully predictable, an appropriate mechanism needs to be put in place to compensate actors for their (data) contributions (Badewitz et al. 2020; Martens et al. 2020). Finally, data providers often do not know which data with which properties (especially in terms of format, latency and quality, etc.) are needed by other actors, so tools for describing data supply and demand need to be established (Fernandez et al. 2020; Susha et al. 2017).

Figure 1 illustrates the conceptual framework of our work. The conceptual framework forms the basis for the development of our observation protocols as well as for the analysis of the results. It reflects our research questions and the related work section.

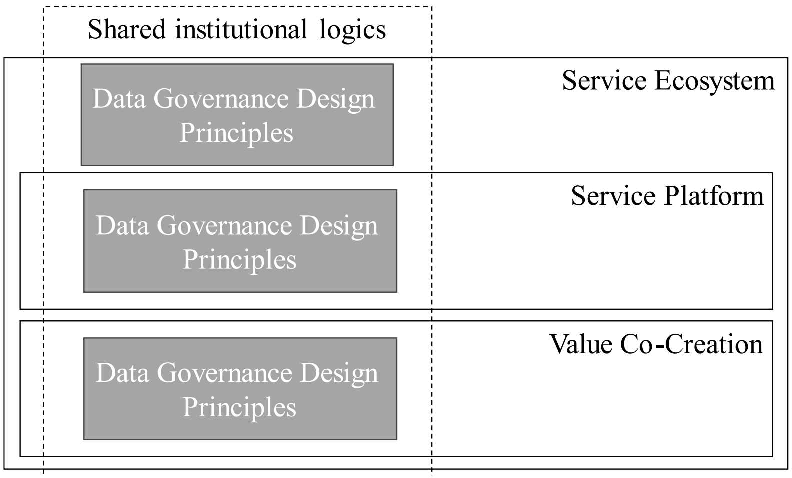

Figure 1: *Conceptual Framework, based on Lusch & Nambisan (2015)*

3 Research Design

We followed a multiple case study design (Yin 2018), in which we accompanied three consortia, each of which included three companies. Each consortium was followed through the establishment of a cooperative business unit as a separate legal entity, and the design and implementation of a supporting data governance. In all cases, the collaborations followed four steps: 1. building the partner network, 2. defining a cooperative goal, 3. establishing a shared unit as a legal entity, 4. prototyping the data sharing solutions for storage and analysis. The timeframes ranged from 9 to 24 months.

To elicit our design principles, we followed the approach of Action Design Research (ADR). In ADR, an organizational-technical solution is developed as an "ensemble artifact" in continuous cycles involving the participating companies and the research consortium (Sein et al. 2011). The ensemble artifact included running pilot systems, each of which was iteratively developed, applied, and evaluated in multiple cycles. Our personal involvement as researchers ensured that the state of the art in data ecosystems, data governance, and SDL was considered in both design and evaluation. It should be emphasized that the (non-funded) companies involved in the design and evaluation functioned as the final authority for the specific solution design, and that all ecosystems continued to collaborate, develop, and adopt the prototypes after the research projects were completed.

For data collection, we primarily conducted workshops with relevant company representatives and at least two researchers. The researchers participated as observers and documented all workshop outcomes in written and photographic protocols. These were later analyzed on the basis of the conceptual framework using Mayring's structured content analysis (Mayring 2014).

A total of 15 case studies were initiated as part of our research activities. However, only three of these led to the formal specification of governance aspects and the institutionalization of a separate legal entity. The fact that only 3 of the 15 initiatives reached this stage is not an unusual ratio. The reasons for this are barriers to data sharing, as described by Fassnacht et al. (Fassnacht et al. 2023). While the three successful cases form the empirical basis for this paper, similar patterns were observed in the other cases. The following sections describe the three cases in more detail:

