Motives and Incentives for Data Sharing in Industrial Data Ecosystems: An Explorative Single Case Study

Joshua Gelhaar
TU Dortmund University
Fraunhofer ISST
joshua.gelhaar@isst.fraunhofer.de

Paul Müller
Fraunhofer ISST
paul.burgwedel@isst.fraunhofer.de

Nils Bergmann
Fraunhofer ISST
nils.bergmann@isst.fraunhofer.de

Rojda Dogan
Fraunhofer ISST
rojda.dogan@isst.fraunhofer.de

Abstract

The increasing connectivity of the business world leads to economic value being created less and less by one company alone, but rather through the exchange and combination of data by various actors in so-called data ecosystems. However, many companies are not yet willing to participate in data ecosystems because they do not see the added value of their participation. This is partly because the motives of data providers do not match the incentives offered to share their data. So far, there are only very few studies that deal with this issue in detail. Therefore, we close this research gap by adopting a conceptual model to the issue of motives and incentives for data sharing and applying it to the industrial data ecosystem Catena-X in a single case study. Through the case study analysis, we can identify seven different motives and eight incentives for data sharing.

Keywords: data ecosystems, data sharing, motives, incentives

1. Introduction

The continuous digitalization of the economy leads to an increasing amount of data. As analytics technologies continue to improve, companies are finding new ways to use data to optimize business processes and develop new business models (Newell & Marabelli, 2015; Veit et al., 2014). However, many companies firstly face the challenge that neither data generation nor data processing is part of their core competencies (Azkan et al., 2021; Weill & Woerner, 2015). Secondly, meaningful data analyses based on machine learning algorithms, for example, often require high-quality data that are not generated within the company itself (Altendeitering et al., 2022; Gregory et al., 2021). These circumstances result in economic value creation taking place less and less in individual companies or traditional value chains (Aaen et al., 2021; Hein et al., 2019). Instead, the business world is becoming increasingly interconnected, with multiple data sources from different organizations being shared and jointly used in cross-sectorial, socio-technical networks referred to as data ecosystems (Geisler et al., 2021; Gelhaar, Groß, & Otto, 2021). Both researchers and practitioners assume that if companies want to reap the benefits of data sharing and remain competitive in the long term, participation in data ecosystems will be less of a choice but rather a necessity in the future (Oliveira & Lóscio, 2018; Thomas & Autio, 2015). The consulting firm Capgemini, for example, predicts that the potential financial benefits of data ecosystems participation can reach up to 9% of a company's annual revenue until 2026 (Capgemini, 2021).

Despite these developments and the resulting potential, many organizations are still hesitant to share their data across companies and to participate in data ecosystems (Kaiser et al., 2019; Prieëlle et al., 2020). One reason for this is that the general research field of data ecosystems is still in its infancy, which leads to a lack of accepted definitions and theories for e.g. the development of business models in data ecosystems (Gelhaar, Groß, & Otto, 2021; Oliveira & Lóscio, 2018). In this context, some authors point to the lack of knowledge about the concrete benefits and added values of cross-organizational data sharing as one of the main obstacles why organizations are currently not motivated to participate in data ecosystems (Azkan et al., 2020; Gelhaar & Otto, 2020; Oliveira et al., 2019). Against this background, incentives for data sharing are increasingly being discussed to motivate companies to

URI: https://hdl.handle.net/10125/103085
978-0-9981331-6-4

participate in data ecosystems, e.g. by sharing their data in them (Gelhaar, Both, & Otto, 2021; Gelhaar, Gürpinar, et al., 2021). To the best of our knowledge, there is no scientific publication to date that deals with and describes in detail the motives and incentives for data sharing in industrial data ecosystems. Based on this research gap, we formulate the following **research question**: *What are motives and incentives for data sharing in industrial data ecosystems?*

To answer the research question, we build on the conceptual model of Leimeister et al. (2009), which has already been influential in IS field, and apply it to the relationship between motives and incentives for data sharing in industrial data ecosystems. Building on this, we conduct an exploratory single case study due to the relatively weak literature on these topics. The use case under investigation is the emerging data ecosystem in the automotive industry "Catena-X". The data basis for the analysis is based on primary data (i.e., expert interviews) as well as secondary data (e.g., press releases, presentations). The rest of this paper is structured as follows: In the next section we outline the theoretical background of data sharing in industrial data ecosystems and data ecosystem governance structures. We also describe the conceptual model mentioned earlier and relevant prior work from the literature. In section 3, we outline our research approach by presenting the case study and our data collection procedures. Subsequently, the results of the case study analysis are presented in section 4. The paper ends with a discussion of the results including implications and limitations of this study.

