10-9-2023

From Challenges to Solution Pathways for Industrial Data Ecosystems - A Socio-Technical Perspective

Marc Brechtel
University of Potsdam, Germany, marc.brechtel@uni-potsdam.de

Dimitri Petrik
University of Stuttgart, Germany, dimitri.petrik@bwi.uni-stuttgart.de

Katharina Hölzle
Fraunhofer Institute of Industrial Engineering, Germany, katharina.hoelzle@iao.fraunhofer.de

From Challenges to Solution Pathways for Industrial Data Ecosystems – A Socio-Technical Perspective

Research Paper

Marc Brechtel[1], Dimitri Petrik[2], and Katharina Hölzle[3]

[1] University of Potsdam, Potsdam, Germany
marc.brechtel@uni-potsdam.de
[2] Graduate School of Excellence Advanced Manufacturing Engineering, Stuttgart, Germany
dimitri.petrik@bwi.uni-stuttgart.de
[3] Fraunhofer Institute for Industrial Engineering, Stuttgart, Germany
katharina.hoelzle@iao.fraunhofer.de

Abstract. Many industrial firms know that inter-firm data sharing holds tremendous potential for the creation of new economic value. However, most of them are not yet willing to share data multilaterally for joint value creation. Consequently, they do not participate in so-called data ecosystems. In our study, we address this issue from a socio-technical perspective and apply a two-step qualitative-empirical research design composed of a problem-centered interview study followed by solution-oriented workshops. We have identified and prioritized the perceived challenges and present first solution pathways. We were able to assign respective responsibilities for needed actions to major roles in data ecosystems. Among others, we can show that the core participants, i.e., data owner and data user, and the governance body need to resolve key challenges mostly in a separate but complementary manner to help industrial data ecosystems emerge.

Keywords: Data Ecosystems, Data Sharing, Socio-Technical Systems

1 Introduction

Data has become a valuable asset for many firms. Inter-firm collaboration on a data level holds tremendous potential (Oliveira et al., 2019; Azkan et al., 2020) to significantly improve or create new business models (Yoo et al., 2010). The added value of digital services highly depends on the accessibility of data sets inside and between firms. Because the required data for offering a value-adding digital service for its intended beneficiaries are often not generated by only one firm, firms are increasingly looking for complementary data sets that other firms could provide (Gelhaar et al., 2021). This is why firms are starting to consider joining networks to share data with peers to collaborate on digital services (Pfähler et al., 2022). The resulting socio-technical networks aim at making data accessible across firms' boundaries are known as data ecosystems (Oliveira et al., 2019). To ensure a secure exchange of sensitive data between organizations, they have the political support of the European Commission

18th International Conference on Wirtschaftsinformatik,
September 2023, Paderborn, Germany

(2020). Adopting the logic of established ecosystem literature (Adner, 2017), data ecosystems are defined by multilateral relationships between different actors with a value proposition that is not decomposable into multiple bilateral relationships.

Literature has identified three major motives for firms participating in a data ecosystem (Bauer et al., 2022): First, existing products and (digital) services can be further improved with respect to productivity and efficiency. Second, firms are increasingly pressured by legislative movements (e.g., Supply Chain Act) demanding higher social and environmental sustainability. Third, data ecosystems provide room for the realization of new data-driven digital services, like collaborative condition monitoring. Here "data is not only shared bilaterally but multilaterally between all actors in an ecosystem" (Gelhaar et al., 2021, p. 127). Nevertheless, firms continue to shy away from sharing data with peers and keep data siloed (Gelhaar et al., 2023). Consequently, the business potential that data ecosystems hold cannot be leveraged (De Prieelle et al., 2020). Early supporters of the data ecosystem idea experience difficulties convincing other actors along the supply chain to join a data ecosystem due to the fear of releasing sensitive data (Bian et al., 2016; Zhiwei et al., 2021). Such obstacles are even recognized by European policy makers (European Commission, 2018). Previous research has shown that it is important to ensure that all potential actors benefit from taking part in a data ecosystem (Gelhaar et al., 2021). Without a clear benefit, firms neglect the idea of data sharing along the supply chain (Gelhaar et al., 2021). This effect is even reinforced when firms consider data as business-sensitive, especially in traditional sectors (Gelhaar et al., 2021). Nevertheless, recent literature emphasizes the socio-technical nature of data ecosystems. For example, Oliveira et al. (2019) emphasize the need for holistic approaches to study data ecosystem evolvement, going beyond the technical challenges. Despite first insights on factors influencing data sharing from a game-theoretical perspective (Kraemer et al., 2021) or focus on data sharing between scientists (Devriendt et al., 2021), further empirical studies are needed to gain knowledge about the socio-technical design of industrial data ecosystems.

This study aims to reason the absence of data ecosystems by identifying and prioritizing the perceived challenges and giving the first indication about solution pathways to move forward the understanding of how to design and scale industrial data ecosystems from a socio-technical perspective. We enrich existing predominantly conceptual work by allocating selected responsibilities to the major roles in a data ecosystem based on empirical data (Oliveira et al., 2019; Otto, Steinbuss, et al., 2019). To get there, we are guided by the following research questions: *What are the critical challenges industrial firms perceive while considering data ecosystems? What solution pathways may help to overcome these challenges and who should take action?* With our study, we empirically identify needed actions to address the prioritized challenges differentiating between the core participants, namely the data owner and the data user, plus the governance body of a data ecosystem in industry.

