Trust me, I'm an Intermediary! Exploring Data Intermediation Services

Conference Paper · September 2023

3 authors:

Julia Schweihoff
Technische Universität Dortmund
9 PUBLICATIONS   11 CITATIONS

SEE PROFILE

Ilka Jussen
Technische Universität Dortmund
10 PUBLICATIONS   15 CITATIONS

SEE PROFILE

Frederik Möller
Technische Universität Braunschweig
69 PUBLICATIONS   677 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   IEDS - Incentives and Economics of Data Sharing View project

Project   Digital Business Engineering View project

COVER SHEET

This is the author's version of the following source:

-Author's Accepted Manuscript-

Trust me, I'm an Intermediary!
Exploring Data Intermediation Services

Julia Christina Schweihoff[1], Ilka Jussen[1], and Frederik Möller[1,2,3]

[1] TU Dortmund University, Chair for Industrial Information Management, Dortmund, Germany
{Julia.Schweihoff,Ilka.Jussen,Frederik.Moeller}@tu-dortmund.de
[2] TU Braunschweig, Information Systems and Data-Driven Enterprise, Braunschweig, Germany
Frederik.Moeller@tu-braunschweig.de
[3] Fraunhofer ISST, Dortmund, Germany
Frederik.Moeller@isst.fraunhofer.de

Abstract. Data ecosystems receive considerable attention in academia and practice, as indicated by a steadily growing body of research and large-scale (industry-driven) research projects. They can leverage so-called data intermediaries, which are mediating parties that facilitate data sharing between a data provider and a data consumer. Research has uncovered many types of data intermediaries, such as data marketplaces or data trusts. However, what is missing is a 'big picture' of data intermediaries and the functions they fulfill. We tackle this issue by extracting data intermediation services decoupled from specific instances to give a comprehensive overview of how they work. To achieve this, we report on a systematic literature review, contributing data intermediation services.

Keywords: Data Sharing, Data Ecosystems, Data Intermediaries

1 Introduction

Ensuring flourishing inter-organizational data sharing requires consensus about what the process entails. For this purpose, the European Commission proposed the *European Strategy for Data*, with the Data Act (DA) as one of its core elements (European Commission, 2022a). Among other things, the DA defines the nature of fair digital environments to spur data-driven innovation and enable novel data-driven services (European Commission, 2022b). Notably, the overarching goal seems to be to build data-sharing ecosystems (i.e., data ecosystems) in which organizations share data in a legally compliant manner to create value (Capgemini Research Institute, 2021). Sharing data is usually associated with a plethora of untapped potential, which, if used adequately, can assist in creating a brighter future and tackling global challenges, such as *sustainable tourism* (World Economic Forum, 2021). Data sharing contributes to developing new services for end consumers and intensifying inter-organizational cooperation and value

creation. To this end, major initiatives, such as the *Catena-X Automotive Network* (Catena-X, 2023), aim to generate a secure and sovereign data ecosystem across their supply chains in the German automotive industry. Novel applications include compliance with the Supply Chain Act, which mandates accountability for supply chain partners (e.g., suppliers), for instance, to prevent child labor (Federal Ministry for Economic Cooperation and Development, 2023). Other goals include making processes along the value chain more transparent and traceable to reduce resource consumption and support circular economy design (Handelsblatt, 2023). However, various barriers prevent effective data sharing between organizations, which are usually associated with a lack of reciprocal trust and transparency (Opriel *et al.*, 2021; Fassnacht *et al.*, 2023). One way to enhance trust in data ecosystems is to install a so-called *data intermediary* that "*facilitates the use of data for other actors*" (Oliveira *et al.*, 2019). These data intermediaries (e.g., data marketplaces or data trusts) can be shaped differently based on the setting they are implemented and the specific needs of the data ecosystems (Oliveira *et al.*, 2019; Richter & Slowinski, 2019). The importance of intermediation services is emphasized in the Data Governance Act (DGA): *"intermediation services between data holders [...] and potential data users, including making available the technical or other means to enable such services; those services may include bilateral or multilateral exchanges of data or the creation of platforms or databases enabling the exchange or joint exploitation of data, as well as the establishment of a specific infrastructure for the interconnection of data holders and data users [...]" (European Commission, 2020).* Given the background of the DA and the DGA and, at the same time, the current push of politics, economy, and research to generate data ecosystems, we propose a classification of data intermediary services in the context of data sharing. Because of the above, we pursue the following research question: *How to classify data intermediary services (data intermediation services) in the context of data sharing?*

To answer this research question, we structure the paper as follows. First, we present the theoretical background focusing on data sharing and the importance of intermediaries. Then Section 3 explains our research design with a detailed description of our literature review. In Section 4, we present our findings from the literature review regarding the services of data intermediaries. Finally, Section 5 concludes our paper with an overview of the contributions and limitations of our research as well as an outlook on further research opportunities.

2   Data Sharing through Data Intermediaries

Sharing data is more than a technical process since it represents a socio-technical transaction between a *data provider* and a *data consumer* that needs to consider an array of decisions (Jussen *et al.*, 2023). One of these decisions is identifying a suitable data ecosystem governance structure, which includes variants, such as peer-to-peer or the implementation of a *data intermediary* (Wernick *et al.*, 2020). Oliveira *et al.* (2019) define intermediary-based data ecosystems as those that depend "*on the presence of data intermediaries in order to generate value from data*" and that it "*is a role that facilitates the use of data for other actors.*" Janssen and Singh (2022) define the data

