3rd Conference on Production Systems and Logistics

Characterization of Relationships in Data Ecosystems

Joshua Gelhaar[1], Felix Becker[1], Tobias Groß[1]
[1]*Fraunhofer ISST, Dortmund, Germany*

Abstract

The importance of data as a strategic resource for the development of innovation is steadily growing. Data-driven value creation increasingly requires cross-company collaboration between various actors with different roles in so-called data ecosystems. So far, however, the existing knowledge in the research field around data ecosystems is still relatively limited. In particular, the relationships and interdependencies between the different actors in a data ecosystem are not well understood yet. To address this research gap, we conduct a structured literature review and interview eleven experts from practice to identify characteristics of relationships between actors in data ecosystems. Among other things, the results show that both tangible characteristics, such as a clear exchange of values, and intangible characteristics, such as trust, are distinguishing features of the relationships between actors in data ecosystems. These study results can serve as a tool for both researchers and practitioners to better understand data ecosystems in general and the relationships and interactions that occur within them.

Keywords

Characterization; Relationships; Data Ecosystems; Data Sharing; Literature Review; Expert Interviews

1. Introduction

The ongoing digitization of the economy and society leads to large amounts of data. Advancing data analytics techniques are harnessing these data volumes and driving changes in existing businesses as well as the emergence of novel business opportunities [1]. At the same time, this development leads to internal company data increasingly being used externally and vice versa, and consequently to organisational boundaries being broken down [2]. These circumstances result in data-driven innovation and economic value creation being less and less created by individual organizations or in traditional value chains [3]. Instead, today's business world is becoming increasingly interconnected, combining various data sources from different organizations in cross-industry, sociotechnical networks – so-called data ecosystems [4]. Both researchers and practitioners argue that now and in the future, ecosystem participation is not a choice but a necessity for companies to take advantage of data sharing and remain competitive in the long term [5,4]. Despite that, many companies still refuse to share their data across companies and thus cannot use the potentials of data ecosystems for their benefits [6]. Especially in traditional sectors such as production and logistics, data is still comparatively rarely shared beyond company boundaries [7]. One rationale for this is seen in the lack of studies and the consequent absence of generally accepted theories and models for data ecosystems [8,4]. [9] conclude in their systematic review of the data ecosystem literature that more research is needed regarding relationships of actors and their characteristics in data ecosystems for theory development. This can be reasoned by the fact that the relationships between the actors create and span the ecosystem and ultimately determine how it functions [10]. In addition, understanding the types of relationships and interactions between the actors is an important step in building and developing a data ecosystem [11]. A better understanding of the actor's roles

and their interrelationships can help to describe data ecosystems more precisely and formally and can serve as a basis for the development of a meta-model or an ontology, for example [12]. To the best of our knowledge, there is no scientific publication yet that deals with the detailed description of the characteristics of relationships in data ecosystems and addresses the research gap mentioned above. To contribute to a deeper understanding of the evolving research field around data ecosystems, we want to address this research gap and answer the following research question in this paper:

Research Question: *How can the relationships of actors in data ecosystems be described and characterized?*

To answer this research question, we develop an overview and description of the characteristics of relationships in data ecosystems. The results are grounded both in the scientific literature, through the conduct of a structured literature review, and in practice, with the help of expert interviews.

The remainder of this paper is structured as follows: After the introduction, we outline the theoretical background of data ecosystems, their roles and relationships, and draw a distinction to existing preliminary work. In Section 3, we outline our research approach and data collection processes. The characteristics of relationships in data ecosystems that we identified are described in section 4. The paper concludes with a discussion of the results and implications for research and practice.

