

5-2-2023

## **DATA GOVERNANCE STRATEGIES FOR DATA PLATFORMS – A MULTIPLE CASE STUDY IN NURSING CARE**

Rahel Jana Gubser

*Charité - Universitätsmedizin Berlin, rahel.gubser@charite.de*

Matthias Schulte-Althoff

*Freie Universität Berlin, matthias.schulte-althoff@fuberlin.de*

Nils Heinemann

*Humboldt Institute for Internet and Society, nils.heinemann@hiig.de*

Jörg Pohle

*Humboldt Institute for Internet and Society, joerg.pohle@hiig.de*

Daniel Fürstenau

*IT University of Copenhagen, daniel.fuerstenau@itu.dk*

Follow this and additional works at: [https://aisel.aisnet.org/ecis2023\\_rip](https://aisel.aisnet.org/ecis2023_rip)

---

### **Recommended Citation**

Gubser, Rahel Jana; Schulte-Althoff, Matthias; Heinemann, Nils; Pohle, Jörg; and Fürstenau, Daniel, "DATA GOVERNANCE STRATEGIES FOR DATA PLATFORMS – A MULTIPLE CASE STUDY IN NURSING CARE" (2023). *ECIS 2023 Research-in-Progress Papers*. 50.

[https://aisel.aisnet.org/ecis2023\\_rip/50](https://aisel.aisnet.org/ecis2023_rip/50)

This material is brought to you by the ECIS 2023 Proceedings at AIS Electronic Library (AISeL). It has been accepted for inclusion in ECIS 2023 Research-in-Progress Papers by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact [elibrary@aisnet.org](mailto:elibrary@aisnet.org).

# DATA GOVERNANCE STRATEGIES FOR DATA PLATFORMS – A MULTIPLE CASE STUDY IN NURSING CARE

*Research in Progress*

Gubser, Rahel, Charité - Universitätsmedizin Berlin, Germany, rahel.gubser@charite.de  
Schulte-Althoff, Matthias, Freie Universität Berlin, Germany, matthias.schulte-althoff@fu-berlin.de  
Heinemann, Nils, Humboldt Institute for Internet and Society, Germany, nils.heinemann@hiig.de  
Pohle, Jörg, Humboldt Institute for Internet and Society, Germany, joerg.pohle@hiig.de  
Fürstenau, Daniel, IT University of Copenhagen, Denmark, daniel.fuerstenau@itu.dk

## Abstract

*It is long established that data from platforms can be useful for deriving patterned insights into people's behavior and conduct. Data platforms are important in fields with limited data availability and strict regulatory and hierarchical structures, such as healthcare and nursing analytics. Hence, we carefully examine three forerunner initiatives in establishing data platforms in the context of nursing care along normative, organizational, and technical dimensions of governance. The cases were selected due to their high level of comparability and to demonstrate three different types of data governance strategies understood as actions to reconcile conflicting interests regarding data and dealing with prevalent data protection law – ranging from strictly processual approaches to the creation of synthetic data. These findings highlight the importance of considering data governance strategies concisely when building data platforms and suggest considerable variety in the configuration of data governance arrangements.*

*Keywords: Data governance, Data platforms, Strategy decision, Nursing data.*

## 1 Introduction

Establishing data governance structures is becoming increasingly important in digital transformation projects in healthcare (Winter & Davidson, 2019), with the growing need to establish accessible data platforms for care research and generating evidence for health analytics and artificial intelligence models (e.g. Bardhan et al., 2020). In particular, evidence-based insights are still largely lacking in the nursing care and hold the promise of a conscious approach to the ongoing 'care crisis', e.g. through the optimisation of processes and documentation (Seibert et al., 2021). For the purposes of this study, we define data platforms as *multi-stakeholder arrangements for the organisation of data storage, processing and sharing*. To date, platform research in healthcare has focused almost exclusively on medical settings, while nursing has been largely neglected. In addition, the secondary use of data reveals the potential for new research purposes. Practical approaches on how to make data accessible are met with extended debates on the social, ethical and economic implications of such a shift towards data

platforms (e.g. Otto et al., 2022). Along these lines, data governance aims to *reconcile the conflicting interests of different stakeholders in terms of value and risk of how data is specifically stored, processed and used* (Grafenstein, 2022). The question of how to strategically integrate data governance in building data platform infrastructure arises from the high sensitivity of data, the vulnerability of patients and the general public interest in the sector (Sabatello et al., 2022).

