prospettiveinORGANIZZAZIONE
Rivista trimestrale di Organizzazione Aziendale
ISSN: 2465-1753

DATA GOVERNANCE IN DATA ECOSYSTEMS: A RESEARCH NOTE

Pubblicato il21 Maggio 2025 di Kazemargi Niloofar, Leonelli Simona, Spagnoletti Paolo, Ceci Federica, Sinaimeri Blerina e Marchesani Filippo

Categoria: Organizational Change and Project Management

In an era where data drives corporate growth, strong governance frameworks are needed to manage data cooperation complexity and hazards. Distributed governance encourages creativity and reduces data cooperation hazards; however, it can suffer of scalability and sustainability. The literature is reviewed to ask crucial questions regarding data governance in such circumstances and offer a targeted research strategy.

Introduction

As data has become one of the most critical assets for securing a competitive advantage in today's business practices, organizations are increasingly integrating different data sources and technologies to obtain better insights (Davenport, 2006). This integration has given rise to data ecosystems where various "actors interact and collaborate to find, archive, publish, consume, or reuse data, fostering innovation, creating value, and

supporting new businesses" (Oliveira et al., 2019). Despite these benefits, many data ecosystems fail to scale up or sustain over time (Jacobides et al., 2024). This highlights the crucial role of data governance and prompts essential questions regarding interorganizational coordination to optimize value creation while mitigating associated risks (Spagnoletti et al., 2024). The literature explores data governance from different perspectives. For instance, some studies delineate various decision domains and objectives (Abraham et al., 2019; Khatri & Brown, 2010), emphasizing the need to align data and IT structures. They also highlight the importance of elevating data governance from an organizational level to an ecosystem level (Abraham et al., 2019). Parmiggiani and Grisot (2020) advocate for a "*practice view*" of data governance, focusing on activities and decisions at the practice level. In this research note, we synthesize key aspects of data governance identified in previous studies and identify research gaps. This allows us to outline a new research agenda aimed at deepening our understanding of the interplay between data governance and data collaboration.

The structure of this paper is as follow. First, we present an overview of the two principal research streams in data governance, highlighting significant insights and discussing perceived gaps in each stream. Second, drawing from our literature review, we pinpoint three main focus areas that could guide future research. We conclude the paper by suggesting possible directions for further investigation.

Data Governance

Governance is allocating decision rights and designing mechanisms to follow those objectives within an organization (Tiwana et al., 2013). The growing body of literature on data governance can be categorized into two main research streams (Paparova et al., 2023; Parmiggiani & Grisot, 2020).

In the first research stream, data governance is generally referred to the protocols, rules, roles, and definitions that elicit an organization's norms or desirable behavior. Khatri and Brown (2010) through a framework derived from IT governance dives systematically into data principles, metadata and access control/fine-grain security models, and privacy/personal information protection services. They underline that data governance ensures the alignment of data operations with business strategy. More recently, Abraham et al. (2019) proposed a conceptual framework that focuses on various data sources, data decisions, and governance mechanisms. However, their approach primarily addresses data governance within the context of IT governance, concentrating on the organizational level. By focusing exclusively on organizational objectives, this framework leaves significant gaps in understanding data governance, especially when data is shared across multiple organizations, and in assessing the broader societal implications of data governance decisions. Therefore, a limitation of this research stream is its reliance on a top-down approach to data governance. This approach offers limited insights, particularly in decentralized settings where no central authority enforces top-down control, and coordination is necessary to standardize new practices (Spagnoletti et al, 2024). Only recently have further studies begun to highlight system-level pathways, focusing on the interaction between institutional functions and ecosystem actors (Abraham et al., 2019; Scholz et al., 2022).