2. Research background

2.1 Data sharing in industrial data ecosystems

Data is often referred to as the "*new raw material of the 21st century*" (Berners-Lee & Shadbolt, 2011). This analogy comes from the fact that data, similar to oil, must go through a value creation process before it can add value. This stepwise value creation process for data is also referred to as the data value chain (Attard et al., 2017). The concept of the data value chain and its phases are sometimes described differently in the literature and divided into various numbers of phases. Nonetheless, it consists at least of the three phases data collection, data interpretation and data exploitation (Gelhaar, Gürpinar, et al., 2021). In the first phase data collection, data is generated and collected (Chen et al., 2014). In this work, we focus on industrial environments where data is often generated automatically by sensors (Azkan et al., 2021). Once generated, data must be processed into information by methods of data analysis, e.g. data mining or machine learning methods (Chen et al., 2014; van den Hoven, 1999). The combination and interpretation of information finally generates added value in the form of knowledge, which can be used, for example, to optimize risks and operations or to improve business models and processes (Gillon et al., 2014). However, the creation of data-based services and products increasingly requires that various data sources from different actors must be combined, which means that the concept of the data value chain can no longer be viewed as a linear process, but rather as a dynamic process with potential loops within it (Aaen et al., 2021; Attard et al., 2017). In addition, the different phases are often carried out by a number of various actors rather than by a single organization, leading to new data-centric forms of collaboration – also referred to as data ecosystems (Lis & Otto, 2020; van den Broek & van Veenstra, 2018). Oliveira et al. (2019) define data ecosystems as "*sociotechnical complex networks in which actors interact and collaborate with each other to find, archive, publish, consume, or reuse data as well as to foster innovation, create value, and support new businesses*".

Based on this definition, a data ecosystem consists of several actors, each of which is an autonomous entity, such as a company, an institution or an individual (Oliveira & Lóscio, 2018). Depending on their motivation and capabilities, actors perform different functions and thus perform one or more roles in the data ecosystem (Azkan et al., 2020). Which roles can exist in a data ecosystem and which are essential for its creation and survival is still debated in the literature (Gelhaar & Otto, 2020; Oliveira et al., 2019). However, based on the essential phases of the data value chain described above, a data ecosystem consists of at least three roles: Data providers who perform data collection activities, analytics service providers who analyze and interpret data, and data user who exploit the data (Aaen et al., 2021; Gelhaar, Gürpinar, et al., 2021). As one actor can perform multiple roles in a data ecosystem, there can be complex interdependencies between actors, leading to situations where actors work cooperatively and competitively at the same time - also known as co-opetition (Guggenberger et al., 2020; Nalebuff & Brandenburger, 1997).

The whole process of exchanging data between the actors in a data ecosystem is often described with the umbrella term "*data sharing*" (Scaria et al., 2018). This understanding then includes both the supply and demand sides of data sharing, i.e., data providers and data users (see e.g. (Jarke et al., 2019)). In this paper, however, we want to focus on the role of data providers and their motivations and incentives to participate in data ecosystems and, consequently, to make their data available to other actors. The rationale for this focus is that data providers own the raw material and thus the

foundation for the entire data value creation process, and consequently a data ecosystem cannot emerge or grow without them and their data (Gelhaar, Gürpinar, et al., 2021; Gelhaar & Otto, 2020). Furthermore, the motivations and incentives of the different roles in a data ecosystem can be very heterogeneous (Aaen et al., 2021; Gelhaar et al., 2022; Oliveira et al., 2019), which is why we believe that a consideration of all roles is beyond the scope of this paper. With this in mind, we follow the more narrow definition of data sharing given by the European Commission (2020) in the Data Governance Act: "*'data sharing' means the provision by a data holder of data to a data user for the purpose of joint or individual use of the shared data, based on voluntary agreements, directly or through an intermediary*".

2.2 Governance of data ecosystems

A data ecosystem can be organized in different ways depending on various factors such as the distribution of power, decision-making or data access rights of the actors involved (Gelhaar, Groß, & Otto, 2021; Lis & Otto, 2021). Two distinctly different forms of organization and governance of data ecosystems mentioned in the literature are keystone player-driven data ecosystems and alliance-driven data ecosystems (Geisler et al., 2021; Otto & Jarke, 2019). In the former case, there is a central player (often referred to as a keystone or orchestrator) who usually controls the data ecosystem, e.g. via a central platform, and can be responsible for providing most of the data and for the growth and success of the ecosystem (Autio, 2021; Gelhaar & Otto, 2020; Oliveira & Lóscio, 2018). Geisler et al. (2021) refer to the B2C platforms operated by Google, Alibaba or Facebook as successful examples of keystone player-driven data ecosystems. In contrast, alliance-driven data ecosystems tend to have a decentralized, distributed organizational form and thus aim for more equitable control and sovereignty of data (Gelhaar, Groß, & Otto, 2021; Otto & Jarke, 2019). Prominent examples of alliance-driven data ecosystems that have emerged recently are the automotive industry data ecosystem "Catena-X" (Catena-X, 2021) or the "Mobility Data Space" in the mobility industry (Mobility Data Space, 2021).

Despite the current trend towards decentralization through the emergence of peer-to-peer technologies such as blockchain, alliance-driven and decentralized data ecosystems are not yet well studied (Gelhaar & Otto, 2020; Gregory et al., 2021). This is especially true for the data governance of these data ecosystems and related mechanisms of participatory interactions, and incentives to participate and influence the dynamics and development of the ecosystem (Cappiello et al., 2019). Against this background, we focus in this paper on an alliance-driven data ecosystem. Nevertheless, we assume, as do Geisler et al. (2021), that many of the motives and incentives discussed can also be relevant in differently organized data ecosystems.

2.3 Conceptual model

The Oxford English Dictionary defines an incentive as "*something, especially money, that encourages a person or organization to do something*". In many areas of life, incentives are used when it comes to encouraging people or organizations to do something in exchange for a reward (Wang et al., 2019). In the field of motivational psychology, for example, research is conducted into how incentives in the form of external factors and situations influence people's behavior (Leimeister et al., 2009). The Motive-Incentive-Activation-Behaviour Model (MIAB) presents these relationships in a simple way (Leimeister et al., 2009; Rosenstiel & Nerdinger, 2011).