Data platforms differ from product and innovation platforms (Schreieck et al., 2022) in terms of design and governance mechanisms, as they enable multi-party data exchange based on their own set of rules and governance mechanisms (Beverungen et al., 2022). These mechanisms highlight the trade-offs between the platform openness and closeness and the value of voluntary horizontal knowledge spillovers (Autio et al., 2018). An essential challenge for platform builders and owners is the choice of a data governance strategy that enables data use in agreement with data protection laws and societal values, such as trust and openness (Beverungen et al., 2022). While public-private partnerships on data and data platforms may be possible, as demonstrated by the example of the Royal Free Trust in the UK and Alphabet's AI company DeepMind Health, these partnerships can create challenging problems (Winter & Davidson, 2019). The literature on data platforms is still sparse and largely lacks a strategic perspective, instead taking either a system design (Otto et al., 2022), organizational (Winter & Davidson, 2017), or a legal (Shabani, 2022) perspective.

This paper explores how data governance is strategically implemented when building data platforms in nursing care. It aims to extend current debates on data governance in the platform economy with a strategic perspective in the context of health and nursing care, given the increasing demand and rise of such data platforms in the field (Wang et al., 2022). We understand the emphasis on strategy as a choice of data platform construction and infrastructure, drawing on Mintzberg and Waters' description of the search for strategy identification as a 'search for consistencies in decision-making behaviour' (1990, p. 286). Despite the growing attention and increasing relevance of data governance issues in the practitioner community and among IS researchers, complex undertakings of data governance projects have often failed in the past (Brous et al., 2016) and data governance strategies in practice are under-researched.

Our research context is Germany, where a federal funding programme for care platform research for the development of AI-based solutions and data repositories was launched in 2021. Using a qualitative case study research approach (Eisenhardt, 1989; Yin, 2017), we identify and analyse governance strategies in the establishment of data platforms for three funded data platform projects in nursing care.

Our contribution to existing research comprises (1) the examination of conflicting interests to be negotiated along normative/legal, organisational and technical differences in governance, thus introducing a broader definition of data governance within platform debates. Doing so, (2) we highlight different approaches to building a data platform in care along five main construct characteristics, thereby extending basic definitions of data platforms. This leads to (3) the identification of different data governance strategies (DGS). By comparing data platform initiatives along key construct characteristics, we highlight strategic challenges and suggest governance options for sustaining data platforms for practitioners and policy makers. Attention to such challenges provides important insights that pave the way for the practical, yet fundamentally legitimate, use and establishment of shared data in care beyond the current insistence on data repositories and data silos.

## **2 Related Research**

In this following chapter, we would like to present the reader a background in related research, orientating along recent debates on data platforms and data governance, as our primary contribution addresses the platform economy and secondary contribution attempts cover debates on data governance.

## 2.1 Data Platforms

To date, the definition and classification of *data* platforms has been limited due to the lack of standardised taxonomies and comparable characteristics. While IS research has covered the topic of *digital* platforms for more than four decades (Constantinides et al., 2018), the focus has been on business-to-consumer product or transaction platforms such as Airbnb and Amazon (Bartelheimer et al., 2022). Traditional multi-sided platforms bring together two or more stakeholder groups and enable them to interact (Hagiu & Wright, 2015). In contrast, data platforms are *multi-stakeholder arrangements to organise data storage, processing and sharing*. They build on existing data sources and infrastructure with the aim of making data interoperable and accessible to create an active ecosystem of complementors and customers (Otto & Jarke, 2019). A popular example in healthcare is the Swiss-based platform MiData, which collects encrypted personal data and selectively granted by data account holders to actively contribute to medical research and clinical trials (MiData, 2023).

The success of platforms often depends on the *organising vision* of the platform developer and operator. Traditional platforms typically need to establish a good reputation in order to attract complementors to interact with their platform and gain market share. For data platforms, specific data-centred aspects need to be explicitly stated in the goals and business outcomes, such as data sovereignty, citizen participation, and privacy protection mechanisms, as noted by Beverungen et al. (2022) in demarcating public data spaces from private digital platforms. Another key issue highlighted by these authors is that of *openness*: while the commercial owner of a traditional platform company dictates openness, public data spaces are generally open to third parties. However, individual data platforms may restrict access to specific user groups or to the public. Furthermore, the *storage architecture* of the data platform can be either decentralised, as is more often the case with data than with the core logic of the platform, or centralised. However, decentralised platform architectures in particular and their respective governance models have not been widely discussed (de Reuver et al., 2018). An *architecture of participation* (O'Reilly, 2004), on the other hand, enables different stakeholders to make meaningful contributions to a platform. This requires appropriate tools, access possibilities and other boundary resources (Ghazawneh & Henfridsson, 2013). Finally, data on platforms provide value, especially through secondary uses beyond the original use in providing services to platform participants (Zuboff, 2019). This enables different *economic models*, such as monetising data for research or commercial use, depending on the choice of revenue model and platform operator.

