Empowering Data Sovereignty: Strategies of Data Intermediaries in Data Ecosystems

Ruben D'Hauwers, August Bourgeus
Imec-SMIT, VUB, Brussel, Belgium
ruben.dhauwers@vub.be, august.bourgeus@vub.be

This study delves into the intricacies of data control mechanisms within (personal) data ecosystems, with the goal of attaining a harmonious balance in data sovereignty between individuals and data providers, facilitated by data intermediaries. Employing a multiple-case study analysis involving eleven data intermediaries, the research utilizes axial coding and triangulation with existing literature to identify dimensions of the novel Taxonomy of Data Control Mechanisms for data intermediaries. These dimensions encompass three meta-dimensions: data access control, power dynamics, and revenue sharing models. These meta-dimensions consist of eight dimensions that can be harmonized to achieve equilibrium in data sovereignty between data subjects and providers. This research contributes to both theoretical comprehension and practical implementation in navigating the complexities of data sovereignty within dynamic data ecosystems.

Keywords:
data sovereignty,
data intermediary
data control
data ecosystem
taxonomy

University of Maribor Press
DOI https://doi.org/10.18690/um.fov.4.2024.13
ISBN 978-961-286-871-0

1 Introduction

In today's data-driven world, the dominance of big tech companies in the data economy has raised concerns about the control and influence they exert over smaller actors (Zuboff, 2015). The concept of data ecosystems has emerged as a promising avenue for fostering collaboration within networks centered around the use of data (S. Oliveira et al., 2019). Particularly within personal data ecosystems, the focus shifts to individuals and their personal data, forming the cornerstone of these ecosystems (Moiso & Minerva, 2012). Moreover, data intermediaries have gained traction, overseeing data governance between data providers and data users (Janssen & Singh, 2022). To ensure the sustainability of personal data ecosystems, a balance in data sovereignty for both data subjects (Hummel et al., 2021) and data providers (Zrenner et al., 2019) is required. This can be facilitated by data intermediaries through access control and data governance mechanisms (Curry, 2020; Gelhaar et al., 2021), empowering data providers and subjects to dictate terms for data usage by data consumers.

This paper addresses the research question: "Through which dimensions of data control mechanisms can data intermediaries influence the equilibrium of data sovereignty between individuals and data providers within personal data ecosystems?" A taxonomy that identifies data control mechanisms within data intermediaries to harmonize data sovereignty for both data providers and subjects is developed, encompassing various data intermediary types. To address this question, a multiple-case study analysis approach (Yin, 2013) is employed, involving eleven data intermediaries. The taxonomy (Nickerson et al., 2013) is developed through axial coding of case studies (Corbin & Strauss, 2008), and triangulated with existing literature (Patton, 1999). The paper begins with a review of literature on data ecosystems, data sovereignty, data intermediaries, and data governance taxonomies. It then outlines the methodology, presents results, discusses implications for academia and real-world applications, and concludes with key findings and future research directions.

2 Literature

2.1 (Personal) Data Ecosystems

Various trends in big data, artificial intelligence, and the Internet of Things have increasingly drawn attention to data ecosystems focusing on data usage (Curry & Sheth, 2018). Within these trends, personal data is often used as the asset traded in data ecosystems (Spiekermann et al., 2015). In the context of business, ecosystems are networks of interacting organizations devoid of hierarchical management, instead bound together by their shared investments, facilitating coordination that eliminate the necessity for individual contractual agreements with each partner (Jacobides et al., 2018). Data ecosystems are business ecosystems (Adner, 2017) that aim to create a focal value proposition with the exchange of data at the center. Data ecosystems are characterized by their connected network structure, the presence of platforms facilitating value creation, and collaborative co-evolution among actors (Schreider 2023, Oliveira et al, 2019). In personal data ecosystems, individuals are integral in exchanging personal data (Ojasalo & Miskeljin, 2020; Spiekermann et al., 2015). A "user-centric" model is advocated, empowering individuals to control the gathering, management, use, and sharing of their data (Moiso & Minerva, 2012). The integration of humans in data ecosystems entails providing them with adequate information and power over their data while maintaining transparency, honesty, and security (Koskinen et al., 2023).In this context, stakeholders in the data ecosystem include data providers, who make data they control available, and data consumers, who receive this data (Otto & Teuscher, 2019). A data subject refers to an individual identifiable through personal data (Scheider et al., 2023). In personal data ecosystems, a data provider could be a private company controlling data or a data subject providing data. This research distinguishes between private data providers and data subjects seeking control over personal data.

