Winter 10-14-2023

Data Ecosystem Governance: A Conceptual Framework

Niloofar Kazemargi
D'Annunzio University of Chieti-Pescara, niloofar.kazemargi@unich.it

Federica Ceci
D'Annunzio University of Chieti-Pescara, f.ceci@unich.it

Data Ecosystem Governance: A Conceptual Framework

Niloofar Kazemargi [1]*[0000-0001-6342-9558] and Federica Ceci[1] [0000-0002-6998-8534]

[1]D'Annunzio University of Chieti, Pescara, Italy

niloofar.kazemargi@unich.it

Abstract. Data need to be created, collected, stored, exchanged, integrated, and processed among a diverse set of actors and infrastructures to create value. Such reliance on other actors leads to the emergence of data ecosystems. Despite the focus on data ecosystems in the literature, little is known about who governs what data activities and how. Data ecosystem governance aims to ensure the alignment of activities with different goals and strategies of ecosystem actors. We contribute to the understanding of data governance by expanding the conceptual model for data ecosystem governance. The framework draws on an extensive review of data governance and ecosystems. We show governance is multi-layer, multi-actor and multi- dimension which creates complexity and interdependencies. The conceptual framework provides a guide for managers to fully understand and implement data ecosystem governance.

Keywords: Data Ecosystem, Governance, Data Sources, Data Activities, Actors, Governance Mechanisms

1 Introduction

In today's business environment, data become a key resource for organizations [1] to make data-driven decisions, improve their business processes, develop new products and services [2], and introduce new business models [3]. Several organizations currently focus on data-driven innovation. For data-driven innovation, organizations require combining different data sources and types. Such interdependencies among a diverse set of actors who produce, provide, use data and data-related resources [4] lead to lead to the emergence of data ecosystems. Data ecosystems are defined as *"socio-technical complex networks in which actors interact and collaborate with each other to find, archive, publish, consume, or reuse data as well as to foster innovation, create value, and support new businesses*"[5].

Along with the growing importance of data, we have witnessed tremendous growth in the formation of data ecosystems to facilitate data-driven innovation. The current data ecosystems have different structures from keystone-centric ecosystems to platform-centric, marketplace- based and decentralized [5, 6]. The primary examples of data ecosystems such as Facebook, TripAdvisor, etc. are governed by a keystone actor - e.g. a few technology/service providers [7]. More recently, data ecosystems with (semi) symmetric power relations between different actors have started to emerge where actors such as governmental institutions and/or an alliance of organizations engage in governing data ecosystems [8].

While governance is needed to orchestrate ecosystem actors [5], little is known about who governs what data-related activities and how. Governance focuses on the allocation of decision rights and the deployment of mechanisms [9]. In data ecosystems, governance entails the allocation of decision rights and accountabilities to ensure the alignment of data-related activities with organizational goals [10]. Data ecosystem governance is complex as different actors engage in providing and using data [5], and influences actors' behavior [11] and consequently data-driven innovation. Further, as data ecosystem actors have different interests and expectations, the allocation of decision rights among actors and enforcement of governance mechanisms become crucial [12] to encourage desirable data-related activities within a data ecosystem.

The studies -see for example [12]- distinguish between data governance at organizational level and call for more studies on data governance at ecosystem level. However, the previous works approach data as organizational assets. Only recently, scholars have started to take into account the characteristics of data and the role of stakeholders in data decisions [13, 14]. In line with this research discourse, in this study, we seek to expand the conceptual model for data ecosystem governance. This study contributes to ongoing research on data governance [12, 15]. Most importantly, we show data ecosystem governance is multi-layer, multi-actor and multi-dimension which creates complexity and interdependencies. On the basis of theoretical analysis, we formulate three propositions concerning the interactions of different constructs of the framework.

In what follows, we present an overview of the data ecosystem and data governance literature and present a conceptual framework for the governance of data ecosystems, and we developed some testable propositions, and close with a conclusion in the final section.

2 Related work

2.1 Data ecosystems

Drawing on the ecosystem structural view [16], an ecosystem entails a set of actors with multilateral relationships whose activities and complementarities contribute to value creation. The multilateral aspect of an ecosystem underlines that the relationships among actors are more complex and not a collection of bilateral relations among ecosystem actors. In an ecosystem, actors rely on other actors' components, assets, or activities. Data ecosystems encompass a set of actors that "directly or indirectly consume, produce or provide data and other related resource" [4]. Data ecosystems create value and therefore are sustainable when data produced or provided by some actors are used and consumed by other ecosystem actors [17].

The most prominent example of data ecosystems is represented by the ones generated by the large digital platforms that generate, collect, and access a large amount of data generated by users. The platform owners are the ones who make decisions about how and to what extent share data generated by users with other actors leading to a monopoly position of platforms, while other actors become increasingly dependent on them to access data [18]. In the last years, however, some efforts to reduce data access barriers are pursued to avoid such monopolistic behaviors and foster data-driven innovation focusing on technical and legal aspects [18], and some went even further in designing infrastructure [19].

Therefore, besides the large digital platform ecosystems, we observe the emergence of data ecosystems with different structures and characteristics. For instance, data ecosystems can have different degrees of openness. In open data ecosystems, data can be accessed with no constraint with the aim of enhancing transparency and supporting decision-making. In contrast, in closed data ecosystems, not all can access data unless they get permission from the platform owner/s [6, 17]. Moreover, data ecosystems can rely on proprietary infrastructure managed and owned by an actor -for instance, social media platforms [1]- or on distributed infrastructure to store, process and exchange data [20].