2 Fundamentals

2.1 Data Ecosystems as Socio-Technical Networks

Prior research conceptualized the ecosystem construct to depict the connections and relations among partners that strive for a strong focal value proposition (Adner, 2017). Drawing on the original concept of business ecosystems (Moore, 1993), ecosystem types reach from innovation ecosystems (Adner, 2006) to digital platform ecosystems (Hein et al., 2020). Data ecosystems are another manifestation of the ecosystem construct. They are defined as distributed socio-technical systems in which autonomous actors interact and collaborate to explore, analyze and interpret data mutually while aiming for joint innovation and value creation (Oliveira et al., 2019). The term socio-technical systems goes back to Emery and Trist (1960). They coined this term to "describe systems that involve a complex interaction between humans, machines and […] environmental aspects" (Baxter & Sommerville, 2011, p. 5). Thus, transferred to today's data ecosystems the original logic of interdependent technical and social elements still holds true and is frequently applied in the information systems (IS) discipline (Sarker et al., 2019).

The success of data ecosystems relies on its actors' active engagement (Heimstädt et al., 2015). The rationale is that more value can be created based on shared data than if each ecosystem actor computes his own data only. Thus, data ecosystems are seen as a driver for innovation in which the data sovereignty of each actor needs to be preserved (Otto, Lis, et al., 2019). Compared to centralized digital platforms known to promote power asymmetries, known as the winner-takes-it-all phenomenon (Van der Aalst et al., 2019), applications functionalities can be implemented in so called connectors to increase the level decentralization in data ecosystems (Otto & Jarke, 2019).

Taking a socio-technical perspective helps to better manage and achieve positive outcomes of information systems (Bostrom & Heinen, 1977). Recent work related to data ecosystems stresses that the consideration of the interdependence of socio-technical dimensions is critical for their success (Oliveira et al., 2019). Hence, following the traditional socio-technical system literature (Brandt & Cernetic, 1998), successful data ecosystems need to consider the interrelation of technology, people, and the organization while aiming for joint value creation. More recently, some scholars even refer to data as as an additional element of socio-technical systems (Wallace, 2015; Weber et al., 2021). Albeit "it is widely acknowledged that adopting a socio-technical approach to system development leads to systems that are more acceptable to end users and deliver better value. […] such approaches are not widely practiced" (Baxter & Sommerville, 2011, p. 4). This also holds true for the ongoing endeavors to establish data ecosystems in industry and latest research in this field (Brechtel & Petrik, 2023). Scholars have mainly focused on a sub-set of the socio-technical dimensions of data ecosystems, namely technology (Oliveira & Lóscio, 2018). Thus, we argue that a socio-technical view, determining the success or failure of data ecosystems, is required.

2.2 Data Ecosystem Roles

There are various roles actors, i.e., enterprises or institutions, can occupy in data ecosystems. A role is a function related to a set of duties occupied by a data ecosystem actor (Oliveira et al., 2019). Typically, a data ecosystem constitutes at least two roles, namely *data owners* and *data users* (Oliveira et al., 2019). These two roles need to be filled by actors every time data is exchanged – making them core participants (Otto, Steinbuss, et al., 2019). However, additional roles can be filled in a data ecosystem (Oliveira et al., 2019). Another important role beside the core participants is responsible for creating the rules and policies. The International Data Spaces Association refers here to the *governance body* (Otto, Steinbuss, et al., 2019). *Data owners* are defined as legal entities creating data and/or executing control over it. For example an industrial firm measuring parameters to assess the condition of its pump in the field might act as a data owner. In contrast, a *data user* is a legal entity equipped with the right to use the data owned by different actors (Otto, Steinbuss, et al., 2019). In our example, a data user, might be a maintenance provider making use of the pump data to improve the uptime. The *governance body "*makes sure that only compliant organizations are granted access" (Otto, Steinbuss, et al., 2019, p. 25). The governance body is responsible for standardization, certifications, and governance (Catena-X, 2022b), aiming for the interoperability, data sovereignty, and security of all data ecosystem participants. This body is comparable to the ecosystem designer taking care of the design and implementation, as introduced by Autio (2022) for innovation ecosystems. Data ecosystem founders or standardization institutions may assume this role (Oliveira et al., 2019).

3 Methodology

In our study, we applied a two-step qualitative-empirical research design composed of problem-centered interviews followed by a set of solution pathway workshops.

Step one: The challenges perceived by participating or at least interested stakeholders in data ecosystems, were investigated through problem-centered interview data collection (Döringer, 2020). Our study put the focus on organizations integrated in an industrial supply chain either producing and selling physical goods, or providing services paving the way for inter-firm data sharing. Therefore, we searched for representatives from industrial firms with expertise in data ecosystems or data economy since making machine data available or utilizing such data from other organizations appears to be difficult (European Commission, 2018). Our interview partners were mostly from German incumbent firms, which either actively participate or communicate interest in industrial data ecosystems, such as Catena-X - a data ecosystem dedicated to the automotive industry. SMEs and start-ups made up a smaller share of the data sample. We completed the relevant set of interview data until theoretical saturation, i.e., the iterative, open coding process did not reveal any new challenges, has been reached. In total, we interviewed 31 experts from 26 different firms between June and November 2021 (Table 1, available at bit.ly/interview_workshop_participants). We used a semi-structured format to discuss predefined topics and add follow-up questions to get valuable information related to the research goal. Details of the interviews and their processing have

been described elsewhere (Brechtel & Petrik, 2023), since this study is part of a superordinate research project. In brief, we asked for the perceived challenges related to data ecosystems first, before we touched upon first solution ideas. Every interview was recorded, transcribed and coded. Via axial coding we got from first-order concepts to second order themes (Corbin & Strauss, 1988). To achieve objectivity, group discussions among the researchers were used to reach a consensus during coding. We used Gioia et al. (2013) to aggregate the codes guided by the three dimensions of the socio-technical framework by Brandt & Cernetic (1998) comprising technology, people, and organization. We added value as a fourth dimension, as it is a defining feature of ecosystems (Jacobides et al., 2020). We decided not to create an explicit dimension dedicated to data since data-related aspects are ubiquitios in the other four dimensions.

