7-4-2022

Past, Present and Future of Data Ecosystems Research: A Systematic Literature Review

Daniel Heinz
Karlsruhe Institute of Technology, daniel.heinz@kit.edu

Carina Benz
Karlsruhe Institute of Technology, carina.benz@kit.edu

Marcel Fassnacht
Karlsruhe Institute of Technology, marcel.fassnacht@kit.edu

Gerhard Satzger
Karlsruhe Institute of Technology, gerhard.satzger@kit.edu

Past, Present, and Future of Data Ecosystems Research: A Systematic Literature Review

Completed Research Paper

Daniel Heinz
Karlsruhe Institute of Technology
Karlsruhe, Germany
daniel.heinz@kit.edu

Carina Benz
Karlsruhe Institute of Technology
Karlsruhe, Germany
carina.benz@kit.edu

Marcel Fassnacht
Karlsruhe Institute of Technology
Karlsruhe, Germany
marcel.fassnacht@kit.edu

Gerhard Satzger
Karlsruhe Institute of Technology
Karlsruhe, Germany
gerhard.satzger@kit.edu

Abstract

Data ecosystems are a phenomenon gaining momentum as more and more organizations realize that novel data-based value propositions might require access to external data sources and inter-organizational collaboration to exploit such data assets. The notion of data ecosystems encompasses constellations composed of autonomous actors that consume, produce, or provide data and other related resources (e.g., software, services, and infrastructure) to co-create value. With rapidly growing activity in the academic community, a state-of-the-art mapping of the research field and its existing knowledge base is of particular importance to effectively guide future research. Thus, we conduct a systematic literature review to identify three distinct research streams manifesting current scientific knowledge. Further, we synthesize these insights into an interdisciplinary research agenda to catalyze future studies in the emerging field of data ecosystems in IS research and beyond.

Keywords: Data Ecosystems, Data Sharing, Digital Innovation, Systematic Literature Review, Research Agenda

Introduction

If data is the "new oil" leading the next industrial revolution (Forbes 2019), how can the economy actually get the engine running? Driven by the promise of digital technology, changing expectations of customers, and the competitive threat of fast-scaling digital business models, firms thirst for successful digital innovation and data-driven business models (Hunke et al. 2021; Nambisan et al. 2020; Wiesböck and Hess 2020). Throughout the creation and operation of such new business solutions, *data* serves as the central resource for innovation and value creation (Otto and Jarke 2019). However, to successfully exploit the potential of data-based innovations, organizations often need to get access to complementary external data assets (e.g., data that their customers or business partners own) and, therefore, engage in inter-organizational collaborations (Benz and Seebacher 2018; Weill and Woerner 2015). Depending on the application scenario, such complementary constellations can range from simple supply-demand chains (e.g., data producer, app creator, and end-user) to complex networks of services and infrastructure provisions (e.g., an autonomously operated factory with machines from multiple manufacturers) (Oliveira and Lóscio 2018). Hence, to "get the engine running", we first "need to build pipelines and gas stations in order to get the oil to where it's burned". In this study, we argue that organizations increase their chances of successfully exploiting data-driven business potentials by forming or participating in such target-oriented partner networks—so-called *data ecosystems*.

Data ecosystems are defined as “a set of networks composed of autonomous actors that directly or indirectly consume, produce or provide data and other related resources (e.g., software, services, and infrastructure)” (Oliveira and Lóscio 2018, p. 4). Typically, these actors play one or more distinctive “roles”, which comprise a set of activities and duties, and connect with other actors through relationships that are beneficial for the overall data ecosystem, promoting its self-regulation (Oliveira and Lóscio 2018). As an extension of the established (business) ecosystem construct (e.g., Jacobides et al. 2018), the notion of “data ecosystems” suggests a more specific perspective that emphasizes *data* as the main resource and data-driven value co-creation as the shared goal among all actors (Oliveira and Lóscio 2018). Data ecosystems differ from traditional value chains, particularly due to their networked structure consisting of often wide-ranging actor constellations and mostly non-hierarchical control mechanisms (Adner 2017; Jacobides et al. 2018). Potential application scenarios for data ecosystems are manifold, ranging from intelligent manufacturing (Kusiak 2018) through smart service applications (e.g., Curry and Sheth 2018; Heinz et al. 2022a) to healthcare ecosystems (Khuntia et al. 2017).

Previous studies have contributed to the concept of data ecosystems by elaborating a cohesive definition (Oliveira and Lóscio 2018) and providing multiple frameworks for characterizing the concept (Azkan et al. 2020; Curry and Sheth 2018; Gelhaar et al. 2021a). Yet, with rapidly growing activity in the academic community, a state-of-the-art mapping of the research field and a systematic identification of adjacent knowledge are important to guide future research efforts effectively. This need is further underpinned, as data ecosystems are a phenomenon positioned at the intersection of various research communities (e.g., information systems, service science, computer science, or production management), thus currently exhibiting a high degree of fragmentation. Consequently, we ask: *What is the current state of research related to data ecosystems, and which are avenues for future research?*

To answer this question, we conduct a systematic literature review (vom Brocke et al. 2009; Webster and Watson 2002). We identify three distinct research streams that manifest the current scientific knowledge on data ecosystems and provide linkages to related constructs during this process. By synthesizing these insights into an interdisciplinary research agenda, we aim to catalyze future studies in the emerging field of data ecosystems and offer guidance for the increasing number of joint corporate endeavors in practice. We thereby contribute to the emerging field of data ecosystems and respond to the call for further investigating, developing, and interrelating fundamental concepts of data ecosystems and existing theories (Beverungen et al. 2020; Oliveira et al. 2019).

In the remainder of this article, we first elaborate on the theoretical foundations of data ecosystems and related work, followed by a description of our methodology. Subsequently, we present the existing body of knowledge on data ecosystems along three research streams and, afterward, point out future directions for the field. Our concluding section discusses the implications and limitations of our contribution and wraps up our study.

Foundations and Related Work

In this section, we present the theoretical foundation of our research and discuss related work. First, we elaborate on how our notion of data ecosystems evolves from the well-established idea of “ecosystems” used in management and IS research. Afterward, we briefly present how our work extends existing approaches of unwrapping the data ecosystems concept.

Borrowed from biology, the term *ecosystem* recently became a popular concept to describe the context for interactions among independent firms in discussions of strategy and beyond (Adner 2017; Jacobides et al. 2018). In such an ecosystem, different parties establish a structured multilateral relationship motivated by the collective outcome of non-generic, i.e., to some degree unique, complementarities. Furthermore, in contrast to traditional supply networks, an ecosystem is rather centrally or hierarchically controlled but runs with decision-making processes that are to some extent distributed among its actors (Jacobides et al. 2018). Empirical studies show that especially in the context of innovating digital solutions, composing and participating in business ecosystems becomes an attractive strategic choice (Weill and Woerner 2015). Consequently, the notion of ecosystems also gained popularity among IS scholars in the recent past (e.g., Benz et al. 2021; Hein et al. 2020; Tiwana et al. 2010).

Stemming from the described growing research paradigm of "ecosystems", scholars recently adopted the term "data ecosystem" as a more granular instance (Gelhaar et al. 2021a). Borrowing aspects from different types of ecosystems, the notion of data ecosystems translates the ecosystem's broad goals of innovation and value creation into a more specific application context: leveraging the potential of *data* to deliver innovation and support business in multilateral settings (Oliveira and Lóscio 2018). The work of Oliveira et al. (2019; 2018) played a seminal role in establishing the field of "data ecosystems". As an extension of their previous work on open data ecosystems (Oliveira et al. 2017), they evolved the concept and provided a first cohesive definition of "data ecosystems", which builds on four constructs: actors, roles, relationships, and resources (Oliveira and Lóscio 2018). Based on their findings on the state-of-the-art of data ecosystems research, they point out relevant perspectives on data ecosystems research (e.g., characterization, formation, and management of data ecosystems) and discuss "theory, models, engineering, and solutions" (Oliveira et al. 2019, p. 620) as the most important areas for developing the field. With our study, we respond to the authors' conclusion that the diversity of prevalent conceptual understandings and the ambiguity of the practical design and functioning of data ecosystems impede the evolution of the research field.

In another study, Curry and Sheth (2018) obtain different types of data ecosystems and determine infrastructure, ecosystem governance, systems engineering, and human centricity as the most current challenges for research and practice. In addition, the research group around Otto and Jarke (2019) adopted the notion of "data ecosystems" and conducted multiple studies (Azkan et al. 2020; Gelhaar et al. 2021a; Gelhaar and Otto 2020; Lis and Otto 2020). With these studies, the authors contribute to a more practical understanding via various empirical studies as well as a more detailed theoretical alignment to related concepts such as business ecosystems (Gelhaar and Otto 2020; Moore 1993) or service-dominant logic (Azkan et al. 2020; Lusch and Nambisan 2015). Furthermore, these contributions allow a more precise definition of the idea and its boundaries, emphasizing data as the "focused object" of interaction among actors (Gelhaar et al. 2021a).