Previous studies have underlined that the existing research mainly focuses on intra-organizational data governance [15], and thus further research is needed on governance within ecosystems [21], in particular in data ecosystems [5].

2.2 Data Governance

Governance focuses on the allocation of decision rights and the deployment of mechanisms to ensure the alignment of activities with organizational goals [9]. Previous literature argues the importance of coordination and orchestration among ecosystem actors [22].

One research stream on data governance views data like other organizational assets to ensure the quality and value creation from data. One of the primitive frameworks for data governance was introduced by Khatri & Brown [23] pointing out the main decisions including data principles, metadata, data access, data quality and data lifecycle. Building on IT governance, in this research stream, the focus of governance was mainly on data like other IT artifacts that must be governed and aligned with the organization's strategy. Viewing data as an organizational asset comes with the next questions about roles and responsibilities within an organization [12, 23] to ensure value creation from data. The scope of governance refers to those who are accountable to ensure that data-related activities are aligned with goals and objectives [9, 24].

The existing literature on data governance has mainly focused on the organizational level [12]: how an enterprise handles and uses its organizational assets. However, recently, scholars see a need for shifting the focus from the organizational to the ecosystem level as organizations increasingly rely on sharing and accessing data both within and outside the organizational boundaries [12, 25]. Patterns of governance are associated with how to ensure desirable data-related activities and outcomes by applying data governance mechanisms. Recent studies based on literature review distinguished structural, procedural and relational mechanisms for controlling how data are used and treated within an organization or among organizations [12, 15, 25]. Only recently, Micheli et al. [8] empirically investigated and compared governance mechanisms among different data ecosystems.

The other research stream has conceptualized data governance by emphasizing the unique characteristics of data as digital artifacts. Data like other digital artifacts have peculiar characteristics: they are editable, programmable, and boundaryless [26]. Data can be combined from different sources and/or aggregated over an extended period of time to augment their value. Data become diffused as data flow [27]. Considering organizational boundaries, data can inflow or outflow within and across organizational boundaries. Some data can be acquired or accessed freely [28]. The value of data depends on not only the quality of data but also how they are used and in which context. Data in nature are non-rival which means data can be collected, shared and used by different actors to create value (without data to be consumed) [29].

Taking into consideration the characteristics of data, several studies expanded our understanding beyond viewing data governance as IT governance. Parmiggiani and Grisot [13] describe the importance of bottom-up decisions and the role of actors who work with data for data governance. Paparova et al. [14] illustrate the differences between data governance and IT governance by showing roles, responsibilities and interests of multiple actors engaged in data collection and use shape data governance.

3 Research Method

The goal of the study is to develop a conceptual model for data ecosystem governance. To develop a conceptual model, we conducted a literature review [30] of research on data governance and ecosystems. To identify relevant literature, we used AIS Electronic Library database for the searching phase. For key-based search, we used “data ecosystem*” as a keyword. Our analysis focused on governance in data ecosystems rather than data/information governance to ensure that our literature review provides new insights and expands the work of Abraham et al. [12] and Scholz et al. [15].

We conducted key-based search in September 2022, which resulted in 143 articles. We also included peer-reviewed conference papers. In qualitative assessment, based on their titles and abstracts, we excluded articles that did not focus explicitly or implicitly on data governance or decisions about data management in data ecosystems. We also performed a forward and backward search to identify other relevant articles. In total, we reached 33 articles addressing data ecosystem governance. In an iterative process, we used open and selective coding to gain new insights from the selected articles. We abstracted codes in four layers: data layers, multi-actor, multi-dimension, and governance mechanisms as the main constructs of our conceptual framework.

4 Framework for data ecosystem governance

With the formation and emergence of new data ecosystems, previous studies illustrate challenges in data governance regarding data rights, ownership, coordination, and incentive systems [31, 32]. While previous research streams provide valuable understanding about data governance, hitherto, there has not been yet any holistic framework for data ecosystem governance. Building on those research streams, we developed a framework (see Figure 1) for data ecosystem governance that is multi-layer, multi-actor, and multi-dimension relying on different governance mechanisms to create value from data as discussed below.

