RESEARCH PAPER

Designing a multi-sided data platform: findings from the International Data Spaces case

Boris Otto[1,2] · Matthias Jarke[3,4]

Abstract

The paper presents the findings from a 3-year single-case study conducted in connection with the International Data Spaces (IDS) initiative. The IDS represents a multi-sided platform (MSP) for secure and trusted data exchange, which is governed by an institutionalized alliance of different stakeholder organizations. The paper delivers insights gained during the early stages of the platform's lifecycle (i.e. the platform design process). More specifically, it provides answers to three research questions, namely how alliance-driven MSPs come into existence and evolve, how different stakeholder groups use certain governance mechanisms during the platform design process, and how this process is influenced by regulatory instruments. By contrasting the case of an alliance-driven MSP with the more common approach of the keystone-driven MSP, the results of the case study suggest that different evolutionary paths can be pursued during the early stages of an MSP's lifecycle. Furthermore, the IDS initiative considers trust and data sovereignty more relevant regulatory instruments compared to pricing, for example. Finally, the study advances the body of scientific knowledge with regard to data being a boundary resource on MSPs.

Keywords Multi-sided data platform · Case study research · International Data Spaces · Data sovereignty · Alliance-driven platform · Data resource

✉ Boris Otto
    Boris.Otto@tu-dortmund.de; Boris.Otto@isst.fraunhofer.de

    Matthias Jarke
    jarke@dbis.rwth-aachen.de; matthias.jarke@fit.fraunhofer.de

[1] Chair for Industrial Information Management, TU Dortmund University, Joseph-von-Fraunhofer-Str. 2-4, D-44227 Dortmund, Germany

[2] Fraunhofer ISST, Emil-Figge-Str. 91, D-44227 Dortmund, Germany

[3] Information Systems Group (Informatik 5), RWTH Aachen University, Ahornstr. 55, D-52074 Aachen, Germany

[4] Fraunhofer FIT, Birlinghoven Castle, D-53754 Sankt Augustin, Germany

Introduction

Motivation

Spurred by the rise of technology companies such as Apple and Facebook, digital platforms are receiving increasing attention both in the scientific and in the practitioners' community. Generally speaking, such platforms allow different parties to build complementary products and services (Cusumano 2010; Gawer and Cusumano 2014). Two basic forms of digital platforms can be distinguished: two-sided platforms and multi-sided platforms. Whereas two-sided platforms mediate between two groups of users, e.g. buyers and sellers from a certain domain (Hagiu 2009), multi-sided platforms (MSPs) bring together multiple groups of users (i.e. not just providers and consumers of products and services). Typically, an

MSP is provided by a "keystone" firm owning the platform, while complementors use the platform to provide additional offerings (De Reuver et al. 2018). There is consensus in the research community that MSPs are of sociotechnical nature, as they comprise various technical and organizational facets and multiple forms of interaction of the MSP with its dynamic environment on both technical and organizational level (Tiwana et al. 2010).

Consequently, the evolution of an MSP is a complex, dynamic process, which is not fully understood yet (Staykova and Damsgaard 2017). The lifecycle of an MSP typically comprises three major phases: the first phase is about the platform's design based on a technological innovation; during the second phase, the platform is adopted and used by different user groups; the third phase then comprises platform scaling and growth activities (Staykova and Damsgaard 2015). While current research has been focusing mainly on the adoption and use of MSPs, the design phase – and especially the (very) early stages of the MSP emerging – are still relatively unexplored. This is somewhat surprising, as Gawer (2014) and de Reuver et al. (2018), for example, consider the nascence (or "genesis") of MSPs a very promising and worthwhile area of research.

MSPs facilitate the establishment of business ecosystems, which are formed by the users interacting on the platform. Researchers have examined a number of case studies on MSPs in which a keystone firm owns and governs the platform (Ondrus et al. 2015; Tan et al. 2015); so these are "keystone-driven MSPs". However, the platform landscape is becoming more and more diverse, with other, more complex governance and ownership structures being observable in different domains. De Reuver et al. (2018) have found that in many cases there is no single platform provider, but that the platform is jointly designed and "shaped" by multiple actors. In this context, Tiwana et al. (2010) point to the importance of distinguishing platforms owned by a single firm from platforms characterized by some form of "shared ownership". Shared ownership materializes in multiple organizational forms, among them the alliance. Gawer (2009), for example, identified some MSPs from the supply chain management domain that are "shared among firms that are part of a formal alliance". Such "alliance-driven MSPs", which are characterized by shared ownership and governance (e.g. a joint-venture company or industry association), as well as decentral platform governance models have been neglected by academic research so far (De Reuver et al. 2018).

Therefore, the paper has been motivated by two research gaps: first, the limited understanding and knowledge of the genesis of MSPs; and second, the lack of research regarding MSPs based on shared ownership. As for the latter, the paper focuses on alliance-driven MSPs, given the many alliance-driven MSPs being launched at present in various domains. Three prominent examples can be found in the mobility sector alone: Caruso,[1] a data-brokering platform for the automotive aftermarket; NEVADA,[2] a joint initiative under the umbrella of the VDA (German Association of the Automotive Industry) for sharing car-generated data; and HERE,[3] a geodata service provider jointly owned by automotive OEMs, their suppliers, and technology companies.

Research questions and approach

The paper addresses three research questions addressing the (very) early stages of the platform design phase:

RQ #1: How do alliance-driven MSPs come into existence and evolve?

In contrast to MSPs owned by a single organization, alliance-driven MSPs require specific governance mechanisms (De Reuver et al. 2018; Tiwana et al. 2010) to be established. Of particular interest are the decision-making processes taking place among the different actors during the early stages of the platform's lifecycle. Another area of interest refers to the different roles that the various stakeholders (i.e. users, owners, others) may assume during this phase.

This directly leads to the second research question:

RQ #2: How do the different stakeholder groups use certain governance mechanisms during the early phase of an alliance-driven MSP's lifecycle?

It must be noted that these governance mechanisms, which mainly aim at designing and building the platform, have to be distinguished from instruments used for regulation, which aim at fostering and controlling the adoption and use of the platform. Discussing the regulatory function of digital platforms in general, Boudreau and Hagiu (2009) found that specific research is needed with regard to instruments that go beyond pricing mechanisms. Thus, the third research question is:

RQ #3: How are regulatory instruments designed to foster the adoption and use of an alliance-driven MSP?

In this context, platform boundary resources are useful for studying both governance mechanisms and regulatory instruments. For conceptualizing platform boundary resources, the paper follows Ghazawneh and Henfridsson (2010). It defines boundary resources as resources that

[1] See https://www.caruso-dataplace.com/.
[2] See https://www.vda.de/en/topics/innovation-and-technology/data-security/what-is.html.
[3] See https://www.here.com.

facilitate relationships and interactions between different actors and user groups. The paper thereby takes up demands from de Reuver et al. (2018), for example, who suggest to focus on MSPs' boundary resources for studying the multifold interactions between MSP actors during the different platform lifecycle phases.

The International Data Spaces[4] (IDS) initiative is a joint effort of various international research institutes and industrial enterprises aiming at establishing a decentralized platform for secure and trusted data sharing. As a multi-sided data platform, the IDS represents an extreme case (Ridder 2017; Yin 2014) of an MSP, as the keystone actor here is not a single firm (e.g. a technology company) but a non-for-profit association forming an alliance of multiple organizations that may assume one or more roles on the platform (such as data provider, data consumer, research organization, software/service provider, auditing firm). Furthermore, the platform's focus on facilitating trusted data exchange and maintaining data sovereignty presents a research opportunity that allows examining a boundary resource (i.e. data) that is different from the ones investigated in previous studies. Data sharing and data exchange facilitated through an MSP needs to be governed, which calls for a comprehensive set of regulatory instruments.

The research activities carried out in this case study were guided by the principles of Action Design Research (ADR) (Sein et al. 2011). More specifically, the ADR approach pursued comprised multiple ADR cycles (i.e. project phases), each of which consists of four steps: 1) problem formulation; 2) building, intervention, and evaluation; 3) reflection and learning; and 4) formalization of learning (Sein et al. 2011). Thus, the case study is characterized by extensive interaction and collaboration between researchers and practitioners (Baskerville 1999), which allows for immediate creation and transfer of knowledge within the situational environment of the MSP's emergence.

The remainder of the paper is structured as follows: this introductory section is followed by an analysis of related work in the field of multi-sided platforms; the third section then outlines the research design applied and the research process conducted; the results of the case study are presented in the fourth section; the fifth section then discusses the case study findings and puts them in perspective with existing literature and theory; the paper concludes with a summary of the case study, its contribution to the body of scientific knowledge, remarks on the study's limitations, and an outlook to future research opportunities.

---

[4] The IDS initiative initially started out as "Industrial Data Space". In March 2018, it changed its name into "International Data Spaces" in order to emphasize the cross-industry approach pursued.

Related work

Constituents of an MSP

MSPs are sociotechnical constructs; they are both technical platforms and market intermediaries. Thus, describing an MSP requires both a technical architecture and an ecosystem architecture (Dal Bianco et al. 2014). The technical architecture consists of modules and components some of which remain stable during the platform's lifecycle, while other vary over time (Baldwin and Woodard 2009). For communication between components and for interaction between users, application programming interfaces (APIs) are used. APIs also allow third parties (i.e. complementors) to provide additional services (Ghazawneh and Henfridsson 2015).

The various actors engaging with each other on the platform constitute a business ecosystem (Gawer and Cusumano 2014; Schreieck et al. 2016). The platform owner provides the platform as a service to different user groups (i.e. sides) and to complementors. Many MSPs are provided and owned by a single entity, which is referred to as the "keystone firm". However, recent research has taken up on "multi-actor settings" (De Reuver et al. 2018), as more and more MSPs are based on such an approach. Apparently, control mechanisms differ from case to case. Governance instruments must be in place to coordinate the agents that together provide and own the platform.