Step two: To further refine and add to our findings from the interview series, we organized workshops. We have decided for workshops as they provide room for stimulating discussions among experts (Ørngreen & Levinsen, 2017) for the purpose to map solution pathways and assigning respective responsibilities. The workshop format comprised four steps (i.e., longlist reflection, narrowing down, prioritizing, and assigning responsibilities), incorporating qualitative research to distill consensus withing the group of experts. In doing so, the first three of the four workshop steps draw on a study design described and applied in previous IS research (Schmidt et al., 2001).

In total, one inter-firm and five intra-firm workshops were carried out. The intra-firm workshops were held with firms that already took part in the previous interview study. Thereby we achieved an acceptance rate of approx. 19% on a firm level. For the inter-firm workshop, we used an event where platform and data ecosystem knowledge carriers participated anyhow and applied the same workshop procedure as for the intra-firm setting. At least one researcher moderated the workshops, while another took notes on a shared MIRO board. We started our workshops by introducing the longlist of challenges, and asked each workshop participant to select the top five challenges. Using MIRO functionalities we brought the independent selection of participants together, and conducted dot voting. If a clear ranking could not be derived based on the distribution of votes within the group immediately, we took a decision on prioritization in the big round. Based on the prioritized challenges, we proceeded with each workshop group by elaborating possible solution pathways. Further, we stimulated group discussions among the participating experts (Myers, 2013) regarding the assignment of responsibilities differentiating between the core participants and the governance body of a data ecosystem. The cross-workshop data analysis resulted in a shortlist of challenges (Table 2), enriched by needed actions per challenge and role (Table 3).

4 Results

Starting with a longlist of 29 socio-technical challenges related to collaboration on data in data ecosystem across firms' boundaries (Brechtel & Petrik, 2023), the workshops enabled the researcher team to distill a shortlist of twelve challenges. Below we refer to aspects introduced by Brechtel & Petrik (2023) in section 4.1, but limit ourselves to the shortlisted challenges, before we map solution pathways to them in section 4.2.

4.1 Challenges

The social-technical perspective allows us to differentiate challenges for industrial data ecosystems into technology-, people-, organization- and value-related challenges.

Technology. Our data revealed two prioritized technology challenges in the emergence of industrial data ecosystems. First, industrial firms struggle with the poor interoperability of system landscapes within and across firms' boundaries. We witnessed different data standards and interfaces per application. As a consequence, we see different levels of data structure homogeneity across firms and especially across industries [#26]. According to our data, this is due to the fact that the industries' standardization efforts vary greatly. Second, our workshops highlighted that many firms, but especially SMEs, still suffer from *data availability* in general. Project activities are required to make this data accessible as the following statement underpins [#27]: "Currently, it is still a huge project business to get access to the data at all, with a huge investment."

People. Based on our workshop results, we add four people-related challenges to the shortlist. First, the *lack of knowledge carriers* seems to be an issue. Especially on the management level, firms lack knowledge carriers to assess the opportunities of data spaces as the underlying technologies and thought structures are distinct from anything that has been encountered previously [#27]. Second, the *poor venture-driven mindset* towards the mass adoption of industrial data ecosystems forms another issue. Our data highlights risk-averse behavior on the management level. Interviewee [#7] claimed that today's decision makers consistently seek to observe and replicate use cases. Following him, only a few are early adopters who are willing to establish data ecosystems. Third, a prioritized leadership-related issue is an *unsuitable incentive system*. Our data reveals that management is not incentivized to budget for data ecosystems [#1]. Hence, if at all, most CEOs stick to lip service only. Finally, the barrier of *mistrust* is also evident. As of today, the knowledge is power thought dominates across firms. Even though, more and more industrial firms acknowledge the necessity of obtaining data from others, the great majority still refrains from sharing data – mainly due to trust issues [#9].

Organization. Silo thinking and unclear data responsibility have been prioritized as organizational challenges. Today, most industrial firms stick to developing proprietary verticals aiming to enhance their solution offering by incorporating digital elements [#19]. Hence, *silo thinking* is omnipresent leading to many solutions emerging side by side. Further, missing intra- and inter-firm *data responsibilities* lead to entities claiming data ownership of the same data, even within a firm. The question of who is responsible for the data remains unresolved within most industrial firms [#26].

Value. Most of the prioritized challenges relate to the dimension of value. First, there are doubts about the added business values with current data ecosystem initiatives are perceived revolving mainly around technology. One interviewee [#19] states that "things are simply thrown together, and the business model consideration falls short." Second, firms shy away from becoming early adopters of data ecosystems, as they cannot estimate the value of their data yet. This is because firms are used to pricing products and services, but not data. Adding to this is the fact that data possesses no value per se [#25]: Its value only arises while being contextualized. However, the individual firm that may provide data often lacks knowledge about the problem context which in

turn hinders the determination of the actual data value. Third, *high initial investments* build a hurdle. Preparing to enter the realm of data ecosystems often involves facing substantial financial commitments while not knowing the exact economic benefit [#28]. This is a key reason why many firms are reluctant about making such investments. Finally, the *poor accessibility of use cases* is also among the shortlisted challenges. Decision-makers are waiting to be inspired by accessible use cases as most data ecosystem use cases miss detailed descriptions and are nothing beyond buzzwords [#5].