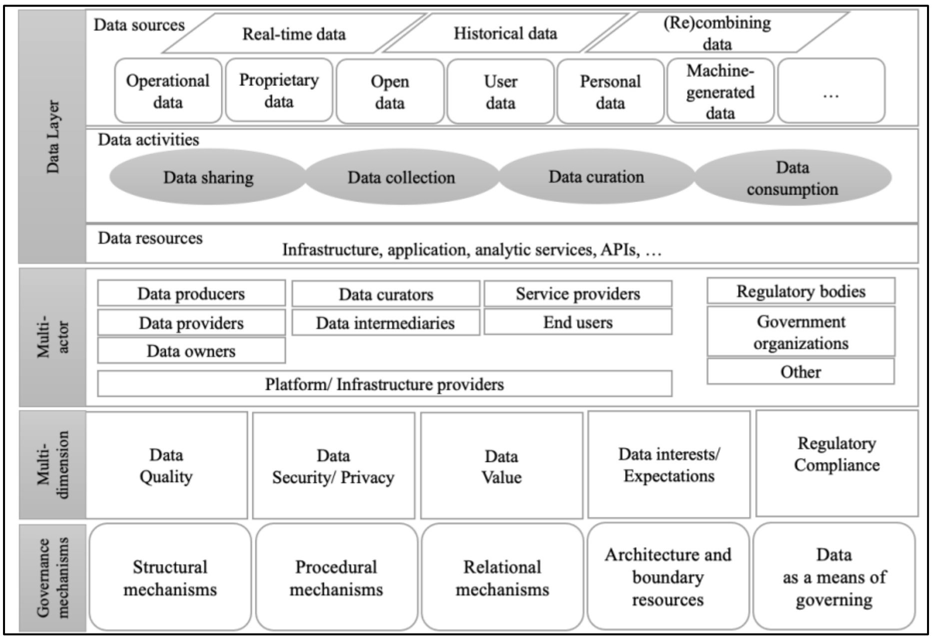

Fig. 1. A conceptual framework for data ecosystem governance

4.1 Data Layer

Data sources

Organizations produce and use different data sources or types. Data governance is essential to ensure the effective use of a diverse set of data sources. Data sources range from operational data, big data including machine-generated data, user data on social media, open data to personal data. Operational data – data for operations within an organization [33]- has been the main target for governance. At the same time, organizations have a long tradition of sustaining their competitive positions by using and controlling proprietary data [34]. Digitization and digital technologies such as enterprise information systems, mobile devices, social media and IoT expand and expand and offer new data sources for value creation [35]. The recent literature has started to elucidate also big data governance [12] including, machine-generated data [36], user data on social media [15], open data [37, 38] and personal data [39]. In the last years, with the emergence of the open data movement[1], organizations opt to share part of their data to be used and modified freely by other actors [40]. Although sharing open data has some benefits such as transparency, there are some risks and challenges associated with profit-maximizing, reusing data for unknown purposes, privacy and security [41]. Data governance needs to consider open data [42] to ensure the alignment of open data strategy with the overall organizational objectives such as digital strategy [41].

Data-related activities

Considering the characteristics of digital data, organizations need to continuously make decisions around data access, data storage, data analysis for data-driven business models [43]. This expands the scope of data ecosystem governance to include not only data as digital artifacts but the alignment of data-related activities with the overall strategy.

Data-related activities are interrelated. Decisions about data collection, curation and consumption [44, 45] influence data collaboration and consequently data-driven innovation. For instance, Parmiggiani and Grisot [13] show that decisions related to data production and use influence data quality and consequently value generated by data. Basole [45] outlines some data are curated by different actors with different expertise and interests (e.g., crowd).

With the open data movement, data sharing become an increasingly important source of digital data streams for data-driven innovation. While the public sector is increasingly engaged in sharing open data, the private sector also took some initiatives. For example, Lufthansa's open API allows third parties to access and use part of their data[2.] The decisions about whether to reveal or not corporate data as open data [37], and its outcomes [38] underline the importance of alignment of data-sharing activities with a long-term organizational strategy around data. At the organizational level, the aim is to maximize the benefit of open data for an organization while limiting the negative consequences of open data. Decisions about data sharing by one actor have direct implications for other actors in value creation (at the ecosystem level).

Thus, data ecosystem governance should include not only what data to govern but also about the data value chain: how data are created, collected, stored, exchanged, integrated, and processed.

Data-related resources

Data creation, collection, storage exchange, and process require digital technologies such as platforms, mobile technologies, cloud services, analytics, artificial intelligence, etc. A single actor may not necessarily possess all data-related resources (such as software and infrastructure), and thus relies on the resources of other actors. For

---

[1] For instance, see European Parliament. (2003). "Directive 2003/98/EC on the re-use of public sector information." Retrieved from https://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:345:0090:0096:en:PDF (visited on 23/11/2018)

[2] Lufthansa Developer Network. URL: https://developer.lufthansa.com/ (visited on 03/10/2020).

instance, a firm can use data generated by users of its online services by relying on cloud services offered by other actors that provide the infrastructure for data processing and exchange [46]. Data-related activities can rely on proprietary infrastructure managed and owned by an actor -for instance, social media platforms [1]- or on distributed infrastructure to store, process and exchange data [20]. The platform-based organizations that are “data made” [47] such as TripAdvisor and Facebook use Application Program Interfaces (APIs) and software development kits (SDKs) to enable other actors to exploit data and create new applications and services. Open data ecosystems use APIs and SDKs not only to enable entrepreneurs to exploit data but also data suppliers to share swiftly their data [48].

4.2 Multi-actor

Early studies on data governance are primarily concerned with how an organization ensures the quality of data and the value creation process [23]. By emerging new opportunities for value creation from data, organizations seek to benefit from data by combining data within and outside of their organizational boundaries [21]. Unlike traditional organizations, there are increasingly more actors engaged in data-related activities such as data sharing, data collection, data curation and data consumption. This led to an evolution in the literature focusing on the organizational level governance as well as ecosystem level.

In studying big data, van den Broek and van Veenstra [39] outline different governance settings around data at the organizational level. While the market arrangement focuses on dyadic inter-organizational relationships, the other forms of governance such as bazaar, hierarchy and network are at the ecosystem level at which relations among actors cannot be reduced to a sum of bilateral connections. Governance over data-related activities limits behavioral complexity, promotes fair use of data as collective resources, and addresses tensions among actors [39] to facilitate and promote data sharing [10].