Case Study 1 (CS1) - Coolant Management: Coolant is used in metalworking to lubricate and cool tools and workpieces. Heavily contaminated coolant not only causes equipment wear and degrades product quality (e.g., by causing rust), but also potentially creates toxic fumes that pose a risk to life and limb (Evans et al. 2020). The companies in CS1 defined a cooperative value proposition based on a data-driven reorganization of the supply and management of coolant processes. To achieve this, digital services (e.g., continuous monitoring of coolant or optimization of the coolant ordering process) were defined and implemented both technically and organizationally. The associated data and information requirements were defined and the relevant reports and analyses were implemented - initially as a prototype by the researchers and later permanently by a professional software development company. The focus was on status data (e.g., pH values, coolant concentration, coolant temperature) of smart objects relevant to the coolant process (e.g., coolant, coolant tanks, shop floor environment). The digital data was transferred to a data space. Between August 2020 and August 2022, a total of seven workshops were conducted as part of the case study, each lasting two to three hours. In addition, several bilateral discussions were held with each of the companies. The participating companies are shown in Table 1.

Table 1: CS1 Participating Companies

Company	**Employees**	**Main representatives**
Metalworking company	~ 60	Managing Directors
Wholesaler and Service Provider for Coolant	~ 75	Head of Strategic Development, Head of Operational Cooling Lubricant Business
Manufacturer of Cleaning Machines for Coolant	~ 10	Managing Director

Case Study 2 (CS2) - Risk Mitigation in the Wood Processing Industry: In CS2, a sawmill, a wholesaler, and a risk broker worked together to stabilize supply chains and detect critical conditions in assets (mainly fire hazards in the sawmill) to increase operational readiness and promote risk reduction. Smart saw blades and machines measure status data (e.g. temperatures, residual currents, and power consumption volumes) and transfer them to a data space. Prototypical digital services have been implemented with partners to monitor and analyze the electric motors of band saws, shredders, and transport motors. Among other things, an alarm is triggered when certain relevant thresholds are exceeded. This allows risks to be identified at an early stage and damage to be prevented. A total of five workshops were held with the participating companies, as well as several bilateral meetings with each company. All workshops and discussions took place between March and December 2022. Table 2 provides a summary.

Table 2: CS2 Participating Companies

Company	**Employees**	**Participant**
Sawmill	~ 30	Managing Director
Wholesaler	~ 30	Managing Director
Insurance Broker	~ 10	Managing Director

Case Study 3 (CS3) - Improving Energy Efficiency in an Industrial Laundry: The third case study brought together an industrial laundry, a textile machinery manufacturer, and a gas burner manufacturer with the common goal of improving the energy efficiency and machine uptime of laundry finishing processes. Laundry finishing is a particularly energy-intensive process in which large laundry finishing machines use gas burners to quickly and gently heat air to dry laundry items. Given the focus on gas as an energy source at the time, the three companies decided to focus on this process first and investigate the potential for reducing energy consumption. Again, status data was collected from key assets in the processes and transferred to a data space; data from the conveyor system leading to the laundry finishing machine (e.g., current number and quality of laundry items), the laundry finishing machine itself (e.g., target and actual temperature, gas consumption), and the laundry shop floor (e.g., temperature and humidity). The data was then monitored and analyzed to gain insight into the common goal. In total, three workshops were held with the participating companies, as well as several bilateral discussions with each of the companies. All workshops and discussions took place between January and September 2022. Table 3 provides a summary of the participants.

Table 3: CS3 Participating Companies

Company	Employees	Participant
Industrial Laundry	~ 100	Managing Director, Expert for Machine Control Systems
Machine Manufacturer for Textile Machines	~ 25	Managing Director, Head of Construction, Expert for Machine Control System
Manufacturer of Gas Burners	~ 10	Managing Director

4 Results

The case studies served as the basis for the derivation of the design principles (DP) and thus the answer to RQ1. Each design principle was identified in one or more workshops in the context of one or more case studies. Using the conceptual framework, the design principles were derived from the field notes and observation protocols. They are abstractions from the cases and were evaluated in the final workshops. The identified DPs will be structured to answer RQ2.

4.1 Design Principles for Institutionalizing Data Ecosystems

The case studies show that it is important for all partners in the ecosystem to share a common goal, while at the same time aligning their own company-specific goals. In particular, the common purpose is needed to ensure the coherence of all subsequent governance measures. In the textile cleaning case study (CS3), the cooperative value proposition defined by the partners at the end of the second workshop was as follows "*Create a data space that enables ecosystem partners to improve energy efficiency and machine uptime in the laundry finishing process based on data-driven insights.*" Our experience supports the conclusion that this goal is abstract enough to allow the ecosystem to evolve as intended. It also sets the stage for the formulation of more concrete rules and regulations, all of which must feed into the provision of appropriate inputs to the data space, as well as the provision of the data space itself. The other two case studies supported this by providing similar insights.
DP1: Establish a cooperative value proposition for the actors in the data ecosystem.