intermediary as "*a mediator between those who wish to make their data available, and those who seek to leverage that data. The intermediary works to govern the data in specific ways and provides some degree of confidence regarding how the data will be used*". Resulting, these illustrative definitions imply that the intermediary is 'in the middle', between a *data provider* and a *data consumer*. Data intermediaries can have a variety of purposes, such as *ensuring data quality* (e.g., Oliveira & Lóscio, 2018) or *organizing data-sharing transactions* (e.g., Agahari *et al.*, 2021). However, the nature and specifics of data intermediaries remain blurry (Wernick *et al.*, 2020) and highly contextual (Oliveira *et al.*, 2019; Janssen & Singh, 2022).
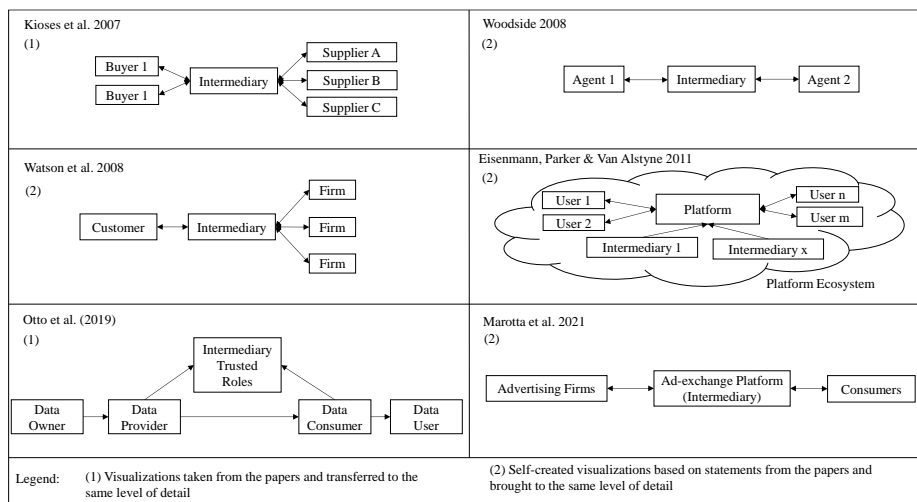
Figure 1. Illustrative conceptualizations of data intermediaries.

Figure 1 shows visualizations of data intermediaries. We derived these visualizations in part from the literature and brought them to a comparable level of detail using existing illustrations from the papers. One major common feature of the visualizations of the different authors is that the intermediary always fulfills some role between at least two other actors. Kioses *et al.* (2007) illustrate intermediaries for data exchange between buyers and suppliers in the example of invoices. Otto *et al.* (2019) also considered a real-world case in their paper, the *International Data Space*. In their visualization, the intermediary is placed between different actors. Intermediaries can take on various sub-roles in their role as "*trusted entities*", such as *broker service provider*, *clearing house, identity provider*, *app store*, and *vocabulary provider* (Otto *et al.*, 2019)[1]. Intermediaries can also appear in the form of *data marketplaces* or *platforms* (Cirullies & Schwede, 2021; Marotta *et al.*, 2022; Sterk *et al.*, 2022). These platforms can be „*created and maintained by one or more intermediaries*" (Eisenmann *et al.*, 2011).

---

[1] For more detailed explanations of the individual roles, we refer to Otto *et al.* (2019).

3   Research Design

To answer the research question outlined above, we explore the literature to extract knowledge about *data intermediation services* in the context of data sharing. Subsequently, our research strategy is a *concept-centric systematic literature review* (Webster & Watson, 2002; vom Brocke *et al.*, 2015). We take a *descriptive approach* as described by Paré *et al.* (2015), which aims to show what the understanding of *data intermediation services* is in the literature and follow the procedure suggested by Templier and Paré (2015) for conducting a literature review (indicated in bold below).

First (1), we defined a set of keywords that we identified as likely to yield relevant literature on data intermediaries. Resulting, we used combinations of "*intermediary*", omitting the prefix *data* to be as comprehensive as possible. From our experience, we found that "*trustee*", "*trusted third-party*", and "*marketplace*" as likely synonyms that could potentially enrich the literature sample. We entangled these keywords with "*data sharing*" or "*data exchange*" accommodating the prevailing blurriness between these concepts (e.g., Jussen *et al.*, 2023). The following example by Koul *et al.* (2018) illustrates the diversity of terms in this field, further substantiating our search strategy: "*The* **data sharing** *process is managed by a* **data marketplace***, a* **trusted third-party** *handling the market participants' request and managing the agreements between them*[2]" (**formulating the problem**).
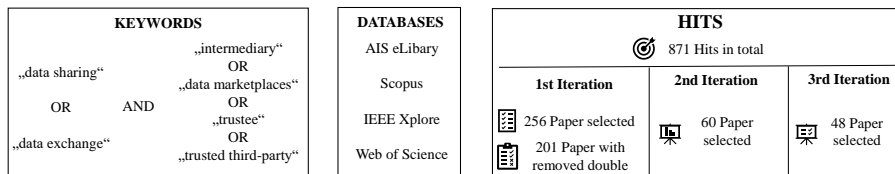

Figure 2. Keywords, databases, filtering process, and hits.

Second (2), we identified relevant databases that contain peer-reviewed literature published in conference proceedings and journals (Levy & Ellis, 2006). To cover conference proceedings, predominantly in the Information Systems (IS) research field, we searched the database of the *AIS*, which hosts high-quality conferences, such as *ECIS* or *ICIS*. We extended the IS focus by searching in *Scopus*, *IEEE Xplore*, and *Web of Science*, which indexed various journals, such as *Electronic Markets* or *BISE* (**searching the literature**).

Third (3), we defined filter mechanisms to focus on those papers that report on data intermediary services or functionalities. Subsequently, to be included, papers must deal with a type of data intermediary, even though the exact wording is not a precondition for inclusion (**screening for inclusion** and **assessing quality**). For example, Abbas *et al.* (2021) investigate business data sharing through data marketplaces, which we classify as a data intermediary, not by name, but by the concept, and subsequently included it in the literature sample. We started with a body of literature containing 871 hits across

---

[2] Highlighting of relevant terms by the authors (bold)

databases, which we reduced (e.g., removing doubles, thematic analysis, accessibility, and language) in three iterations to 48 for in-depth analysis (see Figure 2 for details)[3].

Table 1. Examples of our literature coding.