Often data ecosystems are around keystone organization/s or public authorities where other actors have no (or limited) decision rights on data access, storage, and use. The platform-based organizations (such as TripAdvisor and Facebook) are “data made”: data are generated by user/crowd engagement. Thus, how different types of data are integrated and used is vital to support service innovation and sustain competitive positions [49]. However, this often results in a power asymmetric between actors who engage in data production and the ones who own, control or govern data ecosystems. Somehow related, Otto and Jarke [19] see data ecosystem governance as the effort to ensure the interests of the key stakeholders who own, provide and control platforms and technology, but also as the efforts to coordinate a diverse set of actors who use and share data. This is consistent with the work by Janssen et al. [17] who argue that sustaining data ecosystems depends on the motivation and engagement of not only data providers but also data users.

Previous studies also discuss data ecosystem governance needs to address also incentive structure [9], how to provide a structure and incentive to promote the participation of (new) actors generating, collecting, using, and exchanging data [21, 50].

It is worth noting that, in data layer, tata, data-related activities, and data-related sources are interrelated. For instance, Enders et al., [41] argue that decisions about a single data source and using standards APIs and protocols for data sharing can reduce the time for software development as less effort and time are needed for data curation.

4.3 Multi-dimension

Data quality

One of the key factors for value creation from data is access and use high quality of data. Organizations need to define and implement control points to ensure or improve the quality of generated, collected, integrated, or curated data whether through crowd or experts [45], especially for open data [51]. Evaluation of the quality of data requires a set of definitions, standards, and rules. While some organizations have adopted the organizational level standards, in data ecosystems, shared standards are needed. However, the challenge is managing different data quality standards [52] for different data and different data use.

Data interests/ expectations

Data governance at the ecosystem level needs to address the interests and expectations of a set of actors engage within a data ecosystem [15, 33]. This challenge amplifies as actors' interests and expectations are different and sometimes conflicting [53]. For sustaining data ecosystems, data governance needs to consider data sovereignty[3] to enable actors to negotiate and control the use of data [7]. At the same time, another challenge for creating a sustainable data ecosystem is to ensure that the offered value proposition is aligned with perceived values from the ecosystem. This underlines the importance of a feedback loop among ecosystem actors [54].

Data value

In the literature, a special emphasis is placed on data-driven business models where data and data-related technologies support value creation and capture [55, 56] at the organizational level. For example, Lang et al. [43] study enablers and challenges in different phases of data-driven business model development for incumbent firms. However, it is difficult to show the relationship between data governance and value creation [12, 54].

What distinguishes data governance at the ecosystem level from the organizational level is the focus on protecting the interests of ecosystem actors. The empirical studies show that the lack of control over data-related activities negatively influences value

[3] Data sovereignty refers to "the complete control over stored and processed data and the decision on who is permitted to have access to it. According to GAIA-X: Driver of digital innovation in Europe (2020).

creation [57], and might act as a bottleneck in the formation and development of data ecosystems. Data governance at the ecosystem level needs to address the interests and expectations of a set of actors engage within a data ecosystem [15, 33]. This challenge amplifies as actors' interests and expectations are different and sometimes conflicting [53].

Data Security/ Privacy

Data-related activities and decisions about those activities are distributed among a diverse set of actors [58]. However, the decisions and data-related activities of one actor are not always in favor of other actors' interests [59]: the way one actor stores or processes data may be considered a security and privacy violation by other actors [60]. In some cases, actors have heterogeneous expectations or even ambiguous goals [61] regarding how specific data-related activities must perform. Enders et al. [37] studied decisions related to data sharing and risks and privacy issues associated with releasing open data. Organizations should carefully analyze competitiveness, data misappropriation, innovation opportunity, legal and privacy issues at the organizational level. Similarly, Lang et al. [43] drawing on the resource-based view suggest four main capabilities for developing data-driven business models: data; technologies; organization and monetization capabilities. The lack of pellucidity of how data are accessed, shared, curated, and processed among different actors increases privacy violations which need a more comprehensive overview of privacy issues at both organizational and ecosystem levels [62].

Regulatory compliance

Data ecosystems are embedded in different industry and regulatory landscapes. When there is lax regulatory enforcement, we witness very low Regulatory compliance and norms- for instance with the EU General Data Protection Regulation (GDPR) [63]. In contrast, in highly regulated sectors, data-related activities are required to be in compliance with not only the national/international regulations, but also with sector-specific norms. For instance, in the financial sector as a highly-regulated sector, data sharing among ecosystem actors require compliance management [64].

4.4 Governance Mechanisms

Patterns of governance are associated with applying formal or informal governance mechanisms to ensure desirable data-related activities and outcomes. Building on information technology governance, in a recent literature review, Abraham et al. [12] and Scholz et al. [15] clustered governance mechanisms into structural, procedural, and relational mechanisms.

Structural, procedural, and relational mechanisms are complementarity and often organizations combine different governance mechanisms to effectively and efficiently control how data are used and treated within an organization or among organizations. Structural governance mechanisms define roles and responsibilities and allocate decision rights [12]. While often the focus is on the roles and responsibilities of platform owners, data ecosystem governance needs to define decision rights for other actors engaging in data-related activities [15], particularly in decentralized settings to address power asymmetric [8].

Organizations use different procedural governance mechanisms to ensure the accuracy of data collection, effective use of data and secure data collection and sharing [65]. A distinction is made here between procedural mechanisms at the organizational and the ecosystem levels. Each organization needs to determine its organizational goals for instance through data strategy, how they meet compliance requirements, and address data-related risks [12]. At the ecosystem level, procedural mechanisms such as policies, standards, data agreements, sanction mechanisms need to be applied to guide and reduce the behavior complexity of a heterogenous set of actors [12, 15, 19, 25].