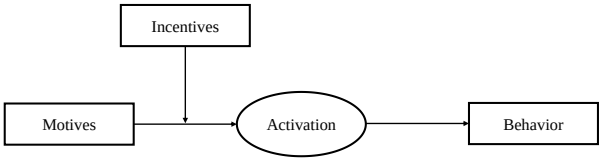

Figure 1. The Motive-Incentive-Activation-Behaviour Model (Leimeister et al., 2009; Rosenstiel & Nerdinger, 2011)

In the MIAB, a motive is generally considered to be the psychological disposition of a person. In a certain situational context, a corresponding motive is activated and then leads to a certain behavior (see Figure 1). Activation means that the individual reacts to perceived innate stimuli (e.g. an innate desire) or external stimuli (e.g. salary, social contacts, etc.) (Leimeister et al., 2009). In this study, we aim to apply the understanding of this simple model and the associated understanding of motives and incentives to the participation of data providers in data sharing in industrial data ecosystems. Reasons for this application are, firstly, that it can be argued that organizations are run by people whose behavior can in turn be explained by the MIAB (Robbins & Judge, 2013). One indicator of this is that trust between the data sharing participants is an essential success factor for a functioning ecosystem, which in turn is influenced by human behavior, among other things (Jagals et al., 2021; van den Broek & van Veenstra, 2018). Secondly, in the context of data sharing, prior publications show that it is useful to distinguish between the concepts of motives and the incentives that address those motives (Azkan et al., 2020; Gelhaar, Gürpinar, et al., 2021). Based on this, in this paper we see motives as reasons and expectations why a data provider would want to share its data in the

data ecosystem. The right incentives in this context are those that address or match data providers' motives for sharing their data.

2.4 Related work

There have already been some efforts in the academic literature to investigate motivations and incentives for data sharing (see Table 1). For example, Azkan et al. (2020) develop in their paper a morphology for data ecosystems based on a structured literature review and the analysis of two case studies and conclude, among other things, that data providers need "*prospects on their benefits to share their data in a data ecosystem environment*". However, the authors do not address what these perspectives and incentives can concretely be. Further preliminary work has been done by Gelhaar et al. (2021), who developed a taxonomy of incentive mechanisms for data sharing in data ecosystems. While the taxonomy gives a good overview of similarities and differences of incentive mechanisms for data sharing, it does not make a distinction between the shared data and the associated domain of the data ecosystem. We believe, however, that a distinction is essential in this context, since, for example, different motives and incentives prevail in the sharing of research data than in the sharing of industrial data (Gelhaar, Gürpinar, et al., 2021; Oliveira et al., 2019). Lastly, the taxonomy is based on a structured literature review and therefore does not include insights from practice. Based on a contract theory approach, Guo et al. (2018) propose the design of an incentive mechanism for data sharing, which aims to consider the interests of both data providers and data users by designing optimal contracts for the data sharing process. Nonetheless, the authors note that agreeing on an optimal contract is very difficult to achieve in reality, which is why their solution is a good theoretical model but does not take into account insights from practice. Xuan et al. (2020) develop a data sharing incentive mechanism model based on game theory using blockchain with smart contracts. However, this is again a very theoretical mathematical model that does not consider any insights from practice and is very difficult to transfer to practice. In this paper we focus on the data provider side, which is why we use the above narrow definition of data sharing. Furthermore, we use a conceptual model that distinguishes between the concepts of motives and incentives, and which is easier to apply to practice. In addition, we place a focus on industrial data, as we believe that motives and incentives can differ greatly from the domain of the data ecosystem. Finally, we want to consider insights from the practice by analyzing the alliance-driven data ecosystem Catena-X in a single case study.

3. Method

Since research on data ecosystems in general, and on the motives and incentives for data sharing in particular, is still relatively new, there are only a few theoretical and empirical findings on these topics so far. Against this background, in this study we choose a single case study approach which, as qualitative research methods in general, is particularly suitable when there is little prior knowledge about a phenomenon in practice (Darke et al., 1998; Eisenhardt & Graebner, 2007). While multi case studies are often used as a basis to generalize practical results and abstract scientific findings, there are also good reasons to conduct a single case study (Lis et al., 2022; Zrenner et al., 2019). These reasons are the possibility to study a complex phenomenon, in a more detailed and focused way, in contrast to the study of several cases (Gerring, 2006). Since there are not yet many alliance-driven data ecosystems that can be investigated and Catena-X is a large and well-known kind of these phenomena, we believe that this is an "extreme" case and that a single-case study is therefore useful and purposeful (Gerring, 2006; Zrenner et al., 2019).

3.1 Case description

Data-driven value creation has become increasingly important in the international automotive industry in the recent years. The exchange of data within supply chains is far from standardized and thus slows down industry-specific cooperation between companies.