While it is beyond the scope of this paper to present the frameworks for characterizing data ecosystems (morphology (Azkan et al. 2020) and taxonomy (Gelhaar et al. 2021a)), we want to highlight how our research approach differs: instead of identifying characteristics to describe a specific data ecosystem or distinct types of data ecosystems, our study pursues the overarching goal to structure existing and future research related to the phenomenon of data ecosystems. For this purpose, we map research streams and point out current problems and gaps as avenues for future research. In doing so, we also delineate our work from Oliveira et al. (2019)'s seminal literature review: first, we acknowledge a shift in data ecosystems research from rather public or governmental initiatives to commercial, and industrial activities and include respective articles. Second, we approach the field of data ecosystems research from a more conceptual point of view instead of settling on the term "data ecosystem". Finally, their article only considers studies published before 2017. Hence, this review can include a much wider database in the emerging research field of data ecosystems.

Methodology

To pursue our goal of structuring existing knowledge in the data ecosystems field, we conducted a systematic literature review following established methods and procedures in IS research (vom Brocke et al. 2009; Webster and Watson 2002). In the following, we briefly introduce and justify our research procedure in five steps: 1) definition of review scope, 2) conceptualization of topic, 3) literature search, 4) literature analysis and synthesis, and 5) research agenda (vom Brocke et al. 2009).

To define the scope of our literature review, we referred to Cooper's (1988) taxonomy: we aim to *integrate* existing *research findings* with a *neutral* perspective. We thereby provide a *conceptual* synthesis of a *representative* set of literature addressing *general scholars* joining the academic debate on data ecosystems. To conceptualize the topic, we reviewed an initial sample of literature to develop our understanding of the underlying phenomenon and decomposed different aspects that together define the most prevailing notion of data ecosystems. We further elaborated on these aspects by collecting synonyms and strongly overlapping descriptive terms in an iterative process (vom Brocke et al. 2009). We generated a search phrase on this basis, which allows us to include adjacent research that uses a different vocabulary for the same phenomenon. The final search term combines the core characteristics of data ecosystems along four areas of interest: 1) *data should* be discussed as it is the key resource in a data ecosystem, 2) the *collaboration* among actors beyond merely transactional data transmission should be reflected, 3) the *inter-organizational* setting should be incorporated, as we focus on value creation through complementarities of organizations, and 4) the idea of an *ecosystem* as the structure or a platform as the focal point among actors should be addressed. We identified multiple synonyms for these concepts throughout our initial literature screening, which we added to our search phrase. Furthermore, we identified terms that combine the latter two partial aspects (e.g., smart service systems). The included phrases can be derived from Figure 1 (I).

Simultaneously, we compiled a list of 62 high-quality outlets for our literature search to ensure academic standards such as articles being peer-reviewed. Thereby, we primarily focused on IS research outlets and added selected journals from related disciplines (primarily economics and production research) where an exploratory review of the overall body of literature revealed potential insights into data ecosystems. As an assessment of quality, we considered both the impact factor as well as a well-accepted ranking of academic outlets. Referring to Web of Science, Scopus, and AISeL as a comprehensive set of databases, we reached an initial sample of 196 articles published by April 2021 after a first filtering process (Figure 1, II).

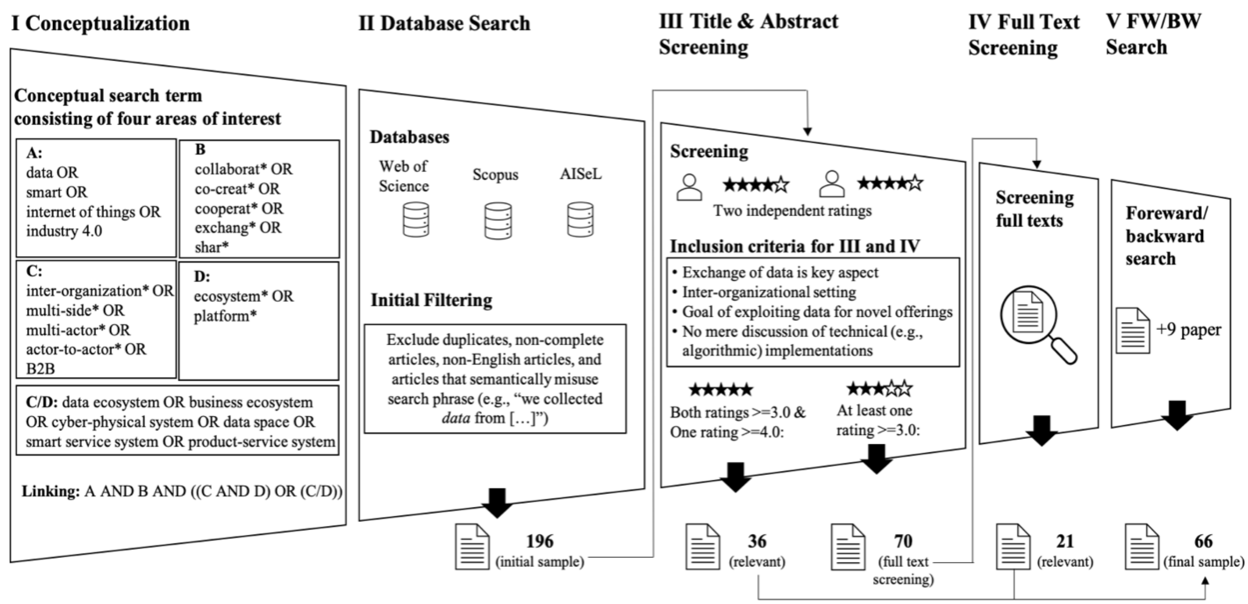

Figure 1 Literature Search Process.

Next, we assessed each article's relevance in line with our conceptual understanding of data ecosystems in a two-stage process, as suggested by Torraco (2005). Based on the articles' title and abstract, two independent researchers rated the articles along a pre-defined set of inclusion criteria (Figure 1, III) on a five-point Likert scale. These criteria were: 1) the exchange of data is a key aspect of the article's research subject, 2) the article examines an inter-organizational setting, 3) the overarching goal of the actors is to exploit data for novel offerings, 4) the article does not merely discuss technical (e.g., algorithmic) implementations. The measurement of the inter-rater agreement using both Cronbach's alpha (0.851) and Krippendorff's alpha (0.718) suggests that the researchers reached a sufficient agreement in their final ratings (Gliem and Gliem 2003; Landis and Koch 1977). At this stage of the screening process, 36 articles were included in the final sample, whereas 70 articles were subject to a second assessment based on their full text (Figure 1, IV). From these 70 articles, 21 others were included in the final sample. Throughout the literature review, we conducted a forward (FW) and backward (BW) search, i.e., looked for further articles that cited or have been cited by articles in our sample contributing to our research question (Webster and Watson 2002)—leading to an addition of nine articles and a final sample of 66 articles (Figure 1, V). Of these, six articles were from the "basket of eight" IS journals; 20 from the IS conferences; 21 from other IS journals, and 19 from the economics and production journals (all included articles are marked with an asterisk (*) in the References section). 6 articles were published during the period 2004-2012 and 17 in 2013-2017. However, most of our sample (n=43) was published from 2018 onwards.

As the next step of our literature review, we analyzed the obtained literature sample for existing constructs, models, and theories from a concept-centric perspective and identified supporting evidence and research gaps (Webster and Watson 2002). For this purpose, we iteratively applied open and selective coding using MAXQDA (Mayring 2004). During the open coding, we identified and paraphrased the core findings of each article regarding their contribution to data ecosystems research. After the first round of open coding, we revisited each code to highlight connections to the academic discourse on data ecosystems.

After the open coding, we apply selective coding, which was guided by cross-validating discussions between the authors to ensure transparency and validity of the results. In this step, we analytically abstracted and aggregated the paper-specific codes into synthesized "themes" and "research streams" as higher-order concepts and more abstract perspectives on the data ecosystems field. For example, Beverungen, Müller, et al. (2019) provide a conceptualization of how smart (connected) products serve as boundary objects in smart service systems. In our analysis, we found that the innovation of smart products and smart services is closely intertwined with the formation of data ecosystems to give rise to value propositions that leverage their capabilities. In line with other reviewed articles, we identify the "evolution of traditional offerings towards smart products and services" as a key theme of data ecosystems research and the process of "innovation to create value from data" as a more abstract research stream in this field.

Since we aim to conduct an integrative literature review that "analyzes the past to prepare for the future" (Webster and Watson 2002), we then developed a research agenda to point out future directions for the field. Referring to the results of our literature review and the underlying sample of articles, we highlight critical yet insufficiently explored research avenues at the intersection of the three identified research streams. By providing linkages to the identified themes and research streams, we aim to present a promising starting point for further research endeavors in IS research and beyond.

Looking Back: Existing Research on Data Ecosystems

In this section, we synthesize the existing body of knowledge by presenting the results of our integrative literature review and linking data ecosystems to related concepts. We shed light on distinct areas of research by introducing and discussing the following three research streams (Figure 2): 1) Creating value from data in innovative offerings ("*Innovating*"), 2) effectively leveraging technology and information ("*Engineering*") and 3) engaging multiple actors within the data ecosystem ("*Collaborating*"). These research streams emerged by aggregating conceptual themes resulting from an iterative coding process throughout our literature review. These themes can also be derived from Figure 2 and are described in more detail in the following subsections.