Relational governance mechanisms encompass increasing awareness and sharing knowledge about data rights and governance among different actors [12, 15]. Establishing trust is another relational governance mechanism. Some data ecosystems are designed based on trust between data providers and platform owners, especially between citizens and public actors [8]. However, relying solely on trust raises a question about the scalability of such data ecosystems since the involvement of other actors is “a voluntary basis” [8].

The architecture of digital infrastructures to exchange data among different actors can be considered a form of governance [66]: for instance, APIs influence what data can be shared with whom and how, and somehow enable/constrain actors’ interactions [11, 48]. Schreieck et al. [67] and Otto and Jarke [19] suggest that data as boundary resources that may facilitate or constrain data exchange and value creation. Thus, a socio-technical perspective must be used to design and govern data ecosystems [50].

Furthermore, some studies point to coordination and control challenges among different actors within data ecosystems. In particular, some studies discuss the importance of the existence of feedback loop to ensure that data ecosystems create value for the main stakeholders [54]. Some data ecosystems have started to monitor data usage[4]. Thus, data can be a means to improve governance or to control/monitor actors’ behaviors.

5 Propositions

5.1 Interaction between actors and data layer

Digitalization and layered-modular architecture have enabled organizations to innovate their business models to create and capture value from data [55]. While some data ecosystems are formed to realize specific value propositions, the objectives might change and evolve as a data ecosystem grows. However, not all data ecosystems are formed with clear objectives (i.e., what benefits they offer to whom).

[4] See for example https://gaia-x.eu/

Creating and capturing value from data require data sources and data-related resources (e.g., cloud infrastructure), skills and capabilities to create, collect, store, exchange, integrate, and process data [36]. With lower entry barriers, more actors opt to engage in data ecosystems taking different roles such as data providers, data users and service providers [5]. Thus, data sources, data-related activities and data-related resources are distributed among different actors who operate autonomously.

An actor's decisions about how to capture data and exploit which data, how to manage data-related resources influence the realization of the overall value proposition. Coordination is needed to understand how each actor create, collect, store, exchange, integrate, and process which data and by using which data-related resources. This creates behavioral complexity as "*behavior is difficult to predict or control*"[68]. Another complexity refers to the fact that some actors play several roles simultaneously within a data ecosystem. Actors also can opt to leave the ecosystem at any time. This means that the configuration of data ecosystems is dynamic due to fluid boundaries.

Proposition 1: In data ecosystems, a shared view of roles and responsibilities reduces behavioral complexity.

5.2 Interaction between actors and mechanisms

In the primary examples of data ecosystems, the keystone actor/s play an important role in implementing governance mechanisms to control the behavior of actors [69]. However, the emerging data ecosystems are more decentralized. In decentralized data ecosystems, different actors engage from the early phase in designing and implementing data ecosystems. Therefore, they engage in (re)negotiating, making trade-offs, defining, designing, promoting, and agreeing on governance mechanisms. Examples of governance mechanisms are policies, standards, procedures, contracts, and measurements to control and monitor the behaviors of a diverse set of actors.

Proposition 2: In data ecosystems, governance mechanisms are designed and implemented through a set of negotiations, agreements, and collective actions.

5.3 Interaction between actors, data layer and mechanisms

The data ecosystems are dynamic and evolve over time. Changes can be associated with ecosystem configuration. Data and data-based services offered by a data ecosystem can attract more actors playing new roles and responsibilities. Changes can also be associated with contextual factors such as regulatory and industry landscape. In the EU, for instance, new regulations and norms are enforced by regulators to ensure fair access and use of data. Changes can be also associated with technological trends such as in data infrastructure. These changes bring new tensions and disputes among different actors in handling data.

Proposition 3: Data ecosystem governance needs to continuously address tensions and disputes among actors due to configurational, contextual, and technological changes.

6 Conclusion

Our conceptual framework serves as a basis for both practitioners and academia to ensure that data ecosystems create value from data. For research implication, this study contributes to ongoing research on data governance [12, 15] by developing a conceptual framework. In this study, we integrated related but scattered literature on data ecosystems and research streams on data. Our conceptual model extends the frameworks introduced by Abraham et al. [12] and Scholz et al. [15], and provides a more comprehensive understanding of data ecosystem governance. We show that data ecosystem governance is multi-layer, multi-actor, and multi-dimension which relies on different governance mechanisms. More specifically, by focusing on the characteristics of data as digital artifacts, we show the importance of activities and actors for data governance [13]. On the basis of theoretical analysis, we also formulate three propositions concerning the interactions of different constructs of the framework. For practical implication, our conceptual framework provides a guide for managers to fully under-stand and implement data ecosystem governance.

The current focus is on data sources such as traditional, big data and open data. However, future research should empirically study data-related activities and their interdependencies. Considering the evolution of data ecosystems over time, data-related activities are increasingly distributed among a diverse set of actors [58]. Thus, to ensure value creation from data, data ecosystem governance should focus on (continuous) alignment of activities with the overall strategy [70]. Moreover, data ecosystem governance includes incentive structures to promote participation and membership of new actors: these could be both monetary and intrinsic motivations [4, 71]. Although studying both the organizational and ecosystem levels might not be feasible in one study, future study can explore the relation and synergies among the levels.