Some scholars have also emphasized the need to shift the focus from the organizational to the ecosystem level (Davidson et al., 2023), especially as organizations increasingly capture and use data across their boundaries. In the second research stream, Information Systems scholars view data like digital artifacts with peculiar characteristics. They shift the focus away from viewing data solely as organizational assets, underling instead the distinctive characteristics of digital data (Abbasi et al., 2016; Jones, 2019; Kallinikos et al., 2013). Consequently, they differentiate between data governance and IT governance (Paparova et al., 2023; Parmiggiani & Grisot, 2020). For instance, Parmiggiani and Grisot (2020) describe the importance of bottom-up decisions (rather than top-down) and the role of actors who actually work with data for data governance at the level. Paparova et al. (2023) made a similar distinction showing the impact of the dynamics of data roles and responsibilities (i.e., "vertical" vs. "horizontal") when data are used for specific or different purposes at the inter-organizational levels. Although this research stream provides valuable insights, their focus remains limited to data-related practices of data science and rarely examine other practices around data.

A New Research Agenda for Data Governance in Data Ecosystems

Having described the two main research streams, we now move on to describing three focus areas that we believe are relevant to spark scholarly debate and guide future inquiries into the transformative potential of data governance in data ecosystems. The three focus areas are (i) the temporal dynamics of interorganizational coordination, (ii) the fluid boundaries of data outcomes, and (iii) the emerging data governance practices.

Temporal dynamics of interorganizational coordination

The existing literature emphasizes the importance of data governance at both organizational and inter-organizational levels (Abraham et al., 2019; Davidson et al., 2023). At the organizational level, data analysts, IT staff, and management, to name a few, handle data processing and decision-making (Abbasi et al., 2016). At the interorganizational level, data governance extends its focus beyond that and encompasses external actors, such as data regulators, data providers and data users (Spagnoletti et al., 2024). However, we know little about the coordination dynamics among actors within an ecosystem. How do ecosystem actors align their governance practices to manage data effectively? How can the value and risks of data be balanced within data ecosystems?

Constantinides et al. (2018) underline that the same set of actors can have different incentives, which creates complexity in aligning their activities with goals and strategies. Such diverse interests and goals create ambiguity related to the roles and responsibilities of each organization. This creates challenges for decision-making authorities in ensuring compliance with established rules. Data flows across different organizations, sectors, and countries with diverse regulations and rules (see, for example, exacerbating tensions among actors). Having conflicting goals and interests, each actor's actions might jeopardize the security of others (Vial, 2023). This makes creating joint value a difficult task (Kazemargi et al., 2023; Spagnoletti et al., 2024).

Power relationships also play a significant role, as some actors may have more control over data. Given the varying interests of actors, effective data ecosystem governance must represent such interests and align value propositions with the ecosystem's perceived value. For instance, some data platforms (e.g., Amazon, Facebook, etc.) collect their users' generated data on a massive scale and use it for their own decision-making processes. Platform owners control how data is shared with third parties and nudge things towards creating monopolies, as content producers, advertisers, or technology companies become dependent on these platforms for access. They create a monopolistic scenario by acting as a data selling agent to other actors or firms. Thus, understanding and managing these diverse perspectives and power imbalances are essential for fostering data collaboration and maximizing the value generated from data. Therefore, studying how coordination mechanisms align the actions of ecosystem actors with heterogeneous interests and responsibilities has relevance (Spagnoletti et al., 2024). Unlike other digital ecosystems, there is a fluidity of roles among actors within data ecosystems. Actors might change their roles based on their needs and capabilities.

For example, a data user—initially using data for analysis or decision-making—might later become a data provider by sharing their data with others in the ecosystem. Additionally, in data ecosystems, an actor might choose to participate for a short time and then decide to withdraw from the ecosystem. For example, an organization might engage in the ecosystem to access a dataset for a project and then opt out once the project is completed. In addition, in a longitudinal study, Aaen et al. (2022) show that while data ecosystems grow and engage new actors, their objectives may diverge or conflict with the original objectives, leading to misalignment and consequently ecosystem failure. This shows that the involvement of actors is not static and might change over time (Spagnoletti et al., 2024). Putting all together, we believe that data governance needs to take into account the temporal dynamics of coordination within and between organizations. This means that data ecosystem governance necessitates ongoing adjustments to keep pace with structural and landscape changes.