Besides such an "internal" governance framework (i.e. the one that is concerned with the group of collective platform owners), a second governance framework is needed, which coordinates interaction between the platform owners and the various users of the platform as well as the complementors. That governance framework specifies decision-making rights with regard to using the platform and the services offered via interfaces (Boudreau and Hagiu 2009; Staykova and Damsgaard 2015; Tiwana et al. 2010). Furthermore, this second governance framework also defines platform access rights and, thus, specifies the platform's degree of openness (De Reuver et al. 2018; Ondrus et al. 2015).

Recent studies have advanced the understanding and knowledge of governance concepts and interaction patterns on such platforms by 1) examining the rules and instruments guiding governance and interaction activities, and 2) taking a boundary resource perspective on the topic. Regarding the former, Tiwana et al. (2010) specified governance and interaction patterns by introducing concepts such as leadership, ownership, and platform rules as elementary platform design constituents. Further research in this area seems useful though, as there is a demand to study not only mechanisms for regulation of MSPs, but also mechanisms for MSP adoption and use (i.e. mechanisms that go beyond the establishment of pricing instruments) (Boudreau and Hagiu 2009).

Lifecycle of an MSP

Previous research on the emergence and evolution of MSPs reached consensus in the sense that – at a high level of abstraction – the lifecycle of an MSP comprises three phases. Henfridsson and Bygstad (2013), for example, identified an evolution path of MSPs going from "innovation" to "adoption" to "scaling". Tan et al. (2015) looked at the ecosystem and the different sides that are attracted to a digital platform, finding that typically a platform matures from a two-sided model without interaction between the sides to a two-sided model with interaction to eventually an MSP model. Taking up on these findings, Tan et al. (2016) examined IT affordance and related activities, leading to the evolution of the ecosystem along the three phases.

The second phase of the lifecycle (i.e. platform adoption and use) has been the subject of investigation of a number of studies. For example, network effects and pricing concepts have been used as theoretical lenses to study the success or failure of MSPs (Evans 2003; Evans and Schmalensee 2013; Weyl 2010). Regarding the preceding (i.e. the design) phase, less research has been conducted, mainly taking an innovation or engineering view (Tan et al. 2016; Tura et al. 2018). Tiwana et al. (2010), for example, studied the relationships between different platform constituents (such as governance and architecture) during this phase. However, detailed insights as to how an MSP comes into existence and evolves during the very early stages of the design phase are missing so far. Given the growing importance of MSPs in the market, more research is needed in this regard (Gawer 2014). This demand is supported by widely accepted insights from the software and systems engineering domain that early design activities – such as conducting a requirement analysis – have the biggest impact on a system's success or failure (Hofmann and Lehner 2001).

Data as a boundary resource of an MSP

As far as the boundary resources are concerned, Henfridsson and Bygstad (2013) argue that this concept is helpful for studying patterns of interaction between the various groups and agents on a digital platform. Boundary resources are resources through which different agents create relationships and interact with each other in order to co-create value (Eaton et al. 2015). Dal Bianco et al. (2014) distinguish between technical and social platform boundary resources. Typical boundary resources are Application Programming Interfaces (APIs) and Software Development Kits (SDKs). Examples for social boundary resources are intellectual property rights and documentation of software services. Furthermore, boundary resources are not stable, but evolve over time. Eaton et al. (2015) coin the notion of "distributed tuning" to describe the process of continuous shaping and reshaping of boundary resources between the different platform actors and users. More recent research has suggested to

increasingly look at such boundary resources of digital platforms as a promising subject of analysis (De Reuver et al. 2018).

How organizations can exchange and share data has long since been an important research topic. The need for companies to exchange and share data has been a major motivation for the development of platforms mediating between suppliers and buyers of goods. Early two-sided data exchange solutions were facilitated by technological standards, such as EDIFACT or ANSI X.12. Gawer (2014) within her integrative platform framework identified traditional buyer-supplier relationships for which data is a technical enabler.

Around the turn of the millennium, electronic marketplaces emerged as intermediaries to reduce the complexity of the increasing need of n:m data exchange (Segev et al. 1999; Timmers 1998), in which data from multiple sources (n) can be bundled and utilized in contextualized presentations to multiple users (m). This intermediary function comprised – among other things – the mapping of the different message schemas of the various standardization initiatives that evolved. Motivated by the success of peer-to-peer-networks in the consumer realm, some researchers explored technologies, and even business models, for peer-to-peer based networks for data exchange in the industrial domain. Technological aspects of peer-to-peer data ecosystems, such as context exchange among different world views of organizations (Goh et al. 1999) or automation of data mappings in heterogeneous settings (Jarke et al. 2014), have been investigated since the late 1990s. In 2005, Franklin et al. (2005) noted the growing richness of digital media and proposed that users should be enabled to create their own "data space", where a free collection of data and media objects could be managed under a user specific network of semantic metadata. In 2010, the notion of the "data lake" was coined (LaPlante and Sharma 2016) and quickly received attention in the practitioners' community. Furthermore, some researchers investigated the role of data within platform based ecosystems (Immonen et al. 2014; Kontokosta 2013; Moiso and Minerva 2012). More recent research has dealt with the upcoming phenomenon of data platforms, mainly encouraged by the discourse around big data (Bharosa et al. 2019; Demchenko et al. 2014; Immonen et al. 2014). These studies focus mainly on platform architecture technology and data flows.

What extant literature is still lacking, however, are studies of data being the key resource on digital platforms. In particular, viewing data as a boundary resource – in order to gain insight into governance frameworks and regulatory mechanisms on digital platforms – is a topic that has not been addressed by the scientific community. More recent studies have addressed this gap in research though. Schreieck et al. (2016), for example, concluded that more research is needed on "how data is used to govern platform ecosystems in practice".

In this context, and in response to the research gaps outlined in the three previous subsections, the paper aims at providing detailed insights into the very early stages of the

design phase of alliance-driven MSPs, using the concept of the boundary resource and putting a special focus on data being the key resource in such a scenario.

Research design

Research context

The context of the research presented in this paper is provided by the IDS initiative. By proposing a reference architecture model for secure and trusted exchange of data that is applicable in various industries, IDS constitutes a blueprint for multi-sided data platforms. As a non-for-profit organization bundling user requirements and providing use cases for testing the IDS Reference Architecture Model, the International Data Spaces Association (IDSA) collaborates closely with the IDS research project, which is led by Fraunhofer and funded by the German Federal Ministry of Education and Research (BMBF). The IDSA consists of more than 90 member organizations from more than fifteen countries, representing different groups of users and stakeholders of the IDS. The list of members comprises user companies acting as data providers and/or data consumers, software and technology vendors, accounting and auditing firms, and non-for-profit organizations (e.g. other industry associations and research organizations). Fraunhofer, as the leading partner of the research activities, is one of the founding members of the IDSA. IDS stands for an alliance-driven MSP, as IDSA members share ownership of the Reference Architecture Model and the design process.

Generally speaking, single-case study research is appropriate for studying extreme cases (Yin 2014) or if a representative and/or complex case is at hand (Donmoyer 2000). The design of the IDS is both unique and complex: first, in contrast to the majority of similar cases, the IDS was not designed and developed by a single keystone entity, but by an alliance of multiple stakeholder groups; second, with its clear focus on data sharing, the IDS allows viewing data as platform boundary resources. To examine the use and management of data in a single-case study setting has been proven useful in the past. Shanks (1997), for example, argued that strategic data planning is a complex phenomenon that must be studied in its environmental context.

Conceptual framework

Following Ridder's (2017) categorization of case study designs, the research presented in this paper is exploratory and aims at closing "gaps and holes" in existing theory. With this purpose, case study research requires a conceptual framework which functions as a theoretical lens for data analysis and interpretation (Dul and Hak 2008, p. 36; Yin 2014). Based on the analysis of the related work, the study uses the conceptual framework described in Table 1.

The platform's ecosystem is formed by the IDS Association; its statutes and rules specify the ecosystem's governance policy, as they define rights and responsibilities related to the design of the IDS Reference Architecture Model and the use of its implementation. The IDS Reference Architecture Model (Otto et al. 2017; Otto et al. 2018) describes the platform as such. The data shared and exchanged via the IDS represents the boundary resource that creates relationships and facilitates interaction between the various actors using the platform. Regulatory instruments guide the adoption and use of the platform. The study uses the conceptual framework to analyze the early design phase of the IDS.

Research process

Generally speaking, ADR is applicable for sociotechnical design artifacts, as they require – in the various phases of the design process – knowledge from the environment (e.g. for requirement identification and artifact evaluation). The paper reports on a longitudinal study covering 3 years and two design cycles regarding the major artifacts, namely the platform architecture and the ecosystem governance model (see Table 2).

The four phases of the research process – namely "Rationale and Requirements", "Institutionalization and Use Cases", "Architecture Design", and "Ecosystem Design" – form the structure along which the case is presented in the following chapter.

Data collection and analysis

The design process used empirical data from the field for identifying design requirements, justifying design decisions, and evaluating design instantiations. Data collection took place in three different forms: first, eleven multilateral workshops (MWs, see Appendix Table 12) were organized to create consensus between different user groups representing different markets (data providers, data users, service/software providers etc.); during the workshops, focus groups served as the main method to derive design requirements and evaluate design artifacts (Tremblay et al. 2010); second, bilateral workshops (BWs, see Appendix Table 13) were organized for creating a more detailed understanding of individual requirements (in particular with regard to input from competing actors); third, use case projects (UCs, see Appendix Table 14) served as an environment to test and evaluate the design artifacts (in particular: the individual components of the IDS architecture) in real-world settings.