2.2 Balancing Data Sovereignty in Data Ecosystems

Big tech companies currently dominate and control the data economy, sidelining smaller actors affecting both individuals and organizations (Knaapi-Junnila et al., 2022; Koskinen et al., 2023). This power dynamic, termed data colonialism (Couldry & Mejias, 2019) or surveillance capitalism (Zuboff, 2015), involves exploiting personal data for profit and impacts less powerful actors by diminishing their

autonomy and control over their data. In this context, the concept of data sovereignty becomes crucial, defining entities' self-determination over their data, encompassing both individuals and businesses (Otto & Teuscher, 2019; Scheider et al., 2023).Regarding data sovereignty of the data subject, different authors advocate the need for individuals to have control over their personal data usage, including determining access and processing purposes, with clarity on data privacy and protection (Hummel et al., 2021). Also self-determined sharing and (monetary) incentivization for the sharing of personal data is required (Lauf et al., 2022). For businesses, enabling data sovereignty of the data subject may contradict with the data sovereignty of the company, as it involves granting data control to data subjects. As companies often perceive (personal) data as an asset crucial for enhancing competitiveness (Gupta & George, 2016), it leads them to prioritize its protection. As granting data control to data subjects potentially compromises their competitive advantage, they may be reluctant to share business-critical data (Tomi Dahlberg & Nokkala, 2017).

To foster sustainable personal data ecosystems, a need arises to balance data sovereignty for both data subjects and data providers. This can be enabled through mechanisms like access control and data control (Zrenner et al., 2019), allowing data providers and data subjects to set terms for data usage by data consumers (Loebbecke et al., 2016; Scheider et al., 2023). This research focuses on identifying data - and access control mechanisms, enabling the balance of data sovereignty for the data subjects and data providers.

2.3 Data Intermediary Models

Data intermediaries, a new model introduced in the Data Governance Act in Europe Act (Regulation (EU) 2022/868, 2022), act as mediators between data providers and users, governing the data and providing confidence in its usage (Janssen & Singh, 2022). Some of these intermediaries facilitate data sovereignty for both data subjects and providers.

Firstly, in enabling data sovereignty for data providers, we consider data sharing pools and data marketplaces. Data-sharing pools (DSPs) (Micheli et al., 2023) are alliances among data providers that share data intending to improve their assets (data products, processes and services) by exploiting the complementarities of the pooled

data. The alliances have a shared purpose, context or application, and are intended to benefit all their participants. Data marketplaces (DM) (Janssen & Singh, 2022; Micheli et al., 2023) serve as platforms facilitating data exchange between buyers and sellers, allowing data sellers to monetize their data while retaining control over its usage and access.

Secondly, various models facilitate data sovereignty for data subjects. In data cooperatives (DC) (Hartman et al., 2020; Micheli et al., 2023), the collective ownership of data is emphasized. They recognize data subjects as vital stakeholders and aim to rectify power imbalances and enable equitable benefit-sharing. Personal Information Management Systems (PIMS) (Micheli et al., 2023; Van Kleek & OHara, 2014) stress individuals' control over their data, countering private companies' influence and empowering users to determine their personal information's usage, fostering a balanced relationship with digital platforms. Additionally, data unions (Micheli et al., 2023) advocate for individuals or groups in the data economy by pooling their data and negotiating fair terms with data buyers, enhancing individuals' bargaining power (Micheli et al., 2023) (Micheli, 2023). Data trusts are based on trust law, which allows data rights holders to delegate control of their data to a trustee (Micheli et al., 2023) based on a legal mechanism that permits the rights of data subjects/holders to be pooled to negotiate terms of use in data subjects' favor (Sadowski et al., 2021).

2.4 Research Gap

The lack of research on data governance in inter-organizational data exchange is evident (Abraham et al., 2019), which similarly applies to data governance within ecosystems. Typologies and taxonomies (Gelhaar et al., 2021; Lis & Otto, 2020) have been developed to address this gap, focusing on data governance within ecosystems, yet research should expand beyond organizational boundaries to encompass personal ecosystem models (Koskinen et al., 2023). While various studies have addressed data sovereignty for data subjects (Scheider et al., 2023) and data providers (Zrenner et al., 2019), there needs to be more research on achieving balance between them within the context of data intermediaries.Thus, the research question for this study is: "Through which dimensions of data control mechanisms can data intermediaries influence the equilibrium of data sovereignty between individuals and data providers within personal data ecosystems?". The focus is on a taxonomy

identifying data control mechanisms data intermediaries can use to balance data sovereignty.