In terms of governance mechanisms, the literature reveals a focus on structural, procedural, and relational governance mechanisms. Thus, one promising avenue for future research is to study how the architecture of infrastructures and data enable new forms of governance, how such new forms of governance complement the existing governance mechanisms, which governance mechanisms are more effective for what types of data ecosystems and under which conditions. Finally, we developed some propositions to analyze the interactions between different concepts and showed complexity in designing and implementing governance.

Acknowledgements

We acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 1409 published on 14.9.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union – NextGenerationEU– Project Title Data4Innovation- Data ecosystem governance toward enhancing data sharing for innovation: implications for organizations P2022HXLBF CUP D53D23017780001–Grant Assignment Decree No. 1376 adopted on 01/09/2023 by the Italian Ministry of Ministry of University and Research (MUR).

This study was also funded by European Union – NextGenerationEU, program" MUR-Fondo Promozione e Sviluppo - DM 737/2021, Project Title "In.Da.Eco." Innovation in Data Ecosystems.

References

1. Alaimo, C., Kallinikos, J., Valderrama, E.: Platforms as service ecosystems: Lessons from social media. J. Inf. Technol. (2019). https://doi.org/10.1177/0268396219881462.
2. Wixom, B.H., Ross, J.W.: How to monetize your data. MIT Sloan Manag. Rev. 58, (2017).
3. Günther, W.A., Rezazade Mehrizi, M.H., Huysman, M., Feldberg, F.: Debating big data: A literature review on realizing value from big data. J. Strateg. Inf. Syst. 26, 191–209 (2017). https://doi.org/10.1016/j.jsis.2017.07.003.
4. Oliveira, M.I.S., Lóscio, B.F.: What is a data ecosystem? ACM Int. Conf. Proceeding Ser. (2018). https://doi.org/10.1145/3209281.3209335.
5. Oliveira, M.I.S., Lóscio, B.F., Lima, G. de F.B.: Investigations into Data Ecosystems: a systematic mapping study. Springer London (2019). https://doi.org/10.1007/s10115-018-1323-6.
6. Gelhaar, J., Groß, T., Otto, B.: A taxonomy for data ecosystems. Proc. Annu. Hawaii Int. Conf. Syst. Sci. 2020-Janua, 6113–6122 (2021). https://doi.org/10.24251/hicss.2021.739.
7. Jarke, M., Otto, B., Ram, S.: Data Sovereignty and Data Space Ecosystems. Bus. Inf. Syst. Eng. 61, 549–550 (2019). https://doi.org/10.1007/s12599-019-00614-2.
8. Micheli, M., Ponti, M., Craglia, M., Berti Suman, A.: Emerging models of data governance in the age of datafication. Big Data Soc. 7, (2020). https://doi.org/10.1177/2053951720948087.
9. Tiwana, A., Konsynski, B., Venkatraman, N.: Information technology and organizational governance: The IT governance cube. J. Manag. Inf. Syst. 30, 7–12 (2013).
10. Lis, D., Otto, B.: Towards a taxonomy of ecosystem data governance. Proc. Annu. Hawaii Int. Conf. Syst. Sci. 2020-Janua, 6067–6076 (2021). https://doi.org/10.24251/hicss.2021.733.
11. Eaton, B., Elaluf-calderwood, S., Sørensen, C., Yoo, Y.: Distributed tuning of boundary resources : the case of Apple ' s iOS service system. MIS Q. 39, 217–243 (2015).
12. Abraham, R., Schneider, J., vom Brocke, J.: Data governance: A conceptual framework, structured review, and research agenda. Int. J. Inf. Manage. 49, 424–438 (2019). https://doi.org/10.1016/j.ijinfomgt.2019.07.008.

