The adoption of data spaces: Drivers toward federated data sharing

Andreas Hutterer
Johannes Kepler University Linz
andreas.hutterer@jku.at

Barbara Krumay
Johannes Kepler University Linz
barbara.krumay@jku.at

Abstract

Data spaces have gained increasing attention, as they allow federated data sharing among and within participants of interoperable data spaces, for the benefit of all. However, data space initiatives are few in number; moreover, data space adoption among organizations is low. Research thus far has mainly focused on technical factors but lacks a more holistic approach that clarifies what drives data space adoption and federated data sharing as main functions. This exploratory study aims to fill this research gap; it identifies 12 drivers developed by 28 interviewed experts, discussing the coding techniques that are most frequently used in grounded theory. The identified drivers contribute to the current knowledge, while also potentially informing data space projects and organizations' decisions regarding data space adoption.

Keywords: Data spaces, Federation, Data Ecosystem, Drivers of Adoption.

1. Introduction

Data has become a strategic resource for competitiveness in the digital economy (Bagad et al., 2021; Gelhaar et al., 2021; Gelhaar & Otto, 2020). Intensive data exchange seems to be necessary for and beneficial to almost all organizations, e.g., data exchange among partners in supply chains (Gelhaar et al., 2021; Heinz et al., 2022). However, issues related to data sovereignty, lack of trust, and added value (Hoßbach-Zimmermann et al., 2023) make interorganizational data sharing a rarity (Bartelheimer et al., 2022; Gelhaar & Otto, 2020). The data space concept was developed to overcome these issues, in particular to clarify technical (e.g., how to integrate heterogenous data) and organizational (e.g., establish trust) issues (Fassnacht et al., 2023; Parvinen, 2020). Technical aspects of data spaces or the formation of a federated platform of data spaces (Beverungen et al., 2022; Otto & Jarke, 2019) as well as adoption of larger concepts such as the supply chain (Nath & Standing, 2010) have been researched, yet the motivations behind data space adoption remain unclear (Hutterer & Krumay, 2022). In addition, most research focus on acceptance (e.g., Davis, 1989) whereas this study investigates the drivers that influence the process of adoption. To close this gap, this study aims at identifying drivers of data space adoption. Due to the scarce knowledge in this area, this study employs an exploratory approach based on expert interviews, to answer the research question: What are the drivers of data space adoption? The paper is structured as follows. First, we provide insights into the current discussion regarding data spaces and their implementation. Next, we describe the methodological approach applied and present the results, focusing on drivers for the adoption of data spaces. In the subsequent discussion section, we elaborate on the results and our approach to answering the research question. Finally, we provide a conclusion, acknowledge this study's limitations, and suggest ideas on further research.

2. Background Information

Currently, data spaces are often described as alliance-driven multi-sided platforms for federated data sharing involving multiple organizations (Otto & Jarke, 2019). The data space concept was introduced in 2005, proposing a new paradigm for data management which focused on the integration of heterogenous data (Franklin et al., 2005). In addition, the technical, legal, and economic environment for multilateral data usage across organizational boundaries is also covered by the data space concept (Kagermann et al., 2021). The terms 'data space' and 'data ecosystems' are often used interchangeably (Capiello et al., 2020; Gelhaar et al., 2021), yet their meanings differ (Hutterer et al., 2023). The concept of data space has been seen as the basis for platform-based data ecosystems (Hutterer & Krumay, 2022) aimed at linking isolated systems so as to share data (Tardieu, 2022).

What makes data spaces attractive is the possibility of exchanging heterogenous data within or outside of the organization. Data spaces overcome the restrictions seen in other approaches (e.g., databases controlled via database management systems), which require defined

URI: https://hdl.handle.net/10125/106926
978-0-9981331-7-1

structures and data formats (Hedeler et al., 2011), while allowing data sharing of heterogenous data stemming from distributed systems (Wang et al., 2016). The focus lies in establishing relationships between heterogeneous data sources, using various technologies for a variety of applications (Guo et al., 2021). To enable this, a data platform serves as an intermediary (Curry et al., 2022) which supports collaboration, feedback, and profiling techniques (Sarma et al., 2009), in addition to data sharing. Some systems are even able to work almost autonomously, e.g., offer services (Dong et al., 2009).

Although organizations recognize the potential of data sharing (e.g., for replenishment or in the supply chain), its adoption is hindered by technical (e.g., integration, querying, or security) and non-technical issues such as unclear benefits, lack of trust, and fear of losing data sovereignty (Fassnacht et al., 2023; Heinz et al., 2022; Hutterer & Krumay, 2022). Some of these issues have also been identified in studies focusing on platform adoption, addressed by information system research (Bartelheimer et al., 2022; Hong et al., 2021). However, expected loss of sovereignty and reduced control over data seem to be very specific in the context of data space adoption (Hummel et al., 2021; Jarke, 2020; Jarke et al., 2019; Kagermann et al., 2021). In particular, organizations fear losing their "self-determination … with regard to the use of their data" (Jarke et al., 2019). Yet measures to preserve sovereignty are at hand (Pettenpohl et al., 2022; Siska et al., 2023). To address these issues, architecture design options (Schleimer et al., 2023; Siska et al., 2023) with diverse design characteristics (Gieß et al., 2023) have been proposed. Centralized architectures (Catena-X, 2022; Drees et al., 2021), characterized by a single node (e.g., Federator) providing core services (e.g., Identity Provider) of the data space, were designed early on (Gieß et al., 2023). However, to overcome issues such as dependence on a central entity, a decentralized architecture, i.e., shared responsibilities based on consensus, was suggested (Pontus-X, 2023). Federated approaches beyond data space architecture go even further, by allowing resources (e.g., data, services, infrastructure) to be shared among and within interoperable data spaces (PrepDSpace4Mobility, 2023). Since each participating data space in a federated approach maintains control over the resources shared, clear governance structures may overcome issues such as a lack of trust or unclear responsibilities (Curry et al., 2022; Torre-Bastida et al., 2022). Thus, decentralized structures (Hutterer & Krumay, 2022; Otto, 2022; Winter et al., 2022) and federated infrastructures based on a data domain and guidelines (Gaia-X Hub Germany, 2022; Siska et al., 2023) have been identified as solutions that preserve sovereignty of the shared data (Hellmeier & von Scherenberg, 2023). Examples of centralized (e.g., IDSA, 2022) and decentralized (e.g., Pontus-X, 2023) architectures exist, as do examples of federated approaches (e.g., Data Space 4.0 Alliance, 2023; PrepDSpace4Mobility, 2023). However, there are only a few fully-fledged examples of federated data space concepts. Among them is Gaia-X in Europe, which provides a fundamental technical infrastructure, a governance framework, and a digital clearing house (Gaia-X AISBL, 2023; Siska et al., 2023).