Example Literature Statement	Coding
"*Due to legal issues, negotiation data must be monitored by a trusted third party (TTP). This allows for independent third-party monitoring of the business transactions (…)"* (Woodside, 2008)	Monitoring of legal, data and business transactions by a trusted third party
„*A Trusted Third Party (TTP), a non-profit organisation, pseudonimises the data to minimise the risk of reidentification. Access and recombination of the data is provided through this service.*" (van den Broek & van Veenstra, 2015)	Pseudonymization of the data, Control data access
"*Some data marketplaces offer an enterprise data marketplace as an additional service. An enterprise data marketplace functions as a private data marketplace that enables organizations to share data (…)."* (van de Ven *et al.*, 2021)	Provide a data infrastructure (e.g., in the form of a data marketplace)

Fourth (4), two authors analyzed the literature using a concept-centric approach to extracting information about data intermediation services (**extracting data**). During the *1st iteration*, we reviewed the *abstracts* and *titles* of the papers and only included German or English papers focusing on "*Data Intermediaries*" within a B2B context. In the *2nd iteration*, we analyzed the papers more closely and excluded papers dealing with the issue of data sharing but not data intermediaries. For example, some papers reported on using the blockchain for data sharing. However, we found it to be a substitute for an intermediary (e.g., Albrecht *et al.*, 2018). The *3rd and final iteration* examined the selected papers for specific aspects mentioned in connection with data intermediation services. Finally, the number of selected papers was 48. Since our search showed theoretical satisfaction after the 3rd iteration, we decided against a further forward/backward search (e.g., Webster & Watson, 2002).

Table 1 shows illustrative literature statements and the corresponding codes. Initially, we started by extracting intuitive elements of the papers, such as visual representations (see Figure 1). To analyze the literature statements systematically, we employed Gioia diagrams (examples for this are Leemann & Kanbach, 2022, and Osterrieder *et al.*, 2020) (**analyzing and synthesizing data**). These types of diagrams are part of grounded theorizing and are 'traditionally' used to analyze, for example, transcribed interviews and generate concepts and theory from them (Gioia *et al.*, 2013). However, we found the procedure highly systematic and helpful for analyzing the literature statements we extracted, which is why we adapted it to our systematic literature review (see Figure 3) and used it to outline a clear thread between literature statements and a class of services.

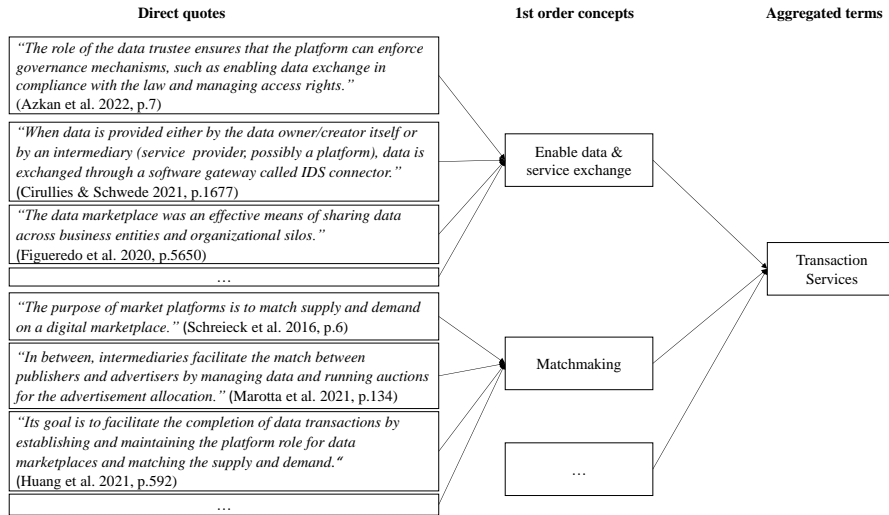[3] List of all analyzed papers is available on request.

Figure 3. Illustrative visualization of the Gioia diagram (based on Gioia *et al.*, 2013).

4   Findings: Data Intermediaries in Data Ecosystems

In the following, we provide a fivefold classification of data intermediation services. This classification is a product of iterative content aggregation, in which we ordered the services we extracted based on their thematic alignment. The resulting field of aggregated terms is the service categories we found prevalent in data intermediaries. At this point in our research, we did not look for mutually exclusive categories, but for the '*big picture*'. Resulting, it is to be expected that these categories overlap conceptually.

Table 2. Data Intermediation Service Types

Service Type	Definition	Distribution	Examples
Transactions Service	Transaction services focus on enabling the general data-sharing process.	18 Papers	E015 Sovity
Governance Service	Governance services aim to regulate the legal aspects of data sharing.	10 Papers	CDQ
Sovereignty Service	Sovereignty services ensure compliance with policies and data ownership.	17 Papers	Sovity GXFS.eu Snowflake
Technology Service	Technology services provide a data infrastructure for sharing and storing data.	26 Papers	GXFS.eu Dawex
Data Service	This service focuses exclusively on the resource of data.	13 Papers	CDQ Catena-X

Below, we will explain each service class in more detail and enrich them through practice examples. Please note each service is highlighted in bold, reflecting the 1st order of concepts from the literature analysis (see Figure 3). The services we find are non-exclusive, and there may be some overlap.

4.1 Transaction Services

Transaction services enable data ecosystems by matching data providers and consumers or supporting activities in the data-sharing process. One primary service of data intermediaries is **to enable data & service exchange.** As visualized in Figure 1, the intermediary takes the position between data providers and data consumers and *enables* the general *exchange of data and services*. Perscheid *et al.* (2020) define the role of the intermediary as "*coordinate users*" and "*enable the exchange of products and services*" to promote collaboration. In addition to facilitating the general exchange of data and services, this also includes bringing together data providers and data consumers so that data exchange can occur in the first place. Bringing the actors together is also classically referred to as **matchmaking**. For example, data marketplaces connect data providers and data consumers in their role as an intermediary. Agahari *et al.* (2021) define this service as "*a matchmaking service between data providers and data buyers, rather than a facilitator of data exchange between the two parties*." This may involve matching demand and supply, but also data providers and data buyers, or different markets (Schreieck *et al.*, 2016; Demchenko & Gommans, 2019; Agahari *et al.*, 2021; Huang *et al.*, 2021). Intermediaries gather information over time, enhancing matchmaking capabilities (Marotta *et al.*, 2022). An example is E015 (2023), an Italian API-Ecosystem and data marketplace that enables application developers to tap into APIs.

There is a broad range of reasons why data sharing fails between organizations. One is the participants' lack of know-how on how to design and implement the data-sharing process (Gelhaar *et al.*, 2021), requiring them to receive **support in facilitating data sharing**. Bastiaansen *et al.* (2019) see the responsibility of the intermediary to provide "*adequate service options for the data sharing supporting processes that match the data provider's business policies*." Companies such as *Sovity* (2023) tackle this issue by providing support services around data sharing (e.g., data space connectors).