To document the results of the various multilateral and bilateral workshops, flipcharts and "brown paper" boards were used. After each workshop session, protocols and

Table 1 Conceptual framework

Theoretical concept	Definition in the case study	Supporting literature	IDS manifestation	RQ Relevance		
				RQ #1	RQ #2	RQ #3
Platform architecture	Conceptual blueprint describing how an extensible, software based system can be partitioned into stable and complementary components, and how these components interact with each other and with the user	(Baldwin and Woodard 2009) (Gawer and Cusumano 2014) (Tiwana et al. 2010)	IDS Reference Architecture Model (in particular: IDS SW Component Architecture)	X		X
Platform boundary resource	Resource allowing different actors to create relationships and interact with each on the platform	(Dal Bianco et al. 2014) (Ghazawneh and Henfridsson 2010) (Schreieck et al. 2016)	Data	X	X	X
Platform design	Process of configuring platform design elements when building the platform	(Peffers et al. 2007; Tura et al. 2018)	IDS design activities from 2015 to 2017	X		
Platform ecosystem	Alignment structure of the actors designing, developing and using the platform	(Adner 2017) (Dal Bianco et al. 2014) (Immonen et al. 2014) (Schreieck et al. 2016) (Tan et al. 2015)	IDS Association		X	
Ecosystem governance	Entirety of rules, responsibilities and decision-making rights affecting the behavior and interaction of the actors of the platform ecosystem	(Schreieck et al. 2016) (Tiwana et al. 2010) (Tura et al. 2018)	IDS ecosystem design		X	
Regulatory instruments	Instruments for fostering and controlling adoption and use of the platform	(Kevin J. Boudreau and Hagiu 2009)	Data sovereignty; interoperability; trust and security			X

Legend:

RQ Research question

X Concept required for answering the respective RQ

Table 2 Research process overview

	2015				2016				2017			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Project phases												
1) Rationale and Requirements	1	1	1									
2) Institutionalization and Use Case Implementation			1	1, 2	1, 2	1, 2						
3) Architecture Design				2	2	2	2	2, 3	3, 4	2	2	2
4) Ecosystem Design					1	1	2	2	2	2	2	3, 4
Milestones												
Base Connector (1st version)							X					
Ecosystem (1st and 2nd version								X			X	
IDS Association					X							
Reference Architecture Model (1st version)									X			
Trusted Connector (1st version)										X		

Legend:

Numbers in "Project phases" represent the respective ADR stages (Sein et al. 2011), i.e. 1) Problem Formulation; 2) Building, Intervention, and Evaluation; 3) Reflection and Learning; and 4) Formalization of Learning

X – Indicates the quarter of the year in which the respective milestone was reached

detailed workshop minutes were sent to the participants to gather feedback and comments.

International Data Spaces: case study results

Project phase #1 – Rationale and requirements

Description

The first phase of the research process was characterized by problem centricity activities, following the concept of entry points for a study in design science research as proposed by Peffers et al. (2007). In a first high-level meeting, stakeholders from selected industries, government, and research agreed that data is increasingly playing an enabler role and turning into a strategic resource in enterprises. Moreover, there was consensus that a platform for data exchange and data sharing had to be driven by the requirements of the platform users, and that data sovereignty of data owners and data providers had to be a central concept of the platform to be developed. Fraunhofer and the BMBF then invited a larger group of stakeholders to a second high-level meeting (see MW1 in Appendix Table 12) in order to confirm the preliminary findings and lay out a roadmap for follow-up activities. In MW1, one participant (the CIO of a large automobile manufacturer) articulated that data security had to be understood as an asset, and that data providers needed to make sure they do not end up only with carrying the cost of data management, but are enabled to take advantage of the benefits of the data economy. Other participants emphasized the need to

take a business perspective on the topic, as pursuing a purely technical approach would not be sufficient.

It was decided to establish a task force to collect requirements and identify possible use cases. In MW2, a first set of requirements was collected. Based on state-of-the-art knowledge regarding data space architectures and data exchange platforms, the industry representatives discussed requirements to be met by a multi-sided data platform in an open workshop session (cf. Table 3). The requirements agreed upon by the majority of participants were documented on a flipchart.

The list of requirements spans a wide range of aspects. It covered both functional requirements, such as integration of components or data flow traceability, and non-functional requirements, such as software component certification or trust.

In addition to the requirements, the task force defined a first set of possible roles relevant for the IDS ecosystem (cf. Table 4). It was agreed that the exchange of data basically should take place between a Data Owner providing the data and a Data User consuming the data, and that an intermediary, called "Broker", should bring together data supply and data demand by providing search and look-up functionality. Furthermore, it was decided that the Broker should be able to monitor data transactions (without being able to read payload information) and offer value-adding data services (e.g. for data analytics). The list of roles was completed by a Certification Agency making sure that each participant, and each software component used for implementing the IDS reference architecture, complies with IDS requirements.

It should be noted that the list of roles does not include any such role as a "platform operator" or "platform provider". The task force agreed that – in line with the conceptual design of

Table 3 Initial set of requirements

ID	Requirement
R1.01	Ecosystem must comprise data providers, data consumers, and intermediaries
R1.02	Onboarding of new participants to the ecosystem must be easy
R1.03	Ecosystem must be open, but participants and software endpoints must be certified
R1.04	Integration of third-party / open data providers must be possible
R1.05	Data rights must be clarified and protected
R1.06	Data heterogeneity must be supported
R1.07	Ecosystem must allow different data to be treated differently, depending on its security classification and its nature as an economic good
R1.08	Data services must be separable into basic services and value-added services
R1.09	Data flow traceability must be possible
R1.10	Data usage conditions must be manageable and remotely enforceable
R1.11	Platform must allow integration and use of sensor data (e.g. from autonomous vehicles)
R1.12	Platform must allow integration and use of existing technologies and standards
R1.13	Platform must be able to work without real-time support
R1.14	Flow of goods and flow of information must be coupled on a permanent basis
R1.15	Trust and security must always be ensured

Table 4 Initial set of roles

Role	Responsibilities
Data Owner	• Provides data
	• Defines usage rights and conditions of use (price etc.)
Data User	• Uses data for internal or external services
Provider/Broker	• Brings Data Owners and Data Users together
	• Provides ready-to-use templates for usage rights, usage, and pricing
	• Monitors data exchange and provides clearing and security
	• Offers on-top services (e.g. big data analytics)
Certification Agency	• Certifies IDS participants and software components

peer-to-peer networks – no central operator or provider was needed. Instead, it was decided that the IDS Reference Architecture Model should be open to be taken up by any market participant, leading to multiple implementations of the IDS. Finally, the task force decided to refine the requirements and design a first version of the IDS Architecture in three use case domains: logistics, automotive/mobility, and healthcare.

Conceptual analysis

Table 5 gives an overview of the conceptual analysis of the "Rationale and Requirements" phase. All concepts, except for ecosystem governance, were addressed in this very early design phase. With four initial roles being identified, this first project phase already addressed the multi-sidedness of the initiative.

The identification of requirements referred to two types of requirements: one the one hand, data sharing and data exchange requirements were defined; on the other hand, requirements for the provisioning and use of the platform itself were identified.

Project phase #2 – Institutionalization and use cases

Description

The second phase of the research process mainly focused on the institutionalization of activities and the analysis and implementation of use cases. The former resulted in the foundation of a non-for-profit association, the IDSA. The work within the IDSA is organized in working groups, two of which are "WG Architecture" and "WG Use Cases and Requirements".

The use case projects provided insights with regard to the data management activities the IDS architecture has to support and the assignment of these activities to the various roles in the business ecosystem (see Table 6).

Data management typically comprises three types of activities with regard to the creation, processing, and use of data: data governance, metadata management, and data lifecycle management. Data governance defines a framework of decision-making rights and processes with regard to the definition, creation, processing, and use of data (Khatri and

Table 5 Conceptual analysis of Phase 1

Theoretical concept	Manifestation in Phase 1
Platform architecture	• Definition of the IDS ecosystem (technical design elements not addressed yet)
Platform boundary resource	• Articulation of requirements for the sharing and use of data by different potential user groups
	• Elaboration of data exchange and data sharing requirements in use cases
Platform design	• Identification of requirements to be met by the platform architecture (see Table 3)
	• Identification and analysis of requirements (driven by search for "common denominator")
Platform ecosystem	• Definition of possible IDS core actors (see Table 4)
Ecosystem governance	• Not addressed yet
Regulatory instruments	• Interoperability (see R1.02: easy onboarding of participants)
	• "Controlled" openness (see R1.03: ecosystem must be open, but participants and software endpoints must be certified)
	• Use of standards (see R1.12: use of existing technologies and standards)

Table 6 Assignment of data management activities to IDS roles

ID	Activity	DO	DU	B	CA	Comment
Data Governance						
A1.01	Determine data usage conditions (to execute data ownership rights)	R, A	–	S	–	
A1.02	Enforce data usage conditions	–	R, A	–	–	Based on distributed usage control principles (see e.g.Pretschner et al. 2006)
A1.03	Ensure data quality	R, A	–	S	–	The data owner is responsible for data quality (Weber et al. 2009)
A1.04	Monitor and log data transactions	S	S	R, A	–	
A1.05	Enable data provenance	S	S	R, A	–	As proposed by e.g. Buneman et al. (2000)
A1.06	Provide clearing services	S	S	R, A	–	
A1.07	Ensure data privacy and data protection	A	S	–	R	Certification ensures that security profiles are met
A1.08	Disguise data ownership (if needed)	S	–	R, A	–	
Metadata Management						
A2.01	Provide vocabularies	S	S	R, A	–	As proposed by e.g. Halilaj et al. (2016)
A2.02	Describe and publish metadata	R, A	–	S	–	
A2.03	Look up and retrieve metadata	–	R, A	S	–	
Data Lifecycle Management						
A3.01	Capture and create data	R,A	–	–	–	
A3.02	Store data	R,A	–	–	–	No central storage activity
A3.03	Enrich and aggregate data	S	R, A	S	–	
A3.04	Distribute and provide data	R, A	–	S	–	
A3.05	Link data	S	S	R, A	–	

Legend:

DO Data owner, *DU* Data user, *B* Broker, *CA* Certification authority

R Responsible, *A* Accountable, *S* Supporting

Brown 2010; Weber et al. 2009). Hence, in the IDS context, data governance comprises also usage rights regarding data shared and exchanged via a multi-sided data platform. Metadata management specifies data about data, comprising both syntactical, semantic and pragmatic information (Sen 2004). It is of particular importance in distributed system environments that activities do not rely on a central data storage instance, but instead allow self-organization of different heterogeneous databases (Franklin et al. 2005; Halevy et al. 2006). Data lifecycle management is concerned with the creation, capturing, processing, enrichment, storage, distribution, and use of data (Ofner et al. 2013).