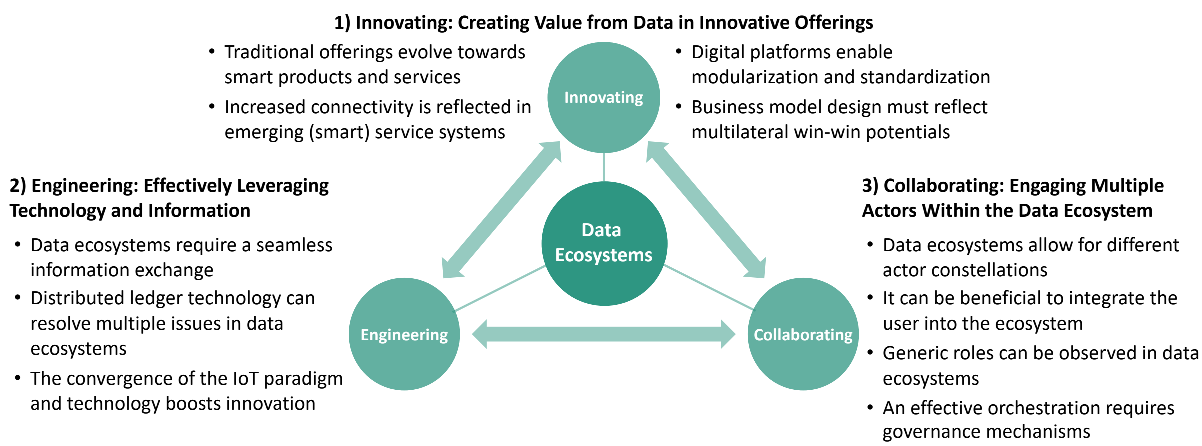

Figure 2 Three Research Streams and Included Themes Contributing to Data Ecosystems.

Innovating: Creating Value from Data in Innovative Offerings

This first research stream operates right at "the heart" of data ecosystems and focuses on its focal value proposition and the novelty of related offerings.

Traditional Offerings Evolve Towards Smart Products and Services. Regarding the focal outcomes of a data ecosystem, multiple articles draw an evolutionary path from mostly physical offerings over digital solutions towards a yet unfolding type of offering (Hirt and Kühl 2018; Tuunanen et al. 2019), today mainly called smart (connected) products and services (Beverungen, Müller, et al. 2019; Zheng et al. 2019). These offerings blend both the physical and virtual world by analyzing data collected via sensor-equipped connected physical objects and create value-in-use through contextual and preemptive services (Tuunanen et al. 2019). Applications of such offerings are manifold, e.g., autonomous driving, remote health care services, and networked manufacturing systems. Furthermore, there are different "value scenarios" for the application of smart products in industrial data ecosystems, where products allow the manufacturer to monitor a device (Saarikko et al. 2016). As a means to evaluate the degree of smartness and "connectedness" of such smart products, Zheng et al. (2019) propose five levels: 1) smart connectivity, 2) smart analytics, 3) digital twin, 4) cognition, and 5) autonomy.

Increased Connectivity Is Reflected in Emerging (Smart) Service Systems. Considering the actual value creation, it becomes evident that the increasing connectivity of such offerings is reflected in more complex actor constellations. Thus, several authors emphasize the emergence of "smart service systems" (e.g., Beverungen, Müller, et al. 2019; Lim and Maglio 2018), where multiple involved actors integrate their resources and activities around smart products and services for mutual benefits. In addition to the multifaceted definition of smart service systems, other authors discuss more application-specific scenarios such as "smart manufacturing (systems)" (Kamble et al. 2020; Kusiak 2018; Osterrieder et al. 2020), "cloud manufacturing" (Kannisto et al. 2020; Moghaddam and Nof 2018) for the manufacturing industry, or services enabled by health information exchanges (Khuntia et al. 2017). To better capture these smart service systems by developing a modeling language, Huber et al. (2019) emphasize the nested structure of such service systems, i.e., "systems of systems" (Porter and Heppelmann 2014; Smedlund et al. 2018), which further increase the coordination and management complexity.

Digital Platforms Enable Modularization and Standardization. As means to reduce complexity, the topics of modularization, standardization, and boundary resources are widely recognized and discussed (Hein et al. 2018). Particularly digital platforms can play an important role in data ecosystems by streamlining information (Kannisto et al. 2020; Khuntia et al. 2017) and acting as an application enablement platform for third-party services (Hein et al. 2019). For example, an empirical study on open banking by Gozman et al. (2018) shows how such platform-centric ecosystems can enhance the separation between service creation via third parties and distribution via the platform providers. Furthermore, boundary resources enable organizations to share information (Rehm and Goel 2013) by ensuring standardization while maintaining interpretive flexibility across organizations (Gosain et al. 2004; Hein et al. 2019). Hence, Beverungen et al. (2020) juxtapose so-called "smart service platforms" as digital boundary objects and smart products as physical boundary objects in smart service systems.

Business Model Design Must Reflect Multilateral Win-Win Potentials. Another relevant topic is the successful business model innovation within data ecosystems. While the collaborative innovation of a data-driven value proposition is the core driver for the formation of a data ecosystem, a profitable business model is key to its survivability (Beverungen et al. 2020). Moreover, it is crucial for such multilateral and "co-opetitive" settings to generate win-win strategies among all actors (Pant and Yu 2018). This challenge is closely linked to the innovation of suitable revenue models, which allow a fair allocation to all parties contributing to the ultimate value proposition (e.g., Pauli et al. 2021). Additionally, application scenarios of data ecosystems allow for revenue models beyond the scope of traditional goods-producing offerings such as subscription-based models (Khuntia et al. 2017). Moreover, Hahn et al. (2016) identify broader forms of "value appropriation" beyond financial payments within an ecosystem, paraphrased as "bundling of services", "lock-in effects", "utilization of downstream capabilities" and "indirect network effects". Closely related, Khuntia et al. (2017) argue that additional service offerings naturally appear with operational maturity within the data ecosystem and need to be structured and bundled appropriately to achieve financial viability.

Engineering: Effectively Leveraging Technology and Information

The second stream comprises efforts to leverage technology and information to create value from data—often referred to as "engineering" in the innovation process (Anke et al. 2020; Böhmann et al. 2014).

Data Ecosystems Require a Seamless Information Exchange. As we understand data as the focal resource in data ecosystems, the processes and activities related to data handling are crucial for successful value creation (Osterrieder et al. 2020). In particular, seamless information exchange from data sources (e.g., smart products) into information systems is important as this facilitates value creation activities in the first place (Papert and Pflaum 2017; Rasouli 2020). We attribute the modularization (Baldwin and Clark 2000) of the overall offering into distinct “service units” rather as an overarching strategy in the innovating research stream, e.g., reflected in the “platform-as-a-service” concept describing a marketplace for third-party-services (“app store”) building on a shared infrastructure (Beimborn et al. 2011). However, *standardization* is even more important within this stream of research. Standardization can involve both the input data itself (e.g., by specifying data formats and quality requirements) (Aulkemeier et al. 2019; Weber et al. 2009) as well as processes and activities related to the transformation of the data into actual value (Hein et al. 2018, 2019). An example of means to enable standardized processes while integrating heterogeneous business parties is the application of adapters (Kannisto et al. 2020; Rasouli 2020).

Distributed Ledger Technology Can Resolve Multiple Issues in Data Ecosystems. Our literature review revealed the potential of different technologies and technology paradigms for data ecosystems. The most frequently considered is cryptography-based distributed ledger technology (DLT) (e.g., Abbas et al. 2020). DLT can be used for immutable data storage (Abbas et al. 2020; Gelhaar and Otto 2020) and to reduce information asymmetry and, thus, overcome trust issues (Kamble et al. 2020). For example, Schulz and Freund (2019) line out how DLT can be used to deal with changing requirements through the advent of industry 4.0 to reduce fraud and increase transparency. Next, Abbas et al. (2020) examine a more narrow application scenario and integrate DLT-based consensus mechanisms into the workflow of maintaining rolling stocks of a railway company to verify the quality of performed maintenance work based on uploaded pictures, and thus, reach a consensus. Both studies reveal that the successful finalization of applied technologies (e.g., regarding practicability or integration capabilities) is a prerequisite for the functioning of such technology-driven solutions.

The Convergence of the IoT paradigm and Technology Boosts Innovation. In many cases, offerings in data ecosystems build on the Internet of Things (IoT) paradigm (Saarikko et al. 2016), which describes an environment where things have identities and virtual representations and use intelligent interfaces to connect and communicate within social, environmental, and user contexts (Beverungen, Müller, et al. 2019). This “smartness” of physical objects (Lim and Maglio 2018) allows new value-creating activities (Saarikko et al. 2016; Weber et al. 2020) and could even “provide the infrastructure of a brave new digital future” (Shim et al. 2019, p. 136). However, translating the technology paradigm into action requires organizations to develop skills to integrate technologies such as ubiquitous sensing and cloud computing into their current practices to enable real-time information processing (Smedlund et al. 2018). Also, the research in our sample describes a disruptive potential of converging IoT and other technologies, such as 1) synergies using DLT in an IoT network to authenticate nodes in a network (Shim et al. 2019), 2) the convergence of IoT, analytics, and artificial intelligence in an upcoming era of “Analytics-of-Things” (Goul 2018), 3) cognitive computing enabling new ways of building and exchanging knowledge across a system (Hirt and Kühl 2018), or 4) decentralized fog/edge computing architectures enabling real-time monitoring and control (Kamble et al. 2020). Overall, contributions in this research stream examine the potential utility of data, information, and digital technologies as boundary resources (Otto and Jarke 2019) to enable or improve processes and activities in data ecosystems. Looking at these efforts, a distinction can be made between research primarily focusing on how technology and information contribute to novel offerings (e.g., IoT-enabled use cases such as smart waste management (Shim et al. 2019), and how they can be used to orchestrate the data ecosystem effectively (e.g., using blockchain-based solutions as a trust-building mechanism for predictive maintenance services (Abbas et al. 2020).