3. Methodology

To answer our research question, we applied an exploratory approach based on principles and coding techniques used in grounded theory (Glaser, 1992). Primarily focusing on the development of a theory based on data, rather than imposing a theory on data from pre-existing knowledge (Glaser, 1992), the approach provides useful guidance for coding various kinds of data (Corbin & Strauss, 2015). Due to the lack of previous studies (Hutterer & Krumay, 2022), we conducted interviews to identify the drivers influencing data space adoption. Experts were selected based on their involvement in interorganizational data sharing or projects related to data spaces. The experts worked in organizations-mainly part of the Gaia-X initiative (Tardieu, 2022)-of different sizes, eight in SMEs, six in mid-sized organizations and 14 in large organizations. The interview guidelines were developed from the research question (Braun & Clarke, 2006), in a way that ensured a structured yet open interview process (Myers, 1997). Interviews were conducted until theoretical saturation was reached, resulting in a sample of 28 experts, all from Central Europe. Table 1 provides on overview of the experts and their roles in their organizations.

Table 1. Experts and Roles

Role	Experts
(Executive) Partner	E16, E22
CDO	E21
CEO	E01, E07, E08, E18
Co-founder & business lead	E23
Data space lead architect	E11
Department head	E15
Deputy CTO	E26
Head of data science	E14
Head of digital business	E27
Head of digitalization	E05, E17
Partner development manager	E03
Project manager	E09, E24, E25
Researcher	E10, E12, E13, E19
Senior BDM	E20
Solution architect	E02, E06, E28
Team lead	E04

The interviews (average duration: 47 minutes) were conducted via Zoom (between October and December 2022), recorded, transcribed, and coded based on coding techniques used in grounded theory (Corbin & Strauss, 2015). Two researchers were involved in the coding process. Open coding created codes from the data, e.g., ‘complex construct is hard to manage’ (E01). In the axial coding process, the already existing codes were grouped based on their similarities, so as to build categories, e.g., the codes ‘complex construct is hard to manage’ (E01) and ‘complexity needs to be manageable’ (E20) resulted in the category “Controllable complexity.” During the coding process, the identified drivers were validated by reference to the existing literature, to avoid inconsistencies and misunderstandings.

4. Results

Based on the interview data, we identified 12 drivers related to data space adoption (controllable complexity, cost clarity, data sovereignty, ecosystem governance, ecosystem readiness, interoperability, mature technology, regulatory certainty, security, technology competence, transparency, and trust). Some drivers evolved directly from the data (e.g., security), whereas others were developed by converting barriers into drivers (e.g., barrier: high complexity—driver: controllable complexity). To enhance the clarity of the drivers, we condensed them to the fullest extent possible. Our approach was to create drivers that were internally as homogeneous as possible, in terms of content, while being distinct from one another. However, due to overlapping content, some of the drivers cannot be unambiguously distinguished from one another. A substantial part of the identified drivers demonstrated congruence with the prevailing literature on adoption behavior. The identified drivers are presented in alphabetical order in a concept matrix (Table 2).

Table 2. Drivers for Data Space Adoption

Expert \ Driver	Controllable complexity	Cost clarity	Data sovereignty	Ecosystem governance	Ecosystem readiness	Interoperability	Mature technology	Regulatory certainty	Security	Technology competence	Transparency	Trust
E01	X	X	X	X	X	X	X	X	X		X	X
E02		X	X	X		X		X				X
E03	X		X	X	X	X	X	X	X	X		X
E04	X	X	X		X	X	X	X		X	X	X
E05	X	X	X	X	X					X	X	
E06	X		X		X	X	X	X	X	X		X
E07	X		X	X		X	X	X	X		X	X
E08	X	X	X	X	X	X	X	X	X	X		X
E09	X	X	X	X	X	X	X	X	X	X		X
E10	X		X		X		X			X	X	
E11	X		X	X	X	X	X	X	X	X	X	X
E12	X		X	X	X	X	X	X	X	X		X
E13	X				X	X	X	X	X	X		X
E14	X		X	X	X		X	X	X	X		
E15	X	X	X	X	X	X	X	X	X	X	X	X
E16	X		X	X	X	X	X	X	X	X		X
E17	X	X			X	X	X		X	X		X
E18	X				X		X	X	X	X		
E19	X				X	X	X	X		X	X	X
E20	X	X	X	X	X	X	X	X	X	X		X
E21	X	X	X	X	X	X	X	X	X	X	X	X
E22	X	X		X	X		X	X	X	X		
E23	X	X	X	X	X		X	X	X	X		
E24	X				X		X	X		X		
E25	X	X		X	X	X	X	X	X	X		X
E26	X			X	X	X	X	X	X	X	X	X
E27	X	X		X	X	X	X		X	X	X	X
E28	X	X		X	X	X	X	X	X	X	X	X
Sum	27	15	18	20	26	21	26	24	22	25	12	21

