The Interplay of Data-Driven Organizations and Data Spaces: Unlocking Capabilities for Transforming Organizations in the Era of Data Spaces

Marius Hupperz
Fraunhofer ISST
marius.hupperz@isst.fraunhofer.de

Anna Gieß
Fraunhofer ISST
anna.giess@isst.fraunhofer.de

Abstract

This research paper highlights the relationship between data-driven organizations and data spaces and focuses on unlocking capabilities that can be used to transform organizations and to remain competitive in the era of data spaces. The increasing availability and diversity of data, as well as advances in technology, have led to the emergence of data spaces. However, to fully leverage these opportunities, organizations must be able to effectively access, process and utilize data from these data spaces. Through an in-depth examination of current literature, this paper explores the capabilities required for organizations to participate in data space activities. The TOE framework was used to structure the derived capabilities. The findings of this research provide insights into the capabilities that organizations and data spaces must consider when looking to co-innovate and realize new business cases. We anticipate that our paper will have significant implications for both practitioners and researchers.

Keywords: Data-Driven Organizations, Data Spaces, Capabilities, TOE Framework.

1. Introduction

Previous studies on digital transformation mainly looked at how organizations need to change their current decision-making processes, focusing on their own data (Berndtsson et al., 2018; Beverungen et al., 2022; Hupperz et al., 2021; Kearny et al., 2016). However, it is becoming increasingly important to understand its effects on various levels, including inter-organizational cooperation and data sharing. With the advent of data space (DS) initiatives, the participation in DSs and the offering of services for DSs are becoming a topic of increasing interest. Examples of these initiatives are the Mobility Data Space[1] or Catena-X[2] and operating companies like Cofinity-X[3] or Sovity[4]. We intend to blaze a trail for aligning the understanding of the digital transformation in the era of DSs in our research paper, as well as to highlight what DSs must offer data-driven organizations (DDO) and how DDOs participate in DSs. Researching the concept and impact of DSs and related ecosystems offers exciting opportunities, e.g., new business cases, in the field of DDOs, for research as well as for practitioners. Therefore, this paper aims to answer the following research question:

What should data-driven organizations and data spaces provide, demand, and expect from one another?

To answer this issue, we propose nine essential key capabilities and their features, which are based on a systematic literature review conducted in accordance with vom Brocke et al. (2015). “The term capabilities emphasized the key role of strategic management in appropriately adapting, integrating, and reconfiguring internal and external organizational skills, resources, and functional competencies toward changing environment” (Teece & Pisano, 1994). The remainder of the paper is structured as follows. The principles of capabilities, as well as several interpretations of DDOs and DSs, are introduced first. Subsequently, the research approach for producing the capabilities and the underlying framework is given. Building on this, individual capabilities are allocated in the TOE framework and described based on their features. The

[1] https://mobility-dataspace.eu
[2] https://catena-x.net/en/
[3] https://www.cofinity-x.com/
[4] https://sovity.de

URI: https://hdl.handle.net/10125/106925
978-0-9981331-7-1

paper finishes with a conclusion of the work's contribution and recommendations for future research.

2. Theoretical Background

In both professional and personal life, data is becoming increasingly important, particularly for industrial organizations' decision-making processes. This leads to a data-driven culture, where data outweighs opinions (Berndtsson et al., 2018). Organizations must embrace this disruptive change process and integrate the knowledge gained from information, which is based on big data, into existing and new processes (Rowley, 2007). Hupperz et al. (2021) describe five key elements of DDOs: Digital transformation, data science, data-driven business models, data-driven innovation and data analytics. Becoming a DDO is a transformational process, including working out a digital strategy, establishing a data-driven culture and interacting with other organizations in a business ecosystem (Hupperz et al., 2021). Data science is about the required IT professionals, who add value to the data, to gain a competitive advantage (Kearny et al., 2016; Vidgen et al., 2017). The work force seizes data analytics to gain business insights from big data (Berndtsson et al., 2018). This leads to data-driven innovation, which enables organizations to make use of their massive amounts of data (Hupperz et al., 2021). Further, data-driven business cases have to be created around the data insights to realize value through data (Guggenberger et al., 2020). These business cases can innovate already-existing processes or create a monetization through business models.

DDOs rely on data to facilitate their decision-making processes. The phrase "co-innovation" thereby represents how innovation is driven by cooperation through data sharing (Gelhaar & Otto, 2020). Data sharing is the domain-independent practice of granting third parties access to other people's data collections (Jussen et al., 2023). In the last years data sharing is gaining momentum, which is accelerated by data ecosystems and data marketplaces (Abbas et al., 2021; Fassnacht et al., 2023; Heinz et al., 2022; Lis & Otto, 2021). Data ecosystems comprise different actors such as producers, suppliers, competitors, and other stakeholders (Iansiti & Levien, 2004; Oliveira et al., 2019). The community aims to develop innovative products or services for customers who are also members of the ecosystem (Gelhaar, Groß, & Otto, 2021). By participating in such target oriented data ecosystems, organizations increase their chances of successfully exploiting data-driven business potentials (Heinz et al., 2022).

To enable organizations to participate in such data ecosystems and to share data, DSs are a crucial technology (Curry, 2020; Otto, 2022a). A DS connects data providers and data users directly through a provided infrastructure (Otto, 2022b). In contrast, a data ecosystem often relies on an actor as a keystone, which typically is a provider for stability and orchestration (Gelhaar & Otto, 2020). DSs, as opposed to central digital platforms, have a federated design, and hence provide new value creation choices based on data ecosystems. As a result, the advent of public DSs will bring an ecosystem perspective to the digital transformation, requiring inter-organizational cooperation as well as the digital transformation of organizations itself (Beverungen et al., 2022). For a better Understanding of what a DS is, we use the following definition: "A DS is a federated, open infrastructure for sovereign data sharing, based on common policies, rules and standards" (Reiberg et al., 2022).