Table 2. A Shortlist of Challenges related to Data Ecosystems

	Challenge Group	Challenge	Prioritized in Workshops
Tech.	Compatibility	Poor interoperability	#3, #5
	Digital Readiness	Poor data availability	#2, #3, #4, #6
People	Knowledge	Lack of knowledge carriers	#2, #5
	Leadership	Poor venture-driven mindset	#1, #2, #4, #6
		Unsuitable incentive system	#1, #2, #6
	Trust	Lack of trust	#5
Org.	Culture	Silo thinking	#4
	Accountability	Unclear data responsibility	#1, #3
Value	Business	Unclear business value	#4, #5, #6
		Unclear data value	#1, #2, #4, #5, #6
	Profitability	Need of high initial investments	#3
	Use Case	Poor accessibility of use cases	#1, #3

4.2 Solution Pathways

In this section, we elaborate on suitable solution pathways to address the challenges using the previously introduced roles in data ecosystems as a structuring element.

Core Participants. Technology-related challenges that require the attention of the core participants, i.e., data owner and data user, concern the *need to take stock of available data* on a firm level. While doing so, labeling and clustering the respective data is mandatory to get toward a firm internal data catalog [#2, WS #4]. Further, workshop group #6 recommended "establishing advanced data management mechanisms" on a firm level. People-related challenges need attention from core participants in at least four cases. First, to address the lack of knowledge carriers the experts recommend establishing a *mentoring program* to take the people with you [#12]. Following workshop group #2 the "digital natives need to train experienced top managers how to cope with data ecosystems." Second, the challenge of the missing venture-driven mindset can only be addressed if firms aim for top management support. To get there, "firms need to ensure *direct access of operational teams* to high-level sponsors with decision-making competence" [WS #4]. Further, data ecosystem supporters within a firm need to *create a sense of urgency* by knowing the right people and identifying the best leverage point [WS #2]. One leverage point might be to install a *right-minded visionary C-level* representative [#7, #11, WS #6]. Third, the core participants need to adjust toward more long-term-oriented thinking. Based on our study results, the currently unsuitable incentive systems need to be replaced by creating and setting *long-term incentivization* [#6] bound to *clear deliverables* [WS #6]. Fourth, the lack of trust needs to be addressed by

adopting a *zero-trust environment*. Here, one of our workshop groups recommends the core participants implement and use only certified solutions to get there [WS #5]. Organization-related challenges that the core participants need to address include silo thinking and unclear data responsibilities. For the former, it needs a cultural mind shift towards "*thinking in terms of the ecosystem* value on not only firm value" [WS #4]. Historically, this ecosystem thinking has been the case with all the consortia firms that are traditionally organized. However, "with respect to data we have lost that a bit. But now we have to get back on track and say let's [share data] under certain rules" [#9]. For the latter, our workshop participants recommend the establishment of a dedicated *data team* [WS #1]. Further, there is a need for a *project organization* over a line organization when it comes to intra- and inter-firm data sharing [#16, WS #3]. Other needed actions identified for the core participants concern unclear business and data value. Participation in lighthouse projects, such as Catena-X, for industrial data ecosystems, may prove beneficial to *demonstrate the added business value* [WS #5]. To better understand the value of data sets, our study shows that the core participants should apply a *case study-driven approach*. Workshop group #2 agreed that "the value of data must be identified over time". Interviewee #2 is convinced that "iterations will help to gain insights that help to interpret the value of data". A thinking that was confirmed in workshop #4. Further, our experts [WS #2] recommend *introducing a data strategy* on the firm level first, before strategic thinking is widened up towards data ecosystems.

Governance Body. Regarding the perceived technology-related challenges, the governance body is required to *set and enforce rules and standards* to improve inter-firm interoperability. Workshop group #3 agreed that the developments and standardization around connectors are crucial. Further, this group agreed that the governance body needs to define standards for how data is best shared - meaning which data format and technical solutions, i.e., systems and interfaces, need to be in place. To improve the given data availability the governance body should *provide clear data requirements*. What is needed is a "to-do list on data for all data ecosystem actors" [WS #4]. Actions by the governance body are also required when it comes to people-related challenges. To improve the level of trust among data ecosystem actors the governance body should foster the idea of a *zero-trust environment* [#26]. Again, the connector plays a crucial role in ensuring data sovereignty [WS #4]. Further, the introduction of *fair data management principles*, such as "previously established certification standards that the ecosystem commits to" [WS #5] may help here. To operationalize this thinking, a *governance model which is managed by a neutral entity* is proposed. Interviewee #11 underpins the need for such a neutral instance by claiming that only then fair conditions are met. To address the organization-related challenge of silo thinking "institutions who push for ecosystem value" [WS #4] are needed. The governance body can be interpreted as such an institution that needs to take responsibility for continuously demonstrating the benefit of *ecosystem thinking* over silo thinking. Value-related challenges ask for actions by the governance body in various fields. First, this body should set and enforce *mechanisms for joint value creation and value capture* among the different data ecosystem actors. Short-term, simple revenue sharing seems to be a conceivable solution approach. Long-term, one workshop group asked for an inter-firm business model [#6]. Further, the governance body "needs to provide support, e.g., by provisioning a *value*