4.1 Controllable complexity

In the interviews, high complexity was very commonly mentioned as a barrier to data space adoption. All but one expert (E02) discussed the complexity of the data space concept during their interviews. Complexity has also been identified in the literature concerning the degree of understanding of technology (Hong et al., 2021). The identified driver refers to the complexity of the concepts in general (E01, E04) and the high complexity of certain concepts (E03). Moreover, complexity was cited as a precondition for participating in data spaces (E04), especially regarding the roles and relationships within central, intermediary, and decentralized approaches (e.g., blockchain). Another aspect of complexity was perceived as obtaining in the establishment of central function services with a central custodian (E06, E11, E12, E20). Complexity was also related to the resources necessary for participation (e.g., “Data space should not be an exclusive tool for companies that have huge software engineering departments”—E04). Parallels of data spaces to service-oriented architectures where noted, but these “failed because connecting these subsystems simply didn’t work. It was too complex, too versatile” (E20). Therefore, to support the adoption of data spaces, controllable complexity would be required.

4.2. Cost clarity

The costs of technology adoption have been consistently linked to the level of effort required for implementation and utilization (De Prieelle et al., 2020). This issue was also represented in the interview data. In the context of data spaces, the adoption and integration of such technologies require various tangible and intangible resources. Our participating experts mentioned costs arising from participation in a data space depending on the organization or project (E15, E20, E22, E27, E28). Unclear cost structures were identified as the main entry barriers (E04, E08, E09, E25). More specifically, membership fees, investment in time, know-how, and hardware, as well as maintenance and updates were mentioned. The added value of participating in data spaces must be economic, in terms of "affordable packages for small organizations" (E27). Other experts (E08, E09) highlighted costs related to technical training. Interestingly, some experts (E09, E25, E27) pointed out that small and medium enterprises (SMEs) must be involved in the implementation of data spaces, which makes cost clarity necessary, since SMEs can hardly deal with constantly changing costs. Regarding costs evolving from data transactions, some experts identified the value of data in general as an issue (e.g., E09). Other experts discussed whether the amount and frequency of data transfer would increase costs ("What if data is shared every millisecond, not Big Data, but with that frequency"—E02), though some doubted this would affect costs (E02, E17, E23, E27). Another participant highlighted the benefit of affordability ("Data access and data transfer is realized in a way that ... is affordable"—E17). However, the experts acknowledge that data spaces contribute to cost reduction, due to increased efficiency (E23) and the integration of heterogeneous data (E28).

4.3. Data sovereignty

Data sovereignty is the "self-determination of individuals and organizations with regard to the use of their data" (Jarke et al., 2019). In the literature, data sovereignty is closely related to digital confidentiality (i.e., the confidentiality of data and the permissions granted by the controlling organization for accessing the network and its data; Massimino et al., 2018). The focus of data sovereignty as a driver is on the significance of the appropriate level of confidentiality (Massimino et al., 2018). Experts acknowledge the preservation of data sovereignty in decentralized data spaces, giving them greater control and ownership over their data (E01, E04, E06, E12, E21, E23), as "sovereignty is only possible in a decentralized ecosystem" (E06). Enforcing data sovereignty requires the implementation of policies and technical measures (E08, E21). Yet the experts expressed their understanding that this is challenging, because "we don't really have the technical means to enforce … data sovereignty … once the data has been transferred out of the organization" (E21). Sovereignty must also be supported by the infrastructure and platform in a bottom-up approach (E06). The experts recognize blockchain technology (E23), usage control (E12), and the federation of the data space (E23) as important for data sovereignty. An interesting aspect seems to be the establishment of a central custodian, to assure the data sovereignty of participants; experts cautioned that this custodian could have too much control and power (E06, E11, E12, E20). Furthermore, security measures, privacy-preserving approaches, and privacy-enhancing technologies are effective in safeguarding data sovereignty (E23). Such approaches include confidential computing and homomorphic encryption, which could be implemented to mitigate security issues (E23).

4.4. Ecosystem governance

Other aspects that were raised during the interviews included governance rules for implementing data spaces. These characterize the intricate dynamics and decision-making procedures within the realm of technology regulation (De Prieelle et al., 2020). The establishment of standardized regulations and guidelines pertaining to the activities of all stakeholders is paramount. This ensures the creation of cohesive concept and shared norms that govern the operation and functioning of data spaces. The experts reflected on their experience regarding data governance in their organizations (e.g., E08 concerning responsibility for data ownership). Eight experts (E14, E15, E20, E21, E22, E25, E26, E28) suggested implementing a common set of rules for data spaces based on agreements among participants. Topics such as standardized data formats and legal frameworks were noted as bases for common rules, whereas ontologies were mentioned as required for the establishment of a trustworthy environment (E05, E12). In particular, standardized formats for the exchange of data between distinct industries (e.g., mobility data in tourism vs. the energy sector) are important for monetizing the value of the data shared (E02, E03, E05). Other aspects of governance include standardization (E03, E07) and even certification of connectors (E07). Furthermore, the role of protocols and their long-term stability (E20) as well as a defined level of uniformity (E03) were mentioned.

Another governance aspect relates to the ownership of and responsibility for the data space itself (E14, E15) and the motivations and goals behind its adoption.