13. Parmiggiani, E., Grisot, M.: Data Curation as Governance Practice. Scand. J. Inf. Syst. 32, 1–38 (2020).
14. Paparova, D., Vassilakopoulou, P., Aanestad, M., Bahus, M.K.: Data Governance Spaces : the Case of a National Digital Service for Personal Health Data. Inf. Organ. 33, 1–41 (2023). https://doi.org/10.1016/j.infoandorg.2023.100451.
15. Scholz, N., Wieland, J., Schäffer, T.: Towards a Framework for Enterprise & Platform Ecosystem Data. In: AMCIS 2022 Proceedings (2022).
16. Adner, R.: Ecosystem as Structure: An Actionable Construct for Strategy. J. Manage. 43, 39–58 (2017).
17. Janssen, M., Charalabidis, Y., Zuiderwijk, A.: Benefits, Adoption Barriers and Myths of Open Data and Open Government. Inf. Syst. Manag. 29, 258–268 (2012).
18. Cabral, L., Haucap, J., Parker, G., Petropoulows, G., Valletti, T., Van Alstyne, M.: The EU Digital Markets Act A Report from a Panel of Economic Experts. (2021). https://doi.org/10.2760/139337.
19. Otto, B., Jarke, M.: Designing a multi-sided data platform: findings from the International Data Spaces case. Electron. Mark. 29, 561–580 (2019). https://doi.org/10.1007/s12525-019-00362-x.
20. Gelhaar, J., Otto, B.: Challenges in the emergence of data ecosystems. Proc. 24th Pacific Asia Conf. Inf. Syst. (2020).
21. De Prieëlle, F., De Reuver, M., Rezaei, J.: The role of ecosystem data governance in adoption of data platforms by Internet-of-Things data providers: Case of Dutch horticulture industry. IEEE Trans. Eng. Manag. (2020).
22. Breidbach, C., Choi, S., Ellway, B., Keating, B.W., Kormusheva, K., Kowalkowski, C., Lim, C., Maglio, P.: Operating without Operations: How Is Technology Changing the Role of the Firm? J. Serv. Manag. 29, 809–833 (2018).
23. Khatri, V., Brown, C. V.: Designing data governance. Commun. ACM. 53, 148–152 (2010). https://doi.org/10.1145/1629175.1629210.
24. Gregory, R.W., Kaganer, E., Henfridsson, O., Ruch, T.J.: IT Consumerization and the Transformation of IT Governance. 42, 1225–1253 (2018). https://doi.org/10.25300/MISQ/2018/13703.
25. Jagals, M., Karger, E.: Inter-organisational data governance: A literature review. ECIS 2021 Res. Pap. 1–19 (2021).
26. Kallinikos, J., Aaltonen, A., Marton, A.: The Ambivalent Ontology of Digital Artifacts. MIS Q. 37, 357–370 (2013).
27. Vassilakopoulou, P., Skorve, E., Aanestad, M.: Enabling openness of valuable information resources: Curbing data subtractability and exclusion. Inf. Syst. J. 29, 768–786 (2019). https://doi.org/10.1111/isj.12191.
28. GroupOGW: Eight principles ofopen government data, Open GovernmentWorking Group. (2007).
29. Krämer, J.: Personal Data Portability In The Platform Economy: Economic Implications And Policy Recommendations. J. Compet. Law Econ. 17, 263–308 (2020). https://doi.org/10.1093/joclec/nhaa030.

30. Webster, J., Watson, R.T.: Analyzing the past to prepare for the future: Writing a literature review. MIS Q. xiii–xxiii (2002).
31. Lis, D., Otto, B.: Data Governance in Data Ecosystems – Insights from Organizations. In: AMCIS 2020 Proceedings. (2020).
32. Susha, I., Janssen, M., Verhulst, S.: Data collaboratives as a new frontier of cross-sector partnerships in the age of open data: Taxonomy development. Proc. Annu. Hawaii Int. Conf. Syst. Sci. 2017-Janua, 2691–2700 (2017). https://doi.org/10.24251/hicss.2017.325.
33. Lee, S.U., Zhu, L., Jeffery, R.: Data Governance for Platform Ecosystems: Critical Factors and the State of Practice. In: Pacific Asia Conference on Information Systems (2017).
34. Markus, M.L.: Datification, Organizational Strategy, and IS Research: What's the Score? J. Strateg. Inf. Syst. 26, 233–241 (2017). https://doi.org/10.1016/j.jsis.2017.08.003.
35. Zuboff, S.: Big other: surveillance capitalism and the prospects of an information civilization. J. Inf. Technol. 30, 75–89 (2015).
36. Pigni, F., Piccoli, G., Watson, R.: Digital Data Streams: Creating value from the real-time flow of big data. Calif. Manage. Rev. 58, 5–25 (2016).
37. Enders, T., Wolff, C., Satzger, G.: Knowing What to Share: Selective Revealing in Open Data. ECIS 2020 - Proc. 28th Eur. Conf. Inf. Syst. (2020).
38. Enders, T., Satzger, G., Wolff, C., Enders, T., Wolff, C.: Why Should I Share ? Exploring Benefits of Open Data for Private Sector Organizations. In: Pacific Asia Conference on Information Systems (PACIS) (2022).
39. van den Broek, T., van Veenstra, A.F.: Governance of big data collaborations: How to balance regulatory compliance and disruptive innovation. Technol. Forecast. Soc. Change. 129, 330–338 (2018). https://doi.org/10.1016/j.techfore.2017.09.040.
40. opendefinition.org.: The Open Definition. Retrieved from https://opendefinition.org/ (2019).
41. Enders, T., Carina, B., Ronny, S., Pamela, L.: How To Implement an Open Data Strategy? Analyzing Organizational Change Processes To Enable Value Creation By Revealing Data. Eur. Conf. Inf. Syst. 1–16 (2020).
42. Zeleti, F.A., Ojo, A.: Open data capability architecture - An interpretive structural modeling approach. Proc. Annu. Hawaii Int. Conf. Syst. Sci. 2020-Janua, 6164–6173 (2020). https://doi.org/10.24251/hicss.2020.754.
43. Lange, H.E., Drews, P., Höft, M.: Realization of Data-Driven Business Models in Incumbent Companies: An Exploratory Study Based on the Resource-Based View. ICIS 2021 Proc. 1–17 (2021).
44. Chua, C., Indulska, M., Lukyanenko, R., Montr, H.E.C., Maass, W., Storey, V.C.: MISQ Research Curation on Data Management Research. 1–12 (2022).
45. Basole, R.C.: Understanding ecosystem data. Proc. Annu. Hawaii Int. Conf. Syst. Sci. 2020-Janua, 5718–5727 (2020). https://doi.org/10.24251/hicss.2020.702.
46. Assunção, M.D., Calheiros, R.N., Bianchi, S., Netto, M.A.S., Buyya, R.: Big Data computing and clouds: Trends and future directions. J. Parallel Distrib.