3 Methdology

Following a taxonomy development methodology (Nickerson et al., 2013), the researchers progressed through two iterations to discern the taxonomy's dimensions. An empirical-to-conceptual (E2C) was performed in the first iteration, deducing dimensions and characteristics from empirical studies (Nickerson et al., 2013). This involved analyzing real-life use cases of existing data intermediaries using a case study analysis method, which allows for examination within authentic contexts (Yin, 2013). The case study methodology is suitable for explanatory and descriptive purposes (Runeson & Höst, 2009), particularly in industry-based scenarios (Verner et al., 2009). Eleven data intermediaries were selected based on the definition: "data intermediaries act as mediators between data providers and users, governing the data and providing confidence in its usage" (Janssen & Singh, 2022). Data intermediaries for this study were chosen using a snowballing approach (Berg, 2006), initially identified through web searches, and reviewing reports in academic and industry literature. Cases encompassing PIMS, data unions, DC, DMs, trusts, and DSPs were analyzed, focusing on their objectives regarding data sovereignty for either data subjects or data providers. This iterative process ensured diverse representation across models, functions, and industries for comprehensive analysis. Refer to Annex 1 for an overview of selected use cases. Qualitative data analysis was conducted using MAXQDA, examining textual descriptions on the intermediaries' websites and whitepapers. Axial coding, combining inductive and deductive thinking (Corbin & Strauss, 2008) was applied iteratively. As different use cases were analyzed, meta-dimensions and dimensions were refined accordingly. In the second iteration, oriented at the conceptual-to-empirical approach (C2E), methodological triangulation (Denzin, 1978; Patton, 1999) was performed to scrutinize the results to increase the credibility of the findings (Patton, 1999). Following the second iteration, all objective and subjective ending conditions were met (Nickerson et al., 2013).

4 Taxonomy of Data Control Mechanisms for Data Intermediaries

Table 1 presents a taxonomy of Data Control Mechanisms for Data intermediaries, identifying modifiable dimensions that shape control and data sovereignty trade-offs among ecosystem actors.

Table 1: Taxonomy of Data Control Mechanisms for Data Intermediaries

Meta-Dimension	Dimension	Characteristics			
Data Access Control	Actor Access Control	Data Subject	Intermediary	Dual	Data Provider
	Data Type	Raw Data	Processed	Inferred or derived	Verified
	Data Access Criteria	User Preference	Collective Policies	Intermediary Policies	Data Provider Relationship	Data Provider Usage Policies
Power Dynamics	Data Storage Location	Data Subject	Intermediary	Data Provider decentral	Data provider central
	Data Processing Location	Data Subject	Intermediary	Data Provider	Data consumer
	Representation	Individual		Group	
	Legal Rights	Enforcement		Compliance	
Revenue Sharing Model		No	Individual interest model		Collective interest model

4.1 Data Access Control

Data access and usage controls ensure data sovereignty, allowing data subjects and providers to effectively regulate access and usage rights (Bussard et al., 2010; Kelbert & Pretschner, 2012; Zrenner et al., 2019).

Actor Access Control determines which actor has the authority to grant access to data, which can be wielded by data subjects or data providers. In PIMS, **data subjects** can control their data, especially in ecosystems with sensitive personal data, like health information (WeAre), or primarily user-generated data (Karamel). Actors' access control can also be managed by the **intermediary**, (e.g. Swash and LunaDNA which handle group-based aggregated data). In certain Data Sharing Pools and Marketplaces, especially in cases with competitive business data (NxtPort, Catena-X), **data providers** often have greater control over access. **Dual actor control** aims for a balanced sharing of control (DjustConnect and DataVillage) in

cases where sensitive personal data is also competitive business data. Data subjects grant access control to data consumers, while data providers can either give consent or open APIs.