2. Research background

2.1 Data ecosystems

The term ecosystem was coined by the biologist Arthur Tansley [13] who proposed a concept to describe the interactions between organisms of different species and their environment as an integrated system [14,15]. Based on this definition, various strands of research have since emerged in which the principles of the biological ecosystem concept are applied to other domains. One of the most popular concepts is the one of business ecosystems which were popularized by James Moore [16]. He uses this concept to describe interacting organizations as an "*economic community*" that aims to produce innovative products and services for customers, which are also part of the business ecosystem [16]. The increasing penetration of digital technologies within the business world has led to the analogy of digital ecosystems, which is seen as a "*digital version*" of business ecosystems [2]. The focus of this study lies on data ecosystems, which can be seen as a special type of digital ecosystems [17]. Following other ecosystem concepts, data ecosystems consist of diverse interactions between multiple actors that contribute to the creation and manipulation of a resource – which in this case are data – through joint activities [4]. On that basis, we see the focus of data ecosystems in the cross-actor generation, processing, sharing, and use of data with the goal to create added value for all actors involved [18,9].

Data ecosystems have certain characteristics that distinguish them from other forms of inter-organizational cooperation, such as traditional value chains or networks [19,15]. One characteristic is the lack of clear ecosystem boundaries, which can lead to varying degrees of interdependencies and relationships among the participating actors and ultimately to a heterogeneous and changing set of members [10]. Another characteristic is referred to as "*co-evolution*" [10]. It describes the condition that the development of one actor can positively affect the development of the other actors, resulting in benefits for all involved [15]. This is also because the ecosystem actors can have cooperative and competitive relationships at the same time – also known as coopetition [15,20].

2.2 Roles and relationships in data ecosystems

Based on the definition above, a data ecosystem consists of multiple actors. An actor is an autonomous entity, such as a company, an institution, or an individual person [4]. Depending on various factors, e.g., the

motivation and capabilities of the actors, they perform different functions in the data ecosystem. A function or activity performed separately in this way is called a role within the ecosystem [18]. An individual actor may in turn perform one or more of these roles in a data ecosystem [21,12]. Exactly which roles can exist in a data ecosystem and which roles are essential is still debated in the literature [9]. However, there is a general agreement that there need to be at least three roles in a data ecosystem [22,23]. First, this is the role of the data provider who is responsible for the generation and collection of data [18,9]. Second, it requires a role that analyses and interprets data which is called analysis service provider [24]. Last, the information gained through data analysis is used by the data user role to generate value from it [25]. In addition to these three roles, the role of the so-called "keystone" actor is also frequently mentioned in the literature [9]. In some data ecosystems, this role may be responsible for providing most of the data as well as promoting the ecosystem, and thus may be instrumental in the ecosystem's growth and success [11,26]. Nevertheless, there are also data ecosystems that have a rather decentralized, distributed organizational form and thus operate without a central actor [8,27]. Instead, these ecosystems are held together by their common goal of shared value creation [11,9]. Since there is already some basic understanding about the roles in data ecosystems and several papers already exist on this topic (see e.g. [4], [9], or [18]), we do not focus on the detailed description of roles in data ecosystems in this paper. Instead, we concentrate on describing and characterizing the relationships between the data ecosystem actors. The reason for this is that the individual actors in a data ecosystem do not generate added value on their own. Rather, the added value arises through the interactions and relationships among each other, such as the exchange and sharing of data [11,12]. Building on [4], we see a relationship in a data ecosystem as an interaction between two data ecosystem actors, which is influenced by their roles and characterized by certain attributes. The detailed description of these characteristics is the goal of this paper.

In the scientific literature on data ecosystems, there is little prior work that has explicitly addressed the relationships and interactions between the different actors of a data ecosystem. Noteworthy in this context, however, are the works of [12], [8], and [21]. Based on a literature review, [12] developed a meta-model describing the basic concepts of data ecosystems and their relationships to each other. However, the authors only identify the essential elements such as actors, roles, relationships, and resources of data ecosystems and do not go into further detail or elaborate on characteristics of the individual elements such as the relationships. The study by [8], who developed a taxonomy for data ecosystems, has a similar focus. The taxonomy contains the essential characteristics and dimensions of data ecosystems but does not include specific characteristics of actor relationships. [21], conversely, examine in more detail the relationships and interactions in what they call data exchange ecosystems. However, the authors base their analysis on graph theory, which leads them to describe only structural features of relationships in data ecosystems, such as the number of relationships. We believe, by contrast, that this does not take into account all the characteristics of relationships, such as trust as an informal characteristic [27].