The fluid boundaries of data outcomes

Information Systems scholars question assumptions about data (Jones, 2019) and underline the unique characteristics of data. Data are referential (Kallinikos et al., 2013; Yoo et al., 2010), non-rival (Krämer, 2020), and create social orders (Beynon-Davies, 2016). Data can be reused multiple times without data being consumed for different purposes (Constantiou & Kallinikos, 2015; Günther et al., 2022; Newell & Marabelli, 2015). Such characteristics of data also render the boundaries of data outcomes (like other digital artifacts) more porous and less stable (Briel & Recker, 2021; Nambisan, 2017). Models and AI systems that provide insights, offer recommendations, make predictions have fluid boundaries as features and value propositions might change (Nambisan, 2017).

Although these characteristics of data provide new opportunities for innovation and collaboration among different actors, they introduce new challenges for data governance. While collected data can be reused for different and unintended data purposes, this can cause privacy issues (Zuboff, 2015). For instance, personal data

can be combined and reused for AI systems (e.g., social scoring) leading not only to invade individual privacy but also to leading to social surveillance (Zuboff, 2015).

The fluid boundaries of data outcomes thus demand a reassessment of data governance by taking into account data reusability to ensure value generation for all engaged actors.

To address negative implications that arise from fluid boundaries of data outcomes, new forms of governance have emerged. These models offer an alternative to centralized data ecosystems by taking into account interests of a broader range of stakeholders (Micheli et al., 2020). These initiatives focus on decentralized data governance and empower data ecosystem' actors to retain control over their data (Möller et al., 2024).

Emerging data governance practices

The current literature distinguishes data governance from IT governance by taking into account the nature of data, and it focuses on data analytics often neglecting the whole spectrum of data-related practices. Therefore, we propose a broader examination of practices to better serve societal needs. The Data Value Chain (DVC) could offer valuable insights. DVC describes how data flows and is utilized to create value from its initial collection to analysis, dissemination, and its ultimate influence on decision-making processes (Watch, 2018).

Data-related practices are distributed throughout the data value chain. This means that every stage, from data creation and collection to storage, processing, analysis, and distribution, involves distinct activities managed by different stakeholders. Each phase requires specialized skills and technologies ensuring that data is effectively utilized and adds value at every step. Data-related practices influence data quality and ultimately decisions derived from data. Data quality is crucial for value creation, requiring organizations to define control points to ensure data reliability. Evaluating data quality necessitates definitions, standards, and rules that must be followed within a data ecosystem.

Also, given that infrastructure maintenance and service procurement practices can potentially influence value creation from data (see, for example, Chengalur-Smith et al., 2010), data governance should shift the focus from organizational efficiency to societal needs. The protection of data within and across information systems is central to effective data governance.

Second, data governance needs to ensure the interoperability of data across different infrastructures. This facilitates seamless data integration, which is crucial for harnessing the potential of diverse data sources and analytical tools, as the management and control of these technologies are distributed across various organizations. Third, ecosystem actors need to comply with and implement regulatory requirements (for instance Data Act) when designing their data services (Davidson et al., 2023). Regulations often require

organizations to demonstrate how they protect user data and handle data in a manner that aligns with legal and ethical standards. For example, cloud service providers and AI tool developers must ensure that their systems are designed to comply with these regulations. Transparency regarding the design and compliance of data services enhances accountability.

Table 1- A new research agenda for governance in data ecosystems

Themes	Short Descriptions	Research Questions
Temporal dynamics of interorganizational coordination	The need for continuous adjustment in data governance strategies	How are tensions manifested in data ecosystems?What antecedents trigger the need for coordination?How does data governance structure influence the alignment of actors' actions? How do coordination mechanisms influence the sustainable growth of data ecosystems?What are the principles by which an organization can align its actions with other organizations? How can inter-organizational data strategies be sustained over time?How do actors contribute to data governance?Which incentive structure ensures the sustainable growth of data ecosystems?How do changes in regulatory policies impact the evolution of coordination mechanisms among organizations in data ecosystems?
The fluid boundaries of data outcomes	The need to address the negative implications posed by data outcomes boundaries	How do data outcomes challenge current data governance frameworks?How do data governance need to address the negative implications posed by fluid boundaries of data outcomes?How should data governance adapt to address AI risks?What implications does data governance have for data reusability?
Emergence of new data governance practices	The need of new data governance practices and their impact	How do data-related practices along the data value chain influence data quality? And to what extent?How do infrastructure maintenance and service procurement practices can be optimized to better serve societal needs?How do decisions about data infrastructures influence value generation?