The industrial partners involved in the process extensively discussed questions related to data ownership. They quickly agreed that data owners should be able to specify usage rights with regard to their data by attaching data usage policies to the data itself. An interesting data management requirement came up in BW7, when a representative from a participating logistics service provider demanded that a broker might need to be able to disguise the identity of a data owner when presenting metadata, and/or to anonymize data before providing it to a data user (i.e. the broker should act as a data service provider also).

It was also consensus that each data transaction had to be logged and monitored in order to enable clearing services (in case a data transaction fails, for example) and data provenance.

With regard to metadata management, the participating industrial partners required that metadata should not be limited to syntactical information about data, but also include data ownership information, general usage conditions, prices for data use, and information about where and how the data can be accessed. Furthermore, it was agreed that the broker should provide vocabulary management services to facilitate easy mapping and linking of data (Halilaj et al. 2016).

Regarding the data storage architecture, it was agreed that data owners should keep control over data storage. All industrial partners opposed against any form of central data storage, due to a lack of trust with regard to data access and data usage on the part of third parties which data owners might be unaware of. As a consequence, it was decided that the IDS architecture should not follow a central data storage approach, such as the concept of the data lake (LaPlante and Sharma 2016; Larson and Chang 2016), for example.

Conceptual analysis

Table 7 gives an overview of the conceptual analysis of the "Institutionalization and Use Cases" phase. The focus of this project phase was on the definition of the ecosystem's governance mechanisms, i.e. the decision-making rights and the data management activities on the platform. These activities allowed for a more detailed specification of the functional requirements to be met by the platform.

Project phase #3 – Architecture design

Description

Based on the overall rationale of the IDS and the requirements for data management activities, researchers and practitioners working together in WG Architecture designed a first version of the IDS Architecture (cf. Fig. 1).

The central software component of this architecture is the IDS Connector, representing a data endpoint that grants participants access to the IDS (Otto et al. 2017).[5] The IDS Connector has three key functions: first, it is responsible for the exchange of data between a data provider and a data consumer (this includes request activities, usage control management and enforcement, mapping, and secure data transmission activities); second, it enables secure and trusted execution of software (i.e. it works as an isolated runtime environment that makes sure both data and software inside the Connector is protected against unauthorized access and manipulation from outside); third, it executes trusted software packages ("Data Apps") that can be retrieved from the IDS App Store.

The App Store was defined in Phase 3 as a new role within the IDS ecosystem. It is responsible for distributing certified Data Apps (i.e. self-contained, self-descriptive software packages that can be deployed inside the IDS Connector). A Data App provides access to data as well as data processing capabilities (Otto et al. 2017).

Information about the data endpoints accessible in the IDS is provided by the Broker, which is responsible for managing a metadata repository (Otto et al. 2017).

Conceptual analysis

Table 8 gives an overview of the conceptual analysis of the "Architecture Design" phase. This phase was dominated by the first cycle of BIE activities (BIE = building, intervention,

---

[5] The first version of the IDS Reference Architecture dModel distinguished between "internal" and "external" IDS Connectors. Internal Connectors were supposed to function as an adapter to internal systems. This concept was abandoned at a later stage, as standard software is available for connecting internal systems to external Connectors.

Springer

and evaluation; i.e. the second stage of the ADR process) regarding the core architecture components (i.e. mainly the IDS Connector). While data served as a functional boundary resource in Phase 2, it served as a technical boundary resource in the third phase.

Project phase #4 – Ecosystem design

Description

Over the course of designing the IDS Reference Architecture Model and its piloting in use cases, it became apparent that the initial architecture design was a good starting point to address the requirements (i.e. use of the IDS Connector; decentral data storage; IDS Broker and IDS App Store as new roles etc.), but obviously was still lacking details needed for implementation. In particular, details about what activities should be performed by which role within the IDS ecosystem had to be defined. In addition, platform activities that were missing in the earlier phases of the research process were identified and specified. One activity of particular importance was identity management to ensure secure authentication and trust among IDS participants. Another task force was established to focus on the design of the business ecosystem. In MW11, the group identified key ecosystem actors and defined an initial set of responsibilities for each of them. Figure 2 shows the first version of the IDS ecosystem design.

In this fourth phase, two additional roles were defined: the Clearing House and the Identity Provider. The Clearing House acts as an intermediary providing clearing and settlement services for all financial and data exchange transactions within the IDS (Otto et al. 2017). The Identity Provider offers services to create, maintain, manage and validate identity information of and for IDS participants.

With the ecosystem's design becoming more and more mature, questions came up regarding the "roll-out" (i.e. the adoption path) of the IDS. In this context, a task force named "Business Modeling and Exploitation" was established, which came up with a sequence of platform features considered instrumental for the launch and adoption of the IDS (see Table 9).

The six main instruments follow a logical order. The first instrument aims at ensuring 1) trust among the different users of the IDS. When trust is achieved, ensuring 2) secure exchange of data and data sovereignty is the next step. These two instruments are required for fostering the emergence of a 3) data ecosystem. For a data ecosystem to run efficiently, 4) interoperability is needed for standardized interaction of ecosystem actors (vocabularies play a key role in this task, as they facilitate the mapping of different data sources and the integration through

Table 7 Conceptual analysis of Phase 2

Theoretical concept	Manifestation in Phase 2
Platform architecture	• Specification of functional requirements for data sharing, data governance, and data use
	• Agreement to use a decentral architecture
Platform boundary resource	• Definition of a set of activities with regard to data as a boundary resource
Platform design	• Foundation of IDSA to institutionalize the ecosystem
	• Establishment of working groups
	• Implementation of use case projects
Platform ecosystem	• No changes compared to Phase 1
Ecosystem governance	• Definition of responsibilities of core actors
	• Specification of decision-making rights regarding data governance and data management activities
Regulatory instruments	• Transfer of instruments from Phase 1 to data governance and data management activities

linked-data representations). On top of data exchange, 5) apps can offer value-adding services using shared data. Finally, 6) data markets emerge on the basis of clearing and billing services (among other things).

Furthermore, domain specific ecosystems were envisaged in different "smart service" domains (e.g. healthcare, mobility, education, or travel). The task force confirmed the decision not to design the IDS Reference Architecture

Model for a single provider, but to allow for as much openness as possible to foster rapid market adoption of the IDS.

Conceptual analysis

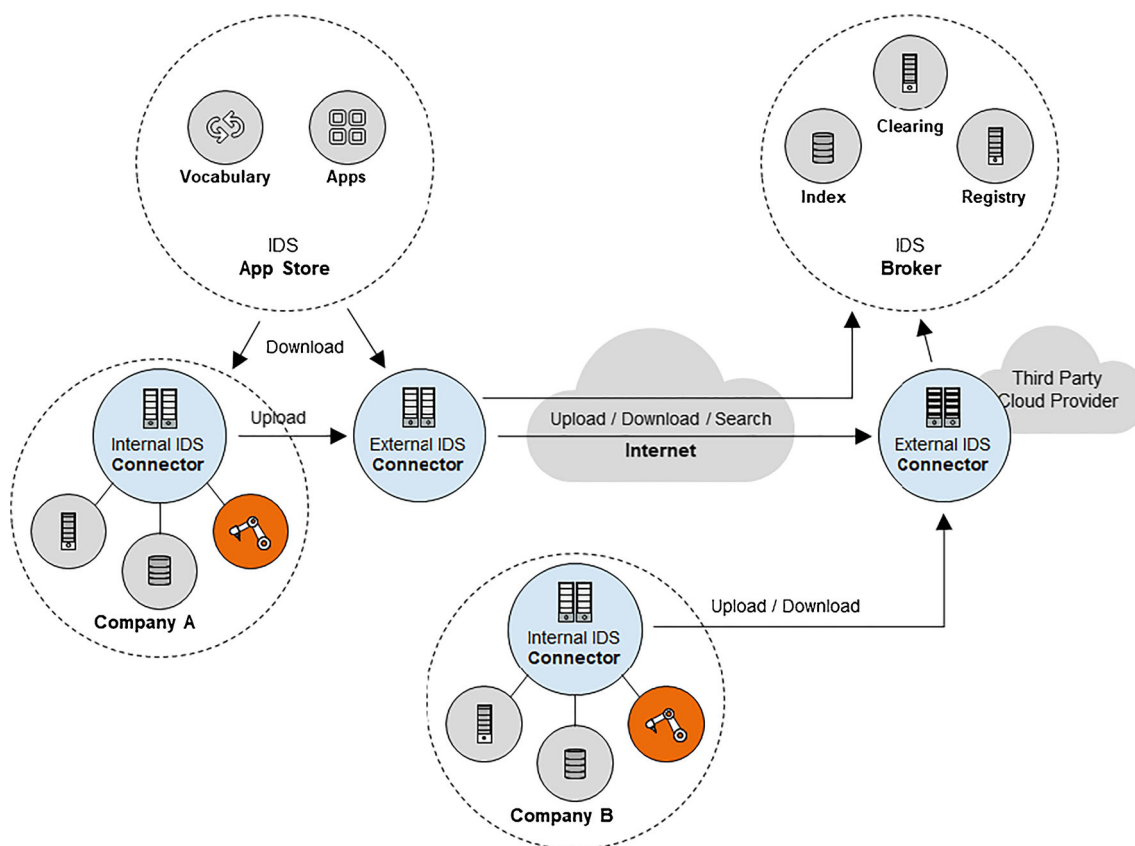Table 10 gives an overview of the conceptual analysis of the "Ecosystem Design" phase. The focus of the activities in this