In all three case studies, the participating companies considered the preservation of their legal autonomy as an irrevocable prerequisite for entering the ecosystem. This is particularly important for companies when a separate legal entity is established to provide

the shared data space on a permanent and binding basis (as was the case in all three cases). It was considered critical that this entity only provide the digital platform infrastructure. The members were strongly opposed to a scenario where the separate legal entity operates independently in the market as a joint venture. CS2 is a telling example: the insurance broker is not only interested in better risk assessment based on shared data to calculate more appropriate risk premiums. It also promotes a reduction in business interruptions at the sawmill, enabling a new type of service for the broker. For this to happen, however, the broker must have full access to the data - but not at the expense of the sawmill's independence, as was made clear in workshops three and four. The data governance framework should preserve the legal autonomy of all partner companies.
DP2: Protect the legal autonomy of the actors in the data ecosystem.

In CS1, partners stated that it was critical to be able to flexibly exit the collaboration at will to avoid lock-in. For the manufacturer of coolant cleaning machines, the ability to bring additional partners into the collaboration was seen as critical to the long-term success of the data ecosystem. Therefore, data governance should address the rules for joining and leaving the collaboration. The ability of the collaboration to scale is important to create and maintain a broad and deep data base for analysis, and this includes the ability to replace partners.
DP3: Provide regulations for the entry and exit of actors in the data ecosystem.

In all three cases, there were initial concerns about the lack of transparency about what the other partners in the ecosystem might do with the data transferred to the data space. For this reason, the initial workshops focused on jointly developing the cooperative value propositions and intended data uses. Transparency about who uses what data in what way and how it benefits them was created through visualization and group discussion. The companies considered status data (e.g., coolant temperature, current temperature in the laundry finisher, etc.) to be less critical than contextual information (e.g., physical location of machines or machine type, etc.). As a result, rules were established that the data space would be used primarily for sharing status data, and other use cases would require separate negotiations.
DP4: Provide rules establishing transparency about the intended use of shared data for the actors in the data ecosystem.

The partners in the three cases argued that the cooperative value propositions can only be realized if all partners make a binding commitment to contribute data of adequate quality. The cases showed that responsibilities for data provision need to be defined quantitatively and qualitatively. Our observations also suggest that in addition to quality, granularity and transfer rate need to be made explicit. Upon closer examination, most partners realized that the time criticality was less than expected. In summary, data governance needs to address the quality, granularity, and transfer rate of shared data.
DP5: Provide regulations that specify data quality, granularity, and transfer rate for shared data in the data ecosystem.

An important question regarding the DP3 regulations is what happens to the data of a partner who leaves the collaboration. Based on the discussions in the workshops, it was determined that the status data that a partner has contributed up to the point of leaving the data ecosystem will remain in the data space. Conversely, contextual information should be immediately deleted or anonymized. However, there are reasons why analytic models, templates for analytic models, benchmarks and results developed or computed with the contributed data can usually remain in the data ecosystem. This DP could be observed in Workshop 3 of CS1 and was repeatedly mentioned in the following workshops.
DP6: Provide regulations that govern the handling of data when an actor leaves the data ecosystem.