Conclusions

In this research note, we emphasize the need for a comprehensive investigation into how actors and societal needs shape data governance, thereby stimulating further research into innovative governance frameworks that foster the responsible and sustainable growth of data ecosystems. We present three sets of non-exhaustive questions that encourage a deeper investigation into the broader implications of data governance.

In the first set, we suggest exploring how the temporal dynamics of interorganizational coordination play a role in shaping the success of data governance frameworks. As actors' roles within ecosystems change, coordination mechanisms must be continuously adjusted to align diverse interests and responsibilities. This is especially important as data flows across organizations, sectors, and even countries, each with its own regulations and governance structures. Future research must investigate how these dynamics can be managed effectively to foster both innovation and security.

In the second set, we suggest that the fluid boundaries of data outcomes present new challenges, particularly in terms of privacy, security, and reusability of data. Governance frameworks need to adapt to address the negative implications of fluid data outcomes, such as privacy breaches or misuse of personal data in AI systems. The emergence of decentralized governance models offers a potential solution by empowering actors to retain control over their data, but further investigation is needed to assess the long-term viability of these approaches.

Finally, emerging data governance practices must expand beyond data analytics to consider the entire spectrum of data-related activities throughout the value chain. This broader examination is crucial for ensuring that data governance frameworks are not only effective within organizations but also responsive to societal needs. As such, future research should focus on how data-related practices influence data quality and value generation, and how infrastructure maintenance and service procurement practices can be optimized for societal benefit.

In conclusion, we call for innovative, responsive, and sustainable governance frameworks that address the complexities of modern data ecosystems. By focusing on interorganizational coordination, fluid data outcomes, and evolving governance practices, future research can significantly contribute to the development of frameworks that balance innovation with responsibility, thereby ensuring the long-term success of data ecosystems.

References

Aaen, J., Nielsen, J. A., & Carugati, A. (2022). The dark side of data ecosystems: A longitudinal study of the DAMD project. European Journal of Information Systems, 31(3), 288-312.

Abbasi, A., Sarker, S., & Chiang, R. H. (2016). Big data research in information systems: Toward an inclusive research agenda. Journal of the association for information systems, 17(2), 3.

Abraham, R., Schneider, J., & Vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. International journal of information management, 49, 424-438.

Beynon-Davies, P. (2016). Instituting facts: Data structures and institutional order. Information and Organization, 26(1-2), 28-44.

von Briel, F., Selander, L., Hukal, P., Lehmann, J., Rothe, H., Fürstenau, D., ... & Wurm, B. (2021). Researching digital entrepreneurship: Current issues and suggestions for future directions. Communications of the Association for Information Systems, 48, 284-304.

Chengalur-Smith, I., Nevo, S., & Demertzoglou, P. (2010). An empirical analysis of the business value of open source infrastructure technologies. Journal of the Association for Information Systems, 11(11), 3.

Constantinides, P., Henfridsson, O., & Parker, G. G. (2018). Introduction—platforms and infrastructures in the digital age. Information systems research, 29(2), 381-400.

Constantiou, I. D., & Kallinikos, J. (2015). New games, new rules: big data and the changing context of strategy. Journal of Information Technology, 30(1), 44-57.

Davenport, T. H. (2006). Competing on analytics. Harvard business review, 84(1), 98.

Davidson, E., Wessel, L., Winter, J. S., & Winter, S. (2023). Future directions for scholarship on data governance, digital innovation, and grand challenges. Information and Organization, 33(1), 100454.

Günther, W. A., Mehrizi, M. H. R., Huysman, M., Deken, F., & Feldberg, F. (2022). Resourcing with data: Unpacking the process of creating data-driven value propositions. The Journal of Strategic Information Systems, 31(4), 101744.