Collaborating: Engaging Multiple Actors Within the Data Ecosystem

This research stream particularly focuses on how actors within a data ecosystem can be engaged to participate successfully and, thus, create value mutually.

Data Ecosystems Allow for Different Actor Constellations. An important consideration regarding data ecosystems is the composition of its actors. While some data ecosystems are open to anyone to participate (Oliveira et al. 2019), many can rather be described as “closed communities” (Gelhaar and Otto 2020, p. 7). Although the collaboration can take place among vertically or horizontally aligned actors, most current initiatives can be characterized as vertical collaboration due to concerns about sharing data with competitors (Gelhaar and Otto 2020). However, technological challenges and novel collaboration potentials induce the integration of tech-savvy third-party service providers and industry-spanning collaboration (Gozman et al. 2018; Otto and Jarke 2019). Today's companies often build on their "social capital" (Patnayakuni et al. 2006) by leveraging existing business relationships to resolve trust issues rather than collaborating with yet unknown partners (Gelhaar and Otto 2020; Pauli et al. 2021). Sensitivity of data—due to both personal information and exposed intellectual property—is an important theme when discussing data ecosystems (Huang et al. 2009; Kieseberg and Weippl 2018). In this respect, Jarke (2017) suggests *data sovereignty* as a key requirement in data ecosystems, which implies a person's or organization's capability of being entirely self-determined regarding one's data. In the context of data ecosystems, this allows actors to exclusively decide about the usage of their data as an economic asset (Bastiaansen et al. 2020). It is a common pattern that the emergence of data ecosystems is largely impacted by a "keystone actor" (Iansiti and Levien 2004) responsible for managing and driving the activities within the ecosystem (Curry and Sheth 2018; Otto and Jarke 2019). Current research controversially debates the advantages and disadvantages of the resulting power imbalance (Oliveira et al. 2019; Otto and Jarke 2019) and discusses more balanced orchestration forms such as "open consortiums" as an alternative (Kannisto et al. 2020; Otto and Jarke 2019).

It can be Beneficial to Integrate the User into the Ecosystem. Unfulfilled customer needs can actually foster the formation of collaborations in ecosystem structures (Rasouli 2020). The exploitation of data enables more individualized—so-called "mass-individualized"—value propositions (Beverungen, Müller, et al. 2019; Zheng et al. 2019), e.g., via autonomous adaptations of products and services ("self-*-capabilities" or self-services). Saarijärvi et al. (2014) describe these recent developments as a "reverse use of customer data", i.e., an evolution from solely using customer data for supporting the provider's internal processes (e.g., supply chain management or targeted marketing) to empowering customers by converting data into usable information for their value creation processes. To effectively design these (self-)services, active user participation in the co-creation process should be pursued (Beverungen et al. 2020; Hein et al. 2019; Zheng et al. 2019). This not only improves the overall offering but also reduces the risk of speculative development (Saarikko et al. 2016). This value co-creation mechanism is a core theme of service innovation and service-dominant logic (Lusch and Nambisan 2015) and is the subject of the work of Azkan et al. (2020), who developed a morphology for data ecosystems along the lines of service-dominant logic.

Generic Roles can be Observed in Data Ecosystems. Following Oliveira and Lóscio (2018), we differentiate between the *actors* of which a data ecosystem is composed and the *roles* as functions played by these actors. Existing research systematically abstracts archetypal roles within different subcategories of data ecosystems (e.g., open data ecosystems (Lindman et al. 2014; Ponte 2015), IoT ecosystems (Papert and Pflaum 2017)), and other specific applications (Gozman et al. 2018). As a rather generalizable categorization for data ecosystems, Anke et al. (2020) propose four primary roles (digital innovator, system integrator, project sponsor, and service operator) and 13 secondary roles (e.g., customer representative, legal advisor, or cloud platform provider), whereas Saarikko et al. (2016) differentiate three roles within an IoT ecosystem: 1) "Enablers" developing and implementing IT solutions to connect products, 2) "engagers" offering (smart) products to utilize these solutions and 3) "enhancers" innovating (third-party) services to create value by exploiting the generated data.

An Effective Orchestration Requires Governance Mechanisms. Finally, a predominant theme in data ecosystems research is the effective orchestration of the overall ecosystem (Breidbach et al. 2018). As the different participants all have their own institutional logic (Hein et al. 2019), it is a crucial challenge to establish governance mechanisms that "appropriately bound participant behavior without excessively constraining the desired level of generativity" (Wareham et al. 2014, p. 1195). For this purpose, a data ecosystem typically employs more or less formalized governance mechanisms (Oliveira et al. 2019; Otto and Jarke 2019, Gelhaar et al. 2021b) that include processes, policies, and structures and allow a comprehensive control of the ecosystem (Lee et al. 2017). A concept of particular interest for orchestrating data ecosystems is *data governance* (Lee et al. 2017), which addresses the means to achieve the desired data sovereignty (e.g., Goul 2018; Markus and Bui 2012; Otto and Jarke 2019). Besides data governance, governance mechanisms also need to address further issues prevalent for ecosystem-like structures, e.g., the decision on including new members and measures how they participate in decision making (Huang et al. 2009; Markus and Bui 2012; Wareham et al. 2014). Despite manifold contributions, identifying effective and suitable governance mechanisms remains a highly topical and important debate in both research and practice, particularly due to the often paradoxical nature of the addressed tensions (Wareham et al. 2014). In the following section, we will introduce this and further unsolved research questions in the field of data ecosystems as part of a research agenda.

Looking Forward: Interdisciplinary Avenues for Future Research

As an integrative literature review should rather "weave the streams of research together to focus on core issues of the field than merely reporting previous literature" (Torraco 2005, p. 362), our second goal is to unfold a research agenda on data ecosystems. In doing so, our literature review allows us to highlight critical yet insufficiently explored research avenues at the intersection of the three identified research streams. Further, we present linkages to the presented literature and, thereby, provide starting points for future studies in IS research (and beyond) to "connect the dots" towards a holistic understanding of data ecosystems (Beverungen et al. 2020; Li et al. 2020). At the end of this section, we summarize the proposed set of research avenues (RAs) for data ecosystems research (Table 1).

Intersection of the Research Streams	**Avenues for Future Research**
Collaborating & Engineering: Engaging Actors through Technology	**RA1:** Shaping governance mechanisms from a socio-technical perspective that enable actors in a data ecosystem to interact with each other effectively **RA2:** Understanding and shaping the role of IT artifacts as (composite) boundary objects between different actors to facilitate collaborative value creation processes **RA3:** Illuminating the distinct role that data plays within a data ecosystem to distinguish data ecosystems from related concepts
Innovating & Collaborating: Creating Collaborative Value from Data	**RA4:** Intensifying insights and measures to enhance actor's engagement to participate in data ecosystem initiatives to co-create value through novel (service) offerings **RA5:** Designing mechanisms to capture and distribute the value of data-driven offerings among the actors of a data ecosystems **RA6:** Examining the holistic and long-term benefits or downsides of data ecosystem initiatives from a (macro-)economic point of view
Engineering & Innovating: Enabling Innovation by Evolving Technology	**RA7:** Leveraging technological concepts to ensure seamless transmission of data across organizational boundaries for the generation of innovative offerings **RA8:** Using technologies as an enabler to overcome intellectual property issues within the data ecosystem and data security issues regarding external parties

Table 1 An Interdisciplinary Research Agenda for Data Ecosystems Research.