In the course of our research project, it became increasingly clear that the step towards a pilot implementation is a relevant challenge. CS1 and CS2 showed that acceptance of the implementation can be increased if the implementation itself is carried out in small steps and cycles as a continuous and PoC process. Participants emphasized that such a process allows an early assessment of whether the value proposition can be realized and what technical and organizational challenges are associated with the later implementation. In addition, the financial risk of implementation is lower in such an environment, and a cyclical approach is helpful for companies because it enables continuous synchronization with parallel digitization initiatives. As mentioned in the description of the case study above, the implementation of the required technical infrastructures was driven in each consortium, initially by the scientists in a prototypical manner, and in the later course of the collaboration in a leading role by a specialized service provider. After initial evaluations by the participating companies, there was a growing willingness to expand the implementation. Accordingly, data governance should provide rules for the continuous development of a proof of concept (PoC). We found that, despite claims to the contrary, the infrastructure required to capture, transfer, and store data of adequate quality is not always available or sufficient. For this reason, the governance frameworks developed in the case studies do not specify the form in which the data is to be collected, while at the same time allowing for at least temporary (semi-) manual entry (e.g. with spreadsheets) in case the automatically collected measurement data does not meet the requirements. For example, in CS1, at first only status data of the shop floor environment of the metalworking company (e.g., temperature and humidity), and of the coolant (e.g., temperature and pH value) could be made available. Later, other smart objects were added, such as coolant tanks (providing fill levels) and, later, coolant concentration measurements. After that, additional contextual information such as visualizations of machine locations, machine types, machine manufacturers, and machine ages were added to the data space. This led to the conclusion that the data ecosystem should be designed to be flexible enough to continuously add new types of data. Only through an iterative approach can acceptance be created for the technical implementation and development of the infrastructure.
DP7: Provide regulations for the continuous development of a PoC in the data ecosystem.

The rules for joint decision making in the context of data sharing became clear in CS1 and 2, especially during the final workshops. To achieve the cooperative value proposition, all partners need to contribute data and resources, so it is important that relevant decisions are made jointly and by majority vote. In CS1 and 2, it was bindingly agreed that fundamental decisions would be made in a general meeting where all members have equal voting rights. Examples of such decisions include responsibility for the implementation and operation of the IS, changes in billing and fee models, or in the underlying software tools and system environments. In terms of data governance, this means that bodies need to be created that allow such decisions to be made. This was particularly evident in CS2. Within the structures created there, individual companies were able to make joint decisions about the integration of additional data sources. Each member had a voice, which ensured trustworthiness for both data sharing and joint development of the data ecosystem.
DP8: Provide regulations ensuring the capability for shared decision making in the data ecosystem.

For the governance structure, it is crucial to clearly define which partner of the data ecosystem takes which role. As the data space is implemented and integrated into day-to-day operations, the need for clearly defined responsibilities becomes paramount. This includes, for example, the maintenance of sensors or data transfer services. A clear assignment of roles is particularly important in the context of liability issues. Roles should be defined and assigned within an independent legal entity of the ecosystem. Roles identified in CS2 (Workshops 3 and 4) include: the data provider contributes status data related to the bandsaw (e.g. temperature, fault currents); the knowledge provider provides industry-specific know-how when needed to perform specific analyses. The decision on which contextual information and industry-specific expertise to contribute is left to the individual company. While the case studies have emphasized the need to consider the different roles, they are subject to the specifics of each data ecosystem and its value proposition. Established role models from the literature can be used as a starting point, see e.g. Böhm et al. 2010; Yoo et al. 2010; Baars et al. 2022. Depending on the value proposition and the partners, the level of detail of the specification can vary.
DP9: Provide regulations to assign roles, responsibilities, and liability to actors of the data ecosystem.

4.2 Grouping and Characterization the Design Principles

To answer RQ2, we grouped and characterized the identified DPs based on our conceptual framework, based on Lusch & Nambisan (2015) (see Figure 1): service ecosystem, service platform, and value co-creation. We see that DP1, DP2, and DP3 support ecosystem formation, DP4, DP5, DP6, and DP7 govern data sharing on the digital platform infrastructure, and DP8, DP9 govern value co-creation activities (see Table 4). Our results suggested another dimension for structuring the design principles, namely a temporal one. Our observations and discussions with the participating companies from the different CS indicated that some of the identified DPs are relatively stable and require only few adjustments - such DPs are referred to as "static elements". Other DPs need to be constantly and quickly adapted to changes, "dynamic elements". We first assigned the DPs individually according to these dimensions and then consolidated the results. DP1, DP2, DP3, DP6, and DP8 were classified as static DPs, while we consider DP5 and DP9 to be dynamic. DP4 and DP7 combine static and dynamic elements (see Table 4). This distinction is also reflected in the types of legal documents developed.