The technology-organization-environment (TOE) framework created by Tornatzky et al. (1990) is used in this study to investigate the three dimensions that determine the capabilities of DDOs and DSs, namely the technological, organizational, and environmental settings (figure 1). Others, like the Leavitt's Model or TOGAF are focusing on organizational design (Saeed & Wang, 2013) or enterprise architecture (Kearny et al., 2016). According to the TOE framework, the adoption and diffusion of a new technology depends on the interaction of its three dimensions (Baker, 2012).

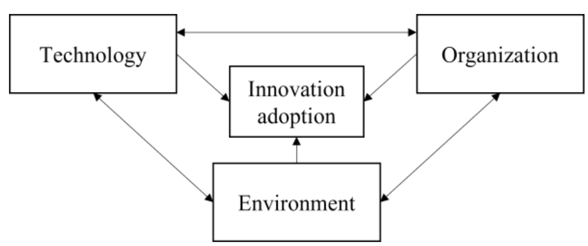

Figure 1. TOE Framework, adopted from Tornatzky et al. (1990)

The interplay of DDOs and DSs cannot merely be a technical endeavor; it also needs to consider organizational and environmental aspects. The TOE framework was chosen, because it is regarded as a comprehensive and holistic framework in analyzing innovation adoption (Gangwar et al., 2015). As DSs are a relatively new topic, organizations can see it as a big challenge to adopt these into their businesses.

3. Research Method

We chose a systematic literature review of the information systems literature to find capabilities of the interplay of DDOs and DSs.

To ensure the quality requirements of appropriate research rigor, breadth, brevity, constituency and clarity (Levy & Ellis, 2006), we followed the approach of vom Brocke et al. (2009) for the systematic literature review. The scope of the literature review is on theories and research findings in the research field of DDOs and DSs, with the purpose of outlining central capabilities that we built using a concept matrix (Webster & Watson, 2002). We aim for a neutral representation of a representative coverage of the literature for general scholars as well as practitioners (vom Brocke et al., 2009). The four literature databases selected were "ACM", "IEEE Xplore", "Scopus" and "AISeL", as these databases include relevant information systems research journals and conference proceedings. We searched the databases for pertinent keywords "data space", "data-driven", and "organization". We employed both American and Oxford English for our keywords to compile a decent literature base. Therefore, we distinguished our search query as follows:

["data space" ∧ "data-driven"] ∨
["data space" ∧ "organization"] ∨
["data space" ∧ "organisation"]

Because there is no clear description of DSs in the context of DDOs, these keywords were chosen. We excluded studies that fell into any of the five exclusion criteria (EC), which were as follows:

- **EC1:** The study is not accessible on the web;
- **EC2:** The study does not present any type of findings or discussion about DDOs or DSs;
- **EC3:** The study does not have one of the keywords in the abstract;
- **EC4:** The study is not written in English;
- **EC5:** The study is a duplicate.

We will use the TOE Framework as guiding dimensions to map the capabilities derived from the structured literature review and allocate the characteristics of DDOs and DSs. We understand capabilities as recurring patterns of actions (Wade & Hulland, 2004) or a coordinated collection of tasks that use the firm's assets as input (Helfat & Peteraf, 2003). The representation will be done with a morphological box and the Nickerson et al. (2009) approach. We chose the morphological box representation since it is widespread method in information systems literature, such as Gelhaar, Groß, and Otto (2021), Lis and Otto (2021) or Rizk et al. (2018).

4. Capabilities of Data-Driven Organizations and Data Spaces

The digital transformation in the era of DSs demands organizations to develop capabilities to contribute to DSs, while at the same time DSs need to develop capabilities to support organizations to enter a DS.

A total of 2174 publications could be identified through the systematic literature review, which led to 647 publications been filtered by title and keywords. After analyzing the abstracts 60 publications went under a full text analysis. We identified 29 relevant studies. As proposed by Webster and Watson (2002),

Table 1. Morphological box of the core capabilities of the interplay of DDOs and DSs.

<table>
<tr><th rowspan="2">Dim.</th><th rowspan="2">Capabilities</th><th colspan="4">Characteristic</th></tr>
<tr><th colspan="2">Data-Driven Organization (DDO)</th><th colspan="2">Data Space (DS)</th></tr>
<tr><td rowspan="4">Technology</td><td>Architecture [56%]</td><td colspan="2">Connector</td><td colspan="2">Reference Architecture</td></tr>
<tr><td>Services [41%]</td><td colspan="2">Data Processing</td><td colspan="2">Service Repository</td></tr>
<tr><td>Data Quality [28%]</td><td colspan="2">Data Availability & Accessibility</td><td>Tools</td><td>Assessment</td></tr>
<tr><td>Interoperability [53%]</td><td colspan="2">Usage Control</td><td>Vocabularies & Ontologies</td><td>APIs</td></tr>
<tr><td rowspan="3">Organiz.</td><td>Business [41%]</td><td colspan="2">Compensation</td><td colspan="2">Data Ecosystem</td></tr>
<tr><td>Governance [38%]</td><td colspan="2">Intra-Organizational</td><td colspan="2">Inter-Organizational</td></tr>
<tr><td>Design [41%]</td><td colspan="2">Domain-Specification</td><td colspan="2">Organizational Structure</td></tr>
<tr><td rowspan="2">Env.</td><td>Trust [53%]</td><td colspan="2">Relationships</td><td>Identification</td><td>Data Sovereignty</td></tr>
<tr><td>Actor [50%]</td><td>Data Provider</td><td>Data User</td><td colspan="2">Intermediaries</td></tr>
</table>

we conducted a backward search, which resulted in three additional papers. In total, 32 papers were examined.