Fig. 1 Initial IDS architecture design (cf. Otto et al. 2016)

Table 8 Conceptual analysis of Phase 3

Theoretical concept	Manifestation in Phase 3
Platform architecture	• Specification of core components (in particular: the IDS Connector), which are made available as first proof-of-concept
Platform boundary resource	• Specification of technical data processing mechanisms
Platform design	• Publishing of 1st version of IDS Reference Architecture Model (Otto et al. 2017)
Platform ecosystem	• Definition of the App Store as a new role
Ecosystem governance	• No changes compared to Phase 2
Regulatory instruments	• No changes compared to Phase 2

phase was on the completion of the platform ecosystem and the regulatory instruments needed to foster adoption and use of the IDS. With regard to the former, the various possible interactions between the ecosystem actors were specified and separated from the technical architecture components. Furthermore, the regulatory instruments form a consensus of the partners in the ecosystem and were developed in the light of IDS rollout and use.

International Data Spaces: case study discussion

Platform architecture

A multi-sided data platform requires clear data governance rules and data management processes for the platform to successfully evolve. A central platform functionality is metadata management. Metadata must include information

about the data owner, about data usage conditions, and about financial aspects (e.g. price of data use). Data provenance (Buneman et al. 2000) is necessary to be able to track the flow of data across multiple nodes of the network (i.e. different actors in the ecosystem).

Another central IDS feature is decentralized data storage. While central approaches, such as data lakes, are of significant value for use cases in which large volumes of data are required (for machine learning purposes, for example), they cannot be considered a reasonable option for critical data goods (which are in the focus of the IDS ecosystem).

In addition, the architecture must allow for distributed usage control including remote policy enforcement (Hilty et al. 2007; Kelbert and Pretschner 2012). Distributed usage control extends the exisiting conceptualization of data governance. The majority of data governance approaches focus on single enterprises and is mainly concerned with defining roles and responsibilities for the management and use of master data (Khatri and Brown 2010; Weber et al. 2009). Distributed
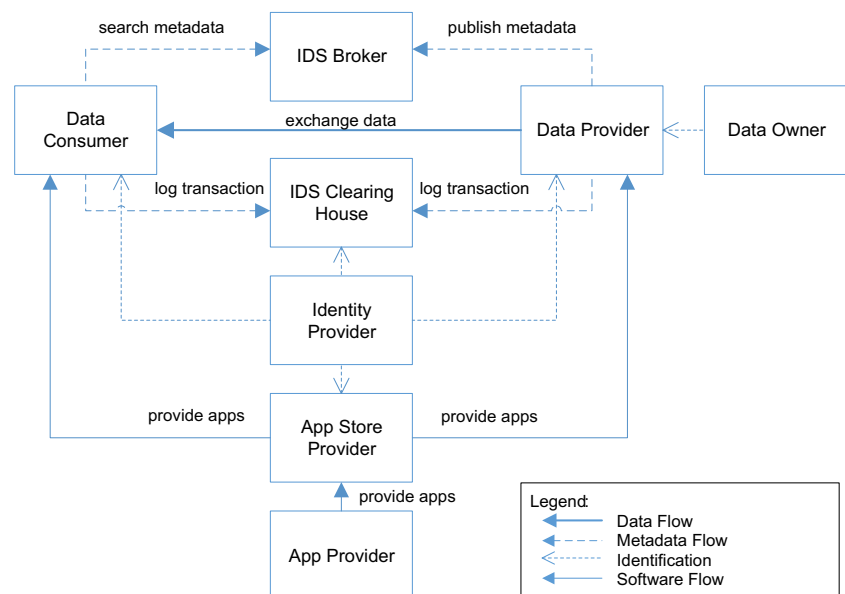
Fig. 2 IDS ecosystem design

Table 9 Instruments for adoption and use of the IDS

1 – Trust	Pr	Co	Br	Id	Ce	Ap	Cl	2 – Security and data sovereignty	Pr	Co	Br	Id	Ce	Ap	Cl
Identity management				X				Authentication, authorization				X			
User certification					X			Usage policy management	X	X					
								Trustworthy communication	X	X		X			
								Technical certification					X		
3 – Data ecosystem	Pr	Co	Br	Id	Ce	Ap	Cl	4 – Standardized interoperability	Pr	Co	Br	Id	Ce	Ap	Cl
Data source description	X							Vocabulary integration			X				
Brokering			X					Data mapping			X				
Vocabulary management	X	X						Inter-cloud linking	X	X					
5 – Value adding apps	Pr	Co	Br	Id	Ce	Ap	Cl	6 – Data markets	Pr	Co	Br	Id	Ce	Ap	Cl
Data processing	X	X				X		Clearing and billing							X
Remote execution	X	X				X		Domain specific ecosystem			X			X	X
								Contract templates			X				

Legend:

X Role involved in functional design principle

Pr Data provider, *Co* Data consumer, *Br* Broker, *Id* Identity provider, *Ce* Certification agency, *Ap* App store, *Cl* Clearing house

usage control, though, can be seen as a means establish data governance in ecosystems.

The results of the IDS case confirm the findings from previous research, e.g. the "network type" of inter-organizational data collaboration observed by van den Broek and van Veenstra (2015), and provide opportunities for future research.

Platform boundary resource

The role of social boundary resources is essential during the early stages of an MSP's design phase, because technical boundary resources do not exist yet during theses stages, and because multiple stakeholder groups need to find consensus about the purpose of the platform. Dal Bianco et al. (2014) point to the importance of knowledge transfer and education measures as social boundary resources during the adoption phase of a typical MSP. In the case of the IDS, however, social boundary resources comprise mainly the organizational structures of the IDS Association (its working groups, task forces etc.) and the "functional view" of data as described in the use cases.

As far as data is concerned, its dual nature as a boundary resource has to be stressed: From a functional perspective, it serves as a social boundary resource, as it supports the identification and analysis of multilateral use cases; regarding "data in transit" functions, it serves as a technical boundary resource as defined by Gawer (2014) and Schreieck et al. (2016). The IDS case points to the importance of data governance and the coordination of shared data management activities in this respect.

Table 10 Conceptual analysis of Phase 4

Theoretical concept	Manifestation in Phase 4
Platform architecture	• Further specification of core components (e.g. the Trusted Connector)
Platform boundary resource	• No changes compared to Phase 3
Platform design	• Completion of ecosystem design "Business Modeling and Exploitation" task force established to prepare and plan IDS adoption and use
Platform ecosystem	• Definition of Clearing House and Identity Provider as new roles
	• Specification of possible interactions between actors
Ecosystem governance	• No changes compared to Phase 3
Regulatory instruments	• Definition of instruments for IDS adoption and use (including interdependencies between them) (see Table 9)

Platform design

The main challenge for an MSP is rapid market entry and fast adoption by a critical mass (Henfridsson and Bygstad 2013). As already mentioned, the lifecycle of an MSP comprises three major phases: innovation, adoption, and scaling (Tan et al. 2015; Tan et al. 2016). In many cases, MSPs are forced to achieve time and adoption advantages with regard to competing MSPs in order to survive. In contrast, one of the biggest hurdles for an alliance-driven MSP is the coordination of the requirements coming from multiple stakeholder groups, as well as the definition of decision-making rights and responsibilities during the early stages of the platform design phase. Thus, adoption – which typically is a key success factor – is not so much of an issue than organizing the innovation process within the alliance. As a consequence, the typical stages which can be observed in an MSP's lifecycle follow a different sequence and are rather characterized by early adoption and innovation and design activities following at a later stage.

Platform ecosystem

The IDS initiative represents a case in which an ecosystem is designed in a joint effort. This is in accordance with the views of those researchers considering a platform ecosystem as the result of a structured design activity (Immonen et al. 2014; Tian et al. 2008). However, opponents of this perspective see ecosystems emerging around a shared value proposition and argue that, due to a range of exogenous factors, ecosystems can only be planned and design to a very limited extent (Adner 2017). Through the participation of various stakeholder groups in the IDS Association, some of the typical endogenous factors, such as competition or technological obsolescence, could be "internalized", and thus being controlled to a certain extent. This means that the ecosystems of alliance-driven MSPs might be better "designable" than more common MSP ecosystems fostered by a keystone firm.

Reflecting these findings against research on the emergence and adoption of open-source ecosystems, such as Fiware (Rodriguez et al. 2018), or against the shift of the OPC Foundation towards an open-source model (Palm et al. 2015–2015), is a promising approach to further advance these propositions.

Ecosystem governance

Coordination of a joint innovation process during the early stages of the platform design phase requires extensive interaction and consultation between the different sides involved.

This stands in contrast with the emergence of keystone-driven MSPs, which tend to develop interaction between sides during later stages of the platform's lifecycle (Tan et al. 2015; Tan et al. 2016).