Opinions regarding the role of the data space operator were mixed. While some experts focused more on technical aspects (e.g., "actor who is also necessary to build up a data space from a technical point of view"—E14), others (e.g., E23) stressed the establishment of autonomous ecosystems, i.e., neutral spaces that are not controlled by any party are necessary to encourage data exchange. The data space is also seen as a marketplace where organizations can publish their data, though with specific requirements (E05, E09, E23, E26, E28). As a baseline, ecosystem governance must assure access to a data space, so as to make access and equal rights meaningful (E22).

4.5. Ecosystem readiness

Our interview data shows that ecosystem readiness is an important driver from the participants' perspectives. Ecosystem readiness is conceptualized as the willingness exhibited by various actors within an ecosystem to embrace and acknowledge the benefits and practicality associated with the adoption of new technology (Toufaily et al., 2021). In the context of data spaces, a crucial aspect of successful implementation and the attraction of new participants involves the imperative of raising awareness and adequately informing ecosystem actors about the advantages inherent in these systems. Accordingly, participants stressed that a core ecosystem to promote collaboration and build trust among partners in data exchange is required (E12, E20, E21, E26, E28). Readiness consists in reaching a critical mass of data sets and use cases before a data space can become truly operational (E03, E04). The experts agreed that data spaces are joined for two reasons: understanding the benefits of data sharing and market pressure. One expert (E03) mentioned that "simply joining and looking through the data will not result in a use case" (i.e., will not make it beneficial). Participants acknowledged that decentralized data spaces and compute-to-data approaches can save time and resources—for instance, by automating repetitive data transfer (E23, E28). As an example, a data space project was mentioned involving the entire automotive industry sharing data along the entire supply chain (E20).

Moreover, "understanding the value of data makes a company powerful" (E20). Given that readiness also relates to actors' willingness to share data, laggers—who consume but do not share data—are a threat to data spaces (E20, E21). A few experts (E01, E14, E17, E19) mentioned that data spaces are constructed in the context of a specific topic, rather than clustering all the data available in an organization. What became evident are the differences between sectors (E05, E08, E09, E10, E14, E16, E17, E19, E23, E25, E26, E27) based on company size (E26). In particular, SMEs will need to exert comparably greater effort without guaranteed outputs (E26), whereas larger organizations from more traditional fields are more likely to face structural (E25) and cultural (E22) challenges (e.g., to enhance competitiveness). Differences regarding readiness are also seen when comparing organizations in Europe with the United States and Asia (E23), in virtue of the fact that Europe currently has a smaller market share and stronger dependencies on other regions, leading to a loss of added value. Since the availability of ready-made data spaces is currently limited (E16, E19, E21), adoption is more difficult in smaller markets with limited resources. However, "networking between all relevant players in a domain is equally important" (E01). Data spaces may be particularly beneficial, because they structure "participants horizontally, at an equal level" (E19). Other aspects mentioned in this context are usability and the degree of technology adoption necessary (E14, E27), which also influence ecosystem readiness.

4.6. Interoperability

The data show mixed perspectives concerning interoperability and its importance in data spaces. Interoperability is the capacity of a component or system to seamlessly and simultaneously operate with multiple IT service providers, irrespective of variations across those providers (Gebregiorgis & Altmann, 2015). Participants stated that it is particularly important to integrate diverse capabilities of data spaces into their respective organizations but also between data space ecosystems. Establishing agreements on data and policy interoperability and implementing international standards are crucial to ensuring the maximum efficiency of data spaces (E26). Drawing from their experience, experts underscored software interoperability (E05) and technical interoperability between different ecosystems (E23) as crucial to interoperability, although semantic interoperability exists (E23). Currently, interoperability is represented through the trusted framework of self-description (E20). The focus of interoperability rests on translator and receiver services (E20). However, technological standards must reflect the distinctiveness of various data space to achieve interoperability (E26). Among the experts, some expressed doubts regarding interoperability between data space frameworks (E08, E28), explaining that "data spaces are standalone pilot projects … interoperability is given through the connectors" (E08).

4.7. Mature technology

According to the participants, the data space concept is still in its early stages and somewhat immature. In the literature, researchers have stated that technology attains maturity only after being accessible and available for an extended period (Toufaily et al., 2021). Data spaces, however, are currently regarded as emerging technologies in the nascent stages of their lifecycle. Experts were surprised that the technology has not been fully developed (E11, E22): "We assumed the technology to be finished but is still being developed" (E25). Thus, although the underlying technology is adequately developed (E11, E14, E20), other parts of the space have yet to be used to their full capacity. One expert describes data spaces as "still very much bits and pieces" (E11). Parts of the technology, such as connectors, were characterized as non-performant (E05, E15) or not yet finished (E14). Based on personal experience, one expert (E14) noted that "certain topics just haven't finished developing, simply because of the maturity level or the concept." However, some participants acknowledged that technology fulfills the requirements for existing use cases (E19, E22, E26) and progress is being made in the development of data space technologies (E23). Furthermore, experts recognized the necessity to adapt to technology early on, even when it has not been fully developed, if it has the potential to become widely used (E11, E22). Experts suggested integrating further functionalities beyond sharing data, as diverse datasets may occur (E26).