Some see the responsibilities of data intermediaries to **resolve conflicts.** Conflict resolution is a pivotal service to enable long-term and sustainable data sharing. Woodside (2008) describes this service as "*independent third-party monitoring of the business transactions, and for resolution of conflicts between parties*". The ability of the intermediary enhances and requires its trustworthiness (Bastiaansen *et al.*, 2019; Bastiaansen *et al.*, 2020). Figure 4 visualizes the basic logic of transaction services. The various actors are in contact with the intermediary, which takes its conceptual place 'in the middle,' enabling and supporting data exchange.
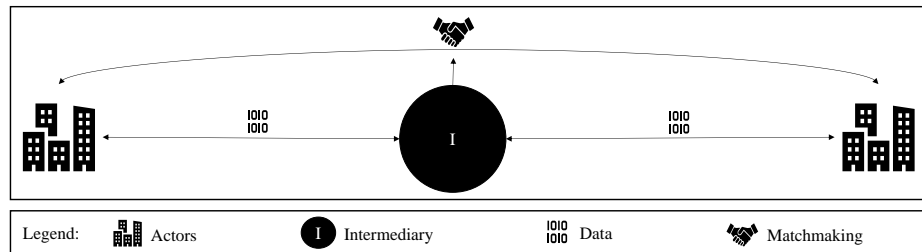

Figure 4. Visualization of transaction services.

4.2 Governance Services

Second, we found services relating to organizing **governance** in the ecosystems. These services provide legal agreements on an ecosystem level and observe their compliance. First, the governance services include the **definition of legal agreements**. Azkan *et al.* (2022) highlight that *" the role of the data trustee ensures that the platform can enforce governance mechanisms, such as enabling data exchange in compliance with the law and managing access rights*." Thus, there are different ways in which these governance specifications can be designed. For example, it can be a data management and governance framework (Demchenko & Gommans, 2019), but also market rules (Huang *et al.*, 2021) or a contract that includes data usage/sharing agreements (Noorian *et al.*, 2014; Bastiaansen *et al.*, 2019; Bastiaansen *et al.*, 2020). In these cases, it is up to the intermediary to protect the data of the actors (Watson *et al.*, 2008). This aspect leads to the second service in the governance services area: **observing and auditing legal compliance**. Woodside (2008) sees a trusted entity's activity as "*monitors the interaction for auditing and legal purposes*." For example, it aims to prevent manipulation and ensure transparency (Demchenko & Gommans, 2019; Agahari & Reuver, 2022). This service is also closely related to conflict resolution. Conflicts can be counteracted by monitoring compliance with legal requirements (Woodside, 2008). An example of governance services is provided by CDQ (2023), providing data quality rules.

In addition to monitoring legal compliance, intermediaries can **monitor interactions** between actors (Woodside, 2008). This can be ensured, for example, through role interaction protocols (Kurtz *et al.*, 2019). Figure 5 visualizes the governance service. The actors have each entered legal agreements with the intermediary and now share their data monitored by the data intermediary.
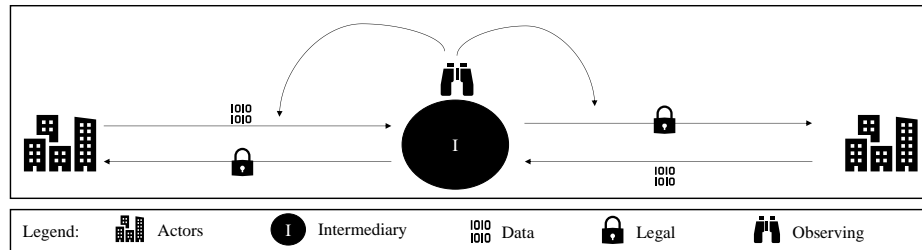

Figure 5. Visualization of governance services.

4.3 Sovereignty Services

Data sovereignty services ensure compliance with policies (e.g., access or usage control) that uphold the self-determination of data ownership by the data provider (e.g., Zrenner *et al.*, 2019). For data intermediaries, we propose four data sovereignty services. A key aspect of this service is **to ensure data sovereignty**. Bastiaansen *et al.* (2019) see the role of the intermediary in supporting data sovereignty and Bastiaansen *et al.* (2020) highlight that "*restricting proliferation of the metadata to their 'home'*

intermediary roles that a data provider or data consumer has subscribed to, will prevent such loss of control and will therefore increase the level of data sovereignty." It becomes clear that the service of providing data sovereignty to an intermediary is decisive for "*enabling data exchange and integrating/involving data markets into B2B relations*" (Demchenko & Gommans, 2019). One example for providing data sovereignty as a service is Sovity (2023). Their approach focus on "*open-source software, establishes a self-determinded control of data usage, to gain full value out of data*" (Sovity, 2023). A core service regarding the implementation of data sovereignty is **anonymizing and pseudonymizing data**. This should minimize the risk of reidentification (van den Broek & van Veenstra, 2015). Susha *et al.* (2017) refer to this service as "*"mediated revealing" (i.e. aggregating and anonymizing datasets before transfer to the user)*" creating a win-win situation for both the data provider and data user. In addition to anonymization and pseudonymization, **certification** is a key component of the sovereignty services. According to Bastiaansen *et al.* (2019), this strengthens the trust function. For example, certification can be implemented by using a decryption key (Sonehara *et al.*, 2011). GXFS.eu (2023) is an example of a certification service. The core content of its activity represents the creation of a federated system in which data can be shared. Access to this system is regulated by an authentication & authorization service. This is intended to enable a "*trusted, decentralized and self-sovereign manner*" without the need for a "*central source of authority*" (GXFS.eu, 2023).