As far as organizational structures are concerned, some similarities between the IDS case and keystone-driven MSPs can be identified – particularly regarding the aspect of platform openness. Openness refers to restrictions regarding the use, development, and commercialization of a platform (Boudreau 2010; Eisenmann et al. 2009). Ondrus et al. (2015) showed that openness typically changes across the lifecycle of an MSP. As a non-for-profit organization, the IDS Association is open to everyone, but nonetheless uses governance mechanisms to align the behavior and interests of its various stakeholders. Openness differs with regard to the stakeholder groups addressed and, thus, can be observed on a provider, a technology, and a user level (Ondrus et al. 2015). Openness is carefully controlled by the platform owner in order to steer the platform's adoption and use into the desired direction. Typically, openness can be seen on a technology level – as in the IDS case – on platforms that are open for open-source development after a certain level of maturity has been reached, and in order to compete against other platforms in the same market.

Regulatory instruments

In contrast to keystone-driven MSPs, on which pricing is among the main instruments regulating adoption and use of the platform, the design and development phase of the IDS was mainly concerned with instruments other than pricing. The IDS considers trustworthiness of participants and data sovereignty of data owners and data providers as key instruments regulating the adoption and use of the platform. Thus, the IDS alliance considered non-pricing instruments more important than pricing instruments. The latter were envisaged only in terms of providing IDS Data Apps through the IDS App Store – and in facilitating data markets further down the road, which would need clearing of data transactions and billing.

While it is apparent to a certain extent that pricing played no major role during the early IDS design, the findings of the IDS case study confirm the results of previous research. Boudreau and Hagiu (2009), for example, observed a wide array of legal, technological, informational and other instruments for regulating MSPs.

The regulatory instruments in the IDS case were directly derived from the user requirements specified – and, thus, mutually agreed upon by the parties involved. Being an outcome of a task force that was open to all members, these instruments were not imposed by a certain group of members, but represent a consensual ecosystem governance

approach. In fact, the structure of the IDS ecosystem is a result of these regulatory instruments, as the instrument's operationalization required the definition of dedicated roles (e.g. Certification Agency or Clearing House). In contrast to keystone-driven MSPs, all aspects of ecosystem governance – i.e. platform ownership, leadership, and rules (Tura et al. 2018) – have been combined in one organizational entity: the IDS Association. While typical MSPs develop regulatory instruments as the platform is being adopted and used, the IDS case displayed some form of "ex-ante" regulation before the platform was fully developed and implemented.

Conclusion

Summary of results

The IDS case represents a unique and – in many regards – an extreme case of an MSP: first, it has been designed and developed by a multi-stakeholder alliance, and not by a single entity or keystone firm; second, the subject of analysis as presented in this paper is the early platform design phase, and not – as in many other studies – the platform adoption and use phase; and third, through its nature as a multi-sided data platform, the IDS case delivers deep insights regarding the use of data as a platform boundary resource. Table 11 juxtaposes the IDS case as an example of an alliance-driven MSP in contrast to typical keystone-driven cases.

With regard to RQ #1 (i.e. how alliance-driven MSPs come into existence and evolve), the IDS case suggests that the typical evolutionary stages of MSPs do not explain the emergence of alliance-driven MSPs (and their evolution during the early platform design phase). In contrast to the lifecycle of a keystone-driven MSP consisting of 1) the phase of innovation, 2) the phase of adoption, and 3) the phase of scaling, the foundation of the alliance itself is based on adoption in the first place. After the adoption of the platform concept has led to the institutionalization of the alliance, innovation activities start as a consensual process in the second phase of the lifecycle.

Regarding RQ #2 (i.e. how different stakeholder groups use certain governance mechanisms during the early phase of an alliance-driven MSP's lifecycle), it has to be noted that shared platform ownership differs from other ecosystem based approaches, such as open-source communities. In contrast to such other approaches, alliance-driven MSPs like the IDS rest on a formal institutionalization and demand execution of ownership rights. Thus, governance mechanisms do not follow a hierarchical institutionalization, but require ecosystem coordination. Examples of such governance mechanisms are the statutes and rules of the IDS Association, the organization of working groups and task forces, and the establishment of collaborative processes (e.g. for the development of the IDS Reference Architecture Model). Furthermore, the IDS case suggests a more differentiated conceptualization of MSP ecosystems. Literature predominantly understands

Table 11 Juxtaposition of keystone-driven and alliance-driven MSP design

Theoretical concept	Keystone-Driven	Alliance-Driven
Platform architecture	Architecture determined by goals of keystone firm	Architecture determined by shared interest of multiple owners (leading to decentral data storage, for example)
Platform boundary resource	Mainly technical boundary resources (APIs, SDKs etc.), supported by "social" boundary resources (e.g. training for developers) (Dal Bianco et al. 2014)	Data (IDS specific) as a boundary resource of "dual" nature, i.e. requiring both technical processing and functional use; many social boundary resources, such as working groups, task forces etc.
Platform design	Core developed by platform owner, then extended by complementors	Consensus oriented design process with focus on "common denominator"
Platform ecosystem	1) Innovation, 2) Adoption, 3) Scaling (Henfridsson and Bygstad 2013)	1) Adoption, 2) Innovation, 3) Scaling
Ecosystem governance	Start with limited number of sides and limited options for interaction between them, then increase number of sides and options for interaction (F. Tan et al. 2016)	Start with complex ecosystem (i.e. multi-stakeholder setting), then reduce to core ecosystem and extend it later on depending on roll-out requirements
Regulatory instruments	Mainly pricing instruments, accompanied by non-pricing instruments	Dominated by non-pricing instruments (as outlined in Table 9); integration of pricing instruments scheduled for scaling phase; data governance

MSP ecosystems as being composed of the platform owner and different user groups/sides. The IDS case shows, however, that there is an "inner" ecosystem (formed by the alliance members) and an "outer" ecosystem (composed by the alliance itself and by non-alliance, i.e. non-member, user groups and sides).

As for RQ #3 (i.e. how regulatory instruments are deployed to foster the adoption and use of an alliance-driven MSP), it is this interaction between the inner and the outer ecosystem that is regulated by such instruments. Trust, security, data sovereignty, and interoperability have been major drivers of the IDS initiative, and are therefore also reflected by the design of possible interactions between data users and data providers in the IDS. The IDS case suggests that common or shared interests that lead to the creation of an alliance-driven MSPs do not per se focus on purely economic or even profit maximizing goals of the platform provider. Instead, by searching for a "common denominator" (see Table 5), alliance-driven MSPs are – at least to a certain extent – of an infrastructural nature, which has an impact on the design of the regulatory instruments.

Contribution to the body of knowledge

The IDS case study presented in this paper aims at advancing the body of scientific knowledge with regard to the early stages of an alliance-driven MSP's design phase. In doing so, it lays a foundation for the conceptual understanding of alliance-driven MSPs as such. The findings from the case study suggest that – in contrast to keystone-driven MSPs – different evolutionary paths can be pursued during the early stages of an MSP's lifecycle. These findings pave the way for a more differentiated analysis of the evolution of MSPs – which is even more useful as the variety of ownership models for MSPs is likely to grow in the future.

Furthermore, the study is among the few that focus on the early platform design phase. It thereby yields detailed insights regarding the coordinating and consensus-finding activities during these early stages of the lifecycle. In the case of the IDS, adoption of the platform purpose and value proposition took place before the platform itself was designed. Thus, the typical sequence of 1) innovation, 2) adoption, and 3) scaling with regard to MSPs does not apply to the IDS case.

Moreover, the IDS case contributes to the body of scientific knowledge by putting the focus on data as a boundary resource of an MSP. The dual nature of data as both a social and a technical boundary resource requires data governance rules and shared data management processes, which need to be distributed between the different user groups (i.e. platform sides). The strategic nature of data has led to a complex ecosystem structure including specific roles for data provenance and data sovereignty. Moreover, the findings add to the understanding of data governance requirements in inter-organizational data sharing platforms in general.

Besides its contribution to research, the paper may have an impact on the current debate in the practitioners' community. The IDS case yields in-depth insights regarding the establishment of an alliance-driven MSP. In doing so, it may be useful for current and future endeavors both on a public level – e.g. the European Union's strategy towards a European Data Space (European Commission 2018) – and on a private level, such as the NEVADA initiative by the VDA (2016).

Limitations and further research opportunities

Limitations of the study stem mainly from its research design. A single-case study can only produce preliminary results with regard to a phenomenon of interest. The findings should be triangulated and validated using additional cases both with regard to alliance-driven and keystone-driven MSPs.

Furthermore, it has to be noted that the IDS case study examined the early stages of the design phase of an alliance-driven MSP only. The IDS initiative is gaining more and more momentum, and use and scaling activities are currently on their way. Thus, the efficacy of the initiative as such, as well as of the regulatory instruments chosen, has yet to be examined.

Further research opportunities exist in multiple areas. For example, the phenomenon of data-centric MSPs is still largely unexplored. Given the growing importance of data as a strategic enterprise resource, and the increasing emergence of commercial data-centric MSP offerings from software and technology providers, there will be an increasing demand for theory analyzing and explaining these phenomenon. Furthermore, the preliminary findings from the IDS case study should be examined further, both from an economic and an engineering perspective. The exploratory analysis of regulatory instruments should be taken further towards investigating the interdependence among these instruments and between them and market oriented instruments (such as pricing). With regard to the design of data-centric MSPs in particular, the data governance rules and data management processes designed in the IDS research project can be considered a promising starting point for more detailed investigations.