4.8. Regulatory certainty

According to the interview data, experts experience the current situation as rather uncertain, as regards regulations. Regulatory uncertainty is characterized by the absence of a comprehensive legal and regulatory framework (Toufaily et al., 2021). Policies and regulations combine to play a pivotal role in shaping the utilization of data spaces, by supporting the establishment and proliferation of data space initiatives. However, the experts discussed both the establishment of regulations (E02, E09, E10) and the fear of creating an overregulated situation (E02, E04, E06, E08, E09, E13). In Europe, legislation, such as the Supply Chain Act, the Digital Product Passport, and the Digital Service Act (E01, E08, E20, E23, E24), are considered influential in the establishment of data spaces. E08 conveyed the expectation that the Digital Product Passport would be influential in the further development of data spaces, but others noted the EU Data Act and the EU Governance Act as positive developments but also noted challenges and controversies around the concept of data infrastructure in data spaces (E01, E20, E23, E24). One expert (E01) even expressed that "the Data Governance Act and Digital Service Act create framework conditions for the development of data spaces." Some experts advocated for more pressure, i.e., regulations that would force organizations in particular industries to adopt data spaces, so as to improve the situation (E08, E23). Others suggested that industry associations, for example, should incentivize adoption, as seen in the mobility sector (E20). Beyond this, political support for data space initiatives (e.g., E02 mentioned Gaia-X and Catena-X) to promote standardization and secure data handling is necessary (E02, E09). However, participants also noted that regulatory restrictions make a full industrial rollout challenging (E02, E04, E06, E08, E09, E13). Overall, the experts expressed fears that a legal framework integrating data protection rules (e.g., the GDPR) and antitrust law (E09, E21) might be too complex; they emphasized the need for simple rules that can be easily understood (E09, E20, E21), to assure regulatory certainty.

4.9. Security

Another driver directly observable from the data is security, especially concerning data protection. In the literature, security pertains to the holistic protection of data within a platform (De Prieelle et al., 2020). The majority of our participating experts ($n = 22$) agreed that data spaces (e.g., Gaia-X) should provide a secure environment for data, one that ensures the protection of sensitive information against unauthorized access. By implementing robust security measures, data spaces can establish a reliable guidance for maintaining data integrity and confidentiality, thereby instilling trust among stakeholders. As E01 reported, "It is clear from existing data spaces that security-by-design must be considered from the beginning." Even measures to prevent industrial espionage were mentioned in this context (E14, E27). The experts addressed topics such as control over data by a central platform (E17), problems evolving from low security of connectors (E06), trusted computing models (E16), and closed chain of trust to ensure security (E26)—also acknowledging the drawbacks of blockchain technology for establishing security (E07). Connectors between data spaces were mentioned as constituting a security issue, because the data needs to be temporarily stored in these connectors, which often follow policies that differ from those of the connected data spaces (E26). In addition, the translator and resolver services that are necessary for interoperability are vulnerable to attacks (E20). Accordingly, a complete digital rights management chain for data sharing would be best (E02). As E08 mentioned, participants "have to be able to assume that this is simply a very secure environment." The level of security inheres mainly in the technology, including edge computing and hybrid solutions (E02), for tackling security issues. Beyond technical security issues, security concerns vary depending on the type of data and nature of the privacy problem, and compliance with standards is viewed as advantageous (E01, E06, E17). Ways to establish security include standardization and certification of data space technologies (E07, E20, E26), establishing data trustees (E21), and a focus on decentralization (E04).

4.10. Technology competence

Another driver of data space adoption discernable from the data is technology competence, initially coded as data literacy. However, in the literature, 'technology competence' refers to the internal technological capabilities of an organization, encompassing the organization's proficiency in effectively utilizing and managing technology, specifically in the context of data spaces (Zhu et al., 2006). We decided to align our terminology with that of the literature. Organizations face challenges regarding finding the right place to implement innovative ideas (E03), competing priorities (E17), lack of awareness (E10, E21), and insufficient digitalization (E10, E14), when aiming at participating in a data space. Deficiencies in internal technical competence, such as inadequate data management within organizations, was addressed by many experts (E04, E05, E08, E09, E11, E12, E14, E22, E24, E25, E26, E27). Organizations often lack processes for weighing the value of data against risks and, therefore, fear possible industrial espionage (E14, E17). Moreover, organizations cannot identify optimal exploitation and monetization of data (E06, E09, E11, E15, E16, E18, E19, E22, E23, E26, E28). Some experts stressed that organizations must be able to identify their data assets, ensure they adhere to standards, and have good metadata (E05, E14, E25). However, the availability of skilled workers (E09) and the required training at all levels of an organization (E28) can be challenging. Solutions such as low-code, no-code applications (E04), and usability (E14, E27) fitting employees' skills would be required.

Another reported challenge involves the awareness of benefits, especially in the current, rather uncertain economic situation (E09), although awareness for some benefits (i.e., promoting the collective use of data; E14, E16, E21) seems to exist. A further sign of technology competence is how well the concept of data spaces is understood or how much it "needs to be explained" (E14, E27). However, it was mentioned that higher-level management may be more open to initiatives (E14, E15), especially when research institutions are involved (E03). Another characteristic of this driver is the expected internal effort required by organizations, such as adapting IT services (E14, E27). Therefore, a specific level of technical competence is required to drive the adoption of data spaces.

4.11. Transparency

Transparency emerged from the data focusing on discoverability and visibility. In the literature, transparency is defined as the degree to which technology is perceived as an advancement in terms of transparency compared to its predecessors (Toufaily et al., 2021). Experts view data spaces as enhancements aimed at achieving greater clarity and visibility in various transactions, thereby fostering an environment of increased transparency (E01, E04, E10, E20, E24, E27). The advantage of data spaces lies in "the whole transparency behind such a data space" (E10). Aside from private players, the public sector is seen as contributing to transparency by making data available (E28). Data availability and traceability based on a metadata catalog are useful for data discoverability (E01, E24, E27), possibly implemented via blockchain technology (E04). A "catalog function forms the basis for matchmaking between data providers and users" (E01). Aside from the catalog, identifying, access management, auditing, and logging are also attributes of transparency (E20). Data discoverability requires the development of a common data model and ontologies (E12). Developing a common data model or semantics, including data formats that may or may not be domain-specific, results in a common language (E19). Yet sectoral specifics must be addressed "by developing … additive language" (E19). However, developing a common ontology for different data spaces with different requirements for semantics is especially challenging. One possible solution is the establishment of general and specific (e.g., industry-specific) metadata descriptions per data space (E20).