The key capabilities and characteristics of the interplay of DDOs and DSs are represented as a morphological box in table 1. Through the coding process of the literature nine capabilities and 22 characteristics were derived. Both, implicitly and explicitly concepts were considered in the identification and aggregation of the morphological box. The coding process was supported by a concept matrix[5], like suggested from Webster and Watson (2002). The capabilities architecture, services, data quality, and interoperability are allocated to the technology dimension. The organizational dimension comprises the business, governance, and design capabilities. Finally, the trust, and actor capabilities are part of the environmental dimension. In addition, the percentage of literature citing each capability is indicated by the numbers in brackets in table 1.

4.1. Technology

The first dimension is technology, which focuses on the data provisioning by the organizations and within the DS as well as the architecture for data sharing.

The capability **architecture** describes the *reference architecture* of DSs and the participation of DDOs including (DS) components and infrastructure. The architecture capability addresses the main technical component that a DDO uses to share its data within the ecosystem, namely a *connector*, for peer-to-peer data transfer. The used architecture affects, among other things, data security, data control, and trust between the ecosystem actors (Gelhaar, Groß, & Otto, 2021). The technical capabilities an organization needs for this are a storage system and a processing engine on the data consumer side, which can be fulfilled by a connector as well (Munoz-Arcentales et al., 2019). A connector provides a trusted environment that enables the achievement of data access and usage control (Brost et al., 2018; Janev et al., 2021; Munoz-Arcentales et al., 2020; Otto, 2022b). The connector represents all devices running services, which produce or consume data, or which provide an ecosystem managing functionality. In addition, services on connectors can represent infrastructure services for the management of the ecosystem, such as for service downloads or service discovery (Brost et al., 2018). This leads to one of the main weaknesses of data sharing in DSs. Even if private sector organizations are willing to share data, many data sharing activities are hindered by a lack of *reference architectures*, data models, technological interoperability, and data compatibility among organizational systems (Fassnacht et al., 2023; Janev et al., 2021; Li et al., 2021; Pullmann et al., 2017). DSs need a 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 (Otto & Jarke, 2019). This underlines the clear need of DSs to provide a reference architecture model (Moller et al., 2021; Otto, 2022b). As mentioned above, DSs need to describe how participants should build a connector to enable the data exchange with other participants. Additionally, the broker component is a directory for all data made available through connectors (Brost et al., 2018). For this purpose, the broker component stores the metadata in the DS to allow for searching and querying the data (Brost et al., 2018; Mrityunjay & Jain, 2011). We presume that the use of a central broker instance may not always be necessary if queries can be sent directly from one connector to another.

On the side of DDOs the capability **services** comprise the characteristics of *processing data*, including gathering and exchanging. Data that already exists or is generated during service use must be further processed into information and knowledge in order to gain insights and create value (Hupperz et al., 2021; Rizk et al., 2018). If the data exchange cannot be managed by the organization itself, such as in the case of small organizations, the organization may need a third-party provider of data services. Additionally, services on a connector can include infrastructure services for a dedicated DS application (Brost et al., 2018). DSs must encourage direct interactions between service providers and customers, to offer services to make data usable, deliver data-based insights, provide data-based recommendations and enable novel ways to conduct business (Beverungen et al., 2022). The service a DS needs to provide is *a service repository*, allowing for service registration. Regarding data exchange and registration, computational techniques for exchanging and registering data while protecting personal data privacy and security are required (Geisler et al., 2022). In regard to service repositories DSs need to provide support services that facilitate the DS usage and application development within a DS (Curry &

[5] https://owncloud.fraunhofer.de/index.php/s/JtI9izNanRgL8Gt

Tuikka, 2022). A service repository should allow organizations to download a service to retrieve data from a corresponding service of a third party (Brost et al., 2018; Janev et al., 2021). The repository could contain for instance analytic services (Pullmann et al., 2017; Zeiner & Unterberger, 2021), data storage services (Shah et al., 2021), or support services to anonymize data or define usage restrictions (Geisler et al., 2022).

Within the **data quality** capability, a DDO needs to ensure *data availability and accessibility.* Data quality is a concept comprising multiple dimensions which represent different characteristics of data (Wang & Strong, 1996). Also, it is the foundation for every important activity within an organization (Altendeitering et al., 2022). However, organizations often lack in availability and accessibility of data (Gelhaar, Gürpinar, et al., 2021; Otto & Jarke, 2019). Within organizations as well as while sharing data with other organizations, data must be made traceable and usable. This can often be attributed to metadata, which describe e.g., the content, methods, and origins of data (Fassnacht et al., 2023; Janev et al., 2021). Data quality is also a central aspect in DSs. DSs need to offer *data quality tools* for participating organizations and continuous *data quality assessment* of the data available in the DS. DSs should provide standard data quality tools for describing, querying, and assessing data quality values (Chou & Jiang, 2022). This requires the creation of a data quality model comprised of data quality characteristics and measures (Geisler et al., 2022). Data quality assessment is pertinent to ensure that the quality and integrity of the data is high (Basole, 2020).