Table 1. Overview of related publications ("+" = considered; "-" = not considered)

Source	Narrow definition of data sharing?	Distinction between motives and incentives?	Based on a conceptual model?	Focus on industrial data?	Insights from practice?
Azkan et al. (2020)	-	+	-	+	+
Gelhaar, Gürpinar, et al. (2021)	+	+	-	-	-
Guo et al. (2018)	-	-	+	+	-
Xuan et al. (2020)	+	-	+	-	-
Goal of this paper	+	+	+	+	+

The Catena-X Automotive Network e.V. was founded in 2021 with the vision of creating a consistent data exchange for all participants in the automotive value chain. Catena-X has set itself the goal of establishing a global network based on European values. Therefore, the alliance creates data value chains to connect network partners – from small and medium-sized enterprises (SMEs) to global players. The focus is on benefits and added value for each participant in the network while maintaining data sovereignty in accordance with the standards of the European Union. For this reason, Catena-X extends on the results developed by Gaia-X, which is a federation of data infrastructure based on European values (Gaia-X AISBL, 2022). Launched by a total of 17 founding members, the Catena-X association has grown to 84 members (Catena-X, 2021).

The association includes big players such as BMW AG, Deutsche Telekom AG, Robert Bosch GmbH, and SAP SE, but also SMEs. The Catena-X initiative is built on three pillars (Catena-X, 2021): First, a research and development project funded by the German government involving the founding members. Second, the non-profit association Catena-X, which is open to all interested partners. The association represents the standardization and governance body to standardize, certify and establish trust. It thus represents the alliance of the data ecosystem described above, in which decisions are made democratically. The third pillar is to become a future decentralized business environment for operating companies of the developed software and services. Within the Catena-X consortium, 10 initial use cases have been defined to demonstrate the added value of data exchange between stakeholders in the automotive industry. These use cases address different issues in the automotive industry such as sustainability or traceability of components or modular production. All use cases have in common that data must be exchanged across companies for their implementation and therefore no single company can realize them alone (Catena-X, 2021). For that reason, Catena-X has set itself the goal of connecting up to 1,000 partners to the data ecosystem by the end of 2022, which is why the topic of motives and incentives for data sharing is also an important one in Catena-X (Catena-X, 2021).

3.2 Data collection & analysis

To gain insights into practice, i.e., motives and incentives for data sharing in industrial data ecosystems, we conducted a series of seven interviews with different experts. Before conducting the interviews, we created an interview guide to ensure that all conversations covered a similar range of topics, characterizing the interview method as semi-structured (Merton & Kendall, 1946). The interview questions were created based on the research question as well as on the previously explained research background, the conceptual model, and the relevant preliminary work presented. The interviews were conducted remotely via Microsoft-Teams, recorded, transcribed, and anonymized. The selected experts come from various organizations which are all part of the Catena-X consortium. The interviews were conducted between November 2021 and February 2022 and had a duration between 45 and 70 minutes. The selection process took care to interview representatives from various stages in the automotive supply chains, i.e., OEMs, Tier-1 suppliers, Tier-2 suppliers etc. (see Table 2).

Table 2. Overview of the interviewed experts

#	Organization	Position	work experience [years]
E1	Tier-n supplier	Senior Manager Digitalization	>20
E2	OEM	Senior Project Manager Catena-X	>17
E3	Tier-1 supplier	Vice President Data & Analytics Governance	>25
E4	Tier-2 supplier	Principal Advisor Strategic Alliances and Governmental Affairs	>25
E5	Tier-1 supplier	General Manager Catena-X	>20
E6	OEM	Project Manager Catena-X	>10
E7	Tier-1 supplier	Senior Manager	>15

This selection process ensured the consideration of different views within the supply chain on the motives and incentives for data sharing. The interviewees were asked questions about, among others, why their organization decided to join Catena-X, what motives their organization has for sharing data with others, as well as what incentives they need to actually share data. For our data analysis of the interview transcripts we followed a qualitative context analysis approach and used the software MAXQDA (Krippendorff, 2013). Furthermore, to support our findings in this work, we follow the recommendation of (Pratt, 2008) and cite power and proof quotes in the presentation of our results, i.e. quotes that strongly underline the point we are trying to make. To further improve the data quality and increase the explanation of our research problem through triangulation we combined our primary data from the interviews with secondary data, i.e., newspaper articles, press releases and video presentations. This allowed us to validate the plausibility of our

observations and insights from the expert interviews (van Dyck et al., 2021). In total we analyzed 25 news articles, 42 press releases, 4 blog articles and 5 presentation videos.

4. Results

Our data allowed us to analyze what motives and incentives data providers have to share their data in the Catena-X data ecosystem. In the following, we first describe general observations we made regarding motives and incentives for data sharing in our data analysis. Subsequently, we divided the identified motives and incentives, following Gelhaar, Gürpinar, et al. (2021), into the three categories economic motives and incentives, legal motives and incentives and social motives and incentives (see Table 3).

4.1 General observations

In our analysis, we first made some general observations about motives and incentives of data sharing, which could not be assigned to a category of motives but are nevertheless relevant for the analysis.

First, we found, both in the evaluation of the interviews and in the analysis of the empirical data, that the most added value of data sharing from data providers lies in the individual Catena-X use cases and not in the participation in the data ecosystem itself. Expert 4 said, for example, "*Well, I think the added value in such an ecosystem comes from the use cases that are possible in the ecosystem.*" Based on this observation, we have identified that, especially at the beginning of Catena-X, the focus is on the development of data-based services that can be used by as many companies as possible in the supply chains of the automotive industry and thus offer them added value. These are, for example, services that enable data exchange via shared digital twins and thus simplify the demand and capacity management along the entire supply chain. In this context, it is interesting to note that there will be services open source for all participants free of charge as well as proprietary solutions from software providers for which a fee is charged. By making essential services open source for all Catena-X participants, as many companies as possible are given incentives in the form of services to offer their data in the data ecosystem.