4.12. Trust

The trust in data consists in the belief that participants will abstain from anticipated or unanticipated actions that may result in harm (De Prieelle et al., 2020). Trust plays a pivotal role in establishing a dependable and secure environment. Participants can be relied upon to act responsibly, uphold data integrity, and adhere to mutually agreed-upon terms, thereby fostering an atmosphere of trustworthiness and reliability. "Trust is key to the success of data spaces and the sharing of data between different actors in a domain" (E01). During our interviews, measures for establishing trust were mentioned, such as regulated trust through public key infrastructure, self-sovereign identities (SSIs), or tokens (E21, E23, E26). More specifically, SSIs with verifiable credentials are often used in data spaces, as they "enable digital verification of credentials in real time" (E23). This requires a common trust anchor and a suitable governance structure, to ensure trust between different SSI environments (E07, E23). Again, blockchain was mentioned (E17) as a potential mechanism for creating trust in the ecosystem. Other approaches, such as trusted computing infrastructures (E23) or data trustees (E16), may constitute viable alternatives to centralized trust models. One expert explained that trust in a data space depends on its technological approach, including a secure environment and meeting technical challenges (E08). Other experts focused on a central authority that can establish trust quickly (E23), such as a certification authority (E11), though its scalability was questioned (E11). Yet this central entity also poses a single point of failure (E23) and corrupts the enabling of a decentralized approach (E06). As one expert stated, "Especially in multilateral environments or ecosystems, there is simply no trust at all" (E06).

5. Discussion

The aim of this study consists in identifying the drivers for data space adoption. Based on 28 qualitative interviews and coding techniques used in grounded theory (Glaser, 1992), we were able to identify and describe 12 drivers influencing data space adoption. Our findings contribute to the current knowledge on data space adoption as presented in the concept matrix (Table 2). Thus, our findings offer a comprehensive perspective on the drivers of data space adoption. Moreover, our research may serve as a guideline for data space initiatives and their development to ensure they are attractive to potential participants but also for organizations' decisions towards data space adoption.

Among the drivers, technology competence, or data literacy, stands out, as it focuses on the organizational capabilities needed for the adoption of a data space. This means that it is the only driver that can be influenced by the organization to achieve the required level or fit for participating in a data space, mainly focusing on organizational readiness and technological skills. This is interesting, as most technology adoption models focus on system characteristics, such as usability and utility (Davis, 1989; Venkatesh et al., 2003). However, the technical competence of employees and the organization has been addressed in the context of digitalization to achieve maturity (Blanka et al., 2022).

Other drivers identified focus more on the data-sharing capacities of data spaces, particularly the trusted sharing of data among participants (Gronlier et al., 2023). This includes data sovereignty, ecosystem governance, security, transparency, and trust. Interestingly, these drivers are dominated by technological aspects (e.g., technological measures in the context of security). This is in line with the current academic discussion, in which technological aspects prevail (Hutterer & Krumay, 2022). Moreover, security, data sovereignty, and trust share some characteristics and even reinforce each other. For example, trusted environments established via security measures enable data sovereignty. The current literature, however, compares data sovereignty to data privacy, focusing on data richness and data usage agreements as the bases for data sovereignty (Jarke et al., 2019). This directly targets ecosystem governance (by establishing rules) and transparency (by focusing on traceability and catalogs to achieve data richness). As a first attempt, we propose that these five drivers focus on the main function of data spaces: data sharing.

The remaining six drivers (complexity, cost clarity, ecosystem readiness, interoperability, mature technology, and regulatory certainty) are also conditions for data space adoption. They share characteristics, such as reliability, scalability, and robustness in a wider sense. Overall, these drivers address how data spaces in general should be developed. These drivers address topics such as data space architectures (e.g., decentralized autonomous organizations), technological aspects, such as connectors for interoperability (Giussani & Steinbuss, 2023), and the influence of political initiatives and regulations (Kagermann et al., 2021). Organizations expect data spaces to address all six drivers before they participate.

This study offers valuable insights, especially for various data space projects, to lay the foundations for deployment scenarios (Steinbuss et al., 2023). Based on the 12 drivers, developers may be able to design data spaces that fit specific conditions, thus further supporting data space adoption. In particular, the drivers focused on data sharing capacity, as well as the six identified conditions, need to be refined by data space providers and projects.

6. Conclusion

This study identifies 12 drivers influencing data space adoption. Since this is, to the best of our knowledge, the first attempt to identify these drivers, our research may provide a foundation for further investigation. Although we selected field experts across sectors, the current knowledge regarding the data space concept remains limited. Future research endeavors should strive to expand upon our findings, validating and further defining the drivers, particularly regarding their relationships and dependencies, while exploring the comparative effectiveness of different drivers in data space adoption. However, the evolving and limited number of data spaces may pose challenges for this evaluation.

7. References

Bagad, P., Mitra, S., Dhamnani, S., Sinha, A. R., Gautam, R., & Khanna, H. (2021). Data-Sharing Economy: Value-Addition from Data meets Privacy. *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, 1105–1108.