To share data, organizations need to reach **interoperability** through common standards (Moller et al., 2021). Broadly considered there are three types of data access methods, including APIs, export/downloads/data dumps, and web scrape (Basole, 2020). The data usage control architecture must enable the policy creation, verification, and enforcement. This allows sharing data, and managing identity, access, and *usage control* (Munoz-Arcentales et al., 2019). Crucial for this is standardization, to ensure the understanding between all the participants involved in the architecture, for managing both usage control and identity (Munoz-Arcentales et al., 2020). For interoperability DSs need to provide *vocabularies, ontologies,* and *APIs.* A domain-specific ontology is frequently used to detect schema matches with the ontology when integrating several diverse data sources. Another method is to use a domain-specific vocabulary to establish a common understanding regarding the naming of attributes within a DS (Janowicz et al., 2014; Moller et al., 2021; Pullmann et al., 2017). Specific technological, legal, or organizational criteria that actors must follow based on their function in the ecosystem are examples of rules and regulations used in DS governance. For example, publishing related metadata to data assets using standard vocabularies, according to specified regulations, and implementing specific technical architectures (Torre-Bastida et al., 2022). The semantic layer's implementation might range from incorporating taxonomies and metadata schema to more in-depth knowledge engineering methodologies such as ontologies and semantic deduction (Anjomshoaa et al., 2022). The standardization of these interfaces plays a huge role in building DSs. Organizations need to make sure the standardization is satisfying others, so they can use the provided data. Additionally, the DS needs to provide standardized APIs for data access (Zeiner & Unterberger, 2021). The quality and extent of the APIs can vary greatly, with some having very rich data extraction capabilities while others are more limited (Basole, 2020). In its most basic version, the DS includes search, catalog, and query functions, as well as API support. These APIs serve as the interface for external application or cloud services (Exposito Jimenez & Zeiner, 2018; Zeiner & Unterberger, 2021), and they support and provide relevant visualization options (Zeiner & Unterberger, 2021).

4.2. Organization

The second dimension is organization, it describes the organizational design and the data insights to be discovered.

The first capability is **business**, comprising *compensation* for DDOs for co-innovated data products and services. To achieve novel data based value propositions organizations might require access to external data sources and inter-organizational collaboration (Heinz et al., 2022). Data-driven business models are less created by a single organization, but rather they are developed by multiple organizations (Hupperz et al., 2021). For DDOs it is a constant challenge to co-innovate more sophisticated solutions that might be able to solve heretofore unsolved problems that feature high complexity, while at the same time the cooperation among diverse participants can be inefficient. To find data-driven solutions for business problems, DDOs need to back up their innovation potential by speeding up their digital transformation, whether or not they co-innovate with others (Beverungen et al., 2022). Additionally, provided data-driven services offer value by giving the ability to monitor, control or optimize processes utilizing various degrees of autonomy (Werling et al., 2022). This requires a compensation for data providers. Gelhaar, Gürpinar, et al. (2021) point out five possible forms of reward: A payment (money), compensation through distributed ledger technologies (virtual assets), exchange for service, reputation for the data provider, and no reward (e.g. government agencies or non-profit organizations). Data-driven innovation opens business insights to organizations which have been previously undiscovered (Hupperz et al., 2021). The business capability for DS is characterized by the *data ecosystem* approach allowing new business model designs (Janev et al., 2021). As the DS is influencing the organization's business goals, they should be aligned. The DDO should define and agree on the DS initiative's approach and scope of aims. Furthermore, the organization's and the ecosystem's progress toward particular DS objectives should be monitored (Curry & Tuikka, 2022). Business, certification, and utility models must be developed to certify the monetary worth of transferred and converted data based on data quality values and domain specific features (Geisler et al., 2022). Economic value is less created by a single organization, so that co-innovation, besides the possibility of inefficiency, is much more needed (Gelhaar & Otto, 2020). Traditional offerings evolve towards smart products and services, increased connectivity is reflected in emerging (smart) service systems, and digital platforms enable modularization and standardization. Therefore, business model design must reflect multilateral win-win potentials, so that every organization involved in the co-innovation gets their stake (Heinz et al., 2022).

Data **governance** takes a huge role in organizations and between organizations. Enabling data governance is critical for the right exchange and sharing of data in accordance with the companies' plans and business models (Geisler et al., 2022; Shah et al., 2021). Lis and Otto (2020) describe internal data governance as *intra-organizational* and external data governance as *inter-organizational*. The intra-organizational data governance which organizations need to establish, consists of a governance model including roles and responsibilities to improve data quality, the management of resources across a single organization and the formalization of guidelines for data resources (Lis & Otto, 2020). If organizations are willing to share their data, the data governance rules must be secured at any given time and moreover monitoring and control techniques must be provided (Munoz-Arcentales et al., 2020). Besides, the intra-organizational data governance cannot fulfill the need for a governance model in a DS (Lis & Otto, 2020; Otto & Jarke, 2019). Within a DS a "second" governance, the inter-organizational data governance, is needed, which coordinates the interactions between the DS and the participants. Additionally, organizations need to know which role they want to take and which implications are arising from certain collaborations (Otto & Jarke, 2019). Within a DS the control over the data resources is crucial (Gelhaar, Groß, & Otto, 2021). This underlines the need for usage control in terms of data protection, security concepts and the ability to establish policies in a DS (Beverungen et al., 2022).