Collaborating & Engineering: Engaging Actors through Technology

To deal with the complex challenges of digital innovation, our literature review suggests an urgent need for organizational collaboration in engineering novel solutions. First, we emphasize the importance of further developing suitable mechanisms for data governance. Despite different scholars already having analyzed and characterized challenges faced by organizations and derived initial best practices (e.g., Lee et al. 2017; Lis and Otto 2020), the field still lacks more detailed studies on useful rules and data management processes (Otto and Jarke 2019). Research on data governance can take a two-sided view on "data"—it can be the reason for governance mechanisms in the first place (e.g., data usage rights), but it can also be used as a means to govern an ecosystem effectively (Schreieck et al. 2016). Thus, we encourage research that addresses prevalent issues regarding data sharing and ownership in data ecosystems (Li et al. 2020). There are different related practical questions, e.g., contractual issues (Goul 2018) or the appropriate ecosystem architecture and decision-making processes (e.g., Otto and Jarke 2019). Hence, research can leverage synergic effects between the "engineering" and "collaborating" research streams, as technical solutions might help reflect data governance principles in the actual data handling workflow. In this light, we posit the following research avenue:

RA1: *Shaping governance mechanisms from a socio-technical perspective that enable actors in a data ecosystem to interact with each other effectively*

Second, we pick up the idea of boundary objects (Star 2010) commonly used in the context of digital platforms (Hein et al. 2019; De Reuver et al. 2018). This idea is closely tied to the inter-organizational context of data ecosystems and presents a useful approach to characterize the benefits of technological solutions as facilitators of value co-creation in ecosystems (Otto and Jarke 2019). Rehm and Goel (2013) argue that IS research should not only discuss individual artifacts but rather "clusters or environments of boundary and marginal objects" (Rehm and Goel 2013, p. 10), which they introduce as "composite boundary objects". As such portfolios of IT artifacts provide fruitful ground regarding inter-organizational innovation activities, we encourage further analysis of their role in data ecosystems.

RA2: *Understanding and shaping the role of IT artifacts as (composite) boundary objects between different actors to facilitate collaborative value creation processes*

Third, it seems promising for research to grasp better the distinct role of data serving as a data ecosystem's key resource. In data ecosystems, data is both the starting point for the technical value creation processes as well as the catalyst for collaboration in the ecosystem. Otto and Jarke (2019) stress this "dual nature" of data in the context of multi-sided platforms spanning a data ecosystem. Following the authors, data serves both as a technical and social boundary resource, requiring both technical processing and functional use within a data ecosystem. Therefore, a data ecosystem architecture must describe data governance rules (social) as well as shared data management processes (technical) regarding different user groups. This view seems like a promising approach to emphasize distinct characteristics of data ecosystems and, thus, allow a more precise definition of the overall concept.

RA3: *Illuminating the distinct role that data plays within a data ecosystem to distinguish data ecosystems from related concepts*

Innovating & Collaborating: Creating Collaborative Value from Data

As a second intersection, data ecosystems research should dedicate further attention to bringing together research on mechanisms for inter-organizational collaboration and research on the innovation activities of an ecosystem's actors. Therefore, we suggest the following three avenues for future research: first, today's research still lacks to combine findings on potential offerings and innovation processes with the implications of a change in structure towards more collaborative constellations (Anke et al. 2020). As previously discussed, it is important that actors actively participate in the ecosystem to benefit from existing complementarities in data ecosystems (Beverungen et al. 2020). However, many companies—especially small and medium enterprises (SMEs)—haven't yet developed the necessary capabilities (e.g., know-how or infrastructure) to participate in a data ecosystem and are rather cautious about opening up to this new type of collaboration (Deloitte Insights 2019; Heinz et al. 2022b). To address these concerns, it is a critical task for drivers of a data ecosystem initiative that they enable and engage their potential partners in order to unfold the full potential of their network. Despite the importance of this task, the existing body of knowledge yet lacks to provide insights or effective measures such as design knowledge for easily accessible and engaging data-sharing measures (Beverungen et al. 2020). Also, research should systematically analyze the concerns and reasons for organizations' cautiousness regarding data-sharing initiatives and develop communicable and practicable mitigation actions to dispel their skepticism (Aulkemeier et al. 2019; Kusiak 2018). An example of such mitigation actions could be proof-of-concept projects serving as "experienceable prototypes" before committing to initiatives on a larger scale (McKinsey Analytics 2020).

RA4: *Intensifying insights and measures to enhance actor's engagement to participate in data ecosystem initiatives to co-create value through novel (service) offerings*

Second, a frequent point of discussion is the *profitability* of data ecosystem initiatives, which we see as an important gap in the reviewed literature. In particular, our review suggests that traditional firms struggle to realize not only feasible but profitable novel collaborative data-driven offerings, determine the actual value-in-use, and thus, experience difficulties in establishing suitable revenue models (Pauli et al. 2021). Despite existing studies on value-capturing mechanisms for data-driven services (e.g., Schüritz et al. 2017), little is known about the overall realization of profitable data ecosystem initiatives. Against this background, the application of technological capabilities can act as an accelerator to support novel data-driven value capturing mechanisms by allowing feasible analytical models to allocate created value among partners in the ecosystem, e.g., by enabling subscription-based service offerings offering a pay-per-part business model. In this context, smart technologies enable an individual, efficient billing process to allow an automated allocation of cash flows. Further studies should, therefore, address this issue and provide insights on generic value capturing mechanisms for data ecosystems and the role of technological advancements in this context.

RA5: *Designing mechanisms to capture and distribute the value of data-driven offerings among the actors of a data ecosystems*

Finally, research yet lacks to prove that the evolution towards more inter-organizational settings is beneficial for an economy and society in the first place. Thus, further data ecosystems research should argue how such collaborations can contribute to the overall economic success, i.e., incorporate a macroeconomic view on data ecosystems initiatives. This might convince political decision-makers to provide more public funding or take other measures such as legal regulations regarding access to data to ensure fair competition to incentivize such endeavors. As participating or forming a data ecosystem is a high-risk project and requires long-term investments, such measures might help to ensure that also SMEs join the table with large corporations, and thus, accelerate their transformation and unfold their innovation potential (Heinz et al. 2021; Heinz et al. 2022b; Kusiak 2018).

RA6: *Examining the holistic and long-term benefits or downsides of data ecosystem initiatives from a (macro-)economic point of view*

Engineering & Innovating: Enabling Innovation by Evolving Technology

Regarding future research on enhancing the engineering capabilities and further evolving them towards enabling applicable innovations in data ecosystems, we tie with the presented findings in the previous section and propose two avenues for future research. First, we have already emphasized the importance of seamless information exchange among organizations. Despite notable efforts, our review suggests that many promising technological means still lack the maturity to be implemented in real-world application scenarios (Abbas et al. 2020; Schulz and Freund 2019). Moreover, data ecosystems research yet fails to provide an overview and discussion of different technological approaches such as DLT (Meroni et al. 2019), federated machine learning (Hirt and Kühl 2018), and ubiquitous sensing combined with cloud computing (Martin et al. 2021; Smedlund et al. 2018). Further, the emergence of decentralized technologies as an enabler to leverage data ecosystems is rarely in the focus of existing research, although it is already partly used in emerging data ecosystems and data initiatives. Further research on actually applying these technologies to innovate novel offerings in different application scenarios should be a research priority in the future. Also, practice-oriented research should accompany the ongoing process of creating and establishing reference architectures, including standards and protocols on the inter-organizational transmission of data. In doing so, these studies should reflect not only technical and legal requirements but also consider socio-organizational requirements for successful innovation in data ecosystems.

RA7: *Leveraging technological concepts to ensure seamless transmission of data across organizational boundaries for the generation of innovative offerings*

Second, issues of protecting data are increasingly rising and hamper novel collaborative service offerings. For one thing, data security is a very relevant requirement to protect critical information from external parties of the ecosystem. Further, current research lacks offering insights into how actors in a data ecosystem can jointly innovate and operate data-driven offerings while preserving their individual intellectual property. For both of these challenges, we enquire about further research projects building on the technological approaches presented in the "engineering" research stream and linking these with application scenarios as presented in the "innovating" subsection. For example, the data ecosystems field yet lacks research that critically evaluates the affordances and constraints of DLT and cryptography in the context of data ecosystems. Another exemplary technology is federated machine learning, which allows for collaboratively training and improving predictive models without any data transfer and exchange between entities within the ecosystem. Despite research on these technologies in IS and related disciplines being manifold, linking these insights with the particular challenges of data ecosystems is missing in today's body of knowledge and, thus, is a promising avenue for future research.

RA8: *Using technologies as an enabler to overcome intellectual property issues within the data ecosystem and data security issues regarding external parties*

Conclusion

This paper set out to review existing research on data ecosystems, which we identified as the focal concept of an emerging and highly relevant research field (Gelhaar et al. 2021a; Oliveira et al. 2019). By conducting an integrative and systematic literature review (vom Brocke et al. 2009; Torraco 2005), we were able to make two main contributions: First, we identified three streams of research (*innovating* novel products and services, *engineering* by leveraging technology and information, *collaborating* among multiple actors within a data ecosystem) and their key themes. In analyzing the literature, it became evident that data ecosystem research exhibits a high degree of fragmentation and that it lacks structural clarity to direct and focus future research. Due to the advanced availability of research on data ecosystems in the research community of information systems, compared to other research communities (e.g., production management), our results are mainly drawn from literature from this field. However, the transferability of the results is certainly given due to the phenomenon being positioned at the intersection of various research communities (e.g., service science, production management). Second, we used the review to extract key questions for further research on data ecosystems along the intersection of three research streams. In doing so, we particularly encourage research to take an interdisciplinary approach and "connect the dots" with contributions spanning across the identified demarcation lines.