Appendix

Table 12 Multilateral workshops (MWs)

ID	Date and Duration	Location	Purpose	Participants	Results/Remarks
MW1	02/15/2015, 3 h	Berlin (DE)	Reflect on IDS rationale; conceptualization and action plan	• 30 high-level industry representatives • 10 high-level representatives from federal ministries and agencies • 18 high-level Fraunhofer representatives	• Rationale confirmed • Agreement to set up a task force for further specification of the concept
MW2	03/27/2015, 3 h	Munich (DE)	Discuss technical options, use cases, input from industry, and action plan	• 19 high-level industry representatives • 3 high-level representatives from federal ministries and agencies • 11 high-level Fraunhofer representatives	• Prioritization of action plan rejected • Mandate given to Fraunhofer to focus on use cases and common requirements
MW3	06/09/2015, 6 h	Dortmund (DE)	Identify logistics use cases	• 19 business managers and architects from industry • 5 Fraunhofer experts	• Transparent supply chain identified as topic of mutual interest • High-level requirements documented
MW4	06/16/2015, 5 h	Dortmund (DE)	Identify automotive/mobility use cases	• 8 business managers and architects from industry • 5 Fraunhofer experts	• High-level requirements documented
MW5	09/23/2015, 3.5 h	Berlin (DE)	Present use case analysis; announce foundation of IDSA	• 30 high-level representatives from industry • 9 high-level representatives from federal ministries and agencies • 17 high-level Fraunhofer representatives	• Requirements, use cases, and plan forward approved • Start of research project announced • Foundation of IDSA announced
MW6	01/26/2016, 3 h	Berlin (DE)	Constitutional meeting of the IDSA	• Representatives of founding members	• Association founded • Use cases acknowledged • Work plan acknowledged
MW7	04/28/2016, 7 h	Dortmund (DE)	Combined meeting of "WG Architecture" and "WG Use Cases and Requirements"	• 32 representatives from industrial partners (both business and IT architects and business analysts) • 10 representatives from research organizations	• Working groups constituted • First work packages distributed
MW8	03/09/2016, 3 h	Berlin (DE)	Materials Data Space meeting	• High-level representatives from industry, research, and politics	• Basic idea of Materials Data Space presented • Positioning as a verticalization of the IDS acknowledged
MW9	02/28/2017, 7 h	Waldkirch (DE)	Meeting of "WG Architecture"	• 21 representatives from industrial partners (mainly business and IT architects) • 6 representatives from research organizations	• First version of the IDS Reference Architecture Model discussed and approved
MW10	05/30/2017, 6 h	St. Augustin (DE)	Meeting of "WG Architecture"	• 16 representatives from industrial partners (mainly business and IT architects) • 7 representatives from research organizations	• Requirements for secure IDS Connector specified • Ecosystem and software component architecture specified • Information model discussed
MW11	08/27/2017, 6 h	Dortmund (DE)	Meeting of "Exploitation and Business Modeling" task force	• 21 representatives from industrial partners (mainly business analysts) • 3 representatives from research organizations	• Overall governance model specified • Ecosystem roles described

Legend: *IDS* International data spaces, *DE* Germany, *IT* Information technology, *IDSA* International data spaces association, *WG* Working group

Table 13 Bilateral workshops (BWs)

ID	Date and Duration	Industry	Focus Topics
BW1	08/05/2015, 3 h	Pharmaceuticals	• Real-life evidence • Smart maintenance • Supply chain transparency
BW2	08/12/2015, 3 h	Automotive	• Traffic management • Mobility
BW3	08/14/2015, 3 h	Retail	• Autonomous supply chain transparency • Traceability
BW4	08/18/2015, 2 h	Industrial machinery and services	• Inbound logistics management
BW5	08/20/2015, 4 h	Sensor technology	• Secure data supply chain
BW6	08/24/2015, 4 h	Logistics and mobility	• Data transparency across the value chain
BW7	08/25/2015, 2 h	Logistics and mobility	• Traceability of transport goods
BW8	08/25/2015, 2,5 h	Pharmaceuticals	• Real-life evidence • Supply chain transparency
BW9	08/28/2015, 7 h	Automotive	• Yard management • Global supply chain transparency
BW10	09/16/2015, 5 h	Electronics	• High-performance supply chain
BW11	09/16/2015, 3 h	Automotive	• Supply chain transparency
BW12	09/17/2015, 5 h	Chemicals	• Transparency of transported goods

Table 14 Use case projects (UCs)

ID	Duration	Industry	Use Case	Ecosystem
UC1	Q3/2015 – Q2/2017	Steel and technology	• Inbound transport logistics and time-window optimization at docking stations • Improved and efficient flow of information between logistics service providers and	• OEM • Transportation service providers • Logistics service providers
UC2	Q4/2016 – Q1/2018	Automotive	• Exchange of supply network and inventory information for end-to-end supply chain and risk management	• OEM • Tier-1 supplier
UC3	Q2/2017 – Q1/2018	Automotive	• Multilateral exchange of steel quality information • Mapping services • Data marketplace	• OEM • Tier-1 supplier • Intermediary (Broker)
UC4	Q4/2016 – Q1/2018	Healthcare	• Privacy-preserving home telecare infrastructure for post-hospital treatment and chronic disease monitoring • Data exchange between mobile sensor streams and patient history documents	• Hospitals • Private practices • Telecare organizations • Health insurances

Legend: *OEM* Original equipment manufacturer

References

Adner, R. (2017). Ecosystem as structure. *Journal of Management, 43*(1), 39–58. https://doi.org/10.1177/0149206316678451.

Baldwin, C. Y., & Woodard, C. J. (2009). The architecture of platforms: A unified view. In A. Gawer (Ed.), *Platforms, markets and innovation* (pp. 19–44). Cheltenham: Edward Elgar.

Baskerville, R. L. (1999). Investigating information systems with action research. *Communications of the AIS, 2*(3), Article 19.

Bharosa, N., Janssen, M., Klievink, B., & Tan, Y-h (2013). Developing multi-sided platforms for public-private information sharing. In *Proceedings of the 14th Annual International Conference on Digital Government Research* (pp. 146–155). https://doi.org/10.1145/2479724.2479747.

Boudreau, K. (2010). Open platform strategies and innovation: granting access vs. devolving control. *Management Science, 56*(10), 1849–1872. https://doi.org/10.1287/mnsc.1100.1215.

Boudreau, K. J., & Hagiu, A. (2009). Platform rules: multi-sided platforms as regulators. In A. Gawer (Ed.), *Platforms, markets and innovation* (pp. 163–191). Cheltenham: Edward Elgar.

Buneman, P., Khanna, S., & Tan, W.-C. (2000). Data provenance: some basic issues. In *Proceedings of the International Conference on Foundations of Software Technology and Theoretical Computer Science* (pp. 87–93). https://doi.org/10.1007/3-540-44450-5_6.

Cusumano, M. (2010). The evolution of platform thinking. *Communications of the ACM, 53*(1), 32–34. https://doi.org/10.1145/1629175.1629189.

Dal Bianco, V., Myllarniemi, V., Komssi, M., & Raatikainen, M. (2014). The role of platform boundary resources in software ecosystems: a case study. In *Proceedings of the 2014 IEEE/IFIP Conference on Software Architecture* (pp. 11–20). https://doi.org/10.1109/WICSA.2014.41.

De Reuver, M., Sørensen, C., & Basole, R. C. (2018). The digital platform: a research agenda. *Journal of Information Technology, 33*(2), 124–135. https://doi.org/10.1057/s41265-016-0033-3.

Demchenko, Y., Laat, C. de, & Membrey, P. (2014). Defining architecture components of the big data ecosystem. In *Proceedings of the 2014 International Conference on Collaboration Technologies and Systems* (pp. 104–112). https://doi.org/10.1109/CTS.2014.6867550.

Donmoyer, R. (2000). Generalizability and the single-case study. In R. Gomm, M. Hammersley, & P. Foster (Eds.), *Case study method: key issues, key texts* (pp. 45–77). London: SAGE.

Dul, J., & Hak, T. (2008). *Case study methodology in business research*. Oxford: Butterworth-Heinemann.

Eaton, B., Elaluf-Calderwood, S., Sørensen, C., & Yoo, Y. (2015). Distributed tuning of boundary resources: the case of Apple's iOS service system. *MIS Quarterly, 39*(1), 217–243. https://doi.org/10.25300/MISQ/2015/39.1.10.

Eisenmann, T. R., Parker, G., & van Alstyne, M. W. (2009). Opening platforms: how, when and why? In A. Gawer (Ed.), *Platforms, markets and innovation* (pp. 131–162). Cheltenham: Edward Elgar.

European Commission (2018). Towards a common European data space. Brussels. Retrieved from European Commission website https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:0232. Accessed 19 May 2019.

Evans, D. S. (2003). The antitrust economics of multi-sided platform markets. *Yale Journal on Regulation, 20*, 325–381.

Evans, D., & Schmalensee, R. (2013). *The antitrust analysis of multi-sided platform businesses*. Cambridge: National Bureau of Economic Research.

Franklin, M., Halevy, A., & Maier, D. (2005). From databases to dataspaces. *ACM SIGMOD Record, 34*(4), 27–33. https://doi.org/10.1145/1107499.1107502.

Gawer, A. (2009). Platform dynamics and strategies: from products to services. In A. Gawer (Ed.), *Platforms, markets and innovation* (pp. 45–77). Cheltenham: Edward Elgar.

Gawer, A. (2014). Bridging differing perspectives on technological platforms: Toward an integrative framework. *Research Policy, 43*(7), 1239–1249. https://doi.org/10.1016/j.respol.2014.03.006.

Gawer, A., & Cusumano, M. A. (2014). Industry platforms and ecosystem innovation. *Journal of Product Innovation Management, 31*(3), 417–433. https://doi.org/10.1111/jpim.12105.

Ghazawneh, A., & Henfridsson, O. (2010). Governing third-party development through platform boundary resources. In *Proceedings of the 31st International Conference on Information Systems* (pp. 1–18).