Based on these findings, it also became clear that Catena-X does not currently envisage that there will be providers of raw data. Instead, in most cases, the data is pre-processed or anonymized so that it can be used by certain services. Consequently, the data provider generally receives a data-based service in return for its data. Expert 1 said about this "*...data can only be used to have added value in an application. I don't know of any discussion yet where raw data is sold.*" In addition, most of the current participants in Catena-X do not want to be solely providers of data in the future, but rather want to help shape the ecosystem and the services that are created there. Expert 2 expressed this, for example, as follows: "*We clearly want to design, use our network, and use this transfer to get involved. We believe that if we were to position ourselves just as a data consumer, Catena X would be a bit weaker.*"

Lastly, we noted that there is an "unwritten law" in Catena-X that everyone is expected to share data. Expert 5 said, for example: "Since it always revolves around data, the principle always applies that what you demand in terms of data you are also prepared to give. There may be companies that produce less data than others, but I think it's a theoretical problem that you only benefit." For us, this shows that the alliance of the data ecosystem defines certain governance rules, which state that a certain willingness to share data is expected, or a participant cannot take part in the data ecosystem.

Table 3. Overview of motives and incentives for data sharing in industrial data ecosystems

Category	Motive	Incentive
Economic	Cost savings	Data-based services that lead to the improvement of internal processes
	Transparency	Data on the origin and authenticity of components and materials
	Differentiation	Data for new data-based business models
		Competitive advantage through information exchange and learning effects
	Strategic partnerships	Measures to enable and foster cross-company data sharing
	Direct compensation	Money payment for pre-processed data
Legal	Compliance	Data-based services that help to comply with legal requirements
Social	Customer demands for sustainability	Data-based service which can calculate the carbon dioxide footprint

4.2 Economic motives and incentives

One of the most frequently mentioned economic motives for data sharing in Catena-X is the possibility of **cost savings**. In Catena-X data providers are motivated to share their data and in return receive data-based services that lead to the improvement of internal processes and thus to cost savings. Examples of this are services that enable permanent detection and immediate feedback of abnormalities in components based on field, production, and quality data. The use of these analysis results can help to react early to problems in the production chain starting from the n-tier supplier to the OEM and can thus prevent the costs of a possible recall action. Expert 2 said, for example, "*If we can realize some of our use cases, such as Circular Economy, then we can save an incredible amount of money, e.g., in material costs*".

Another motive we identified in the data analysis is the motive of **transparency**. In this regard, Catena-X has formulated the goal of establishing connected data chains along the entire value chain of the automotive industry. This can only be made possible if every participant in the value chain shares their corresponding data. Consequently, the premise is that data providers provide needed data and get back the data they need in return. While increased transparency supports other processes, such as the aforementioned recall activities, and thus overlaps with the cost-saving motive, data providers have further incentives for transparency. On the one hand, each data provider is expected to provide data on the origin of its components and raw materials, thus providing data on the authenticity of the components. On the other hand, this data, and the transparency it creates should help to prevent counterfeiting. In addition, the end-to-end data chains created in this way are expected to help make the entire supply chain more transparent for the individual participants to increase the resilience of the entire supply chain. Expert 7 said the following on this: "*I would say in a simplified way: data exchange provides transparency in many cases and transparency can lead to more efficiency in the second step but can also have other advantages.*"

A further economic motive we identified is **differentiation**. In this motive, data providers hope to gain a competitive advantage over their competitors by sharing their data in the data ecosystem. This motive can be addressed by two different incentives. On the one hand, data providers can be offered other data in exchange for their shared data. This data can then be used by the data provider, for example, to build their own data-based business models on it. Expert 4 said: "*Another point is that the ecosystem can be used as another sales channel. So, we can better reach our existing customers and attract additional customers via new digital services.*" On the other hand, the data analysis shows that the active participation of data providers in the Catena-X industrial data ecosystem is already seen as an opportunity for differentiation. This is explained by the fact that, as written above, every Catena-X participant is expected to share data. This implies that every participant is forced to a certain extent to participate in the development of standards for e.g., data formats and data sharing technologies. The incentive for data providers here is the differentiation through possible learning effects in the development of the required data sharing standards. Expert 1 explains this as follows: "*That is why it is interesting for us as a supplier to be able to differentiate ourselves on the one hand and on the other hand also to develop new standards or new protocols for data exchange*".

Another motive that emerged in our data analysis was that of new **strategic partnerships**. In this dimension, data providers are motivated by their participation in Catena-X, that is, by sharing their data, to form new strategic partnership. In practice, this can happen, for example, when a data provider offers its data on a data marketplace and this data is requested by a data user with whom the data provider has not previously shared any data. This transaction can then lead to further business relationships and partnerships. This motive is addressed in Catena-X by implementing both technical and organizational measures in the data ecosystem that ensure transparency and interoperability between the different ecosystem participants. One example of this in Catena-X is the joint development of standards for the description of metadata. This enables as many data providers as possible to describe and offer their data in a standardized way so that it can be viewed by many possible data users on a data marketplace, for example. Expert 5 commented: „*...small and medium-sized companies can use Catena-X to put themselves in the „shop window ": „This is who we are, this is what we can do, these are our capacities". ...Catena-X will lead to new supply chains being formed. The earlier you participate, the better it is...*". The last identified economic motive is that of **direct compensation**. This motive is addressed by the incentive of direct monetary payment to the data provider for its shared data. As described in Section 4.1, due to privacy concerns, in most cases this will be the sale of pre-processed data rather than raw data. Expert 6 explained: "*We also think about the fact that a role in the system can be a data seller. However, I see data selling more as an "ultima ratio". We should try to close the data chains via use cases so that everyone can benefit from it*." We conclude that although direct payment for data provided is possible, it is rare due to various challenges, such as the economic valuation of the data.