Data Type concerns the type of data consumers receive access to, which helps mitigate the sensitivity of the data (Abrams, 2014; Kugler & Plank, 2022). **(Raw) data** are inherently sensitive for both data subjects and providers, representing the lowest level of control. Providing **processed data** (anonymized, pseudonymized or aggregated data) reduces sensitivity for both parties (LunaDNA's aggregated genomic data). Sharing **inferred or derived** data alone, without divulging algorithms or raw data, is a method of selective sharing based on a need-to-know basis (Datavillage and Swash sIntelligence). **Verification** by the data provider (Karamel verified diplomas) transfers control to both the verifying agent and the data consumer, ensuring data validity.

Data Access Criteria defines how access can be granted to various actors within an ecosystem and are different types of attribute –based (Gupta et al., 2018), relationship-based (Gates, 2007) or role-based (Ferraiolo & Kuhn, 1992) access control mechanisms. The actor that sets the access control criteria holds the highest data sovereignty. **User preference**-based access is set by the data subject, including preferences regarding the data consumer's identity, the data, and the context at the time of the access request. It can happen through individual consent (DjustConnect, Karamel) or automatically (Consent-o-matic). In **Collective policies**, the rights of data subjects/holders can be pooled, and the terms of use for the data determined in the suppliers' interests (datatrusts.uk), often based on democratic decision-making in data cooperatives (Midata.coop). **Intermediary-based** access control hinges on participation within a defined ecosystem, and it is determined by policies set by the intermediary (data sharing regulations set by NxtPort). **Data provider relationship**-based access control fosters controlled data sharing among trusted partners (DataVillage). In intricate ecosystems like Catena-X, **data provider usage-based** access control is established through legally binding policies set by the data provider, delineating the conditions under which data consumers can utilize data assets.

4.2 Power Dynamics

Data sharing within data ecosystems is heavily influenced by power imbalances among companies (Li & Lin, 2008). Data intermediaries can utilize control mechanisms like storage and processing location, representation, and legal rights to manage these power dynamics.

Data Storage Location plays a crucial role in determining control and power dynamics within digital ecosystems (Gelhaar et al., 2021; Scheider et al., 2023). In the **data provider central** scenario (tech giants like Google), control lies firmly in the hands of the data holder. Conversely, in **data subject decentral models** (Karamel and WeAre), personal data is stored within individual-centric pod systems hosted by neutral data storage hosts. In **intermediary-central storage models**, trusted intermediaries store personal data on behalf of users, allowing them to exert power by aggregating data rights across multiple individual data subjects (Midata.coop, Swash and LunaDNA). In **decentralized data provider storage** (Catena-X), data is distributed across multiple data provider locations. Through connectors and interoperability standards, interoperable sharing is facilitated.

Data Processing Location, also referred to as the distribution of intelligence (Ballon, 2007), is an essential consideration, particularly given the rising utilization of AI and data mining. Within data ecosystems, this concept pertains to the specific allocation of processing power, control, and functionality across the system. Data may undergo processing **locally at the data subject's** end. While not exemplified in our sample, this scenario is common in applications such as privacy assistants (Morel & Fischer-Hübner, 2023). In an **intermediated** context, secure collaboration spaces are created where the algorithm is securely located and data remains encrypted (DataVillage), ensuring confidentiality on both the data and the algorithm within the collaboration. Alternatively, data can be processed at the **data provider**'s end, where aggregated data is processed, or algorithms are run to generate insights (DjustConnect or NxtPort). Conversely, raw data can be provided to the **data consumer**, who then processes the data or runs their algorithms, typically observed in large, dominant companies like Facebook and Google.

Representation within a collective can enable data subjects to exert greater control over their data (Delacroix & Lawrence, 2019)**.** Thus, data intermediaries may facilitate a "communal approach to data sharing," involving the entire community in decision-making regarding data rights (Ho & Chuang, 2019). This representation by the data intermediary can take a **collective** or in a group form, where data from various subjects are pooled to enhance power, (Midata.coop, LunaDNA and Swash). Alternatively, representation can be **individual**, where data subjects independently determine data usage without leveraging community power (WeAre, Karamel, and Consent-o-matic).

Legal Rights can be entrusted by a data subject to a data intermediary (World Economic Forum, 2022), serving as a legal mechanism that consolidates the rights of data subjects/holders and determines data usage terms in their favor (Sadowski et al., 2021). This arrangement can enhance protection against privacy infringements and unethical handling of personal data (Micheli et al., 2023). Among data intermediaries, we note a distinction: some actively advocate for the **enforcement** of data subjects' legal rights (Datatrusts.uk). Conversely, others primarily focus on ensuring legal **compliance** without actively enforcing legal rights, by aligning data handling with legal standards (DjustConnect, NxtPort).