catalog for data exchange" [WS #4] which is based on an ecosystem-wide alignment indicating the value of each data point. Another aspect mentioned has been that the governance body should consider making the early adopters offer *initial cost coverage* until critical mass is reached [WS #3]. As SMEs often lack the resources to actively participate in data ecosystems, they need *special attention* from the governance body. This is especially true when considering that SMEs bridge the gap between major actors at the beginning and end of the automotive supply chain [WS #5]. Finally, we see that the governance body needs to better *describe and modularize use cases* to make them more accessible to the mass market. More specifically, it is recommended to introduce a syntax for the description of use cases related to industrial data ecosystems. Thereby the what and why from a firm and ecosystem perspective need to be especially addressed. "We need to move away from pure buzzwords. Instead, we should introduce a syntax for the use case description […]" [#18].

Table 3. Assignment of Needed Actions per Challenge and Role

	Challenge	Solution	Needed Action	CP*	GB**
Tech.	Poor interoperability	Standards	Set & enforce rules/standards		x
	Poor data availability	Data Cataloging	Take stock of intra-firm data	x	
			Provide data requirements		x
People	Lack of knowledge carriers	Mentoring	Introduce mentoring program	x	
			Establish an intra-firm data ecosystem community	x	
	Poor venture-driven mindset	Top management support	Ensure direct reporting of project teams to sponsors	x	
			Create a sense of urgency	x	
			Install right-minded C-level	x	
	Unsuitable incentive system	Long-term orientation	Set long-term incentivization	x	
			Define clear deliverables	x	
	Lack of trust	Data sovereignty	Establish & adopt zero-trust environment	x	x
			Introduce fair data mgmt. principles		x
			Introduce governance model managed by a neutral entity		x
Org.	Silo thinking	Cultural mind shift	Establish ecosystem thinking	x	x
	Unclear data responsibility	Data team	Establish intra-firm data team	x	
			Introduce a project org.	x	
Value	Unclear business value	Value demonstration	Demonstrate added value	x	
			Introduce mechanisms for value creation & value capture		x
	Unclear data value	Data Valuation	Intra-firm data valuation based on case studies	x	
			Introduce a data strategy	x	
			Provide a data valuation catalog		x
	Need of high initial investments	Subsidy	Provide initial cost coverage		x
			Take care of SME requirements		x
	Poor accessibility of use cases	Use case accessibility	Describe & modularize use cases		x

*CP = Core Participants | **GB = Governance Body

5 Discussion

Our study examines the challenges perceived by (potential) data ecosystem actors and indicates what is needed to overcome them. The assignment of responsibilities allows us to derive advice for the roles and activities that data ecosystems require.

The governance body should primarily focus on finding solutions for the remaining technology-related challenges. Our research confirms and adds to the obligations that come with the role of the governance body, conceptualized earlier, for example by Oliveira et al. (2019). We find support for the argument that the governance body needs to take responsibility for standardization inside a data ecosystem (Catena-X, 2022b). Thus, based on our study results we welcome the decision by the European Commission to invest in federated cloud infrastructures in the period of 2021 to 2027 (European Commission, 2020). Further, we stress that the governance body needs to set and enforce rules "to encourage and to control the participation of actors" (Oliveira et al., 2019, p. 608). These rules need to be "commonly understood when it comes to articulating usage conditions on the data" (Otto, 2022, p. 13). Further, our data confirms that the governance body needs to provide data-related requirements and specify technical components, i.e., the connector each data ecosystem actor needs to operate. As a major building block of data ecosystems, the connector serves as the "interface between the internal systems […] and the [data] ecosystem itself" (Pettenpohl et al., 2022, p. 32).

The core participants need to primarily find answers to the perceived people- and organization-related challenges. Our data indicates that it is paramount that the core participants need to take action to address people- and organization-related challenges. Considering people as enablers, we endorse the argumentation by Kitsios et al. (2017) that the core participants, esp. data providers, need to inform and educate about the significance and value of shared data. Our workshop findings indicate that mentoring and community building within firms may be a viable way to compensate the lack of knowledge carriers. This finding is complementary to the idea of a network of data stewards from across data-intensive organization, fostered by the European Commission (2020). Besides, top management support needs to be ubiquitous in the data-owning and the data-using organization. In line with previous research, we "have identified top management support in an organization as a critical factor that is positively related to technology adoption" (Wang & Lo, 2020, p. 3). Even though both researchers and practitioners assume that data ecosystems will be an imperative rather than an option for the competitiveness of industrial firms (Oliveira & Lóscio, 2018; Gelhaar et al., 2023), we can show that the core participants still need to better understand the long-term undertaking of data ecosystems. Finally, our study results underpin that trust among the core participants is central to leveraging the idea of data sharing (Gelhaar et al., 2021; European Commission, 2020). Hence, they should have an intrinsic motivation to work towards a zero-trust environment aiming for data sovereignty: before data is shared among data owners and data user, the respective entities need to be authenticated, authorized, and continuously validated through set standards and mechanisms (Mehraj & Banday, 2020). From the perspective of the organization, we support the arguments by Martin et al. (2021), claiming that ecosystem thinking by all data ecosystem actors is crucial to break existing data silos. In this regard, our data points to the

need for a cultural mind shift. Further, we have indications that the core participants need to assign "a team responsible for external data sharing" (Zhiwei et al., 2021, p. 32) to ensure clarity regarding data ownership on a firm level.