According to Gelhaar, Gürpinar, et al. (2021) there are four categories that can be considered for the **design** of a DS. These *domain-specifications* are scientific, government, industry, and personal data domain. Organizations participating in a DS activity need to pick the appropriate domain for them. Often the appropriate domain for an organization is already predefined by the organization itself. Scientific DSs are for researchers to share their research data, governmental DSs are for data provided by public agencies or government units to create value for citizens or companies, in industrial DSs companies or private organizations provide data from e.g. IT systems, sensors or physical devices, and personal DSs are for data collected by and about people e.g. volunteered data or observed data (Gelhaar, Gürpinar, et al., 2021). The organizational design describes the relationships, interactions, and organization of a DS. Gelhaar, Groß, and Otto (2021) describe four forms of the *organizational structure*: Keystone-centric, platform-centric, marketplace-based, and decentralized. The keystone-centric DS is governed by a leader who usually captures most value (Piller et al., 2021). In a platform-based DS a provider offers the infrastructure and services to support the sharing and usage of data (Azkan et al., 2020). A marketplace-based structure provides, beside a technical platform, additional components and functions, e.g. business models, applications, rules and services for data sharing, as part of the data ecosystem infrastructure. Additionally, DSs may be decentralized, without a central actor, providing the infrastructure and services needed in a federated way (Gelhaar, Groß, & Otto, 2021).

4.3. Environment

The third dimension is environment. This dimension focuses on establishing a trusted environment for data sharing as well as the roles and actors within.

According to Gelhaar and Otto (2020) building **trust** between ecosystem actors is one of the biggest cooperative challenges in the emergence of data ecosystems. Trust builds the foundation of data sharing and facilitates new *relationships* between organizations (Haak et al., 2018; Munoz-Arcentales et al., 2020; Pullmann et al., 2017). Furthermore, diversity and equality in data sets shared, transferred, and processed via data-driven pipelines should be pursued (Geisler et al., 2022). To show that an organization is a trustworthy partner the common rules and the inter-organizational data governance of the DS (see 4.2) must be followed. Additionally, organizations must have the performance to react quick and efficient to the accomplished data usage control policies and have the flexibility to adapt to the specific requirements of different data sharing scenarios (Munoz-Arcentales et al., 2020). Flexibility becomes even more important, when thinking about the fact that DSs allow different and new actor constellations, while nowadays trust is mostly build by leveraging existing business relationships (Heinz et al., 2022). To achieve trust, data usage control techniques need to be established in DSs (Munoz-Arcentales et al., 2019). For these it is important how the data provider can retain ownership and sovereignty over the data and thus control its further use (Gelhaar, Gürpinar, et al., 2021). *Data sovereignty* itself is a core aspect of DSs (Munoz-Arcentales et al., 2019). In a broader sense, data sovereignty is the decision making authority to determine the usage of one's own data and the results obtained from it (Aydin & Bensghir, 2019). Data sovereignty enables the control and optimization of data access, thus ensuring a secure, reliable, and transparent approach for data access and exchange between different parties (Beverungen et al., 2022; Lis & Otto, 2020). The level of assurance given to data owners over their data is an important capability of the DS (Curry & Tuikka, 2022). By providing expressive legal frameworks for data exchange, such as legal references, duties, licenses, and ethical principles, trust is established. Enforcing data protection and ownership safeguards privacy, sovereignty, and legal compliance with licenses and rules for data sharing, exchange, and processing. To ensure trust, participants need to *identify* themselves in a DS. For this purpose the operator of a DS needs to provide the identity provider component (Brost et al., 2018; Otto, 2022b). Additionally, the identity provider works as an administrator for the DS and locates new participants in the DS (Mrityunjay & Jain, 2011).

In data ecosystems different **actor** constellations can be observed. Generic roles can be used to facilitate the beneficial and intended integration of new users into an ecosystem (Heinz et al., 2022; Otto, 2022b). For data sharing two actors are mandatory, the *data provider* and the *data consumer* (Pullmann et al., 2017; Shah et al., 2021). The data owner and the data user can be added to these roles (Munoz-Arcentales et al., 2020). Both, the data owner and data provider, as well as the data user and data consumer can also be one agent (Pullmann et al., 2017). We consider these actors as typical roles for DDOs in data sharing environments. On the side of the DSs a wide range of actors can be observed. Otto and Jarke (2019) describe DSs as multi-sided platforms, which are sociotechnical constructs, being technical platforms and market *intermediaries*. Actors on the DS side can be technical providers as well as market intermediaries. In a DS an identity provider verifies all actors involved in the architecture and also provides all the characteristics related to the identity management (Munoz-Arcentales et al., 2020). In addition to that, Munoz-Arcentales et al. (2020) mention a data controller, defined by the GDPR, being responsible with the data provider to guarantee the protection of the data owners' rights. Azkan et al. (2022) describe seven further types of intermediaries, being service, data infrastructure, and app store providers as well as ecosystem orchestrator and data trustee.

5. Conclusion and Outlook

The interplay of DDOs and DSs presents a unique opportunity for transforming organizations in the era of data spaces. This paper developed a morphological box to capture and organize the capabilities for data sharing in DSs. These capabilities were split into characteristics of DDOs and DSs and show what to provide, demand, and expect from one another.

We can draw various conclusions about theory and practice from our findings. In terms of scientific contributions, we aim to close the gap between seeing the emergence of DSs and looking at the digital transformation of organizations differently. We consider the interplay of DDOs and DSs as a cross-industry changing event in the upcoming years. Our work suggests a way to better understand the relationship between DSs and DDOs. Therefore, the developed morphological box represents characteristics of what DDOs and DSs need to provide, demand, and expect from each other. Each capability presents an exceptional opportunity for in-depth research projects, which are valid on their own, but connected and contextualized through our morphological box mapped on the TOE framework. While our findings are high-level, they enable looking at DDOs and DSs with reduced complexity. That, naturally, both is a limitation through the abstraction of detail and a contribution through its general and harmonizing scope. The findings of this study also have implications for practitioners, on the side of organizations as well as DS providers. Our work, which provides a connected view on DDOs and DSs, can serve as a starting point for utilizing DS activities in the digital transformation process. For example, organizations interested in joining a DS, get an overview of the technological, organizational, and environmental topics to focus on. The same applies to initiatives establishing a new domain-specific DS.