4.3 Revenue Sharing Model

The **revenue sharing model** is the extent to which revenues are shared within the ecosystem can facilitate fair profit-sharing among its members (Lauf et al., 2022), potentially leading to a more equitable distribution of benefits for the data subjects. In scenarios where data subjects **do not receive a share of revenue**, profits generated from data sales remain unallocated. Conversely, when revenue sharing occurs, profits from data sales are redistributed, promoting equitable sharing. The **individual-interest model** (Fox, 2020) allows each data subject to receive a portion of revenues based on the amount of shared data. This may involve data subjects receiving stocks, leading to potential monetary returns (LunaDNA), or redistributing revenues generated from data sales (Swash). In contrast, the **collective-interest model** (Fox, 2020) involves community trusts, directing value redistribution collectively towards specific groups.

5 Discussion

This paper makes a significant contribution to both the data sovereignty literature (Hummel et al., 2021; Zrenner et al., 2019) and the data governance literature (Gelhaar et al., 2021; Lis & Otto, 2020) by introducing a taxonomy that serves two primary purposes: Firstly, it facilitates the mapping of various data governance mechanisms aimed at achieving a balance of data sovereignty between data subjects and data providers. Secondly, it provides the building blocks for developing data governance models for data intermediaries. This framework involves the creation of mechanisms for data control, which are closely linked to the governance models implemented by these intermediaries.

Additionally, this paper contributes to the literature on data intermediary models (Janssen & Singh, 2022; Micheli et al., 2023) by uncovering the different mechanisms for data intermediaries to achieve a balance of data sovereignty tailored to the specific purposes of the data intermediary. The purpose of the data intermediary can be control over sensitive or competitive data (Hummel et al., 2021; Zrenner et al., 2019) equitable benefit sharing (Fox, 2020; Lauf et al., 2022) or decentralizing power in the data economy towards the data subject (Zuboff, 2015).First, if the purpose is to effectively manage personal sensitive or competitive data, the control mechanisms depend on two key factors: the company's data competitiveness level (Enders et al., 2020; Kugler & Plank, 2022) and the individual's data sensitivity (Belen Saglam et al., 2022). In this case, data access management is the major control mechanism which aims to balance data sovereignty and mitigate data sensitivity during sharing. Depending on data sensitivity and competitiveness, control varies. Data providers hold the most control in the case of low-sensitivity, high-competitiveness data (e.g., Catena-X), data subjects in the case of highly sensitive, low-competitive data (e.g., Karamel), and dual mechanisms are needed for highly sensitive, highly competitive data (e.g., DjustConnect). Second, value redistribution can be the primary purpose of the data intermediary, as seen in data unions where intermediaries empower data subjects by aggregating their data to create value for consumers. The revenue sharing control mechanism is crucial to enable this purpose. Third, if the purpose is to decentralize power in the data economy, the focus shifts to empowering data subjects and rebalancing dynamics. The power dynamics mechanics involves forming user groups in data unions and DCs to collectively negotiate data access and

centralize storage and access control with PIMS, giving individuals greater control. Moreover, in data trusts, legal rights are enforced by the data intermediary.

6 Conclusion

Various entities like PIMS, data unions, data trusts, data cooperatives, data pools, and DMs aim to achieve a balanced data sovereignty between individuals and data providers, facilitated by mechanisms which enable data sovereignty. These mechanisms were examined in this paper, leading to a taxonomy with three meta-dimensions: data access control, power dynamics, and revenue sharing models, identifying eight dimensions essential for reaching an equilibrium in data sovereignty, depending on the purpose of the data intermediary. These dimensions enable intermediaries to tailor control mechanisms for sensitive data protection, changing power dynamics towards the user, and equitable benefit sharing. Various entities like PIMS, data unions-, trusts and - cooperatives, DSPs and DMs have different data sovereignty balance points, with control mechanisms facilitating this purpose.

This research contributes to theoretical understanding of data governance, data ecosystems and data intermediary literature. The practical applications enable data intermediaries to navigate data sovereignty complexities within evolving data ecosystems by providing data control mechanism building blocks. Limitations to the research include the case study depth and breadth; deeper analysis can reveal underlying control mechanism characteristics, while broader case studies can support evaluating the taxonomy's validity. Future research could augment the taxonomy with value creation and governance dimensions and model different intermediary types using taxonomy dimensions.