3. Research approach

To answer the research question formulated above, we aim to describe the characteristics of relationships in data ecosystems in a structured and concise way. For this, we conducted a structured literature review that is based on standardized and accepted guidelines in our research field (see section 3.1). To incorporate insights from practical reality into the research findings of this paper, we also conducted a series of eleven expert interviews (see section 3.2). In this way, we were able, on the one hand, to define and describe an initial number of characteristics through the literature analysis. On the other hand, we were able to validate and specify the characteristics identified in the expert interviews on the basis of the literature.

3.1 Literature review

We conducted a structured literature review following the approach described by [28] and the guidelines by [29]. Based on the research question formulated above, we chose Scopus as our scientific literature database because it contains more than 25,100 titles from more than 5,000 international publishers and thus promises very good results for our field of interest as it indexes the most relevant journals and conference proceedings [30]. We used the search string "data ecosystem" OR "data-driven ecosystem" OR "data-based ecosystem" as these terms are used synonymously by some authors [8]. The results were limited to English-language literature and peer-reviewed only. As a result, we received 457 as the initial set of publications. Within this first set, we reviewed the titles, abstracts, and keywords of the articles to check whether a hit fit the research scope. If the content of an article remained unclear, we examined the whole paper. We eliminated publications that did not fit our research scope or had no relevance with data ecosystems and the relationships between their roles. For example, we excluded papers dealing with the so-called Big Data Ecosystem (see e.g. [31]) as these often describe the software ecosystem around Hadoop and therefore have a different focus than the data ecosystems we intend to study. Eventually, this resulted in 64 relevant papers. In a second iteration, we complemented the literature set with a forward- and backward-search, as suggested by [28], to identify additional relevant publications. This led to the consideration of 12 additional useful articles. Consequently, a total of 76 articles were considered for the literature review. Figure 1 provides an overview of the literature search process. We analysed the publications thoroughly to extract scientific insights about the key features and characteristics of relationships between actors in data ecosystems from the body of literature.

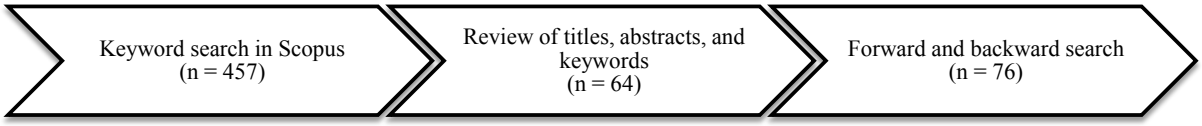

Figure 1: The structured literature review process

3.2 Expert interviews

To gain insights into the real-world environment, i.e., the relationships between actors within data ecosystems, we conducted a series of eleven 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 interviews as semi-structured [32]. All interviews were conducted remotely via Internet communication tools, recorded, transcribed, and anonymized. The analysis of the transcripts followed a qualitative context analysis [33]. To support our findings in this work, we follow the recommendation of [34] 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.

The selected interviewees come from various industries and functions such as management, project and operations managers, and consultants (see Table 1). The selection process took care to interview representatives from organizations that actively take one or more roles in a data ecosystem as well as representatives from organizations that advise other companies on how to engage in data ecosystems. This ensured that both an inside view and an outside view of the characteristics of relationships between roles in data ecosystems were considered. The interviewees were asked questions about the general understanding of data ecosystems, the relationships and interdependencies between the participating ecosystem roles, as well as the potentials and challenges that can arise from these relationships. These questions were particularly concerned with understanding how and why organizations establish relationships with other actors in data ecosystems.