4.3 Legal motives and incentives

The only motive from the legal category that we identified in our data analysis is that of compliance. The background to this motive is the increasing number of legal directives and laws for the automotive industry, such as the Supply Chain Act, which require data to be shared between supply chain actors. Catena-X wants to address this by jointly developing data-based services that help to implement these guidelines in a legally compliant manner. To use this service, users are expected to share the relevant data required by other actors to comply with the regulations. Expert 3 expressed this as follows: "*In Catena-X, regulation also plays a big role. We must do certain things, like comply with the Supply Chain Act, and that requires data exchange...*".

4.4 Social motives and incentives

The social motive for data sharing that we identified in our data analysis is that of increasing **customer demands for sustainability**. This can be seen, for example, in the fact that more and more customers want to know how much carbon dioxide was generated in the production of their car, and ultimately the carbon footprint of a car increasingly influences their purchase decision. Consequently, to address this demand, OEMs need to calculate the carbon dioxide footprint across the entire supply chain of a car. On the one hand, this again requires that all supply chain participants involved provide the relevant data. On the other hand, it requires data-based services that process this data and calculate the carbon dioxide produced. These services are jointly developed in Catena-X and made available to the actors, in particular to support SMEs in the supply chain that do not have the capabilities to develop these services themselves. Expert 2 said, for example: "*In Catena-X we try to address big questions, especially in the area of sustainability. For example, how can we prove the reduction of the CO2 footprint or how can we increase the share of recycled material?*".

5. Discussion and Conclusion

This paper was motivated by the research question of what motives and incentives data providers must share their data in industrial data ecosystems. Through a single case study investigation of the Catena-X industrial data ecosystem, we were able to identify 7 motives and 8 incentives for data sharing. The following implications for theory and practice can arise from these findings: First, our work makes **scientific contributions** by enhancing the general understanding of the emerging and still unexplored research field of industrial data ecosystems. Particularly as the number of (individual) case studies on industrial data ecosystems in the literature is still quite limited so far. Against this background, we believe that our study results extend the existing knowledge about motives and incentives for data sharing in industrial data ecosystems and thus contribute to a better general understanding of these complex issues. Furthermore, other researchers can use our results as a basis for the development of management and engineering methods for industrial data ecosystems, which are still missing in the scientific literature (Oliveira et al., 2019). Specifically, the study results can be used, for example, to develop design principles for incentive mechanisms as part of a data governance framework for data ecosystems (Möller et al., 2020). Secondly, our study results also offer diverse **practical contributions**. On the one hand, the results can help companies to better understand the data ecosystems in which they are already active in general and the incentive mechanisms that prevail there in particular. This enables practitioners to actively shape the data ecosystems and incentive mechanisms to their own advantage and to derive greater and more individual benefit from sharing their data. In addition, data users can better understand the motives of potential data providers and design appropriate incentives for data sharing. On the other hand, the study results can be used by emerging data ecosystems, especially those that are organized in an alliance-driven way, to design the data ecosystem and its incentives in such a way that as many data providers as possible are motivated to share their data in it.

However, there are naturally some **limitations** to consider when interpreting our results. First, general limitations lie in the nature of single case studies, i.e., in the limited transferability of results due to the focus on the case-specific situation of Catena-X. Therefore, the present individual case study should not be seen as a quantifiable or statistical evaluation of companies, but rather as a study with an exploratory character (Lis et al., 2022). Second, the application of the Motive-Incentive-Activation-Behaviour Model allows for a simple representation of the relationships between motives and incentives. Nevertheless, this simplicity limits the model's ability to generate complex or measurable outcomes. Finally, it should be considered that Catena-X is still under development and therefore some aspects around the topic of incentives for data sharing have not yet been finally clarified or will change over time. For example, we could not identify any cultural motivating factors for data sharing which according to (Gelhaar, Gürpinar, et al., 2021) may exist in other data ecosystems. However, these limitations also open **future research** directions. First, it would be

interesting to investigate whether a ranking can be established between the identified motives and their incentives. Secondly, it could be useful to explore how our findings can be transferred to data ecosystems in other sectors. For example, the study findings could be tested and validated through the use of multiple case studies or cross-sector surveys (van Dyck et al., 2021). Lastly, it would be of interest to see how the motives and corresponding incentives for data sharing may change over time, as Catena-X is still in its emergence phase and a data ecosystem has different life cycles with different characteristics.

Acknowledgment

This research was supported by the German Federal Ministry of Education and Research (BMBF) in the context of the IEDS project (no. IEDS001) and by the Fraunhofer Cluster of Excellence "Cognitive Internet Technologies". The authors are responsible for the content of this publication.