Both, the governance body, and the core participants must take responsibility when it comes to value-related challenges. The interplay of actions taken by the governance body and the core participants is crucial. First, to tackle the prioritized challenge of unclear business value, both need to find the shoulder to meet the *quid pro quo* requirement: every actor who contributes to the data ecosystem, must also get something in return (Gelhaar et al., 2021). Second, we add to prior research regarding the need for joint data valuation mechanisms, e.g., considering benefit-based or combinatorial valuation approaches (Azkan et al., 2022). Our data says that a case study-driven approach by the core participants may help the governance body to get to a data valuation catalog valid for all data ecosystem actors. In line with the project of Transfer-X (Sauer & Pontes, 2023), aiming for the engagement of SMEs in data ecosystems through an easily accessible information and knowledge platform, our data adds that the governance body should take care of the SMEs and their requirements to get towards mass adoption. Otherwise, projects like Catena-X, are likely to fail, as roughly 70 percent of the targeted actors are SMEs (Catena-X, 2022a). Our data show that the governance body needs to make use cases accessible to current and future actors to further a joint value perspective among all data ecosystem actors. This would contribute towards "a shared understanding of the ecosystem's operations and goals" (Gelhaar et al., 2021, p. 131).

6 Contribution

This study contributes to the understanding of industrial data ecosystem and indicates what is needed to pave their mass adoption. Our study serves as the foundation to leverage socio-technical design know-how on data ecosystems. As our findings can be used for different abstraction levels of data ecosystem design, we argue that our work can be useful for researchers, practitioners, and even policy makers.

Taking the theoretical perspective, to the best of our knowledge, our study is one of the first that goes beyond conceptually designating data ecosystems as socio-technical systems. We adopt the logic of socio-technical system theory (Brandt & Cernetic, 1998; Baxter & Sommerville, 2011) to identify what is needed to engineer industrial data ecosystems sustainably. While researching the perceived challenges and giving examples for solution pathways based on our empirical data, we have found considerable proof that the multi-dimensional socio-technical perspective helps to analyze and manage the complexity of data ecosystems. Following Baxter and Sommerville (2011), we are convinced that the results of our socio-technical system analysis can form the basis of increased value perception regarding data ecosystems and subsequently improve the overall acceptance rate. Our structured approach to mapping prioritized challenges and selected solutions to the responsible actors across four system dimensions highlights the potential of researching data ecosystems from a socio-technical perspective in the future. We extend the predominantly conceptual knowledge on data ecosystem engineering (Oliveira et al., 2019; Otto, Steinbuss, et al., 2019) by empirically revealing

which responsibilities go along with a certain role. We also contribute empirical knowledge to the experimental insights presented by Kraemer et al. (2021). To our knowledge, our study is the first to empirically elicit which actors occupying major roles need to accept which responsibilities for stabilizing and nurturing the ecosystem.

For practitioners, we provide explanations for the absence of data ecosystems in industry. The shortlist of challenges helps to understand why industrial data ecosystems are not subject to mass adoption yet. At the same time, our study proposes needed actions on how to overcome these challenges based on an aggregation of practitioner views. By differentiating the major roles actors can occupy in a data ecosystem about the responsibilities for actions, practitioners benefit from knowing who needs to do what to make data ecosystems flourish. Our study gives support to the European data strategy (European Commission, 2020). Based on our study results, we recommend European policy makers to further support the establishment of zero-trust environments in a business context to let core participants trust each other.

7 Limitations & Outlook

The results of our study are subject to limitations that should be considered. First, the generalizability of our results should be treated with caution. The sample was drawn from mostly German firms that are either actively participating or communicating interest in data ecosystems, which may result in a biased sample as data ecosystem averse firm positions have not been considered at all. Second, with respect to the needed actions and their allocation to the major roles it is worth mentioning that not all data ecosystem roles are considered. Third, concerning the needed actions, we focused on the prioritized challenges only. Addressing non-prioritized challenges could be also beneficial for the establishment of industrial data ecosystems. In that context, the number of conducted workshops can be explained by the novelty of data ecosystems, which are still in a seminal stage (Oliveira et al., 2019). Nevertheless, due to the use of the socio-technical lens, we are convinced that our results will prove helpful for both researchers and practitioners. The constraints outlined above indicate opportunities for research. Future research can further enrich and validate our findings by including data ecosystem averse firm positions as well. Thereby, an international comparison of positions seems to valuable as well. Narrowing down the scope to a single case study may help to identify already implemented solutions and still unaddressed challenges of the ones identified. Further, we propose considering the other roles data ecosystem actors (Oliveira et al., 2019) can occupy while allocating the needed actions. To mitigate another limitation of our study, we recommend not leaving out the non-prioritized challenges elaborated by Brechtel & Petrik (2023). We suppose that each single challenge as well as their complex interrelations can be decisive in the competition between ecosystems for valuable data owners. Regarding future research, we see the chance to map suitable solution pathways for these as well.

References

Autio, E. (2022) ‘Orchestrating ecosystems: A multi-layered framework’, *Innovation*, 24(1), 96–109.

Adner, R. (2006) ‘Match Your Innovation Strategy to Your Innovation Ecosystem’, *Harvard Business Review*, 84(4), 98–107.

Adner, R. (2017) ‘Ecosystem as Structure: An Actionable Construct for Strategy’, *Journal of Management*, 43(1), 39–58.