7 Annex: overview of data intermediaires and mapping of data control mechanisms

		Djust Connect	We Are	Karamel	Catena-X	NxtPort	Swash	LunaDNA	Consent-o-Matic	Data Village	Midata.coop	Datatrusts.uk
Data Intermediary Type		DM	PIMS	PIMS	DM	DM	Data Union	Data Union	PIMS	DSP	DC	Data trust
Data access Control	**Actor Access Control**	Dual	Data Subject	Data Subject	Data provider	Data Provider	Intermediary	Intermediary / Data subject	Data Subject	Dual	Data subject	Intermediary
	Data type	Derived	Derived *	Raw data / Derived/Verified	Raw data / Processed*	Raw data /Derived	Processed	Processed/ Derived	Raw data*	Derived	Raw/Derived *	Raw*
	Data Access Criteria	User preference	User preference	User preference	Data provider Usage Policy	Intermediary Policies	Intermediary policies	Intermediary policies	User preference	Relationship / User preference	Collective policies	Collective policies
Power Dynamics	**Data Storage**	Data provider decentral	Data subject decentral	Data subject decentral	data provider decentral	data provider decentral	Intermediary central	Intermediary central	Data provider decentral*	Data Provider decentral	Intermediary central	Intermediary central *
	Processing	Data provider*	Data consumer*	Intermediary / Data provider	Data provider / Data consumer	Data provider / Data consumer	Intermediary	Intermediary / Data consumer	Data consumer*	Intermediary	Data consumer	Data provider/ consumer *
	Representation	Individual	Individual	Individual	Individual	Individual	Group	Group	Individual	Individual	Group	Group
	Legal Rights	Compliance	Compliance	Complaince	Compliance	Compliance	Enforcement	Enfotcement	Enforcement	Compliance	Compliance	Enforcement
Revenue Sharing Model		No	No	No	No	No	Individual interest model	Individual interest model	No	No	No	No

* In these instances, the characteristics were determined from the context or were not conclusive based on the available information accessible to the researchers.

References

Abraham, R., Johannes Schneider, & vom Brocke, J. (2019). Data Governance: A conceptual framework, structured review, and research agenda. International Journal of Information Management, 49, 424–438. https://doi.org/10.1016/j.ijinfomgt.2019.07.008

Abrams, M. (2014). The Origins of Personal Data and its Implications for Governance. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2510927

Adner, R. (2017). Ecosystem as a structure: An actionable construct for strategy. Journal of Management, 43(1), 39–58.

Ballon, P. (2007). Business Modelling Revisited: The Configuration of Control and Value. The Journal of Policy, Regulation and Strategy for Telecommunications, Information and Media, 9(5), 6–19.

Belen Saglam, R., Nurse, J. R. C., & Hodges, D. (2022). Personal information: Perceptions, types and evolution. Journal of Information Security and Applications, 66, 103163. https://doi.org/10.1016/j.jisa.2022.103163

Berg, S. (2006). Snowball Sampling—I. In Encyclopedia of Statistical Sciences. John Wiley & Sons, Ltd. https://doi.org/10.1002/0471667196.ess2478.pub2

Bussard, L., Neven, G., & Preiss, F.-S. (2010). Downstream Usage Control. 2010 IEEE International Symposium on Policies for Distributed Systems and Networks, 22–29. https://doi.org/10.1109/POLICY.2010.17

Corbin, J. M., & Strauss, A. (2008). Basics of Qualitative Research (3rd ed.): Techniques and Procedures for Developing Grounded Theory.

Couldry, N., & Mejias, U. A. (2019). Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject. Television & New Media, 20(4), 336–349. https://doi.org/10.1177/1527476418796632

Curry, E. (2020). Fundamentals of Real-time Linked Dataspaces. In E. Curry (Ed.), Real-time Linked Dataspaces: Enabling Data Ecosystems for Intelligent Systems (pp. 63–80). Springer International Publishing. https://doi.org/10.1007/978-3-030-29665-0_4

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

Delacroix, S., & Lawrence, N. D. (2019). Bottom-up data Trusts: Disturbing the 'one size fits all' approach to data governance. International Data Privacy Law, 9(4), 236–252. https://doi.org/10.1093/idpl/ipz014

Denzin, N. (1978). Sociological methods: A sourcebook. McGraw-Hill.