Table 1: Overview of the interviewed experts

#	Organization / Industry	Position	Duration [min:ss]
E1	Industry association	IT Specialist	42:15
E2	Conglomerate	Senior Director	47:39
E3	Automotive	IT Manager	32:24
E4	Consulting	Consultant	27:21
E5	Consulting	Consultant	26:53
E6	IT Consulting	Senior Consultant	46:25
E7	Industry association	Innovation Advisor	39:13
E8	Conglomerate	Senior Principal Engineer	40:21
E9	Manufacturing	Principal Consultant	36:12
E10	Academy of sciences	Scientific Officer	33:37
E11	Manufacturing	Board Member	44:28

4. Results

In this section, we present the results of our methodological approach. Overall, we were able to identify the six characteristics Value Stream, Co-Creation, Interdependence, Trust, Intention, and Relevance. These characteristics are explained in detail in the following:

Value Stream – Every relationship in a data ecosystem is based on a value stream. By definition, within data ecosystems, these value streams focus on the exchange and sharing of data. However, both the literature and the expert interviews made it clear that data providers do not necessarily have to receive data in return for their data offering (see e.g. [22]). Instead, data providers can be compensated for their data in various ways. Expert 10 put it as follows: "*And in this data ecosystem, the people interact in some way and then have some kind of interaction with each other. Most of the time, the data goes from one person to another, who perhaps pays money for it or can offer a service for it…*". However, the simplest form of compensation is money, as money is easy to handle and applicable in many situations [35]. At the same time paying money implies the challenge of valuing data in monetary terms [36]. This is one of the reasons why data providers often receive a data-driven service in return for their data instead of money [18]. An example that was mentioned by some experts was the use case in which a machine operator makes his data available to a service provider and receives a predictive maintenance service in return. Consequently, the value stream between two actors in a data ecosystem consists of the exchange of data, money, or services. It is usually easy to determine which value is exchanged between the roles, but it is often difficult to say whether they have the same (monetary) value.

Co-Creation – The second identified characteristic of relationships in data ecosystems is the strong degree of collaboration between the actors to jointly create value. In the literature, this is also referred to as value co-creation [18,3]. This is due to the fact that in data ecosystems, various actors converge, who can offer different data, and because they may have varying data analysis capabilities [11,9]. Only through close collaboration between these ecosystem actors value can ultimately be created for all involved stakeholders [37]. In the interviews with the industry experts, it became clear that in all data ecosystems, one actor cannot take on all the necessary roles and tasks, and therefore the collaboration of different actors is inevitable for the realization of most use cases. "*In a data ecosystem companies are engaged in value-adding activities based on data and jointly develop new service-providing products or services*" said expert 5. In turn, expert 1 emphasizes that this collaboration is particularly relevant for small and medium-sized enterprises as their

size makes them dependent on other companies to be able to exploit the potential arising from data ecosystems. The degree and quality of collaboration between two actors are difficult to quantify. However, an indication of a high degree of collaboration could be the frequency of communication, i.e. the exchange of data and information between the actors [21].

Interdependence – Related to the characteristic co-creation is the characteristic interdependence. It describes the degree to which different actors in the ecosystem must interact and share data to perform their roles and tasks [38]. Expert 4 has expressed this as follows: "*Characteristic of a data ecosystem is that the actors are dependent on each other in some way. For some companies, this dependence on others may be stronger than for others.*" Analogous to other ecosystem types, this interdependence can vary [8]. Close and highly dependent relationships are called tightly coupled [39]. Loosely coupled refers to relationships that are less formal and close, suggesting a higher degree of openness in the ecosystem [15]. For example, [21] showed in their study that the same roles in a data ecosystem usually have little or no relationship to each other and are therefore relatively independent. The authors explain their findings with the fact that the same roles in a data ecosystem can be in competition with each other and therefore do not want to exchange data with each other, for example. However, a dependency does not have to be bilateral. [26] point out, for example, that in some cases service providers can be very dependent on certain data providers, and therefore, there may be more of a one-sided dependency.

Trust – It became apparent in both the literature review and the expert interviews that trust is a central characteristic of any relationship in a data ecosystem. [40] state, for example, "*Trust is the fundamental component of all relationships in a data sharing ecosystem*". In data ecosystems, this usually involves trust in how other actors handle the data of the data providers and whether they adhere to the agreed terms of use [18]. Expert 2 highlighted the following in this regard: "*But that means you have to trust your supplier not to, let's say, play fast and loose with your data and pass it on to some other competitor and say to them, "look how they do it in their factory".*" However, it is difficult to say how trust takes shape in a relationship since trust is an informal factor that can be difficult to measure [17,27]. Nevertheless, theory and practice agree that relationships in data ecosystems are often built on pre-existing business relationships [11]. Furthermore, in a data ecosystem, various measures such as technological design decisions and formal agreements can be taken to increase trust in the data exchange and thus in the relationships between the actors [27,6]. Examples of this are the use of technologies that allow data usage conditions to be defined and enforced, such as the International Data Spaces [11], or the use of blockchain technologies that can replace trust in an intermediary [41].