Azkan, C., Möller, F., Meisel, L., & Otto B. (2020) ‘Service dominant logic perspective on data ecosystems - A case study based morphology’, In: Quammah, A.E. (Ed.): Proceedings of the *European Conference on Information Systems*. Marrakesh, Morocco.

Azkan, C., Gür, I., Hupperz, M., Gelhaar, J., Gieß, A., Groß, T. et al. (2022) *Anreizsysteme und Ökonomie des Data Sharings - Handlungsfelder des unternehmensübergreifenden Datenaustausches und Status quo der deutschen Wirtschaft*. Dortmund: Fraunhofer Institut für Software und Systemtechnik.

Bauer, K., Bienel, M., Drost, M., Grüner, S., Haße, H., Hill, M. et al. (2022) *Multilateral data sharing in industry – Concept using ‘collaborative condition monitoring’ as a basis for new business models*. Berlin: Federal Ministry for Economic Affairs and Climate Action.

Baxter, G. & Sommerville, I. (2011) ‘Socio-technical systems: From design methods to systems engineering’, *Interacting with Computers*, 23(1), 4–17.

Bian, W., Shang, J., & Zhang, J. (2016) ‘Two-way information sharing under supply chain competition’, *International Journal of Production Economics*, 178, 82–94.

Bostrom, R.P., & Heinen, J.S. (1977) ‘MIS Problems and Failures: A Socio-Technical Perspective. Part I: The Causes’, *MIS Quarterly*, 1(3), 17–32.

Brandt, D. & Cernetic, J. (1998) ‘Human-centred appraoches to control and information technology: European experiences’, *AI & Society*, 12(1–2), 2–20.

Brechtel, M. & Petrik, D. (2023) ‘Converging challenges: industrial data ecosystems and a vital business model’, In: Baumann, S. (Ed.); *Handbook on Digital Platforms & Ecosystems in Manufacturing*. Cheltenham: Edward Elgar Publishing (handbook forthcoming). Accepted chapter available at https://www.researchgate.net/publication/371641004_Converging_challenges_industrial_data_ecosystems_and_a_vital_business_model.

Catena-X (2022a) *Catena-X Automotive Network*, https://catena-x.net/fileadmin/user_upload/Vereinsdokumente/Catena-X_UEbersicht.pdf. Accessed: 06.03.2023

Catena-X (2022b) *Catena-X Operating Model Whitepaper*, https://catena-x.net/fileadmin/user_upload/Publikationen_und_WhitePaper_des_Vereins/CX_Operating_Model_Whitepaper_02_12_22.pdf. Accessed: 06.03.2023.

Corbin, J. & Strauss, A. (1988) *Basics of qualitative research - Techniques and procedures for developing grounded theory*. Thousand Oaks: SAGE Publications Inc.

De Prieelle, F., De Reuver, M., & Rezaei, J. (2020) ‘The Role of Ecosystem Data Governance in Adoption of Data Platforms by Internet-of-Things Data Providers: Case of Dutch Horticulture Industry’, *IEEE Transactions on Engineering Management*, 69(4), 940–950.

Devriendt, T., Borry, P. & Shabani, M. (2021) ‘Factors that influence data sharing through data sharing platforms: A qualitative study on the views and experiences of cohort holders and platform developers’, PLoS ONE., 16(7), e0254202.

Döringer, S. (2020) ‘The problem-centred expert interview - Combining qualitative interviewing approaches for investigating implicit expert knowledge’, *International Journal of Social Research Methodology*, 24(3), 265–278.

European Commission (2018) *Guidance on sharing private sector data in the European data economy*, Communications from the Commission to the European Parliament, the Council, the European economic and social Committee and the Committee of the Regions, Brussels.

European Commission (2020) *A European strategy for data*, Communications from the Commission to the European Parliament, the Council, the European economic and social Committee and the Committee of the Regions, Brussels.

Emery, F.E. & Trist, E.L. (1960) 'Socio-technical systems', In: Churchman, C.W. & Verhulst, M. (Eds.); *Management Science Models and Techniques.* 2nd ed. Oxford: Pergamon Press.

Gelhaar, J., Bergmann, N., Müller, P., & Dogan, R. (2023) 'Motives and Incentives for Data Sharing in Industrial Data Ecosystems: An Explorative Single Case Study', In: Bui, T.X. (Ed.); Proceedings of the *Hawaii International Conference on System Sciences*. Maui, Hawaii.

Gelhaar, J., Both, J., & Otto, B. (2021) 'Requirements For Incentive Mechanisms In Industrial Data Ecosystems', In: Herberger, D., & Hübner M. (Eds.); Proceedings of the *Conference on Production Systems and Logistics*. Hannover, Germany.

Gioia, D.A., Corley, K.G., & Hamilton, A.L. (2013) 'Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology', *Organizational Research Methods*, 16(1), 15–31.

Heimstädt, M., Saunderson, F., & Heath, T. (2015) 'Conceptualizing Open Data Ecosystems: A timeline analysis of Open Data development in the UK', In: Edelmann, N., & Parycek, P. (Eds.); Proceedings of the *International Conference for E-Democracy and Open Government*. Krems, Austria.

Hein, A., Schreieck, M., Riasanow, T., Setzke, D., Wiesche, M., Böhm, M., & Krcmar, H. (2019) 'Digital platform ecosystems', *Electronic Markets*, 30(1), 87-98.