Naturally, our study has limitations that must be noted when interpreting the results. Firstly, the concept around DSs is continually developing. Therefore, our morphological box should be understood as a time-bound snapshot which must be updated periodically to remain relevant. As a result, new capabilities may emerge, or old ones may lose relevance. Secondly, the data collection itself is subject to interpretation. Thus, other researchers may infer different capabilities depending on their influences, preferences, and biases. Thirdly, the results are only based on literature research. The third limitation also brings us to future research possibilities. A possible next step could be to go through the morphological box with practitioners and to evaluate our theoretical results. Another possible research opportunity building on our research could be the derivation of archetypical patterns or maturity levels, which could be beneficial for measuring the extent to which organizations are able to participate in DS activities.

In summary, the interplay between DDOs and DSs has the potential to drive transformational change across industries. The capabilities described in this paper help readers understand what DDOs must contribute to DSs to participate, as well as what DSs must provide to encourage the sharing of data.

Acknowledgements

This work was supported by the German State Government of North Rhine-Westphalia (NRW) and is part of the research project "MDSxNRW – Connection of mobility data from NRW to the Mobility Data Space" (25.25-MDSxNRW).

References

Abbas, A. E., Agahari, W., van de Ven, M., Zuiderwijk, A., & Reuver, M. de (2021). Business data sharing through data marketplaces: A systematic literature review. *Journal of Theoretical and Applied Electronic Commerce Research*, *16*(7), 3321–3339. https://doi.org/10.3390/jtaer16070180

Altendeitering, M., Dübler, S., & Guggenberger, T. M. (2022). Data quality in data ecosystems: Towards a Design Theory. *AMCIS 2022 Proceedings*, Article 3.

Anjomshoaa, A., Elvira, S. C., Wolff, C., Pérez Baún, J. C., Karvounis, M., Mellia, M., Athanasiou, S., Katsifodimos, A., Garatzogianni, A., Trügler, A., Serrano, M., Zappa, A., Glikman, Y., Tuikka, T., & Curry, E. (2022). Data platforms for data spaces. In E. Curry, S. Scerri, & T. Tuikka (Eds.), *Data spaces* (pp. 43–64). Springer International Publishing. https://doi.org/10.1007/978-3-030-98636-0_3

Aydin, A., & Bensghir, T. K. (2019). Digital data sovereignty: Towards a conceptual framework. In *2019 1st International Informatics and Software Engineering Conference (UBMYK)* (pp. 1–6). IEEE. https://doi.org/10.1109/UBMYK48245.2019.8965469

Azkan, C., Möller, F., Ebel, M., Iqbal, T., Otto, B., & Poeppelbuss, J. (2022). Hunting the treasure: Modeling data ecosystem value co-creation. *ICIS 2022 Proceedings*, Article 14. https://aisel.aisnet.org/icis2022/entren/entren/14

Azkan, C., Möller, F., Meisel, L., & Otto, B. (2020). Service dominant logic perspective on data ecosystems: A Case Study Based Morphology. *Proceedings of the 28th European Conference on Information Systems (ECIS)*. https://aisel.aisnet.org/ecis2020_rp/65

Baker, J. (2012). The technology–organization–environment framework. In Y. K. Dwivedi, M. R. Wade, & S. L. Schneberger (Eds.), *Integrated Series in Information Systems: Vol. 28. Information systems theory: Explaining and Predicting Our Digital Society, Vol. 1* (Vol. 28, pp. 231–245). Springer New York. https://doi.org/10.1007/978-1-4419-6108-2_12

Basole, R. (2020). Understanding ecosystem data. In T. Bui (Ed.), *Proceedings of the Annual Hawaii International Conference on System Sciences, Proceedings of the 53rd HICSS.* Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2020.702

Berndtsson, M., Forsberg, D., Stein, D., & Thomas, S. (2018). Becoming a data-driven organization. *ECIS 2018 Proceedings*, 1–9.

Beverungen, D., Hess, T., Köster, A., & Lehrer, C. (2022). From private digital platforms to public data spaces: implications for the digital transformation. *Electronic Markets*, *32*(2), 493–501. https://doi.org/10.1007/s12525-022-00553-z

Brost, G. S., Huber, M., Weiß, M., Protsenko, M., Schütte, J., & Wessel, S. (2018). An ecosystem and IoT device architecture for building trust in the industrial data space. In D. Gollmann & J. Zhou (Eds.), *Proceedings of the 4th ACM Workshop on Cyber-Physical System Security* (pp. 39–50). ACM. https://doi.org/10.1145/3198458.3198459

Chou, D., & Jiang, M. (2022). A survey on data-driven network intrusion detection. *ACM Computing Surveys*, *54*(9), 1–36. https://doi.org/10.1145/3472753

Curry, E. (2020). Future research directions for dataspaces, data ecosystems, and intelligent systems. In E. Curry (Ed.), *Real-time linked dataspaces* (pp. 297–304). Springer Nature. https://doi.org/10.1007/978-3-030-29665-0_18

Curry, E., & Tuikka, T. (2022). An organizational maturity model for data spaces: A data sharing wheel approach. In E. Curry, S. Scerri, & T. Tuikka (Eds.), *Data spaces* (pp. 21–42). Springer International Publishing. https://doi.org/10.1007/978-3-030-98636-0_2

Exposito Jimenez, V. J., & Zeiner, H. (2018). Serverless cloud computing : A comparison between "function as a service" platforms. In *Computer Science & Information Technology* (pp. 15–22). Academy & Industry Research Collaboration Center (AIRCC). https://doi.org/10.5121/csit.2018.80702

Fassnacht, M., Benz, C., Heinz, D., & Leimstoll, Jannis, Satzger, Gerhard (2023). Barriers to data sharing among private sector organizations. *Proceedings of the Hawaii International Conference on Systems Sciences (HICSS-56)*.