Ghazawneh, A., & Henfridsson, O. (2015). A paradigmatic analysis of digital application marketplaces. *Journal of Information Technology, 30*(3), 198–208. https://doi.org/10.1057/jit.2015.16.

Goh, C. H., Bressan, S., Madnick, S., & Siegel, M. (1999). Context interchange: new features and formalisms for the intelligent integration of information. *ACM Transactions on Information Systems, 17*(3), 270–293. https://doi.org/10.1145/314516.314520.

Hagiu, A. (2009). Two-sided platforms: product variety and pricing structures. *Journal of Economics & Management Strategy, 18*(4), 1011–1043. https://doi.org/10.1111/j.1530-9134.2009.00236.x.

Halevy, A., Franklin, M., & Maier, D. (2006). Principles of dataspace systems. In *Proceedings of the 25th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* (pp. 1–9). https://doi.org/10.1145/1142351.1142352.

Halilaj, L., Petersen, N., Grangel-González, I., Lange, C., Auer, S., Coskun, G., & Lohmann, S. (2016). Vocol: an integrated environment to support version-controlled vocabulary development. In *Proceedings of the 20th International Conference on Knowledge Engineering and Knowledge Management* (pp. 303–319). https://doi.org/10.1007/978-3-319-49004-5_20.

Henfridsson, O., & Bygstad, B. (2013). The generative mechanisms of digital infrastructure evolution. *MIS Quarterly, 37*(3), 907–931. https://doi.org/10.25300/MISQ/2013/37.3.11.

Hilty, M., Pretschner, A., Basin, D., Schaefer, C., & Walter, T. (2007). A policy language for distributed usage control. In *Proceedings of the European Symposium on Research in Computer Security* (pp. 531–546). https://doi.org/10.1007/978-3-540-74835-9_35.

Hofmann, H. F., & Lehner, F. (2001). Requirements engineering as a success factor in software projects. *IEEE Software, 18*(4), 58–66. https://doi.org/10.1109/MS.2001.936219.

Immonen, A., Palviainen, M., & Ovaska, E. (2014). Requirements of an open data based business ecosystem. *IEEE Access, 2*, 88–103. https://doi.org/10.1109/ACCESS.2014.2302872.

Jarke, M., Jeusfeld, M., & Quix, C. (2014). Data-centric intelligent information integration—from concepts to automation. *Journal of Intelligent Information Systems, 43*(3), 437–462. https://doi.org/10.1007/s10844-014-0340-5.

Kelbert, F., & Pretschner, A. (n.d.) Towards a policy enforcement infrastructure for distributed usage control. In *Proceedings of the 17th ACM symposium on Access Control Models and Technologies* (pp. 119–122). https://doi.org/10.1145/2295136.2295159.

Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM, 53*(1), 148–152. https://doi.org/10.1145/1629175.1629210.

Kontokosta, C. E. (2013). Energy disclosure, market behavior, and the building data ecosystem. *Annals of the New York Academy of Sciences, 1295*, 34–43. https://doi.org/10.1111/nyas.12163.

LaPlante, A., & Sharma, B. (2016). *Architecting data lakes: data management architectures for advanced business use cases*. Sebastopol: O'Reilly Media.

Larson, D., & Chang, V. (2016). A review and future direction of agile, business intelligence, analytics and data science. *International Journal of Information Management, 36*(5), 700–710. https://doi.org/10.1016/j.ijinfomgt.2016.04.013.

Moiso, C., & Minerva, R. (2012). Towards a user-centric personal data ecosystem: the role of the bank of individuals' data. In *Proceedings of the 16th International Conference on Intelligence in Next Generation Networks* (pp. 202–209). https://doi.org/10.1109/ICIN.2012.6376027.

Ofner, M. H., Straub, K., Otto, B., & Oesterle, H. (2013). Management of the master data lifecycle: a framework for analysis. *Journal of Enterprise Information Management, 26*(4), 472–491. https://doi.org/10.1108/JEIM-05-2013-0026.

Ondrus, J., Gannamaneni, A., & Lyytinen, K. (2015). The impact of openness on the market potential of multi-sided platforms: a case study of mobile payment platforms. *Journal of Information Technology, 30*(3), 260–275. https://doi.org/10.1057/jit.2015.7.

Otto, B., Auer, S., Cirullies, J., Jürjens, J., Menz, N., Schon, J., & Wenzel, S. (2016). *Industrial Data Space: Digitale Souveränität über Daten*. Munich: Fraunhofer, Industrial Data Space e.V.

Otto, B., Lohmann, S., Auer, S., Brost, G., Cirullies, J., Eitel, A., . . . Wenzel, S. (2017). *Reference Architecture Model for the Industrial Data Space*. Munich, Berlin: Fraunhofer, Industrial Data Space e.V.

Otto, B., Lohmann, S., Steinbuß, S., & Teuscher, A. (2018). *IDS Reference Architecture Model (Version 2.0).* Berlin, Munich: International Data Spaces Association, Fraunhofer.

Palm, F., Gruner, S., Pfrommer, J., Graube, M., & Urbas, L. (2015). Open source as enabler for OPC UA in industrial automation. In *Proceedings of the 20th Conference on Emerging Technologies & Factory Automation* (pp. 1–6). https://doi.org/10.1109/ETFA.2015.7301562.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems, 24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302.

Pretschner, A., Hilty, M., & Basin, D. (2006). Distributed usage control. *Communications of the ACM, 49*(9), 39–44. https://doi.org/10.1145/1151030.1151053.

Ridder, H.-G. (2017). The theory contribution of case study research designs. *Business Research, 10*(2), 281–305. https://doi.org/10.1007/s40685-017-0045-z.

Rodriguez, M. A., Cuenca, L., & Ortiz, A. (2018). FIWARE open source standard platform in smart farming - a review. In *Proceedings of the IFIP Advances in Information and Communication Technology. Collaborative Networks of Cognitive Systems* (Vol. 534, pp. 581–589). https://doi.org/10.1007/978-3-319-99127-6_50.

Schreieck, M., Wiesche, M., & Krcmar, H. (2016). Design and governance of platform ecosystems - key concepts and issues for future research. In *Proceedings of the 24th European Conference on Information Systems* (Paper 76).

Segev, J., Gebauer, J., & Färber, F. (1999). Internet-based electronic markets. *Electronic Markets, 9*(3), 138–146. https://doi.org/10.1080/101967899359021.

Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS Quarterly, 35*(1), 37–56. https://doi.org/10.2307/23043488.

Sen, A. (2004). Metadata management: past, present and future. *Decision Support Systems, 37*(1), 151–173. https://doi.org/10.1016/S0167-9236(02)00208-7.

Shanks, G. (1997). The challenges of strategic data planning in practice: an interpretive case study. *The Journal of Strategic Information Systems, 6*(1), 69–90. https://doi.org/10.1016/S0963-8687(96)01053-0.

Staykova, K. S., & Damsgaard, J. (2015). A typology of multi-sided platforms: the core and the periphery. In *Proceedings of the 23rd European Conference on Information Systems* (Paper 176). https://doi.org/10.18151/7217486.

Staykova, K. S., & Damsgaard, J. (2017). Towards an integrated view of multi-sided platforms evolution. In *Proceedings of the 38th International Conference on Information Systems* (Paper 13).

Tan, B., Lu, X., Pan, S. L., & Huang, L. (2015). The role of IS capabilities in the development of multi-sided platforms: the digital ecosystem strategy of Alibaba.com. *Journal of the AIS, 16*(4), 248–280. https://doi.org/10.17705/1jais.00393.

Tan, F. T. C., Tan, B., & Pan, S. L. (2016). Developing a leading digital multi-sided platform: examining IT affordances and competitive actions in Alibaba.com. *Communications of the AIS, 38*, 738–760. https://doi.org/10.17705/1CAIS.03836.

Tian, C. H., Ray, B. K., Lee, J., Cao, R., & Ding, W. (2008). BEAM: A framework for business ecosystem analysis and modeling. *IBM Systems Journal, 47*(1), 101–114. https://doi.org/10.1147/sj.471.0101.

Timmers, P. (1998). Business models for electronic markets. *Electronic Markets, 8*(2), 3–8. https://doi.org/10.1080/10196789800000016.

Tiwana, A., Konsynski, B., & Bush, A. A. (2010). Platform evolution: coevolution of platform architecture, governance, and environmental dynamics. *Information Systems Research, 21*(4), 675–687. https://doi.org/10.1287/isre.1100.0323.

Tremblay, M. C., Hevner, A. R., & Berndt, D. J. (2010). Focus groups for artifact refinement and evaluation in design research. *Communications of the AIS, 26*, 599–618.

Tura, N., Kutvonen, A., & Ritala, P. (2018). Platform design framework: conceptualisation and application. *Technology Analysis & Strategic Management, 30*(8), 881–894. https://doi.org/10.1080/09537325.2017.1390220.

Van den Broek, T., & van Veenstra, A. F. (2015). Modes of governance in inter-organizational data collaborations. In *Proceedings of the 23rd European Conference on Information Systems* (Paper 188).

VDA (2016). *Access to the vehicle and vehicle generated data.* Berlin. Retrieved from VDA website https://www.vda.de/en/services/Publications/access-to-the-vehicle-and-vehicle-generated-data.html. Accessed 19 May 2019.

Weber, K., Otto, B., & Österle, H. (2009). One size does not fit all: a contingency approach to data governance. *Journal of Data and Information Quality, 1*(1), 1–27. https://doi.org/10.1145/1515693.1515696.

Weyl, E. G. (2010). A price theory of multi-sided platforms. *American Economic Review, 100*(4), 1642–1672. https://doi.org/10.1257/aer.100.4.1642.

Yin, R. K. (2014). *Case study research: design and methods* (5th ed.). Los Angeles, London, New Delhi, Singapore, Washington, DC: SAGE.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.