This study offers theoretical *implications* that contribute to a deeper understanding of data ecosystem research and, thus, to the ongoing debate on co-creating and realizing value from data in digital offerings. We compiled and structured a comprehensive set of research on the concept of data ecosystems and synthesized the existing body of knowledge, as such a review has so far been omitted. The identified research streams and themes might help both specialized researchers and also a broader scientific audience to grasp the idea of "data ecosystems" and reproduce the state-of-the-art of the related academic debate. Furthermore, the results of the literature review point out the linkage of data ecosystems to related concepts, which might stimulate the transfer of knowledge in the future. With the proposed research agenda, we want to catalyze further research on data ecosystems.

Our research certainly comes with some *limitations*. First, we can hardly argue for completeness regarding the representation of the body of literature on data ecosystems in our review. Despite using a conceptual and systematic sampling approach, the identified literature sample is limited to the search phrase and set of chosen outlets. In particular, the computer science discipline seems underrepresented compared to its contribution potential. Second, even though we incorporated established research methods for conducting a literature review (vom Brocke et al. 2009; Webster and Watson 2002) and conducting qualitative content analysis (Mayring 2004), the analysis and interpretation of our data are somewhat influenced by the subjective assessment of the authors. Finally, the proposed avenues for future research were derived primarily from a theoretical point of view. Thus, we aim to triangulate the proposed research agenda with an upcoming interview study to catalyze further research on data ecosystems in close dialogue with practitioners.

Overall, we see great potential in the rapidly growing field of data ecosystem research. We aim that our work contributes to a better understanding of data ecosystems and that it introduces this important topic to a broader audience and, thereby, inspires scholars to join the academic debate. The proposed research agenda constitutes a promising starting point for further research endeavors in IS research and beyond. Based on future evidence, we as scholars may be able to better understand the implications of collaborating in data ecosystems, and thus, assist the economy and society to "get the engine running".

Acknowledgements

This work has been supported by the German Federal Ministry of Education and Research through the research project "bi.smart" (grant no. 02J19B041).

References

* Abbas, Y., Martinetti, A., Moerman, J. J., Hamberg, T., and van Dongen, L. A. M. 2020. "Do You Have Confidence in How Your Rolling Stock Has Been Maintained? A Blockchain-Led Knowledge-Sharing Platform for Building Trust between Stakeholders," International Journal of Information Management (55). (https://doi.org/10.1016/j.ijinfomgt.2020.102228).

Adner, R. 2017. "Ecosystem as Structure: An Actionable Construct for Strategy," Journal of Management (43:1), pp. 39–58. (https://doi.org/10.1177/0149206316678451).

* Anke, J., Poeppelbuss, J., and Alt, R. 2020. "It Takes More than Two to Tango: Identifying Roles and Patterns in Multi-Actor Smart Service Innovation," Schmalenbach Business Review, pp. 599–634. (https://doi.org/10.1007/s41464-020-00101-2).

* Aulkemeier, F., Iacob, M.-E., and van Hillegersberg, J. 2019. "Platform-Based Collaboration in Digital Ecosystems," Electronic Markets (29:4), pp. 597–608. (https://doi.org/10.1007/s12525-019-00341-2).

* Azkan, C., Möller, F., Meisel, L., and Otto, B. 2020. "Service Dominant Logic Perspective on Data Ecosystems - A Case Study Based Morphology," ECIS 2020 Proceedings.

Baldwin, C. Y., and Clark, K. B. 2000. Design Rules: The Power of Modularity, (Vol. 1.), MIT Press.

* Bastiaansen, H. J. M., Kollenstart, M., Dalmolen, S., and van Engers, T. M. 2020. "User-Centric Network-Model for Data Control with Interoperable Legal Data Sharing Artefacts," in PACIS 2020 Proceedings.

Beimborn, D., Miletzki, T., and Wenzel, S. 2011. "Platform as a Service (PaaS)," Business and Information Systems Engineering (3:6), pp. 381–384. (https://doi.org/10.1007/s12599-011-0183-3).

Benz, C., Riefle, R., and Schwarz, C. 2021. "Co-creating Value in B2B Platform Ecosystems–Towards a Deeper Understanding of the Emergence and Nature of Actor Engagement," 16th International Conference on Wirtschaftsinformatik, pp 236-242. Springer, Cham.

Benz, C., and Seebacher, S. 2018. "Open innovation in ecosystems–a service science perspective on open innovation," International Conference on Exploring Service Science, pp. 112-124. Springer, Cham.

* Beverungen, D., Breidbach, C. F., Poeppelbuss, J., and Tuunainen, V. K. 2019. "Smart Service Systems: An Interdisciplinary Perspective," Information Systems Journal (29:6, SI), pp. 1201–1206. (https://doi.org/10.1111/isj.12275).

* Beverungen, D., Kundisch, D., and Wünderlich, N. 2020. "Transforming into a Platform Provider: Strategic Options for Industrial Smart Service Providers," Journal of Service Management. (https://doi.org/10.1108/JOSM-03-2020-0066).

* Beverungen, D., Müller, O., Matzner, M., Mendling, J., and vom Brocke, J. 2019. "Conceptualizing Smart Service Systems," Electronic Markets (29:1), pp. 7–18. (https://doi.org/10.1007/s12525-017-0270-5).

Böhmann, T., Leimeister, J. M., and Möslein, K. 2014. "Service Systems Engineering," Business & Information Systems Engineering (6:2), pp. 73–79. (https://doi.org/10.1007/s12599-014-0314-8).

* Breidbach, C., Choi, S., Ellway, B., Keating, B. W., Kormusheva, K., Kowalkowski, C., Lim, C., and Maglio, P. 2018. "Operating without Operations: How Is Technology Changing the Role of the Firm?," Journal of Service Management (29:5), pp. 809–833. (https://doi.org/10.1108/JOSM-05-2018-0127).

vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., and Cleven, A. 2009. "Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process," in ECIS 2009 Proceedings.

* Busquets, J. 2010. "Orchestrating Smart Business Network Dynamics for Innovation," European Journal of Information Systems (19:4), pp. 481–493. (https://doi.org/10.1057/ejis.2010.19).

Cooper, H. M. 1988. "Organizing Knowledge Syntheses: A Taxonomy of Literature Reviews," Knowledge in Society (1:1), Springer, p. 126. (https://doi.org/10.1007/BF03177550).

* Curry, E., and Sheth, A. 2018. "Next-Generation Smart Environments: From System of Systems to Data Ecosystems," IEEE Intelligent Systems (33:3), IEEE, pp. 69–76. (https://doi.org/10.1109/MIS.2018.033001418).

Deloitte Insights. 2019. Data Ecosystems: How Third-Party Information Can Enhance Data Analytics, pp. 66–72.

* Elkoumy, G., Fahrenkrog-Petersen, S. A., Dumas, M., Laud, P., Pankova, A., and Weidlich, M. 2020. "Secure Multi-Party Computation for Inter-Organizational Process Mining," Lecture Notes in Business Information Processing (387), pp. 166–181. (https://doi.org/10.1007/978-3-030-49418-6_11).

* Faber, A., Rehm, S.-V., Hernandez-Mendez, A., and Matthes, F. 2019. "Collectively Constructing the Business Ecosystem: Towards Crowd-Based Modeling for Platforms and Infrastructures," Lecture