Gangwar, H., Date, H., & Ramaswamy, R. (2015). Understanding determinants of cloud computing adoption using an integrated TAM-TOE model. *Journal of Enterprise Information Management*, *28*(1), 107–130. https://doi.org/10.1108/JEIM-08-2013-0065

Geisler, S., Vidal, M.-E., Cappiello, C., Lóscio, B., Gal, A., Jarke, M., Lenzerini, M., Missier, P., Otto, B., Paja, E., Pernici, B., & Rehof, J. (2022). Knowledge-driven data ecosystems toward data transparency. *Journal of Data and Information Quality*, *14*(1), 1–12. https://doi.org/10.1145/3467022

Gelhaar, J., Groß, T., & Otto, B. (2021). A taxonomy for data ecosystems. In T. Bui (Ed.), *Proceedings of the Annual Hawaii International Conference on System Sciences, Proceedings of the 54th HICSS.* Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2021.739

Gelhaar, J., Gürpinar, T., Henke, M., & Otto, B. (2021). Towards a taxonomy of incentive mechanisms for data sharing in data ecosystems. *PACIS 2021 Proceedings*, Article 121. https://aisel.aisnet.org/pacis2021/121

Gelhaar, J., & Otto, B. (2020). Challenges in the emergence of data ecosystems. *PACIS 2020*.

Guggenberger, T. M., Möller, F., Boualouch, K., & Otto, B. (2020). Towards a unifying understanding of digital business models. *PACIS 2020 Proceedings*(70).

Haak, E., Ubacht, J., van den Homberg, M., Cunningham, S., & van den Walle, B. (2018). A framework for strengthening data ecosystems to serve humanitarian purposes. In M. Janssen, S. A. Chun, V. Weerakkody, A. Zuiderwijk, & C. Hinnant (Eds.), *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age* (pp. 1–9). ACM. https://doi.org/10.1145/3209281.3209326

Heinz, D., Benz, C., Fassnacht, M., & Satzger, G. (2022). Past, present and future of data ecosystems research: A Systenatic Literature Review. *PACIS 2022 Proceedings*, Article 46. https://aisel.aisnet.org/pacis2022/46

Helfat, C., & Peteraf, M. (2003). The dynamic resource-based view: capability lifecycles. *Strategic Management Journal*, *24*(10), 997–1010. https://doi.org/10.1002/smj.332

Hupperz, M., Gür, I., Möller, F., & Otto, B. (2021). What is a data-driven organization? *Americas Conference on Information Systems*.

Iansiti, M., & Levien, R. (2004). *The keystone advantage: What the new dynamics of business ecosystems mean for strategy, innovation, and sustainability*. Harvard Business School Press.

Janev, V., Vidal, M. E., Endris, K., & Pujic, D. (2021). Managing knowledge in energy data spaces. In J. Leskovec (Ed.), *ACM Digital Library, Companion Proceedings of the Web Conference 2021* (pp. 7–15). Association for Computing Machinery. https://doi.org/10.1145/3442442.3453541

Janowicz, K., Hitzler, P., Adams, B., Kolas, D., & Vardeman II, C. (2014). Five stars of linked data vocabulary use. *Semantic Web*, *5*(3), 173–176. https://doi.org/10.3233/SW-140135

Jussen, I., Schweihoff, J., Dahms, V., Möller, F., & Otto, B. (2023). Data sharing fundamentals: Characteristics and definition. In *Proceedings of the 56th Annual HICSS: January 3-6, 2023*.

Kearny, C., Gerber, A., & van der Merwe, A. (2016). Data-driven enterprise architecture and the TOGAF ADM phases. In *SMC: 9-12 Oct. 2016* (pp. 4603–4608). IEEE. https://doi.org/10.1109/SMC.2016.7844957

Levy, Y., & Ellis, T. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science: The International Journal of an Emerging Transdiscipline*, *9*, 181–212. https://doi.org/10.28945/479

Li, K., Yuan, L., Zhang, Y., & Yue, Y. (2021). Reducing redundancy in data organization and arithmetic calculation for stencil computations. In B. R. de Supinski (Ed.), *ACM Digital Library, Proceedings of the International Conference HPC* (pp. 1–15). Association for Computing Machinery. https://doi.org/10.1145/3458817.3476154

Lis, D., & Otto, B. (2020). Data governance in data ecosystems: Insights from Organizations. *AMCIS 2020 Proceedings*, Article 12. https://aisel.aisnet.org/amcis2020/strategic_uses_it/strategic_uses_it/12

Lis, D., & Otto, B. (2021). Towards a taxonomy of ecosystem data governance. In T. Bui (Ed.), *Proceedings of the Annual Hawaii International Conference on System Sciences, Proceedings of the 54th HICSS.* Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2021.733

Moller, J., Jankowski, D., & Hahn, A. (2021). Towards an architecture to support data access in research data spaces. In *2021 IEEE 22nd International Conference on Information Reuse and Integration for Data Science: IRI 2021 : virtual conference, 10-12 August 2021 : proceedings* (pp. 310–317). IEEE. https://doi.org/10.1109/IRI51335.2021.00049

Mrityunjay, S., & Jain, S. (2011). A survey on dataspace. *International Conference on Network Security and Applications*.