The last identified service in the data sovereignty services framework is **data access control**. In this case, the intermediary has *"control over how data and service flows between platform participants"* (Kurtz *et al.*, 2019). This can happen for a wide variety of reasons. For example, the intermediary may have *"an incentive to strategically and selectively reveal certain consumer data but not others, regulation of data exchanges may actually increase consumers' welfare and/or firms' payoffs"* (Marotta *et al.*, 2022). A side effect of this control is to strengthen the privacy of individual actors and is intended to prevent data manipulation (Roa *et al.*, 2017; Agahari & Reuver, 2022). For instance, Snowflake (2023) provides a platform for secure data sharing, including controlled data access. Figure 6 visualizes the sovereignty service. It shows the data exchange between two actors via an intermediary. The intermediary ensures data sovereignty as well as anonymization and pseudonymization. Access to the data is also controlled. Two scenarios are shown in the figure. On the left, the actor is granted access, while on the right, access is denied.
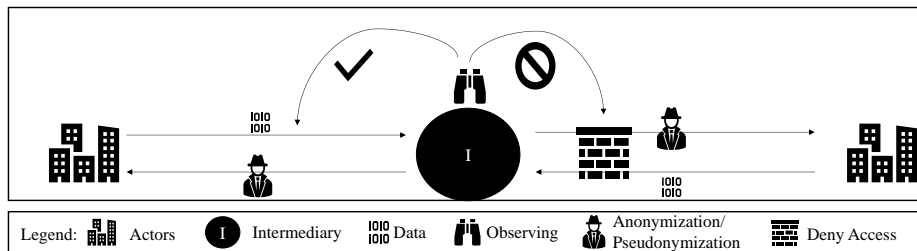

Figure 6. Visualization of sovereignty services.

4.4 Technology Services

One service range deals with providing **technology**, such as infrastructure, or ensuring standardization in data sharing. As already mentioned, a decisive obstacle to data sharing is the lack of know-how regarding the implementation of data sharing (Gelhaar *et al.*, 2021). In addition to implementation at the organizational level, this point often also relates to the technical level, such as the data infrastructure or a standardized sharing process. **Standardization** is vital in data sharing because it enables data exchange by creating interoperability among stakeholders (Bastiaansen *et al.*, 2019). For example, Bastiaansen *et al.* (2020) state that "*from a usercentric perspective, the focus of standardization must be on accessibility of the functionality provided by the intermediary organizations providing services on the legal concepts of data sharing agreements and usage contracts*". The aforementioned GXFS.eu (2023) project offers certification services as well as standardization. Besides implementing standardization, the intermediary can **provide the data infrastructure**. According to Agahari *et al.* (2021), the goal of this service is "*to ease the burden for data providers and consumers*". This service can be implemented in the form of software, but also hardware, other data-related services (e.g., cloud services), the provision of a platform (Öksüz, 2014; Fuerstenau & Auschra, 2016; Schreieck *et al.*, 2016; Demchenko & Gommans, 2019; Kurtz *et al.*, 2019; Schmidt, 2022; Woroch & Strobel, 2022). For example, the data infrastructure service can provide data storage (Watson *et al.*, 2008; Kurtz *et al.*, 2019; Azkan *et al.*, 2022; Bergman *et al.*, 2022). The provision of infrastructure via an intermediary "*will allow for trustworthy and transparent trading and exchange of user data among applications*" (Noorian *et al.*, 2014). An example of this service is the company Dawex (2023). They provide a data platform through which companies can build their data marketplace to share data with their customers, partners, or within the company.

Figure 7 visualizes the technology service. The intermediary assumes the central role by providing a platform, for example, a data marketplace, through which the various players can share their data. This platform can include other services, such as taking care of governance and sovereignty issues so that the platform's actors can rely on standardized processes and not face the obstacle of implementing the technical implementation of data sharing.
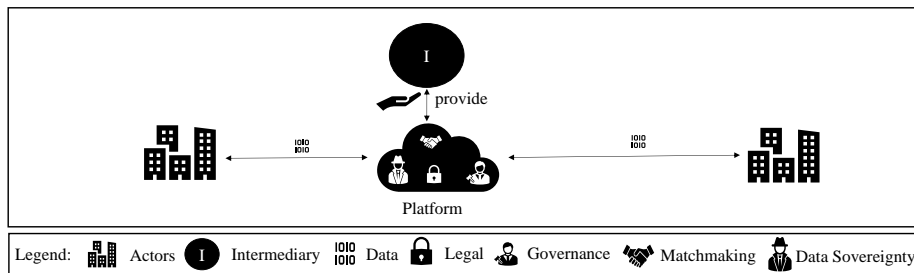

Figure 7. Visualization of technology services.

4.5   Data Services

Data Services are the final service dimension we identified in our analysis. This service focuses exclusively on the resource of data. On the one hand, there is a focus on aggregating the necessary data and ensuring that this data has specific quality characteristics.

A key aspect of data sharing is the **quality** as well as the **integrity** of the data (Winter & Davidson, 2017). Agahari *et al.* (2021) highlight that "*data marketplaces offer complementary applications and services such as data visualizations, data valuation, and data analytics*". The importance of ensuring data quality is also highlighted by (Demchenko & Gommans, 2019) that "*data traded as economic goods must possess/demonstrate such properties as measurable quality, identifiability, veracity, non-rivalry/reusability, privacy, compliance and comply with the FAIR principles*". Ensuring data quality can play a critical role in conflict resolution (Bastiaansen *et al.*, 2019) and motivate participants to participate in data sharing (Huang *et al.*, 2021). One example from the field is CDQ (2023), which offers companies a platform, security standards, and data quality assurance. This is guaranteed by the fact that the data platform "*provides access to 70+ trustworthy open and premium data sources and offers easy data mapping and system integration*". One goal of intermediaries (e.g., in the form of a data marketplace) is to make data "*more accessible to potential users*" (Figueredo *et al.*, 2020). This **data aggregation** service goes hand in hand with quality assurance since the choice or provision of data (sources) to the consumer also determines the possible quality of the data (Winter & Davidson, 2017). Catena-X (2023) offers an example of such a service to aggregate data in the automotive industry and make it accessible to participants in the ecosystem. Figure 8 visualizes that the intermediary enables data exchange between different actors. By checking the data provided by the actors, the data intermediary ensures the required data quality. In doing so, the intermediary can draw on data provided by the actors and publicly available data, aggregate them and make them available.
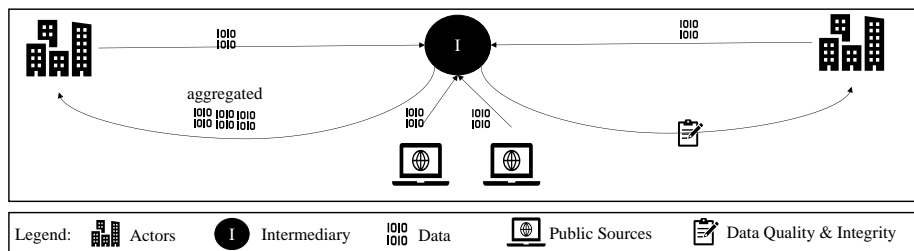

Figure 8. Visualization of data services.