Table 4: Data Governance Design Principles

#	Data Governance Design Principles	CS	SDL	Static/ Dynamic
DP1	Establish a cooperative value proposition for the actors in the data ecosystem	CS1, CS2, CS3	Service Eco-system	Static
DP2	Protect the legal autonomy of the actors in the data ecosystem	CS1, CS2, CS3		Static
DP3	Provide regulations for the entry and exit of actors in the data ecosystem	CS1		Static
DP4	Provide rules establishing transparency about the intended use of shared data for the actors in the data ecosystem	CS1, CS2, CS3	Service Platform	Static/ Dynamic
DP5	Provide regulations that govern the handling of data when an actor leaves the data ecosystem	CS1, CS2, CS3		Dynamic
DP6	Provide regulations that govern the handling of data when an actor leaves the data ecosystem	CS1		Static
DP7	Provide regulations for the continuous development of a PoC in the data ecosystem	CS1, CS2,		Static/ Dynamic
DP8	Provide regulations ensuring the capability for shared decision making in the data ecosystem	CS1, CS2	Value Co-Creation	Static
DP9	Provide regulations to assign roles, responsibilities, and liability to actors of the data ecosystem	CS1, CS2, CS3		Dynamic

5 Discussion

Essential aspects of a governance structure in the context of a data ecosystem have not been adequately addressed in existing work (Weber et al. 2022; Pretzsch et al. 2020; Baars et al. 2021; Gelhaar & Otto Boris 2020). Furthermore, our literature review supports the conclusion that more precise recommendations are still lacking. This paper contributes to this research gap by abstracting insights from three case studies and deriving DPs that guide the design of data governance. By using SDL as a theoretical lens, we were able to place our findings within a larger theoretical construct. In this way, data governance measures contribute to the establishment of a shared institutional logic and thus provide the framework for collaboration in the sense of SDL. Previous work has pointed out that collaboration in data ecosystems is fluid and often not based on fixed structures (Gelhaar & Otto Boris 2020). To address this, this paper provides a specification of DPs that distinguishes between static and dynamic DPs. First initiatives in science and practice have started to work on the selection of adequate legal forms for data ecosystems (Reiberg et al. 2022; Pretzsch et al. 2020; Weber et al. 2022). The static and dynamic design principles identified in this paper support such activities by delineating aspects that need to be considered. A relevant line of further research is a deeper exploration of DPs and their mapping to different legal forms.

6 Conclusion and Limitations

The paper provides a contribution to the service innovation framework in the context of institutionalizing data governance aspects in data ecosystems. This is relevant for both academia and practice. For the scientific community, the paper adds value by identifying nine design principles based on the discussions on data governance in data ecosystems. The identified dynamic and static design principles can be used as a basis for further research to determine which design principles can be mapped to which documents of different legal forms. The added value of our paper for practice is that it provides a basis for companies to decide which governance aspects they should address in a data ecosystem.

The paper has several limitations: First, as an ADR project, the researcher was involved in the observed social system, which carries a risk of bias (Cassell & Symon 2004). Second, the DPs are not free of overlap. Third, as a qualitative study, the proposed DPs and conclusions are of a preliminary nature and would require further rigorous testing - we make no claim to completeness regarding the DPs. Fourthly, the companies in the study are mainly located in southern Germany, which may lead to a geographical bias. The chosen research design does not yet allow us to assess how important individual DPs are for the institutionalization of data ecosystems. Further cases will be initiated to answer this question. In addition, these cases will also help to increase the resilience of the DPs.

7 Acknowledgment

The case studies described in this paper were conducted in the context of the research project “Digitale Datenräume zur Kooperation von KMUs unter Einsatz von KI zur Schaffung von Wettbewerbsvorteilen gegenüber ausländischem Wettbewerb, der auf Datenplattformen setzt”, which was funded by the Ministry of Economy, Labor and Tourism Baden-Württemberg.

References

Abraham, R., Schneider, J. & Vom Brocke, J. (2019), Data governance: A conceptual framework, structured review, and research agenda, *in* 'International Journal of Information Management', **49**, pp. 424–438.

Adner, R. (2017), Ecosystem as structure: An actionable construct for strategy, *in* 'Journal of Management', **43**(1), pp. 39–58.