Enders, T., Wolff, C., & Satzger, G. (2020). Knowing What to Share: Selective Revealing in Open Data. European Conference on Information Systems. https://www.semanticscholar.org/paper/Knowing-What-to-Share:-Selective-Revealing-in-Open-Enders-Wolff/04ba82f4971c731287e2e576c2be23bf6f35c577

Ensign, P. C., & Hébert, L. (2009). Competing explanations for knowledge exchange: Technology sharing within the globally dispersed R&D of the multinational enterprise. The Journal of High Technology Management Research, 20(1), 75.

Ferraiolo, D., & Kuhn, R. (1992). Role-Based Access Controls. In Proceedings of the 15th National Computer Security Conference (pp. 554–563). National Institute of Standards and Technoloty. https://doi.org/10/13/rolebased-access-controls/final

Fox, K. (2020). The Illusion of Inclusion—The "All of Us" Research Program and Indigenous Peoples' DNA. New England Journal of Medicine, 383(5), 411–413. https://doi.org/10.1056/NEJMp1915987

Gates, C. (2007). Access Control Requirements for Web 2.0 Security and Privacy.

Gelhaar, J., Groß, T., & Otto, B. (2021). A Taxonomy for Data Ecosystems. Hawaii International Conference on System Sciences.

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. Information & Management, 53(8), 1049–1064. https://doi.org/10.1016/j.im.2016.07.004

Gupta, M., Patwa, F., & Sandhu, R. (2018). An Attribute-Based Access Control Model for Secure Big Data Processing in Hadoop Ecosystem. Proceedings of the Third ACM Workshop on Attribute-Based Access Control, 13–24. https://doi.org/10.1145/3180457.3180463

Hartman, T., Kennedy, H., Steedman, R., & Jones, R. (2020). Public perceptions of good data management: Findings from a UK-based survey. Big Data & Society, 7(1), 2053951720935616. https://doi.org/10.1177/2053951720935616

Ho, C., & Chuang, T.-R. (2019). Governance of Communal Data Sharing. Institute of Network Cultures. http://ir.sinica.edu.tw/handle/201000000A/55961

Hummel, P., Braun, M., & Dabrock, P. (2021). Own Data? Ethical Reflections on Data Ownership. Philosophy & Technology, 34(3), 545–572. https://doi.org/10.1007/s13347-020-00404-9

Jacobides, M. G., Cennamo, C., & Gawer, A. (2018). Towards a theory of ecosystems. Strategic Management Journal, 39(8), 2255–2276. https://doi.org/10.1002/smj.2904

Janssen, H., & Singh, J. (2022). The Data Intermediary (SSRN Scholarly Paper 4070971). https://papers.ssrn.com/abstract=4070971

Kelbert, F., & Pretschner, A. (2012). Towards a policy enforcement infrastructure for distributed usage control. Proceedings of the 17th ACM Symposium on Access Control Models and Technologies, 119–122. https://doi.org/10.1145/2295136.2295159

Knaapi-Junnila, S., Rantanen, M. M., & Koskinen, J. (2022). Are you talking to me? – Calling laypersons in the sphere of data economy ecosystems. Information Technology & People, 35(8), 292–310. https://doi.org/10.1108/ITP-01-2021-0092

Koskinen, J., Knaapi-Junnila, S., Helin, A., Rantanen, M. M., & Hyrynsalmi, S. (2023). Ethical governance model for the data economy ecosystems. Digital Policy, Regulation and Governance, 25(3), 221–235. https://doi.org/10.1108/DPRG-01-2022-0005

Kugler, P., & Plank, T. J. (2022). Coping with the Double-Edged Sword of Data-Sharing in Ecosystems. Technology Innovation Management Review.

Lauf, F., Scheider, S., Bartsch, J., Herrmann, P., Radic, M., Rebbert, M., Nemat, A., Langdon, C. S., Konrad, R., Sunyaev, A., & Meister, S. (2022). Linking Data Sovereignty and Data Economy: Arising Areas of Tension. Wirtschaftsinformatik 2022 Proceedings. https://aisel.aisnet.org/wi2022/it_for_development/it_for_development/19

Lis, D., & Otto, B. (2020). Data Governance in Data Ecosystems – Insights from Organizations.