5   Contributions, Limitations, and Outlook

The paper has multiple **contributions**. Primarily, we give an overview of services a data intermediary can fulfill. We see this as highly useful against the background of the DA and DGA and mandating the installment of neutral data intermediaries. While there is research reporting on several data intermediaries, taking a big-picture approach and

extracting services holistically is a novel contribution. For **research**, this overview can help others to go into more detail about specific services and complement our findings. Indeed, the findings can act as a link between different streams of literature investigating demarcated types of intermediaries, such as *data marketplaces* (e.g., Abbas *et al.*, 2021), *data trusts* (e.g., Stachon *et al.*, 2023) or roles in *data space* (e.g., Otto *et al.*, 2019; Otto & Jarke, 2019). In particular, we took a broader approach, decoupling the services from specific instances. Consequently, they can be used flexibly and as a repertoire of services and functionalities tailored to specific scenario requirements. Additionally, this approach can inspire researchers dealing with a particular data intermediary (e.g., data trust) to consider novel services not from that domain and potentially integrate them. For **practice**, we see a relevant issue with the DGA and DA that potentially requires practitioners to accommodate data intermediaries to offer specific services in a specific way. While, at this point, our research does not compare the service spectrum with requirements from legislation, this could be a potential next step for practitioners aiming to implement data intermediaries or data ecosystems in general.

The paper has several **limitations**. Currently, the research builds on a systematic literature review, which we complemented by finding single practice examples (see above). This means that we can have missed articles, that the keywords we chose excluded potentially relevant literature, and that we analyzed the literature based on our interpretation and experience. Given that data intermediaries are relevant in practice (e.g., Advaneo GmbH, 2023), and European Legislation (e.g., European Commission, 2020) facilitates their implementation, we expect a growing body of research and projects to emerge in the near future. Subsequently, we provide a snapshot of current research and expect it to be needing updates frequently.

There are multiple ways forward for **new research**. Primarily, we see methodological complementation as particularly interesting. For example, by integrating more data sources (e.g., interviews), the field of data intermediaries can be explored more in-depth and broader. In particular, conducting an interview study with data intermediary users and operators should be highly valuable to explore relevant issues that are hard to conceptualize from the literature, such as the *neutrality* of intermediation services. Data intermediary research requires interdisciplinary research since the DGA and DA might mandate specific functions, and incorporating neutrality can touch on data intermediary business models. Resulting, what comes into play is the role of data intermediaries investigated from an IS research perspective, legal, and business. This entails researching business models for data intermediaries since they are prerequisites for long-term competitive sustainability. The categories as they are now are non-mutually exclusive and have overlapping characteristics. Exploring whether there are characteristics (e.g., conflict resolution) that should be exclusive to one category is the next step in our research.

6   Acknowledgements

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. Available from: https://doi.org/10.3390/jtaer16070180.

Advaneo GmbH (2023) *Advaneo Data Marketplace*. Available from: https://www.advaneo-datamarketplace.de/database/en/ [Accessed 9 March 2023].

Agahari, W., Dolci, R. & Reuver, M. de (2021) Business Model Implications of privacy-preserving technologies in data marketplaces: The case of multi-party computation. In: *Proceedings of the 29th European Conference on Information Systems*, A Virtual AIS Conference.

Agahari, W. & Reuver, M. de (2022) Rethinking consumers' data sharing decisions with the emergence of multi-party computation: an experimental design for evaluation. In: *Proceedings of the 30th European Conference on Information Systems*, Timisoara: Romania.

Albrecht, S., Reichert, S., Schmid, J., Strüker, J., Neumann, D. & Fridgen, G. (2018) Dynamics of Blockchain Implementation – A Case Study from the Energy Sector. In: *Proceedings of the 51st Hawaii International Conference on System Sciences*, Hawaii: USA.

Azkan, C., Möller, F., Ebel, M., Iqbal, T., Otto, B. & Pöppelbuss, J. (2022) Hunting the Treasure: Modeling Data Ecosystem Participant Value. In: *Proceedings of the 43rd International Conference on Information Systems*, Kopenhagen: Denmark.

Bastiaansen, H., Dalmolen, S., Kollenstart, M. & Punter, M. (2019) Infrastructural Sovereignty over Agreement and Transaction Data ('Metadata') in an Open Network-Model for Multilateral Sharing of Sensitive Data. In: *Proceedings of the 40th International Conference on Information Systems*, Munich: Germany.

Bastiaansen, H., Dalmolen, S., Kollenstart, M. & van Engers, T.M. (2020) User-Centric Network-Model for Data Control with Interoperable Legal Data Sharing Artefacts. In: *Proceedings of the 24th Pacific Asia Conference on Information Systems*, Dubai: UAE.

Bergman, R., Abbas, A.E., Jung, S., Werker, C. & Reuver, M. de (2022) Business model archetypes for data marketplaces in the automotive industry: Contrasting business models of data marketplaces with varying ownership and orientation structures. *Electronic Markets*, 32(2), 747–765.

Capgemini Research Institute (2021) *Data Sharing Masters: How smart organizations use data ecosystems to gain an unbeatable competitive edge*. Available from: https://www.capgemini.com/de-de/insights/research/data-sharing-masters/ [Accessed 22 February 2023].

Catena-X (2023) *Catena-X*. Available from: https://catena-x.net/en/ [Accessed 25 January 2023].

CDQ (2023) *CDQ*. Available from: https://www.cdq.com/ [Accessed 25 January 2023].

Cirullies, J. & Schwede, C. (2021) On-demand Shared Digital Twins – An Information Architectural Model to Create Transparency in Collaborative Supply Networks. In: *Proceedings of the 54th Hawaii International Conference on System Sciences*, Hawaii: USA.

Dawex (2023) *Data Exchange Platform*. Available from: https://www.dawex.com/de/ [Accessed 7 March 2023].