Azkan, C., Möller, F., Meisel, L. & Otto, B. (2020), Service dominant Logic Perspective on Data Ecosystems - a Case Study based Morphology, *in* 'Proceedings 28th European Conference on Information Systems'.

Baars, H., Tank, A., Weber, P., Kemper, H. G., Lasi, H., & Pedell, B. (2021), Cooperative approaches to data sharing and analysis for industrial internet of things ecosystems, *in* 'Applied Sciences', **11**(16), 7547.

Baars, H., Weber, P. & Tank, A. (2022), Institutionalizing Analytic Data Sharing in SME Ecosystems. - A role-based Perspective, *in* 'Proceedings 55th Hawaii International Conference on System Sciences'.

Badewitz, W., Kloker, S. & Weinhardt, C. (2020), The Data Provision Game: Researching Revenue Sharing in Collaborative Data Networks, *in* 'Proceedings 22th IEEE Conference on Business Informatics (CBI), IEEE.

Blaschke, M., Riss, U., Haki, K. & Aier, S. (2019), Design principles for digital value co-creation networks: a service-dominant logic perspective, *in* 'Electronic Markets' **29**, pp. 443–472.

Böhm, M., Koleva, G., Leimeister, S., Riedl, C. & Krcmar, H. (2010), Towards a generic value network for cloud computing, *in* 'International workshop on grid economics and business models', Springer, pp. 129–140.

Braud, A., Fromentoux, G., Radier, B. & Le Grand, O. (2021), The road to European digital sovereignty with Gaia-X and IDSA, *in* 'IEEE network', **35**(2), pp. 4–5.

Cassell, C. & Symon, G. (2004), Essential guide to qualitative methods in organizational research, SAGE.

De Reuver, M., Ofe, H., Agahari, W., Abbas, A. E. & Zuiderwijk, A. (2022), The openness of data platforms: a research agenda, *in* 'Proceedings 1th International Workshop on Data Economy', pp. 34-41.

Evans, R. D., Hooijman, J. J. & van der Veer, J. M. (2020), Advances in cooling and lubrication for high speed machining, *in* 'High Speed Machining', Elsevier, pp. 97–125.

Fassnacht, M., Benz, C., Heinz, D., Leimstoll, J. & Satzger, G. (2023), Barriers to Data Sharing among Private Sector Organizations, *in* 'Proceedings 56th Hawaii International Conference on Systems Sciences'.

Fernandez, R. C., Subramaniam, P. & Franklin, M. J. (2020), Data market platforms: Trading data assets to solve data problems, *in* 'Proceedings of the VLDB Endowment', **13**(12), pp 1933–1947.

Gelhaar, J. & Otto B. (2020), Challenges in the Emergence of Data Ecosystems, *in* 'Proceedings 23th Pacific Asia Conference on Information Systems'.

Hein, A., Scheiber, M., Böhm, M., Weking, J., Rocznik, D. & Krcmar, H. (2018), Towards a Design Framework for Service Platform Ecosystems, *in* 'Proceedings 26th European Conference on Information Systems'.

Kannan, V., Mathew, S. K. & Lehner, F. (2022), Digitally Enabled Shrimp Farming: A Service-Dominant Logic View, *in* 'Proceedings Americas Conference on Information Systems'.

Lehrer, C., Wieneke, A., vom Brocke, J. A. N., Jung, R. & Seidel, S. (2018), How big data analytics enables service innovation: materiality, affordance, and the individualization of service, *in* 'Journal of Management Information Systems', **35**(2), pp. 424–460.

Levy, Y. & Ellis, T. J. (2006), A systems approach to conduct an effective literature review in support of information systems research, *in* 'Informing Science Journal', **9**, pp. 181–213.

Lis, D. & Otto, B. (2021), Towards a taxonomy of ecosystem data governance, *in* 'Proceedings 54th Hawaii International Conference on System Sciences', pp. 6067-6076.

Lusch, R. F. & Nambisan, S. (2015), Service innovation: A service-dominant logic perspective, *in* 'MIS quarterly' **39**(1), pp. 155–176.

Martens, B., de Streel, A., Graef, I., Tombal, T. & Duch-Brown, N. (2020), Business-to-Business data sharing: An economic and legal analysis, JRC Digital Economy Working Paper, No. 2020-05, European Commission, Joint Research Centre (JRC).