Günther, W. A., Mehrizi, M. H. R., Huysman, M., Deken, F., & Feldberg, F. (2022). Resourcing with data: Unpacking the process of creating data-driven value propositions. The Journal of Strategic Information Systems, 31(4), 101744.

Jacobides, M. G., Cennamo, C., & Gawer, A. (2024). Externalities and complementarities in platforms and ecosystems: From structural solutions to endogenous failures. Research Policy, 53(1), 104906.

Jones, M. (2019). What we talk about when we talk about (big) data. The Journal of Strategic Information Systems, 28(1), 3-16.

Kallinikos, J., Aaltonen, A., & Marton, A. (2013). The ambivalent ontology of digital artifacts. MIS quarterly, 357-370.

Kazemargi, N., Spagnoletti, P., Constantinides, P., & Prencipe, A. (2023). Data control coordination in cloud-based ecosystems: the EU GAIA-X ecosystem. In Research Handbook on Digital Strategy (pp. 289-307). Edward Elgar Publishing.

Khatri, V., & Brown, C. V. (2010). Designing data governance. Communications of the ACM, 53(1), 148-152.

Krämer, J. (2021). Personal data portability in the platform economy: Economic implications and policy recommendations. Journal of Competition Law & Economics, 17(2), 263-308.

Micheli, M., Ponti, M., Craglia, M., & Berti Suman, A. (2020). Emerging models of data governance in the age of datafication. Big Data & Society, 7(2), 2053951720948087.

Möller, F., Jussen, I., Springer, V., Gieß, A., Schweihoff, J. C., Gelhaar, J., ... & Otto, B. (2024). Industrial data ecosystems and data spaces. Electronic Markets, 34(1), 41.

Nambisan, S. (2017). Digital entrepreneurship: Toward a digital technology perspective of entrepreneurship. Entrepreneurship theory and practice, 41(6), 1029-1055.

Newell, S., & Marabelli, M. (2015). Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of 'datification'. The Journal of Strategic Information Systems, 24(1), 3-14.

Oliveira, M. I., Barros Lima, G. D. F., & Farias Lóscio, B. (2019). Investigations into data ecosystems: a systematic mapping study. Knowledge and information systems, 61, 589-630.

Paparova, D., Aanestad, M., Vassilakopoulou, P., & Bahus, M. K. (2023). Data governance spaces: the case of a national digital service for personal health data. Information and Organization, 33(1), 100451.

Parmiggiani, E., and Grisot, M. (2020). Data Curation as Governance Practice. Scandinavian Journal of Information Systems, 32(1), 1–38.

Scholz, N., Wieland, J., and Schäffer, T. (2022). Towards a Framework for Enterprise & Platform Ecosystem Data. AMCIS 2022 Proceedings.

Spagnoletti, P., Kazemargi, N., Constantinides, P., and Prencipe, A. (2024). Data Control Coordination in the Formation of Ecosystems in Highly Regulated Sectors. Journal of the Association for Information Systems,

Preprints. 167.

Tiwana, A., Konsynski, B., & Venkatraman, N. (2013). Information technology and organizational governance: The IT governance cube. Journal of management information systems, 30(3), 7-12.

Vial, G. (2023). Data governance and digital innovation: a translational account of practitioner issues for IS research. Information and Organization, 33(1), 100450.

Watch, O. D. 2018. The Data Value Chain: Moving from Production to Impact. Data2X. Https://Opendatawatch. Com/Publications/the-Data-Value-Chain-Moving-from-pro Duction-to-Impact.

Winter, J. S., and Davidson, E. (2022). Harmonizing Regulatory Regimes for the Governance of Patient-Generated Health Data. Telecommunications Policy, 46(5), Elsevier Ltd, p. 102285.

Yoo, Y., Henfridsson, O., & Lyytinen, K. (2010). Research commentary—the new organizing logic of digital innovation: an agenda for information systems research. Information systems research, 21(4), 724-735.

Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. Journal of information technology, 30(1), 75-89.