Intention – Every actor in an ecosystem has an incentive and motivation for their participation [15,21]. A relationship in an ecosystem can only develop if the motivation of both actors matches together [42,4]. However, this is not trivial, as the motivation of an actor is not always clear for the other actors, can be contradictory in some areas due to competition, or can also change over time [22,11]. A relationship in a data ecosystem is therefore characterized by a constant tension between the motivations of the actors. Several experts emphasized in the interviews that a relationship between two actors can only emerge and develop if the added value of the relationship, i.e. how they can benefit from the collaboration, is clear to both sides. Expert 11 states in this regard in the interview: "*If in the end not everyone benefits, companies would not be interested. This is one of the basic principles of the industry: "I don't do anything I can't profit from".*"

Relevance – Looking at a data ecosystem as an interconnected system, it can be noted that some relationships are more important to the emergence and functioning of the ecosystem than others. This can be explained by the fact that some actors bring resources such as data or services into the ecosystem that cannot be provided at all or only with difficulty by other actors [18]. The relationships that these actors have with other actors, especially if these are relationships between important actors, may therefore be more relevant to the functioning of the ecosystem than others [11,21]. Typically, these are the relationships between the data providers, who offer much of the relevant data, and the analysis service providers, who offer the core service

of the data ecosystem [18,43]. Conversely, some relationships provide more supportive services to the ecosystem and are easier to replace. Examples of support services are data quality evaluation or matchmaking services [43]. A frequently mentioned support service is also the so-called infrastructure provider [18,17]. However, it became clear in the interviews that this service is difficult to categorize uniformly in terms of its relevance, since the infrastructure can have a high or rather low relevance depending on the protection requirements of the shared data. For example, expert 2 said "*The technology used for data sharing can be important in some data ecosystems, but not in others. For example, when it comes to possible areas of competition, i.e. company X may not be able to run its products over an architecture from company Y.*" For this reason, among others, we believe that the relevance of a relationship in a data ecosystem is highly context-dependent and can only be determined through qualitative analysis.

5. Conclusion

This study deals with the description and characterization of relationships between actors in data ecosystems. A precise understanding of the relationships is important as the interactions between the actors span the data ecosystem and it is only through collaboration that the added value of the ecosystem is created. Against this background, this study developed characteristics which describe relationships in data ecosystems in a structured and concise way.

Several implications for research and practice arise from the results of this work. From a **scientific perspective**, our results contribute to the still relatively young data ecosystem literature. While a few previous studies have dealt with the description and characterization of data ecosystems in general, the results of this study go deeper by describing the relationships between actors in more detail and more specifically. Consequently, the results, derived from the scientific knowledge base and expert interviews, can help to expand the existing body of knowledge and specify the common understandings and definitions of data ecosystems. Ultimately, this can be a further step towards developing fundamental and comprehensive theories of data ecosystems that do not currently exist in the scientific literature [8,9].

From a **managerial perspective**, the study results can be used by practitioners as a starting point to better understand the relationships of the data ecosystem in which the organization already participates. Based on this, relationships could be better managed and actively shaped. Ultimately, the findings of the study can help organizations develop relationships with other actors to build new data ecosystems to realize the potential of cross-organizational data sharing in data ecosystems.

However, our study is naturally subject to certain **limitations** which must be considered when interpreting the results. First, the steady progress of digitization and the still young age of the data ecosystem literature leads to evolving concepts and definitions around data ecosystems. Linked to this lack of commonly accepted theories is the challenge of distinguishing data ecosystems from related ecosystem concepts such as digital and platform ecosystems and associated forms of collaboration such as alliances and networks. Furthermore, data collection and analysis are also subject to certain limitations. On the one hand, additional articles could be found as data sources by using further scientific databases besides Scopus. On the other hand, the statements from the expert interviews are limited in their generalizability as other experts might give different answers. Finally, the analysis of the literature, as well as the interviews is subject to some interpretation, which means that other researchers may identify different characteristics depending on their influences, preferences, and biases.