References

Aaen, J., Nielsen, J. A., & Carugati, A. (2021). The dark side of data ecosystems: A longitudinal study of the DAMD project. *European Journal of Information Systems*, 1–25.

Altendeitering, M., Dübler, S., & Guggenberger, T. (2022). Data Quality in Data Ecosystems. In *Twenty-eighth Americas Conference on Information Systems*.

Attard, J., Orlandi, F., & Auer, S. (2017). Exploiting the Value of Data through Data Value Networks. In *Proceedings of the 10th International Conference on Theory and Practice of Electronic Governance - ICEGOV '17*.

Autio, E. (2021). Orchestrating ecosystems: a multi-layered framework. *Innovation: Organization & Management*, 1–14.

Azkan, C., Iggena, L., Möller, F., & Otto, B. (2021). Towards Design Principles for Data-Driven Services in Industrial Environments. In *Proceedings of the 54th Hawaii International Conference on System Sciences*.

Azkan, C., Möller, F., Meisel, L., & Otto, B. (2020). Service Dominant Logic Perspective on Data Ecosystems - A Case Study based Morphology. In *Proceedings of the 28th European Conference on Information Systems*.

Berners-Lee, T., & Shadbolt, N. (2011). *There's gold to be mined from all our data*. https://www.thetimes.co.uk/article/theres-gold-to-be-mined-from-all-our-data-s3qffdkz7kq

Capgemini. (2021). *Data Sharing Masters: How smart organizations use data ecosystems to gain an unbeatable competitive edge.*

Cappiello, C., Gal, A., Jarke, M., & Rehof, J. (2019). Data Ecosystems: Sovereign Data Exchange among Organizations: Report from Dagstuhl Seminar 19391. *Dagstuhl Reports*, *9*(9), 66–134.

Catena-X. (2021). *Catena-X Automotive Network: Homepage*. https://catena-x.net/; accessed 5-5-2022

Chen, M., Mao, S., & Liu, Y. (2014). Big Data: A Survey. *Mobile Networks and Applications*, *19*(2), 171–209.

Darke, P., Shanks, G., & Broadbent, M. (1998). Successfully completing case study research: combining rigour, relevance and pragmatism. *Information Systems Journal*, *8*(4), 273–289.

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory Building From Cases: Opportunities And Challenges. *Academy of Management Journal*, *50*(1), 25–32.

European Commission. (2020). *Proposal for a Regulation of the European Parliament and of the council on European data governance (Data Governance Act)* (COM/2020/767 final).

Gaia-X AISBL. (2022). *Gaia-X Website*. https://www.gaia-x.eu/; accessed 6-1-2022

Geisler, S., Vidal, M.-E., Cappiello, C., Lóscio, B. F., Gal, A., Jarke, M., Lenzerini, M., Missier, P., Otto, B., Paja, E., Pernici, B., & Rehof, J. (2021). Knowledge-driven Data Ecosystems Towards Data Transparency. *Journal of Data and Information Quality (JDIQ)*, *14*(1).

Gelhaar, J., Becker, F., & Groß, T. (2022). Characterization of Relationships in Data Ecosystems. In *Proceedings of the Conference on Production Systems and Logistics: CPSL 2022.*

Gelhaar, J., Both, J., & Otto, B. (2021). Requirements For Incentive Mechanisms In Industrial Data Ecosystems. In *Proceedings of the Conference on Production Systems and Logistics: CPSL 2021*.

Gelhaar, J., Groß, T., & Otto, B. (2021). A Taxonomy for Data Ecosystems. In *Proceedings of the 54th Hawaii International Conference on System Sciences*.

Gelhaar, J., Gürpinar, T., Henke, M., & Otto, B. (2021). Towards a Taxonomy of Incentive Mechanisms for Data Sharing in Data Ecosystems. In *PACIS 2021 Proceedings*.

Gelhaar, J., & Otto, B. (2020). Challenges in the Emergence of Data Ecosystems. In *Proceedings of the 24th Pacific Asia Conference on Information Systems*.

Gerring, J. (2006). *Case study research: Principles and practices*. Cambridge University Press.

Gillon, K., Aral, S., Lin, C.-Y., Mithas, S., & Zozulia, M. (2014). Business Analytics: Radical Shift or Incremental Change? *Communications of the Association for Information Systems*, *34*.

Gregory, R. W., Henfridsson, O., Kaganer, E., & Kyriakou, H. (2021). Data Network Effects: Key Conditions, Shared Data, and the Data Value Duality. *Academy of Management Review*.

Guggenberger, T. M., Möller, F., Haarhaus, T., Gür, I., & Otto, B. (2020). Ecosystem Types in Information Systems. In *Proceedings of the 28th European Conference on Information Systems*.

Guo, B., Deng, X., Guan, Q., Tian, J., & Zheng, X. (2018). An Incentive Mechanism for Cross-Organization Data

Sharing Based on Data Competitiveness. *IEEE Access*, *6*, 72836–72844.

Hein, A., Weking, J., Schreieck, M., Wiesche, M., Böhm, M., & Krcmar, H. (2019). Value co-creation practices in business-to-business platform ecosystems. *Electronic Markets*, *29*(3), 503–518.

Jagals, M., Karger, E., & Ahlemann, F. (2021). Is Trust Shapeable? Design Requirements for Governing Sharing Networks. In *Proceedings of the 16th International Conference on Design Science Research in Information Systems and Technology*.