Bartelheimer, C., zur Heiden, P., Lüttenberg, H., & Beverungen, D. (2022). Systematizing the lexicon of platforms in information systems: A data-driven study. *Electronic Markets*, *32*(1), 375–396. https://doi.org/10.1007/s12525-022-00530-6

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

Blanka, C., Krumay, B., & Rueckel, D. (2022). The interplay of digital transformation and employee competency: A design science approach. *Technological Forecasting and Social Change*, *178*, 121575. https://doi.org/10.1016/j.techfore.2022.121575

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Capiello, C., Gal, A., Jarke, M., & Rehof, J. (2020). *Data Ecosystems: Sovereign Data Exchange among Organizations (Dagstuhl Seminar 19391)*. 69 pages. https://doi.org/10.4230/DAGREP.9.9.66

Catena-X. (2022). *Catena-X Operating Model Whitepaper*. https://catena-x.net/fileadmin/user_upload/Publikationen_und_WhitePaper_des_Vereins/CX_Operating_Model_Whitepaper_02_12_22.pdf

Corbin, J. M., & Strauss, A. L. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (Fourth edition). SAGE.

Curry, E., Scerri, S., & Tuikka, T. (2022). Data Spaces: Design, Deployment, and Future Directions. In E. Curry, S. Scerri, & T. Tuikka (Eds.), *Data Spaces* (pp. 1–17). Springer International Publishing. https://doi.org/10.1007/978-3-030-98636-0_1

Data Space 4.0 Alliance. (2023). *Data Space 4.0*. https://manufacturingdataspace-csa.eu/

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 319–340.

De Prieelle, F., De Reuver, M., & Rezaei, J. (2020). The Role of Ecosystem Data Governance in Adoption of Data Platforms by Internet-of-Things Data Providers: Case of Dutch Horticulture Industry. *IEEE Transactions on Engineering Management*, 1–11. https://doi.org/10.1109/TEM.2020.2966024

Dong, X. L., Halevy, A., & Yu, C. (2009). Data integration with uncertainty. *The VLDB Journal*, *18*(2), Article 2.

Drees, H., Kubitza, D. O., Lipp, J., Pretzsch, S., & Langdon, C. S. (2021). *Mobility Data Space – First Implementation and Business Opportunities*. 27th ITS World Congress, Hamburg.

Fassnacht, M., Benz, C., Heinz, D., Leimstoll, J., & Satzger, G. (2023). *Barriers to Data Sharing among Private Sector Organizations*. Hawaii International Conference on Systems Sciences (HICCS).

Franklin, M., Halevy, A., & Maier, D. (2005). From databases to dataspaces: A new abstraction for information management. *ACM SIGMOD Record*, *34*(4), Article 4. https://doi.org/10.1145/1107499.1107502

Gaia-X AISBL. (2023). *Gaia-X Framework*. https://docs.gaia-x.eu/framework/

Gaia-X Hub Germany. (2022). *Was ist ein Datenraum?* Gaia-X Hub Deutschland c/o acatech – Deutsche Akademie der Technikwissenschaften. https://gaia-x-hub.de/wp-content/uploads/2022/10/20220914_White_Paper_22.1_Definition_Datenraum_final.pdf

Gebregiorgis, S. A., & Altmann, J. (2015). IT Service Platforms: Their Value Creation Model and the Impact of their Level of Openness on their Adoption. *Procedia Computer Science*, *68*, 173–187. https://doi.org/10.1016/j.procs.2015.09.233

Gelhaar, J., Groß, T., & Otto, B. (2021). *A Taxonomy for Data Ecosystems*. Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2021.739

Gelhaar, J., & Otto, B. (2020). Challenges in the Emergence of Data Ecosystems. *Pacific Asia Conference on Information Systems (PACIS)*, 175.

Gieß, A., Möller, F., Schoormann, T., & Otto, B. (2023). *Design options for data spaces*. Thirty-first European Conference on Information Systems (ECIS 2023).

Giussani, G., & Steinbuss, S. (2023). *Data Connector Report* (5.0). International Data Spaces Association. https://doi.org/10.5281/ZENODO.7711222

Glaser, B. G. (1992). *Basics of grounded theory analysis*. Sociology Press.

Gronlier, P., Hierro, J., & Steinbuss, S. (2023). *Technical Convergence*. Data Space Business Alliance. https://data-spaces-business-alliance.eu/wp-content/uploads/dlm_uploads/Data-Spaces-Business-Alliance-Technical-Convergence-V2.pdf

Guo, J., Cheng, Y., Wang, D., Tao, F., & Pickl, S. (2021). Industrial Dataspace for smart manufacturing: Connotation, key technologies, and framework. *International Journal of Production Research*, 1–16. https://doi.org/10.1080/00207543.2021.1955996

Hedeler, C., Belhajjame, K., Paton, N. W., Fernandes, A. A., Embury, S. M., Mao, L., & Guo, C. (2011). Pay-as-you-go mapping selection in dataspaces. *Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data*, 1279–1282.

Heinz, D., Benz, C., Fassnacht, M., & Satzger, G. (2022). Past, Present and Future of Data Ecosystems Research: A Systematic Literature Review. *Pacific Asia Conference on Information Systems (PACIS)*.

Hellmeier, M., & von Scherenberg, F. (2023). *A delimitation of data sovereignty from digital and technological sovereignty*. Thirty-first European Conference on Information Systems (ECIS 2023).