However, the aforementioned limitations also point to possibilities for **future research** paths. By combining existing role descriptions and the results of this work, an ontology for data ecosystems could be developed that represents the essential roles and their relationships to each other in a structured way. An ontology would create a shared understanding of the topic of data ecosystems and be a further step towards developing a comprehensive theory [9]. Furthermore, it would be interesting to investigate how some characteristics can

be built or developed. For example, what can be done if the intentions of two actors do not match? This may require incentive mechanisms in the data ecosystem that motivate actors to actively participate in the ecosystem and share their data [17,42].

6. References

[1] Brynjolfsson, E., McAfee, A., 2017. The Business of Artificial Intelligence: What it can — and cannot — do for your organization. Harvard Business Review 95, 3–11.

[2] Lis, D., Otto, B., 2020. Data Governance in Data Ecosystems - Insights from Organizations, in: AMCIS 2020 Proceedings.

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

[4] 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.

[5] Capgemini, 2021. Data Sharing Masters: How smart organizations use data ecosystems to gain an unbeatable competitive edge, 56 pp. https://www.capgemini.com/wp-content/uploads/2021/07/Final-Web-Version-of-Report-Data-Ecosystems.pdf.

[6] 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 Trans. Eng. Manage., 1–11.

[7] Priego, L.P., Osimo, D., Wareham, J., 2019. Data sharing practice in Big Data ecosystems. Esade Working Paper Nº 273, 38 pp.

[8] Gelhaar, J., Groß, T., Otto, B., 2021. A Taxonomy for Data Ecosystems, in: Proceedings of the 54th Hawaii International Conference on System Sciences.

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

[10] Jacobides, M.G., Cennamo, C., Gawer, A., 2018. Towards a theory of ecosystems. Strat Mgmt J 39 (8), 2255–2276.

[11] Gelhaar, J., Otto, B., 2020. Challenges in the Emergence of Data Ecosystems, in: Proceedings of the 24th Pacific Asia Conference on Information Systems (PACIS).

[12] Oliveira, M.I., Oliveira, L.E., Batista, M.G.R., Lóscio, B.F., 2018. Towards a Meta-model for Data Ecosystems, in: Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age, Delft, Netherlands, pp. 1–10.

[13] Tansley, A.G., 1935. The Use and Abuse of Vegetational Concepts and Terms.". Ecology 16 (3), 284–307.

[14] Chapin III, F.S., A. Matson, P., M. Vitousek, P., 2014. The ecosystem concept, in: Chapin III, F.S. (Ed.), Principles of Terrestrial Ecosystem Ecology. Springer, Berlin, Germany, pp. 3–22.

[15] 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.

[16] Moore, J.F., 1993. Predators and Prey: A New Ecology of Competition. Harvard Business Review 71 (3), 75–86.

[17] 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.

[18] 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.

[19] Adner, R., 2017. Ecosystem as Structure. Journal of Management 43 (1), 39–58.

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

[21] Hayashi, T., Ishimura, G., Ohsawa, Y., 2021. Structural Characteristics of Stakeholder Relationships and Value Chain Network in Data Exchange Ecosystem. IEEE Access 9, 52266–52276.

[22] 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, Dubai, UAE.

[23] van den Homberg, M., Susha, I., 2018. Characterizing Data Ecosystems to Support Official Statistics with Open Mapping Data for Reporting on Sustainable Development Goals. IJGI 7 (12).

[24] Azkan, C., Iggena, L., Gür, I., Möller, F.O., Otto, B., 2020. A Taxonomy for Data-Driven Services in Manufacturing Industries, in: Proceedings of the 24th Pacific Asia Conference on Information Systems (PACIS).

[25] Curry, E., 2016. The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches, in: Cavanillas, J.M., Curry, E., Wahlster, W. (Eds.), New Horizons for a Data-Driven Economy. Springer International Publishing, Cham, pp. 29–37.