Notes in Business Information Processing (363), pp. 158–172. (https://doi.org/10.1007/978-3-030-26169-6_8).
Forbes. 2019. "Data Is the New Oil - And That's a Good Thing." (https://www.forbes.com/sites/forbestechcouncil/2019/11/15/data-is-the-new-oil-and-thats-a-good-thing/?sh=66ec19097304, accessed March 15, 2021).
Gelhaar, J., Groß, T., and Otto, B. 2021a. "A Taxonomy for Data Ecosystems," 2021 54th Hawaii International Conference on System Sciences (HICSS), pp. 6113–6122. (https://doi.org/10.24251/hicss.2021.739).
Gelhaar, J., Gürpinar, T., Henke, M., and Otto, B. 2021b. "Towards a Taxonomy of Incentive Mechanisms for Data Sharing in Data Ecosystems," PACIS 2021 Proceedings.
* Gelhaar, J., and Otto, B. 2020. "Challenges in the Emergence of Data Ecosystems," PACIS 2020 Proceedings.
Gliem, J. A., and Gliem, R. R. 2003. "Calculating, Interpreting, and Reporting Cronbach's Alpha Reliability Coefficient for Likert-Type Scales," in Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education.
* Gosain, S., Malhotra, A., and El Sawy, O. A. 2004. "Coordinating for Flexibility in E-Business Supply Chains," Journal of Management Information Systems (21:3), pp. 7–45. (https://doi.org/10.1080/07421222.2004.11045816).
* Goul, M. 2018. "APC Forum: Poised between 'a Wild West of Predictive Analytics' and 'an Analytics of Things Westworld Frontier,'" MIS Quarterly Executive (17:4), pp. 333–347.
* Gozman, D., Hedman, J., and Sylvest, K. 2018. "Open Banking: Emergent Roles, Risks & Opportunities," ECIS 2018 Proceedings.
* Hahn, C., Huntgeburth, J., and Zarnekow, R. 2016. "Leverage Once, Earn Repeatedly—Capabilities for Creating and Appropriating Value in Cloud Platform Ecosystems," Lecture Notes in Information Systems and Organisation (24), pp. 143–164. (https://doi.org/10.1007/978-3-319-62636-9_10).
* Harlamova, M., and Kirikova, M. 2018. "Trust Handling Framework for Networks in Cyber Physical Systems of Industry 4.0," Lecture Notes in Business Information Processing (330), pp. 29–43. (https://doi.org/10.1007/978-3-319-99951-7_3).
* Hein, A., Böhm, M., and Krcmar, H. 2018. "Platform Configurations within Information Systems Research: A Literature Review on the Example of IoT Platforms," MKWI 2018 - Multikonferenz Wirtschaftsinformatik, pp. 465–476.
Hein, A., Schreieck, M., Riasanow, T., Setzke, D. S., Wiesche, M., Böhm, M., and Krcmar, H. 2020. "Digital Platform Ecosystems," Electronic Markets (30:1), pp. 87–98. (https://doi.org/10.1007/s12525-019-00377-4).
* Hein, A., Weking, J., Schreieck, M., Wiesche, M., Böhm, M., and Krcmar, H. 2019. "Value Co-Creation Practices in Business-to-Business Platform Ecosystems," Electronic Markets (29:3), pp. 503–518. (https://doi.org/10.1007/s12525-019-00337-y).
Heinz, D., Benz, C., Bode, J., Hunke, F., and Satzger, G. 2022a. "Exploring the Potential of Smart Service Systems: A Multi-Actor View on Affordances and Their Actualization," Journal of Service Management Research (SMR).
Heinz, D., Hunke, F., and Breitschopf, G. F. 2021. "Organizing for Digital Innovation and Transformation: Bridging Between Organizational Resilience and Innovation Management," 16th International Conference on Wirtschaftsinformatik, pp 548-564. Springer, Cham.
Heinz, D., Park, H.-R., Benz, C., Satzger, G. 2022b. "Innovating Smart Product-Service Systems in Manufacturing SMEs: Current Practices, Affordances, and Constraints," 24th IEEE Conference on Business Informatics (IEEE CBI 2022). (https://doi.org/10.5445/IR/1000146476).
* Hirt, R., and Kühl, N. 2018. "Cognition in the Era of Smart Service Systems: Inter-Organizational Analytics through Meta and Transfer Learning," ICIS 2018 Proceedings.
* Huang, P., Ceccagnoli, M., Forman, C., and Wu, D. J. 2009. "When Do ISVs Join a Platform Ecosystem? Evidence from the Enterprise Software Industry," ICIS 2009 Proceedings.
* Huber, R. X. R., Püschel, L. C., and Röglinger, M. 2019. "Capturing Smart Service Systems: Development of a Domain-Specific Modelling Language," Information Systems Journal (29:6), pp. 1207–1255. (https://doi.org/10.1111/isj.12269).
Hunke, F., Heinz, D., and Satzger, G. 2022. "Creating customer value from data: foundations and archetypes of analytics-based services," Electronic Markets. (https://doi.org/10.1007/s12525-021-00506-y).

Iansiti, M., and Levien, R. 2004. The Keystone Advantage: What the New Dynamics of Business Ecosystems Mean for Strategy, Innovation, and Sustainability., Boston: Harvard Business Press.
Jacobides, M. G., Cennamo, C., and Gawer, A. 2018. "Towards a Theory of Ecosystems," Strategic Management Journal (39:8), pp. 2255–2276. (https://doi.org/10.1002/smj.2904).
* Jarke, M. 2017. "Data Spaces: Combining Goal-Driven and Data-Driven Approaches in Community Decision and Negotiation Support," Lecture Notes in Business Information Processing (293), pp. 3–14. (https://doi.org/10.1007/978-3-319-63546-0_1).
* Kamble, S. S., Gunasekaran, A., Ghadge, A., and Raut, R. 2020. "A Performance Measurement System for Industry 4.0 Enabled Smart Manufacturing System in SMMEs- A Review and Empirical Investigation," International Journal of Production Economics (229). (https://doi.org/10.1016/j.ijpe.2020.107853).
* Kannisto, P., Hästbacka, D., and Marttinen, A. 2020. "Information Exchange Architecture for Collaborative Industrial Ecosystem," Information Systems Frontiers (22:3), pp. 655–670. (https://doi.org/10.1007/s10796-018-9877-0).
* Khuntia, J., Mithas, S., and Agarwal, R. 2017. "How Service Offerings and Operational Maturity Influence the Viability of Health Information Exchanges," Production and Operations Management (26:11), pp. 1989–2005. (https://doi.org/10.1111/poms.12735).
* Kieseberg, P., and Weippl, E. 2018. "Security Challenges in Cyber-Physical Production Systems," Lecture Notes in Business Information Processing (302), pp. 3–16. (https://doi.org/10.1007/978-3-319-71440-0_1).
* Kipper, L. M., Furstenau, L. B., Hoppe, D., Frozza, R., and Iepsen, S. 2020. "Scopus Scientific Mapping Production in Industry 4.0 (2011–2018): A Bibliometric Analysis," International Journal of Production Research (58:6), 2-4, pp. 1605–1627. (https://doi.org/10.1080/00207543.2019.1671625).
* Kohtamäki, M., and Rajala, R. 2016. "Theory and Practice of Value Co-Creation in B2B Systems," Industrial Marketing Management (56), pp. 4–13. (https://doi.org/10.1016/j.indmarman.2016.05.027).
* Kusiak, A. 2018. "Smart Manufacturing," International Journal of Production Research (56:1–2), pp. 508–517. (https://doi.org/10.1080/00207543.2017.1351644).
Landis, J. R., and Koch, G. G. 1977. "The Measurement of Observer Agreement for Categorical Data," Biometrics, pp. 159–174.
* Lee, S. U., Zhu, L., and Jeffery, R. 2017. "Data Governance for Platform Ecosystems: Critical Factors and the State of Practice," PACIS 2017 Proceedings.
* Lee, Y. W., Madnick, S. E., Wang, R. Y., Wang, F. L., and Zhang, H. 2014. "A Cubic Framework for the Chief Data Officer: Succeeding in a World of Big Data," MIS Quarterly Executive (13:1).
Levy, Y., and Ellis, T. J. 2006. "A systems approach to conduct an effective literature review in support of information systems research," Informing Science (9).
Li, A. Q., Kumar, M., Claes, B., and Found, P. 2020. "The State-of-the-Art of the Theory on Product-Service Systems," International Journal of Production Economics (222). (https://doi.org/10.1016/j.ijpe.2019.09.012).
* Lim, C., and Maglio, P. P. 2018. "Data-Driven Understanding of Smart Service Systems through Text Mining," Service Science (10:2), pp. 154–180. (https://doi.org/10.1287/serv.2018.0208).
* Lindman, J., Kinnari, T., and Rossi, M. 2014. "Industrial Open Data: Case Studies of Early Open Data Entrepreneurs," 2014 47th Hawaii International Conference on System Sciences (HICSS), pp. 739–748. (https://doi.org/10.1109/HICSS.2014.99).
Lis, D., and Otto, B. 2020. "Data Governance in Data Ecosystems – Insights from Organizations," AMCIS 2020 Proceedings.
Lusch, R. F., and Nambisan, S. 2015. "Service Innovation: A Service-Dominant Logic Perspective," MIS Quarterly (39:1), pp. 155–175. (https://doi.org/10.25300/MISQ/2015/39.1.07).
* Lüttenberg, H., Bartelheimer, C., and Beverungen, D. 2018. "Designing Predictive Maintenance for Agricultural Machines," ECIS 2018 Proceedings.
* Markus, M. L., and Bui, Q. N. 2012. "Going Concerns: The Governance of Interorganizational Coordination Hubs," Journal of Management Information Systems (28:4), pp. 163–198. (https://doi.org/10.2753/MIS0742-1222280407).
Martin, D, Kühl, N., and Satzger, G. 2021. "Virtual Sensors," Business & Information Systems Engineering (63:3), pp. 315-323. (https://doi.org/10.1007/s12599-021-00689-w).
Mayring, P. 2004. "Qualitative Content Analysis," A Companion to Qualitative Research, pp. 159–176.
McKinsey Analytics. 2020. Four Ways to Accelerate the Creation of Data Ecosystems.