Hong, J., Guo, P., Deng, H., & Quan, Y. (2021). The adoption of supply chain service platforms for organizational performance: Evidences from Chinese catering organizations. *International Journal of Production Economics*, *237*, 108147. https://doi.org/10.1016/j.ijpe.2021.108147

Hoßbach-Zimmermann, N., Hantschel, P., Falk, S., Guggenberger, T., & Leyh, C. (2023). *Digitale Geschäftsmodelle und Wertschöpfung in Datenräumen für die Industrie*. Plattform Industrie 4.0. https://www.plattform-i40.de/IP/Redaktion/DE/Downloads/Publikation/Geschaeftsmodelle_Wertschoepfung_Datenraeume.pdf?__blob=publicationFile&v=3

Hummel, P., Braun, M., Tretter, M., & Dabrock, P. (2021). Data sovereignty: A review. *Big Data & Society*, *8*(1), 205395172098201. https://doi.org/10.1177/2053951720982012

Hutterer, A., & Krumay, B. (2022). Integrating Heterogeneous Data in Dataspaces—A Systematic Mapping Study. *PACIS 2022 Proceedings*. Pacific Asia Conference on Information Systems (PACIS). https://aisel.aisnet.org/pacis2022/222

Hutterer, A., Krumay, B., & Mühlburger, M. (2023). What Constitutes a Data space? Conceptual Clarity beyond Technical Aspects. *29th Americas Conference on Information Systems (AMCIS)*. Americas Conference on Information Systems (AMCIS), Panama. https://aisel.aisnet.org/amcis2023/eco_systems/eco_systems/5

IDSA. (2022). *Reference Architecture Model, Version 4.0*. https://docs.internationaldataspaces.org/ids-ram-4/

Jarke, M. (2020). Data sovereignty and the internet of production. *International Conference on Advanced Information Systems Engineering*, 549–558.

Jarke, M., Otto, B., & Ram, S. (2019). Data sovereignty and data space ecosystems. In *Business & Information Systems Engineering* (Vol. 61, Issue 5, pp. 549–550). Springer.

Kagermann, H., Streibich, K.-H., & Suder, K. (2021). *Digital Sovereignty Status Quo and Perspectives*. acatech – National Academy of Science and Engineering.

Massimino, B., Gray, J. V., & Lan, Y. (2018). On the Inattention to Digital Confidentiality in Operations and Supply Chain Research. *Production and Operations Management*, *27*(8), 1492–1515. https://doi.org/10.1111/poms.12879

Myers, M. D. (1997). Qualitative Research in Information Systems. *MIS Quarterly*, *21*(2), 241–242.

Nath, T., & Standing, C. (2010). Drivers of information technology use in the supply chain. *Journal of Systems and Information Technology*, *12*(1), 70–84. https://doi.org/10.1108/13287261011032661

Otto, B. (2022). The evolution of data spaces. In *Designing Data Spaces: The Ecosystem Approach to Competitive Advantage* (pp. 3–15). Springer International Publishing Cham.

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

Parvinen, P. (2020). Advancing data monetization and the creation of data-based business models. *Communications of the Association for Information Systems*, *47*(1), 2.

Pettenpohl, H., Spiekermann, M., & Both, J. R. (2022). International data spaces in a nutshell. *Designing Data Spaces; Springer: Cham, Switzerland*, 29–40.

Pontus-X. (2023). *Pontus-X Documentation*. https://docs.genx.minimal-gaia-x.eu/docs/intro

PrepDSpace4Mobility. (2023). *Preparatory action to establish a common European data space for mobility*. https://mobilitydataspace-csa.eu/wp-content/uploads/2023/07/psf-27-june.pdf

Sarma, A. D., Dong, X. L., & Halevy, A. Y. (2009). Data modeling in dataspace support platforms. In *Conceptual Modeling: Foundations and Applications* (pp. 122–138). Springer.

Schleimer, A. M., Fraunhofer, I., Jahnke, N., & Otto, B. (2023). Architecture Design Options for Federated Data Spaces. *Proceedings of the 56th Hawaii International Conference on System Sciences*. Hawaii International Conference on System Sciences (HICCS). https://hdl.handle.net/10125/103078

Siska, V., Karagiannis, V., & Drobics, M. (2023). *Building a Dataspace: Technical Overview*. Gaia-X Hub Austria. https://www.gaia-x.at/wp-content/uploads/2023/04/WhitepaperGaiaX.pdf

Steinbuss, S., de Roode, M., Papakosta, S., Klinker, P., Punter, M., & Jimenez, S. (2023). *Data Spaces Landscape – Overview and relations of data spaces initiatives, standards, and tools (1.0)*.

Tardieu, H. (2022). Role of Gaia-X in the European Data Space Ecosystem. In *Designing Data Spaces: The Ecosystem Approach to Competitive Advantage* (pp. 41–59). Springer International Publishing Cham.

Torre-Bastida, A. I., Gil, G., Miñón, R., & Díaz-de-Arcaya, J. (2022). Technological Perspective of Data Governance in Data Space Ecosystems. In *Data Spaces: Design, Deployment and Future Directions* (pp. 65–87). Springer International Publishing Cham.

Toufaily, E., Zalan, T., & Dhaou, S. B. (2021). A framework of blockchain technology adoption: An investigation of challenges and expected value. *Information & Management*, *58*(3), 103444. https://doi.org/10.1016/j.im.2021.103444

Venkatesh, Morris, Davis, Fred, & Davis. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425. https://doi.org/10.2307/30036540

Wang, Y., Song, S., & Chen, L. (2016). A survey on accessing dataspaces. *ACM SIGMOD Record*, *45*(2), 33–44.

Winter, J., Frey, A., & Biehler, J. (2022). Towards the Next Decade of Industrie 4.0 – Current State in Research and Adoption and Promising Development Paths from a German Perspective. *Sci*, *4*(3), 31. https://doi.org/10.3390/sci4030031

Zhu, K., Dong, S., Xu, S. X., & Kraemer, K. L. (2006). Innovation diffusion in global contexts: Determinants of post-adoption digital transformation of European companies. *European Journal of Information Systems*, *15*(6), 601–616.