[26] Heimstädt, M., Saunderson, F., Heath, T., 2014. Conceptualizing Open Data ecosystems: A timeline analysis of Open Data development in the UK. Discussion Papers 2014/12, Free University Berlin, School of Business & Economics.

[27] Lis, D., Otto, B., 2021. Towards a Taxonomy of Ecosystem Data Governance, in: Proceedings of the 54th Hawaii International Conference on System Sciences, pp. 6067–6076.

[28] Webster, J., Watson, R.T., 2002. Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly 26 (2), xiii–xxiii.

[29] vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A., 2009. Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process, in: Proceedings of the 17th European Conference on Information Systems, Verona, Italy.

[30] Elsevier, 2020. Content - How Scopus Works. https://www.elsevier.com/solutions/scopus/how-scopus-works/content.

[31] Sumbaly, R., Kreps, J., Shah, S., 2013. The "Big Data" Ecosystem at LinkedIn, in: Proceedings of the 2013 international conference on Management of data - SIGMOD '13, New York, USA, pp. 1125–1134.

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

[33] Krippendorff, K., 2013. Content analysis: An introduction to its methodology, 3rd ed. Sage, Los Angeles, 441 pp.

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

[35] Badewitz, W., Kloker, S., Weinhardt, C., 2020. The Data Provision Game: Researching Revenue Sharing in Collaborative Data Networks, in: 22nd Conference on Business Informatics (CBI), Antwerp, Belgium, pp. 191–200.

[36] Spiekermann, M., Wenzel, S., Otto, B., 2018. A Conceptual Model of Benchmarking Data and its Implications for Data Mapping in the Data Economy, in: Multikonferenz Wirtschaftsinformatik (MKWI). Data driven X - Turning Data into Value, pp. 314–325.

[37] Wilson, B., Cong, C., 2021. Beyond the supply side: Use and impact of municipal open data in the U.S. Telematics and Informatics 58, 101526.

[38] Curry, E., Sheth, A., 2018. Next-Generation Smart Environments: From System of Systems to Data Ecosystems. IEEE Intell. Syst. 33 (3), 69–76.

[39] Rong, K., Lin, Y., Li, B., Burström, T., Butel, L., Yu, J., 2018. Business ecosystem research agenda: more dynamic, more embedded, and more internationalized. Asian Bus Manage 17 (3), 167–182.

[40] Abebe, R., Aruleba, K., Birhane, A., Kingsley, S., Obaido, G., Remy, S.L., Sadagopan, S., 2021. Narratives and Counternarratives on Data Sharing in Africa, in: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency FAccT, pp. 329–341.

[41] Bons, R.W., Versendaal, J., Zavolokina, L., Shi, W.L., 2020. Potential and limits of Blockchain technology for networked businesses. Electron Markets 30 (2), 189–194.

[42] 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, Online. 10.08.2021 – 11.08.2021.

[43] Immonen, A., Ovaska, E., Paaso, T., 2018. Towards certified open data in digital service ecosystems. Software Qual J 26 (4), 1257–1297.

Biography

Joshua Gelhaar, M.Sc. (*1991) studied industrial engineering at TU Dortmund University. Since 2019 he works as a scientist in the department "Data Business" at the Fraunhofer Institute for Software and Systems Engineering ISST in Dortmund, Germany. His main research focus is on data sharing, especially within industrial data ecosystems.

Felix Becker, M.Sc. studied industrial engineering at TU Dortmund University and recently completed his master's thesis in the field of "Interdependencies and potentials of data ecosystems" in cooperation with the Fraunhofer Institute for Software and Systems Engineering ISST in Dortmund. He now works for a manufacturer of logistics systems.

Tobias Groß, M.A. (*1992) studied business management at FH Dortmund University of Applied Sciences and Arts. Since 2021 he works as a scientist in the department "Data Business" at the Fraunhofer Institute for Software and Systems Engineering ISST in Dortmund, Germany. His main research focus is on data sharing, especially investigating incentives for companies to take part in industrial data ecosystems.