* Meroni, G., and Plebani, P. 2018. "Combining Artifact-Driven Monitoring with Blockchain: Analysis and Solutions," Lecture Notes in Business Information Processing (316), pp. 103–114. (https://doi.org/10.1007/978-3-319-92898-2_8).
* Meroni, G., Plebani, P., and Vona, F. 2019. "Trusted Artifact-Driven Process Monitoring of Multi-Party Business Processes with Blockchain," Lecture Notes in Business Information Processing (361), pp. 55–70. (https://doi.org/10.1007/978-3-030-30429-4_5).
* Moghaddam, M., and Nof, S. Y. 2018. "Collaborative Service-Component Integration in Cloud Manufacturing," International Journal of Production Research (56:1–2), pp. 677–691. (https://doi.org/10.1080/00207543.2017.1374574).
Moore, J. F. 1993. "Predators and Prey: A New Ecology of Competition," Harvard Business Review (71:3), pp. 75–86.
Nambisan, S., Lyytinen, K., and Yoo, Y. 2020. Handbook of Digital Innovation, Edward Elgar Publishing.
* Oliveira, M. I. S., Lima, G. de F. B., and Lóscio, B. F. 2019. "Investigations into Data Ecosystems: A Systematic Mapping Study," Knowledge and Information Systems (61:2), pp. 589–630. (https://doi.org/10.1007/s10115-018-1323-6).
* Oliveira, M. I. S., and Lóscio, B. F. 2018. "What Is a Data Ecosystem?," ACM International Conference Proceeding Series (18), pp. 1–9. (https://doi.org/10.1145/3209281.3209335).
Oliveira, M. I. S., Oliveira, L. E. R. A., Lima, G. de F. B., and Lóscio, B. F. 2017. "A Platform for Supporting Open Data Ecosystems," Lecture Notes in Business Information Processing (291), pp. 313–337. (https://doi.org/10.1007/978-3-319-62386-3_15).
Okoli, C., and Schabram, K. 2010. "A guide to conducting a systematic literature review of information systems research," Working Papers on Information Systems. (https://doi.org/10.2139/ssrn.1954824).
* Osterrieder, P., Budde, L., and Friedli, T. 2020. "The Smart Factory as a Key Construct of Industry 4.0: A Systematic Literature Review," International Journal of Production Economics (221). (https://doi.org/10.1016/j.ijpe.2019.08.011).
* Otto, B., and Jarke, M. 2019. "Designing a Multi-Sided Data Platform: Findings from the International Data Spaces Case," Electronic Markets (29:4), pp. 561–580. (https://doi.org/10.1007/s12525-019-00362-x).
* Pant, V., and Yu, E. 2018. "Getting to Win-Win in Industrial Collaboration under Coopetition: A Strategic Modeling Approach," Lecture Notes in Business Information Processing (330), pp. 47–66. (https://doi.org/10.1007/978-3-319-99951-7_4).
* Papert, M., and Pflaum, A. 2017. "Development of an Ecosystem Model for the Realization of Internet of Things (IoT) Services in Supply Chain Management," Electronic Markets (27:2), pp. 175–189. (https://doi.org/10.1007/s12525-017-0251-8).
* Patnayakuni, R., Seth, N., and Rai, A. 2006. "Building Social Capital with It and Collaboration in Supply Chains: An Empirical Investigation," ICIS 2006 Proceedings.
* Pauli, T., Fielt, E., and Matzner, M. 2021. "Digital Industrial Platforms," Business & Information Systems Engineering (63:2), pp. 181–190. (https://doi.org/10.1007/s12599-020-00681-w).
* Ponte, D. 2015. "Enabling an Open Data Ecosystem: Preliminary Findings from the Market," ECIS 2015 Proceedings.
Porter, M. E., and Heppelmann, J. E. 2014. "How Smart, Connected Products Are Transforming Competition," Harvard Business Review (92:11), pp. 64–88.
* Rasouli, M. R. 2020. "An Architecture for IoT-Enabled Intelligent Process-Aware Cloud Production Platform: A Case Study in a Networked Cloud Clinical Laboratory," International Journal of Production Research (58:12), pp. 3765–3780. (https://doi.org/10.1080/00207543.2019.1634847).
* Rehm, S.-V., and Goel, L. 2013. "Composite Boundary Objects in Interorganizational Innovation Activities," ECIS 2013 Proceedings.
De Reuver, M., Sørensen, C., and Basole, R. C. 2018. "The Digital Platform: A Research Agenda," Journal of Information Technology (33:2), pp. 124–135. (https://doi.org/10.1057/s41265-016-0033-3).
* Saarijärvi, H., Grönroos, C., and Kuusela, H. 2014. "Reverse Use of Customer Data: Implications for Service-Based Business Models," Journal of Services Marketing (28:7), pp. 529–537. (https://doi.org/10.1108/JSM-05-2013-0111).
* Saarikko, T., Westergren, U. H., and Blomquist, T. 2016. "The Inter-Organizational Dynamics of a Platform Ecosystem: Exploring Stakeholder Boundaries," 2016 49th Hawaii International Conference on System Sciences (HICSS), pp. 5167–5176. (https://doi.org/10.1109/HICSS.2016.639).

* Sandberg, J., Holmström, J., and Lyytinen, K. 2020. “Digitization and Phase Transitions in Platform Organizing Logics: Evidence from the Process Automation Industry,” MIS Quarterly (44:1), pp. 129–153. (https://doi.org/10.25300/MISQ/2020/14520).

Schreieck, M., Wiesche, M., and Krcmar, H. 2016. “Design and Governance of Platform Ecosystems - Key Concepts and Issues for Future Research,” ECIS 2016 Proceedings.

* Schulz, K. F., and Freund, D. 2019. “A Multichain Architecture for Distributed Supply Chain Design in Industry 4.0,” Lecture Notes in Business Information Processing (339), pp. 277–288. (https://doi.org/10.1007/978-3-030-04849-5_25).

Schüritz, R., Seebacher, S., and Dorner, R. 2017. “Capturing Value from Data: Revenue Models for Data-Driven Services,” 2017 50th Hawaii International Conference on System Sciences (HICSS), pp. 5348–5357. (https://doi.org/10.24251/hicss.2017.648).

* Shehab, E., Fowler, C., Rodriguez Gil, A., Abdalla, H., Darwish, M., Abdulhafed, H., Ahmed, A., Ahouie, H., Alechnovic, A., Paumes, C., Tacchini, E., Urbane, I., and Richardson, M. 2013. “Enhancement of Product Information Collaboration and Access in the Aerospace Industry,” International Journal of Production Research (51:11), pp. 3225–3240. (https://doi.org/10.1080/00207543.2012.754965).

* Shim, J. P., Avital, M., Dennis, A. R., Rossi, M., Sørensen, C., and French, A. 2019. “The Transformative Effect of the Internet of Things on Business and Society,” Communications of the Association for Information Systems (44:1), pp. 129–140. (https://doi.org/10.17705/1CAIS.04405).

* Smedlund, A. 2016. “Digital Health Platform Complementor Motives and Effectual Reasoning,” 2016 49th Hawaii International Conference on System Sciences (HICSS), (B. T. X. Sprague R.H., ed.), IEEE Computer Society, pp. 1614–1623. (https://doi.org/10.1109/HICSS.2016.204).

* Smedlund, A., Ikävalko, H., and Turkama, P. 2018. “Firm Strategies in Open Internet of Things Business Ecosystems: Framework and Case Study,” 2018 51st Hawaii International Conference on System Sciences (HICSS), pp. 1591–1600. (https://doi.org/10.24251/hicss.2018.199).

Star, S. L. 2010. “This Is Not a Boundary Object: Reflections on the Origin of a Concept,” Science Technology and Human Values (35:5), pp. 601–617. (https://doi.org/10.1177/0162243910377624).

* Teixeira, J. A., Mian, S. Q., and Hytti, U. 2016. “Cooperation among Competitors in the Open-Source Arena: The Case of OpenStack,” ICIS 2016 Proceedings.

Tiwana, A., Konsynski, B., and Bush, A. A. 2010. “Platform Evolution: Coevolution of Platform Architecture, Governance, and Environmental Dynamics,” Information Systems Research (21:4), pp. 675–687. (https://doi.org/10.1287/isre.1100.0323).

Torraco, R. J. 2005. “Writing Integrative Literature Reviews: Guidelines and Examples,” Human Resource Development Review (4:3), pp. 356–367. (https://doi.org/10.1177/1534484305278283).

* Tuunanen, T., Kazan, E., Salo, M., Leskelä, R.-L., and Gupta, S. 2019. “From Digitalization to Cybernization: Delivering Value with Cybernized Services,” Scandinavian Journal of Information Systems (31:2), pp. 83–96.

* Wareham, J., Fox, P. B., and Giner, J. L. C. 2014. “Technology Ecosystem Governance,” Organization Science (25:4), pp. 1195–1215. (https://doi.org/10.1287/orsc.2014.0895).

* Weber, K., Otto, B., and Österle, H. 2009. “Ones Size Does Not Fit All - A Contingency Approach to Data Governance,” Journal of Data and Information Quality (1:1). (https://doi.org/10.1145/1515693.1515696).

* Weber, P., Hiller, S., and Lasi, H. 2020. “Identifying Business Potentials within an IoT Ecosystem - An Explorative Case Study in the Industrial Domain,” AMCIS 2020 Proceedings.

Webster, J., and Watson, R. T. 2002. “Analyzing the Past to Prepare for the Future: Writing a Literature Review,” MIS Quarterly (26:2), xiii–xxiii. (https://doi.org/10.2307/4132319).

Weill, P., and Woerner, S. L. 2015. “Thriving in an Increasingly Digital Ecosystem,” MIT Sloan Management Review (56:4), pp. 27–34.

Wiesböck, F., and Hess, T. 2020. “Digital Innovations: Embedding in Organizations,” Electronic Markets (30:1), pp. 75–86. (https://doi.org/10.1007/s12525-019-00364-9).

* Zheng, P., Lin, Y., Chen, C. H., and Xu, X. 2019. “Smart, Connected Open Architecture Product: An IT-Driven Co-Creation Paradigm with Lifecycle Personalization Concerns,” International Journal of Production Research (57:8), pp. 2571–2584. (https://doi.org/10.1080/00207543.2018.1530475).