Comput. 79, 3–15 (2015).
47. Alaimo, C., Kallinikos, J., Valderrama-Venegas, E.: AFTER SOCIAL MEDIA : THE FORMATION OF THE TRIPADVISOR PLATFORM ECOSYSTEM. 1–41.
48. Bonina, C., Eaton, B.: Cultivating open government data platform ecosystems through governance: Lessons from Buenos Aires, Mexico City and Montevideo. Gov. Inf. Q. 37, 101479 (2020). https://doi.org/10.1016/j.giq.2020.101479.
49. Alaimo, C., Kallinikos, J., Valderrama, E.: Platforms as service ecosystems: Lessons from social media. J. Inf. Technol. 35, 25–48 (2020).
50. Heinz, D., Benz, C., Fassnacht, M., Satzger, G.: Past, Present and Future of Data Ecosystems Research: A Systematic Literature Review. PACIS 2022 Proc. (2022).
51. Najafabadi, M.M., Cronemberger, F.A.: Systemic effects of an open government program on data quality: the case of the New York State's Food Protection program area. Transform. Gov. People, Process Policy. (2022). https://doi.org/10.1108/TG-11-2021-0194.
52. Jarvenpaa, S.L., Essén, A.: Data sustainability: Data governance in data infrastructures across technological and human generations. Inf. Organ. 33, (2023). https://doi.org/10.1016/j.infoandorg.2023.100449.
53. Lauf, F., Scheider, S., Bartsch, J.: Linking Data Sovereignty and Data Economy : Arising Areas of Tension. In: Wirtschaftsinformatik 2022 Proceedings (2022).
54. Najafabadi, M.M., Luna-Reyes, L.F.: Open government data ecosystems: A closed-loop perspective. Proc. Annu. Hawaii Int. Conf. Syst. Sci. 2017-Janua, 2711–2720 (2017). https://doi.org/10.24251/hicss.2017.327.
55. Wiener, M., Saunders, C., Marabelli, M.: Big-data business models: A critical literature review and multiperspective research framework. J. Inf. Technol. 35, 66–91 (2020). https://doi.org/10.1177/0268396219896811.
56. Alaimo, C., Kallinikos, J., Aaltonen, A.: Data and value. In: Handbook of Digital Innovation. Edward Elgar Publishing (2020).
57. Lefebvre, A., Schermerhorn, E., Spruit, M.: How research data management can contribute to efficient and reliable science. In: ECIS 2018 Proceedings (2018).
58. Mokobombang, N.N., Gutierrez, J., Petrova, K.: Value-creating Roles Played by the Actors in Open Government Data : A Systematic Literature Review. In: ACIS 2020 Proceedings (2020).
59. Vial, G.: Understanding digital transformation: A review and a research agenda. J. Strateg. Inf. Syst. 28, 118–144 (2019). https://doi.org/10.1016/j.jsis.2019.01.003.
60. Newell, S., Marabelli, M.: Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of "datification." J. Strateg. Inf. Syst. 24, 3–14 (2015). https://doi.org/10.1016/j.jsis.2015.02.001.
61. Benfeldt, O., Persson, J.S., Madsen, S.: Data Governance as a Collective Action

Problem. Inf. Syst. Front. 22, 299–313 (2020). https://doi.org/10.1007/s10796-019-09923-z.
62. Burmeister, F., Kurtz, C., Vogel, P., Drews, P., Schirmer, I., Burmeister, F., Vogel, P., Kurtz, C., Drews, P., Schirmer, I.: Unraveling Privacy Concerns in Complex Data Ecosystems with Architectural Thinking. In: ICIS 2021 Proceedings (2021).
63. Martin, N., Matt, C., Niebel, C., Blind, K.: How Data Protection Regulation Affects Startup Innovation. Inf. Syst. Front. 21, 1307–1324 (2019). https://doi.org/10.1007/s10796-019-09974-2.
64. Dinckol, D., Ozcan, P., Zachariadis, M.: Regulatory Standards and Consequences for Industry Architecture: The Case of UK Open Banking. Res. Policy. 52, 104760 (2023). https://doi.org/10.5465/ambpp.2022.16634abstract.
65. Borgman, H., Heier, H., Bahli, B., Boekamp, T.: Dotting the i and crossing (out) the T in IT governance: New challenges for information governance. In: 2016 49th Hawaii International Conference on System Sciences (HICSS). pp. 4901–4909. IEEE (2016).
66. Constantinides, P., Henfridsson, O., Parker, G.: Platforms and Infrastructures in the Digital Age. Inf. Syst. Res. 7047, 1–20 (2018). https://doi.org/10.1287/isre.2018.0794.
67. Schreieck, M., Wiesche, M., Krcmar, H.: Design and governance of platform ecosystems - Key concepts and issues for future research. 24th Eur. Conf. Inf. Syst. ECIS 2016. (2016).
68. Tiwana, A.: Platform Ecosystems Aligning Architecture, Governance, and Strategy. Newnes (2013).
69. Selander, L., Henfridsson, O., Svahn, F.: Capability search and redeem across digital ecosystems. J. Inf. Technol. 28, 183–197 (2013).
70. Saunders, C., Benlian, A., Henfridsson, O., Wiener, M.: IS Control & Governance. MIS Q. Res. Curations, Ashley Bush Arun Rai, Eds., http//misq.org/research-curations. (2020).
71. Najjar, M.S., Kettinger, W.J.: Data Monetization: Lessons from a Retailer’s Journey. MIS Q. Exec. 12, (2013).