Jarke, M., Otto, B., & Ram, S. (2019). Data Sovereignty and Data Space Ecosystems. *Business & Information Systems Engineering*, *61*(5), 549–550.

Kaiser, C., Stocker, A., & Fellmann, M. (2019). Understanding Data-driven Service Ecosystems in the Automotive Domain. In *AMCIS*.

Krippendorff, K. (2013). *Content analysis: An introduction to its methodology* (3rd ed.). Sage.

Leimeister, J. M., Huber, M., Bretschneider, U., & Krcmar, H. (2009). Leveraging Crowdsourcing: Activation-Supporting Components for IT-Based Ideas Competition. *Journal of Management Information Systems*, *26*(1), 197–224.

Lis, D., Arbter, M., Spindler, M., & Otto, B. (2022). An Investigation of Antecedents for Data Governance Adoption in the Rail Industry—Findings From a Case Study at Thales. *IEEE Transactions on Engineering Management*, 1–18.

Lis, D., & Otto, B. (2020). Data Governance in Data Ecosystems - Insights from Organizations. In *AMCIS 2020 Proceedings*.

Lis, D., & Otto, B. (2021). Towards a Taxonomy of Ecosystem Data Governance. In *Proceedings of the 54th Hawaii International Conference on System Sciences*.

Merton, R. K., & Kendall, P. L. (1946). The Focused Interview. *American Journal of Sociology*, *51*(6), 541–557.

Mobility Data Space. (2021). *Homepage*. https://mobility-dataspace.eu/; accessed 5-7-2022

Möller, F., Guggenberger, T., & Otto, B. (2020). Towards a Method for Design Principle Development in Information Systems. In *Proceedings of the 15th International Conference on Design Science Research in Information Systems and Technology*.

Nalebuff, B. J., & Brandenburger, A. M. (1997). Co-opetition: Competitive and cooperative business strategies for the digital economy. *Strategy & Leadership*, *25*(6), 28–33.

Newell, S., & Marabelli, M. (2015). Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of 'datification'. *The Journal of Strategic Information Systems*, *24*(1), 3–14.

Oliveira, M. I., Barros Lima, G. d. F., & Lóscio, B. F. (2019). Investigations into Data Ecosystems: a systematic mapping study. *Knowledge and Information Systems* (Survey Paper).

Oliveira, M. I., & Lóscio, B. F. (2018). What is a data ecosystem? In *Proceedings of the 19th Annual International Conference on Digital Government Research*, Delft, Netherlands.

Otto, B., & Jarke, M. (2019). Designing a multi-sided data platform: findings from the International Data Spaces case. *Electronic Markets*, *29*(4), 561–580.

Pratt, M. G. (2008). Fitting Oval Pegs Into Round Holes. *Organizational Research Methods*, *11*(3), 481–509.

Prieëlle, F. de, Reuver, M. de, & 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*, 1–11.

Robbins, S. P., & Judge, T. A. (2013). *Organizational Behavior* (15th ed.). Pearson.

Rosenstiel, L. von, & Nerdinger, F. W. (2011). *Grundlagen der Organisationspsychologie: Basiswissen und Anwendungshinweise [Basics of Organizational Psychology]* (7th ed.). Schäffer-Poeschel.

Scaria, E., Berghmans, A., Pont, M., Arnaut, C., & Leconte, S. (2018). *Study on data sharing between companies in Europe*. Luxembourg.

Thomas, L. D. W., & Autio, E. (2015). The processes of ecosystem emergence. *Academy of Management Proceedings*.

van den Broek, T., & van Veenstra, A. F. (2018). Governance of big data collaborations: How to balance regulatory compliance and disruptive innovation. *Technological Forecasting and Social Change*, *129*, 330–338.

van den Hoven, J. (1999). Information Resource Management: Stewards of Data. *Information Systems Management*, *16*(1), 88–90.

van Dyck, M., Lüttgens, D., Piller, F., & Diener, K. (2021). Positioning Strategies in Emerging Industrial Ecosystems for Industry 4.0: A Longitudinal Study of Platform Emergence in the Agricultural Industry. In *Proceedings of the 54th Hawaii International Conference on System Sciences*.

Veit, D., Clemons, E., Benlian, A., Buxmann, P., Hess, T., Kundisch, D., Leimeister, J. M., Loos, P., & Spann, M. (2014). Business Models: An Information Systems Research Agenda. *Business & Information Systems Engineering*, *6*(1), 45–53.

Wang, H., He, D., Yu, J., & Wang, Z. (2019). Incentive and Unconditionally Anonymous Identity-Based Public Provable Data Possession. *IEEE Transactions on Services Computing*, *12*(5), 824–835.

Weill, P., & Woerner, S. L. (2015). Thriving in an Increasingly Digital Ecosystem. *MIT Sloan Management Review*, *56*(4), 27–34.

Xuan, S., Zheng, L., Chung, I., Wang, W., Man, D., Du, X., Yang, W., & Guizani, M. (2020). An incentive mechanism for data sharing based on blockchain with smart contracts. *Computers & Electrical Engineering*, *83*.

Zrenner, J., Möller, F. O., Jung, C., Eitel, A., & Otto, B. (2019). Usage control architecture options for data sovereignty in business ecosystems. *Journal of Enterprise Information Management*, *32*(3), 477–495.