Jacobides, M.G., Cennamo, C., & Gawer, A. (2020) *Distinguishing platforms and ecosystems - Complementarities, value creation, and coordination mechanisms*, https://www.jacobides.com/_files/ugd/0b15b1_7e0678d2815541bb920c77ec41e4d305.pdf. Accessed: 02.03.2023

Kitsios, F., Papachristos, N., & Kamariotou, M. (2017) 'Business Models for Open Data Ecosystem: Challenges and Motivations for Entrepreneurship and Innovation', In: Loucopoulos, P. (Ed.); Proceedings of the *IEEE Conference on Business Informatics*. Thessaloniki, Greece.

Kraemer, J., Stüdlein, N., & Zierke, O. (2021) 'Data as a Public Good: Experimental Insights on the Optimal Design of B2B Data Sharing Platforms', *SSRN Electronic Journal*. (preprint)

Martin, S., Gautier, P., Turki, S., & Kotsev, A. (2021) *Establishment of sustainable data ecosystems: Recommendations for the evolution of spatial data infrastructures*. Luxembourg: Publications Office of the European Union.

Mehraj, S. & Banday, M.T. (2020) 'Establishing a Zero Trust Strategy in Cloud Computing Environment', In: Kalaiarasi, A. (Ed.); Proceedings of the *International Conference on Computer Communication and Informatics*. Coimbatore, India.

Moore, J.F. (1993) 'Predators and prey - A new ecology of competition', *Harvard Business Review*, 71(3), 75–86.

Myers, M.D. (2013) *Qualitative research in business & management*. 2nd ed. London: SAGE Publication Inc.

Oliveira, M.I.S., Lima, G. de F., & Lóscio, B.F. (2019) 'Investigations into data ecosystems - A systematic mapping study', *Knowledge and Information Systems*, 61(2), 589–630.

Oliveira, M.I.S. & Lóscio, B.F. (2018) 'What is a data ecosystem?', In: Janssen, M., Chun, S., Weerakkody, V., Zuiderwijk, A., & Hinnant, C. (Eds.): Proceedings of the *Annual International Conference on Digital Government Research*. Delft, Netherlands.

Otto, B., Lis, D., Jürjens, J., Cirullies, J., Opriel, S., Howar, F. et al. (2019) *Data Ecosystems - Conceptual Foundations, Constituents and Recommendations for Actions*. Dortmund: Fraunhofer Institute for Software and Systems Engineering.

Otto, B., Steinbuss, S., Teuscher, A., Lohmann, S., Auer, S., Bader, S. et al. (2019) *Reference Architecture Model*. Dortmund: International Data Spaces Association.

Otto, B. (2022) 'The evolution of data spaces', In: Otto, B., Ten Hompel, M., & Wrobel, S.: *Designing Data Spaces - The Ecosystem Approach to Competitive Advantage*. Cham: Springer.

Otto, B. & Jarke, M. (2019) 'Designing a multi-sided data platform: findings from the International Data Spaces case', *Electronic Markets*, 29, 561–580.

Ørngreen, R., & Levinsen, K. T. (2017). Workshops as a Research Methodology. *Electronic Journal of E- Learning*, 15(1), 70–81.

Pettenpohl, H., Spiekermann, M., & Both, J. (2022) 'International data spaces in a nutshell', In: Otto, B., Ten Hompel, M., & Wrobel, S. (Eds.): *Designing Data Spaces - The Ecosystem Approach to Competitive Advantage*. Cham: Springer.

Pfähler, K., Baars, H., Hiller, S., Morar, D., & Petrik, D. (2022) 'Data Analytics Services for Additive Manufacturing Ecosystems', In: Davis, G., Brown, S., & Subramani, M.: Proceedings of the *Amercias Conference on Information Systems*. Minneapolis, USA.

Sarker, S., Chatterjee, S., Xiao, X., & Elbanna, A. (2019) 'The sociotechnical axis of cohesion for the IS discipline: Its historical legacy and its continued relevance', MIS Quarterly, 43(3), 695–719.

Sauer, O. & Pontes, U. (2023) *"Transfer-X" paves the way for SMEs into the data economy*, https://www.iosb.fraunhofer.de/en/press/press-releases/2023/transfer-x-transformation-medium-sized-data-economy.html. Accessed: 04.03.2023.

Schmidt, R., Lyytinen, K., Keil, M., & Cule, P. (2001) 'Identifying Software Project Risks: An International Delphi Study', *Journal of Management Information Systems*, 17(4), 5–36.

Van Der Aalst, W., Hinz, O., & Weinhardt, C. (2019) 'Big Digital Platforms: Growth, Impact, and Challenges', *Business & Information Systems Engineering*, 61(6), 645–648.

Wallace, P. (2015) *Introduction to Information Systems*. 2nd ed. Essex: Pearson Education Ltd.

Wang, H.-J., & Lo, J. (2020) 'Factors Influencing the Adoption of Open Government Data at the Firm Level', *IEEE Transactions on Engineering Management*, 67(3), 670–682.

Weber, M., Hacker, J., and Vom Brocke, J. (2021) 'Resilience in Information Systems Research-A Literature Review from a Socio-Technical and Temporal Perspective', In: Valacich, J. (Ed.); *Proceedings of the International Conference on Information Systems*. Austin, USA.

Yoo, Y., Henfridsson, O. and Lyytinen, K. (2010) 'The new organizing logic of digital innovation: An agenda for information systems research', *Information Systems Research*, 21(4), 724–735.

Zhiwei, J., Thieullent, A., Jones, S., Perhirin, V., Baerd, M., Shagrithaya, P. et al. (2021) *Data sharing masters - How smart organizations use data ecosystems to gain an unbeatable competitive edge*. New York: Capgemini Research Institute.