Demchenko, Y. & Gommans, L. (2019) Open Data Market Architecture and Functional Components. In: *Proceedings of the 17th International Conference on High Performance Computing & Simulation.* Dublin:Ireland.

E015 (2023) *E015 digital ecosystem: regione lombardia*. Available from: https://www.e015.regione.lombardia.it/ [Accessed 7 March 2023].

Eisenmann, T., Parker, G. & van Alstyne, M. (2011) Platform envelopment. *Strategic Management Journal*, 32(12), 1270–1285. Available from: https://doi.org/10.1002/smj.935.

European Commission (2020) *REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL: on European data governance (Data Governance Act)*. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0767&from=DE [Accessed 23 March 2022].

European Commission (2022a) *A European Strategy for data*. Available from: https://digital-strategy.ec.europa.eu/en/policies/strategy-data [Accessed 22 February 2023].

European Commission (2022b) *Data Act: Commission proposes measures for a fair and innovative data economy*. Available from: https://ec.europa.eu/commission/presscorner/detail/en/ip_22_1113 [Accessed 11 November 2022].

Fassnacht, M., Benz, C., Heinz, D., Leimstoll, J. & Satzger, G. (2023) Barriers to Data Sharing among Private Sector Organizations. In: *Proceedings of the 56th Hawaii International Conference on System Sciences*, Hawaii: USA.

Federal Ministry for Economic Cooperation and Development (2023) *More fairness in global supply chains – Germany leads the way*. Available from: https://www.bmz.de/en/issues/supply-chains [Accessed 9 March 2023].

Figueredo, K., Seed, D. & Wang, C. (2020) A Scalable, Standards-Based Approach for IoT Data Sharing and Ecosystem Monetization. *IEEE Internet of Things Journal*, 9(8), 5645–5652.

Fuerstenau, D. & Auschra, C. (2016) Open Digital Platforms in Health Care: Implementation and Scaling Strategies. In: *Thirty Seventh International Conference on Information Systems*, Dublin: Ireland.

Gelhaar, J., Gürpinar, T., Henke, M. & Otto, B. (2021) Towards a taxonomy of incentive mechanisms for data sharing in data ecosystems. In: *Proceedings of the 25th Pacific Asia Conference on Information Systems*, Dubai: UAE.

Gioia, D.A., Corley, K.G. & Hamilton, A.L. (2013) Seeking Qualitative Rigor in Inductive Research. *Organizational Research Methods*, 16(1), 15–31. Available from: https://doi.org/10.1177/1094428112452151.

GXFS.eu (2023) *Gaia-X Federation Services GXFS*. Available from: https://www.gxfs.eu/de/ [Accessed 7 March 2023].

Handelsblatt (2023) *Dax-Konzerne und Mittelständler schaffen Cloud-Plattform für Autoindustrie*. Available from: https://www.handelsblatt.com/unternehmen/industrie/catena-x-dax-konzerne-und-mittelstaendler-schaffen-cloud-plattform-fuer-autoindustrie/28954264.html [Accessed 22 February 2023].

Huang, L., Dou, Y., Liu, Y., Wang, J., Chen, G. & Zhang, X. et al. (2021) Toward a research framework to conceptualize data as a factor of production: The data marketplace perspective. *Fundamental Research*, 1(5), 586–594. Available from: https://doi.org/10.1016/j.fmre.2021.08.006.

Janssen, H. & Singh, J. (2022) Data intermediary. *Internet Policy Review*, 11(1). Available from: https://doi.org/10.14763/2022.1.1644.

Jussen, I., Schweihoff, J., Dahms, V., Möller, F. & Otto, B. (2023) Data Sharing Fundamentals: Definition and Characteristics. In: *Proceedings of the 56th Hawaii International Conference on System Sciences*, Hawaii: USA.

Kioses, E., Pramatari, K., Doukidis, G. & Bardaki, C. (2007) Measuring the Business Value of Electronic Supply Chain Collaboration: The Case of Electronic Invoicing. In: *Proceedings of the 20th Bled eConference*, Bled: Solvenia.

Koul, N.A., Sabrina, A. & Morin, J. (2018) Agreements Framework for Data Market Ecosystem. In: *Proceedings of the 13th Pre-ICIS Workshop on Information Security and Privacy*, San Francisco: USA.

Kurtz, C., Wittner, F., Semmann, M., Schulz, W. & Böhmann, T. (2019) The Unlikely Siblings in the GDPR Family: A Techno-Legal Analysis of Major Platforms in the Diffusion of Personal Data in Service Ecosystems. In: *Proceedings of the 52nd Hawaii International Conference on System Sciences*, Hawaii: USA.

Leemann, N. & Kanbach, D.K. (2022) Toward a taxonomy of dynamic capabilities – a systematic literature review. *Management Research Review*, 45(4), 486–501. Available from: https://doi.org/10.1108/MRR-01-2021-0066.

Levy, Y. & Ellis, T. (2006) A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research. *International Journal of an Emerging Transdiscipline*, 9, 181-212. Available from: https://doi.org/10.28945/479.

Marotta, V., Wu, Y., Zhang, K. & Acquisti, A. (2022) The Welfare Impact of Targeted Advertising Technologies. *Information Systems Research*, 33(1), 131–151.

Noorian, Z., Iyilade, J., Mohkami, M. & Vassileva, J. (2014) Trust Mechanism for Enforcing Compliance to Secondary Data Use Contracts. In: *IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications*, Exeter: United Kingdom.

Öksüz, A. (2014) Turning Dark into White Clouds – A Framework on Trust Building in Cloud Providers via Websites. In: *Proceedings of the 20th Americas Conference on Information Systems*, Savannah: USA.

Oliveira, M.I.S., 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. Available from: https://doi.org/10.1007/s10115-018-1323-6.

Oliveira, M.I.S. & Lóscio, B.F. (2018) What is a data ecosystem? In: *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age (dg.o '18)*, New York: USA.

Opriel, S., Möller, F., Burkhardt, U. & Otto, B. (2021) Requirements for Usage Control based Exchange of Sensitive Data in Automotive Supply Chains. In: *Proceedings of the 54th Hawaii International Conference on System Sciences, Proceedings of the Annual Hawaii International Conference on System Sciences*, Hawaii: USA.