Mayring, P. (2014), Qualitative content analysis: theoretical foundation, basic procedures and software solution.

Oliveira, M. I. S. & Lòscio, B. F. (2018), What is a data ecosystem? *in* 'Proceedings 19th International Conference on Digital Government Research: Governance in the Data Age'.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A. & Moher, D. (2021), The PRISMA 2020 statement: an updated guideline for reporting systematic reviews, *in* 'International journal of surgery', **88**.

Perscheid, G., Ostern, N. K. & Moormann, J. (2020), Determining Platform Governance. Framework for Classifying Governance Types, *in* 'Proceedings 41th International Conference on Information Systems'.

Porter, M. E. & Heppelmann, J. E. (2014), How smart, connected products are transforming competition, *in* 'Harvard Business Review', **92**(11), pp. 64–88.

Pretzsch, S., Drees, H. & Rittershaus, L. (2020), MOBILITY DATA SPACE. EIN SICHERER DATENRAUM FÜR DIE SOUVERÄNE UND PLATTFORM ÜBER GREIFENDE BEWIRTSCHAFTUNG VON MOBILITÄTSDATEN, Accessed: 08.03.2023.

Reiberg, A., Niebel, C. & Kraemer; P. (2022), What is a Data Space?, White Paper.

Scholz, N., Wieland, J. & Schäffer, T. (2022), Towards a Framework for Enterprise & Platform Ecosystem Data Governance, *in* 'Proceedings Americas Conference on Information Systems'.

Sein, M. K., Henfridsson, O., Purao, S., Rossi, M. & Lindgren, R. (2011), Action design research, *in* 'MIS quarterly', **31**(1), pp. 37–56.

Shipilov, A. & Gawer, A. (2020), Integrating research on interorganizational networks and ecosystems, *in* 'Academy of Management Annals', **14**(1), pp. 92-121.

Susha, I., Janssen, M., & Verhulst, S. (2017), Data collaboratives as a new frontier of cross-sector partnerships in the age of open data: Taxonomy development, *in* 'Proceedings 50th Hawaii International Conference on System Sciences'

Van den Broek, T. & Fleur van Veenstra A. (2015), Modes of governance in inter-organizational data collaborations, *in* 'Proceedings 23th European Conference on Information Systems'.

Vargo, S. L. & Lusch, R. F. (2004), The four service marketing myths: remnants of a goods-based, manufacturing model, *in* 'Journal of service research' **6**(4), pp. 324–335.

Vargo, S. L. & Lusch, Robert F. (2008), Service-dominant logic: continuing the evolution, *in* 'Journal of the Academy of marketing Science', **36**, pp. 1–10.

Vargo, S. L. & Lusch, R. F. (2016), Institutions and axioms: an extension and update of service-dominant logic, *in* 'Journal of the Academy of marketing Science', **44**, pp. 5–23.

Vom Brocke, J., Simons, A., Niehaves, B., Niehaves, B., Reimer, K., Plattfaut, R. & Cleven, A. (2009), Reconstructing the giant: On the importance of rigour in documenting the literature search process, *in* ‘Proceedings 17th European Conference on Information Systems’.

Weber, P., Hiller, S. & Lasi, H. (2020), Identifying Business Potentials within an IoT Ecosystem – An Explorative Case Study in the Industrial Domain, *in* ‘Americas Conference on Information Systems’.

Weber, P., Werling, M., Tank, A. & Baars, H. (2022), Institutionalisierung digitaler Ökosysteme in der Rechtsform einer Genossenschaft: Case Study im produzierenden Kontext, *in* ‘HMD Praxis der Wirtschaftsinformatik’, **59**(5), pp. 1353–1365.

Weill, P. & Ross, J. W. (2004), IT governance: How top performers manage IT decision rights for superior results, Harvard Business Press.

Yin, R. K. (2018), Case study research and applications. Design and methods, **6**, SAGE.

Yoo, Y., Henfridsson, O. & Lyytinen, K. (2010), Research commentary - the new organizing logic of digital innovation: an agenda for information systems research, *in* ‘Information systems research’ **21**(4), pp. 724–735.