Munoz-Arcentales, A., López-Pernas, S., Pozo, A., Alonso, Á., Salvachúa, J., & Huecas, G. (2019). An architecture for providing data usage and access control in data sharing ecosystems. *Procedia Computer Science*, *160*, 590–597. https://doi.org/10.1016/j.procs.2019.11.042

Munoz-Arcentales, A., López-Pernas, S., Pozo, A., Alonso, Á., Salvachúa, J., & Huecas, G. (2020). Data usage and access control in industrial data spaces:

Implementation using FIWARE. *Sustainability*, *12*(9), 3885. https://doi.org/10.3390/su12093885
Nickerson, R., Varshney, U., Muntermann, J., & Isaac, H. (2009). Taxonomy development in information systems: Developing a taxonomy of mobile applications. *ECIS 2009 Proceedings*(388).
Oliveira, M. I., Barros Lima, G. d. F., & Farias Lóscio, B. (2019). Investigations into data ecosystems: A systematic mapping study. *Knowledge and Information Systems*, *61*(2), 589–630. https://doi.org/10.1007/s10115-018-1323-6
Otto, B. (2022a). The evolution of data spaces. In B. Otto (Ed.), *Designing data spaces* (1st ed., pp. 3–15). Springer International Publishing. https://doi.org/10.1007/978-3-030-93975-5_1
Otto, B. (2022b). A federated infrastructure for european data spaces. *Communications of the ACM*, *65*(4), 44–45. https://doi.org/10.1145/3512341
Otto, B., & Jarke, M. (2019). Designing a multi-sided data platform: findings from the international data spaces case. *Electronic Markets*, *29*(4), 561–580. https://doi.org/10.1007/s12525-019-00362-x
Piller, F., van Dyck, M., Lüttgens, D., & Diener, K. (2021). Positioning strategies in emerging industrial ecosystems for industry 4.0. In T. Bui (Ed.), *Proceedings of the Annual Hawaii International Conference on System Sciences, Proceedings of the 54th HICSS.* Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2021.743
Pullmann, J., Petersen, N., Mader, C., Lohmann, S., & Kemeny, Z. (2017). Ontology-based information modelling in the industrial data space. In *2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)* (pp. 1–8). IEEE. https://doi.org/10.1109/ETFA.2017.8247688
Reiberg, A., Niebel, C., & Kraemer, P. (2022). *What is a data space? White Paper 1/2022*.
Rizk, A., Bergvall-Kåreborn, B., & Elragal, A. (2018). Towards a taxonomy for data-driven digital services. In T. Bui (Ed.), *Proceedings of the Annual Hawaii International Conference on System Sciences, Proceedings of the 51st HICSS.* Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2018.135
Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, *33*(2), 163–180. https://doi.org/10.1177/0165551506070706
Saeed, B. B., & Wang, W. (2013). Organisational diagnoses: a survey of the literature and proposition of a new diagnostic model. *International Journal of Information Systems and Change Management*, *6*(3), Article 58328, 222. https://doi.org/10.1504/IJISCM.2013.058328
Shah, S. I. H., Peristeras, A. V., & Magnisalis, I. (2021). Government big data ecosystem: Definitions, types of data, actors, and roles and the impact in public administrations. *Journal of Data and Information Quality*, *13*(2), 1–25. https://doi.org/10.1145/3425709
Teece, D., & Pisano, G. (1994). The Dynamic Capabilities of Firms: an Introduction. *Industrial and Corporate Change*, *3*(3), 537–556. https://doi.org/10.1093/icc/3.3.537-a
Tornatzky, L., Chakrabarti, A. K., & Fleischer, M. (1990). *The processes of technological innovation*. *Issues in organization and management series*. Lexington Books.
Torre-Bastida, A. I., Gil, G., Miñón, R., & Díaz-de-Arcaya, J. (2022). Technological Perspective of Data Governance in Data Space Ecosystems. In E. Curry, S. Scerri, & T. Tuikka (Eds.), *Data spaces* (pp. 65–87). Springer International Publishing. https://doi.org/10.1007/978-3-030-98636-0_4
Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research*, *261*(2), 626–639. https://doi.org/10.1016/j.ejor.2017.02.023
vom Brocke, J., Simons, A., Niehaves, B., Reimer, K., Plattfaut, R., & Cleven, A. (2009). Reconstructing the giant: On the importance of rigour in documenting the literature search process. *ECIS 2009 Proceedings*, *161*. https://aisel.aisnet.org/ecis2009/161
vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the Association for Information Systems*, *37*. https://doi.org/10.17705/1CAIS.03709
Wade, & Hulland (2004). Review: The Resource-Based View and Information Systems Research: Review, Extension, and Suggestions for Future Research. *MIS Quarterly*, *28*(1), 107. https://doi.org/10.2307/25148626
Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, *12*(4), 5–33. https://doi.org/10.1080/07421222.1996.11518099
Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. In *MIS Quarterly* (Vol. 26, pp. 13–23).
Werling, M., Weber, P., & Tank, A. (2022). Value modeling in IoT ecosystems with a central trusted entity: Qualitative Interviews and Explorative Case Study. *AMCIS 2022 Proceedings*, Article 4. https://aisel.aisnet.org/amcis2022/DataEcoSys/DataEcoSys/4
Zeiner, H., & Unterberger, R. (2021). Time-aware data spaces - A key computing unit in the edge-to-cloud continuum. In M. Younas (Ed.), *2021 International Conference on Future Internet of Things and Cloud: 23-25 August 2021, virtual (online)* (pp. 250–255). IEEE. https://doi.org/10.1109/FiCloud49777.2021.00043