Osterrieder, P., Budde, L. & Friedli, T. (2020) The smart factory as a key construct of industry 4.0: A systematic literature review. *International Journal of Production Economics*, 221, 107476. Available from: https://doi.org/10.1016/j.ijpe.2019.08.011.

Otto, B. & Jarke, M. (2019) Designing a multi-sided data platform: findings from the International Data Spaces case. *Electronic Markets*, 29(4), 561–580. Available from: https://doi.org/10.1007/s12525-019-00362-x.

Otto, B., Steinbuß, S., Teuscher, A. & Lohmann, S. (2019) *Reference Architecture Model: Version 3.0*. Available from: https://www.fraunhofer.de/content/dam/zv/en/fields-of-research/industrial-data-space/IDS-Reference-Architecture-Model.pdf [Accessed 11 September 2021].

Paré, G., Trudel, M.-C., Jaana, M. & Kitsiou, S. (2015) Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, 52(2), 183–199. Available from: https://doi.org/10.1016/j.im.2014.08.008.

Perscheid, G., Ostern, N.K. & Moormann, J. (2020) Determining Platform Governance: Framework for Classifying Governance Types. In: *Proceedings of the 41st International Conference on Information Systems*, Hyderabad: India.

Richter, H. & Slowinski, P.R. (2019) The Data Sharing Economy: On the Emergence of New Intermediaries. *IIC - International Review of Intellectual Property and Competition Law*, 50, 4–29. Available from: https://doi.org/10.1007/s40319-018-00777-7.

Roa, J., Villarreal, P. & Fantinato, M. (2017) Institutional Violence Complaints in Argentina: A Privacy Study. In: *Proceedings of the 50th Hawaii International Conference on System Sciences*, Hawaii: USA.

Schmidt, M.-C. (2022) Determination of Relational Success? On the Interrelation of Digital Industrial Platforms and Buyer-Supplier Relationships. In: *Proceedings of the 26th Pacific Asia Conference on Information Systems, Annual Pacific Asia Conference on Information Systems,* July, 5-9, Taipei-Sydney.

Schreieck, M., Wiesche, M. & Krcmar, H. (2016) DESIGN AND GOVERNANCE OF PLATFORM ECOSYSTEMS - KEY CONCEPTS AND ISSUES FOR FUTURE RESEARCH. In: *Proceedings of the 24th European Conference on Information Systems*, Istanbul: Turkey.

Snowflake (2023) *Snowflake: Customer Stories: Snowflake Data Cloud*. Available from: https://www.snowflake.com/en/why-snowflake/customers/ [Accessed 25 January 2023].

Sonehara, N., Echizen, I. & Wohlgemuth, S. (2011) Isolation in Cloud Computing and Privacy-Enhancing Technologies. *Business & Information Systems Engineering*, 3, 155–162. Available from: https://doi.org/10.1007/s12599-011-0160-x.

Sovity (2023) *sovity - Data Sovereignty as a Service*. Available from: https://sovity.de/de/start/ [Accessed 7 March 2023].

Stachon, M., Möller, F., Guggenberger, T. & Tomcyzk, M. (2023) Understanding Data Trusts. In: *Proceedings of the 31st European Conference 2023*, Kristiansand: Norway.

Sterk, F., Peukert, C., Hunke, F. & and Weinhardt, C. (2022) Understanding Car Data Monetization: A Taxonomy of Data-Driven Business Models in the Connected Car Domain. In: *Proceedings of the 17th International Conference on Wirtschaftsinformatik*, Nürnberg: Germany.

Susha, I., Janssen, M. & Verhulst, S. (2017) Data Collaboratives as a New Frontier of Cross-Sector Partnerships in the Age of Open Data: Taxonomy Development. In: *Proceedings of the 50th Hawaii International Conference on System Sciences*, Hawaii: USA.

Templier, M. & Paré, G. (2015) A Framework for Guiding and Evaluating Literature Reviews. *Communications of the Association for Information Systems*, 37, 112–137. Available from: https://doi.org/10.17705/1CAIS.03706.

van de Ven, M., Abbas, A.E., Kwee, Z. & Reuver, M. de (2021) Creating a Taxonomy of Business Models for Data Marketplaces. In: *Proceedings of the 34th BLED eConference*, Online Conference.

van den Broek, T. & van Veenstra, A.F. (2015) Modes of Governance in Inter-Organizational Data Collaborations. In: *Proceedings of the 23rd European Conference on Information Systems*, Münster: Germany.

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(9), 205–224.

Watson, R.T., Piccoli, G., Brohman, M.K. & Parasuraman, A. (2008) Customer-Managed Interactions: A New Paradigm for Firm-Customer Relationships. *MIS Quarterly Executive*, 4(2), 319–327.

Webster, J. & Watson, R.T. (2002) Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, 26(2), xiii–xxiii. Available from: http://dl.acm.org/citation.cfm?id=2017160.2017162.

Wernick, A., Olk, C. & Grafenstein, M. von (2020) Defining Data Intermediaries - A Clearer View through the Lens of Intellectual Property Governance. *Technology and Regulation*, 65–81. Available from: https://doi.org/10.26116/techreg.2020.007.

Winter, J.S. & Davidson, E. (2017) Investigating Values in Personal Health Data Governance Models. In: *Proceedings of the 23rd Americas Conference on Information Systems, Annual Americas Conference on Information Systems*, Boston: USA.

Woodside, J.M. (2008) Dynamic Healthcare Connectivity and Collaboration with Multi-Agent Systems. In: *Proceedings of the Fourteenth Americas Conference on Information Systems*, Toronto: Canada.

World Economic Forum (2021) *6 ways data sharing can shape a better future*. Available from: https://www.weforum.org/agenda/2021/04/six-ways-data-sharing-can-shape-a-better-future/ [Accessed 13 February 2023].

Woroch, R. & Strobel, G. (2022) Show me the Money: How to monetize data in data-driven business models? In: *Proceedings of the 17th International Conference on Wirtschaftsinformatik*, Nürnberg: Germany.

Zrenner, J., Möller, F., Jung, C., Eitel, A. & Otto, B. (2019) Usage control architecture options for data sovereignty in business ecosystems. *Journal of Enterprise Information Management*, 32(3), 477–495. Available from: https://doi.org/10.1108/JEIM-03-2018-0058.