Loebbecke, C., van Fenema, P. C., & Powell, P. (2016). Managing inter-organizational knowledge sharing. The Journal of Strategic Information Systems, 25(1), 4–14. https://doi.org/10.1016/j.jsis.2015.12.002

Micheli, M., Farrell, E., Carballa, S. B., Posada, S. M., Signorelli, S., & Vespe, M. (2023, August 24). Mapping the landscape of data intermediaries. JRC Publications Repository. https://doi.org/10.2760/261724

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

Morel, V., & Fischer-Hübner, S. (2023). Automating privacy decisions—Where to draw the line?

Nickerson, R. C., Varshney, U., & Muntermann, J. (2013). A method for taxonomy development and its application in information systems. European Journal of Information Systems, 22(3), 336–359. https://doi.org/10.1057/ejis.2012.26

Ojasalo, J., & Miskeljin, P. (2020). PROPOSING A PRELIMINARY CONCEPT FOR A PERSONAL DATA ECOSYSTEM (p. 7459). https://doi.org/10.21125/inted.2020.1992

Otto, B., & Teuscher, S. (2019). International Data Spaces Association—Reference Architecture Model (p. 118). International Data Spaces Association. https://internationaldataspaces.org/wp-content/uploads/IDS-Reference-Architecture-Model-3.0-2019.pdf

Patton, M. (1999). Enhancing the quality and credibility of qualitative analysis. In BMC Health Serv Res (Vol. 34, p. 1208).

Regulation (EU) 2022/868, 152 OJ L (2022). http://data.europa.eu/eli/reg/2022/868/oj/eng

Runeson, P., & Höst, M. (2009). H??st, M.: Guidelines for Conducting and Reporting Case Study Research in Software Engineering. Empirical Software Engineering 14, 131-164. Empirical Software Engineering, 14, 131–164. https://doi.org/10.1007/s10664-008-9102-8

Sadowski, J., Viljoen, S., & Whittaker, M. (2021). Everyone should decide how their digital data are used—Not just tech companies. Nature, 595(7866), 169–171. https://doi.org/10.1038/d41586-021-01812-3

Scheider, S., Lauf, F., Möller, F., & Otto, B. (2023). A Reference System Architecture with Data Sovereignty for Human-Centric Data Ecosystems. Business & Information Systems Engineering, 65(5), 577–595. https://doi.org/10.1007/s12599-023-00816-9

S. Oliveira, M. I., Barros Lima, G. de F., & Farias Lóscio, B. (2019). Investigations into Data Ecosystems: A systematic mapping study. Knowledge and Information Systems, 61(2), 589–630. https://doi.org/10.1007/s10115-018-1323-6

Spiekermann, S., Acquisti, A., Böhme, R., & Hui, K.-L. (2015). The challenges of personal data markets and privacy. Electronic Markets, 25(2), 161–167. https://doi.org/10.1007/s12525-015-0191-0

Tomi Dahlberg, & Nokkala, T. (2017). Willigness to share supply chain data in an ecosystem governed platform—An Interview Study. 32nd Bled eConference Humanizing Technology for a Sustainable Society, 32, 619–638. https://doi.org/10.18690/978-961-286-280-0

Van Kleek, M., & OHara, K. (2014). The Future of Social Is Personal: The Potential of the Personal Data Store. In D. Miorandi, V. Maltese, M. Rovatsos, A. Nijholt, & J. Stewart (Eds.), Social Collective Intelligence (pp. 125–158). Springer International Publishing. https://doi.org/10.1007/978-3-319-08681-1_7

Verner, J. M., Sampson, J., Tosic, V., Bakar, N. A. A., & Kitchenham, B. A. (2009). Guidelines for industrially-based multiple case studies in software engineering. 2009 Third International Conference on Research Challenges in Information Science, 313–324.

World Economic Forum. (2022). Advancing Digital Agency: The Power of Data Intermediaries. World Economic Forum. https://www.weforum.org/publications/advancing-digital-agency-the-power-of-data-intermediaries/

Yin, R. (2013). Case study Research and Applications (5th ed.). Sage Publishing.

Zrenner, J., Möller, F. O., 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. https://doi.org/10.1108/JEIM-03-2018-0058

Zuboff, S. (2015). Big other: Surveillance Capitalism and the Prospects of an Information Civilization. Journal of Information Technology, 30(1), 75–89. https://doi.org/10.1057/jit.2015.5