Synthesising the previous arguments, we characterise data platforms under the constructs of *organising vision, openness, storage distribution, participation* and *economic model*. Since data governance is a particularly central aspect for data platforms to allow generative mechanisms to emerge in their respective ecosystems (cf. Zittrain, 2008), in the next subsection we sketch out a conceptualisation of data governance that may be well suited to address the particularities of data platforms.

## 2.2 Data Governance of Data Platforms

Viewing data as "societal resources rather than (or in addition to) privately held assets" (Davidson et al., 2021) calls for important research and policy questions around data governance, particularly its alignment with different goals, priorities and interests (Rosenbaum, 2010; Winter & Davidson, 2019; Zuboff, 2015). In IS literature, data governance is defined rather narrowly as the largely intra-organizational specification and formalization of decision rights, procedures, and controls to reconcile potential conflicts between increasing the value of data and minimizing data-related costs and risks (Abraham et al., 2019), such as the exercise of authority and control over data management. This understanding refers to situations where organisations aim to deliberately control events and behaviours through certain formalised measures, but it falls short of capturing both inter-organisational relationships and the involvement of third party stakeholders (Davidson et al., 2023).

Following the broader understanding of 'governance' advocated in political science, i.e. not so much as mechanisms of control by one entity over others, but rather as reflexive coordination between entities (Hofmann et al., 2017), 'data governance' is thus to be understood as a matter of coordination between different stakeholders with regard to the collection, use, sharing and/or reuse of data (Madison, 2020). Data governance aims to reconcile conflicting interests, which not only differ between the different stakeholders involved in terms of their value and risks, but the value and risks themselves may change over time depending on how the data is specifically processed and used (Grafenstein, 2022). This contextuality reveals what needs to be negotiated in a data governance process to produce an outcome acceptable to all stakeholders: what data should be processed, by whom, for what purposes, in what way, with what technical systems and methods (cf. Micheli et al., 2020). Data governance thus requires the coordination of the normative, organisational and technological dimensions and their interdependence within a complex socio-technical system (Grafenstein, 2022).

The *normative* dimension refers to the applicable law governing the collection, sharing or re-use of data. Data governance structures and processes should reflect the legal framework in which they operate, as the usability of data depends on the rights, responsibilities and obligations associated with it (Beynon-Davies & Wang, 2019). In addition, the normative dimension includes private orders, such as contractual agreements, but also cultural values and social norms, as they can have a similar regulatory function, and how they are translated into practice, i.e. into norms of decision-making and behaviour on the data platform itself (Kretschmer et al., 2020). The *organisational* dimension consists of all those structures, processes and activities (Zhang et al., 2022) that concretely implement the strategy on how to maximise the value of data and minimise the risks, not least compliance risks. The *technological* dimension is defined by the architectural design consisting of the hardware and software infrastructure for storing and processing the data (Grafenstein, 2022), including technical services provided by third parties.

In summary, we define data platforms as multi-stakeholder arrangements to organise the storage, processing and sharing of data, and data governance as the coordination and (re)negotiation to reconcile conflicting interests along normative, organisational and technological dimensions and their interdependencies with respect to building a data platform. Data governance strategies (DGS) then comprise decision streams, i.e. patterns of doing and deciding that aim at implementing data governance to create a data platform.

### **3 Methods**

This paper aims to shed light on data governance strategies (DGS) in the process of making nursing data available, following a multi-sided, inter-professional platform perspective (e.g. Fürstenau et al., 2021). It follows an abductive approach (Alvesson & Kärreman, 2007; Locke et al., 2008).

#### **3.1 Study Design**

Our overall research design was guided by a multiple case study approach (Eisenhardt, 1989) for in-depth analysis. We aim to generate theory (e.g. Harris & Sutton, 1986) on the relationship between a) data governance in b) the development phase of a data platform in c) the context of nursing care. The aim is to d) develop a theory of the strategies through which data governance is implemented in order to further understand how this shapes the nature of data platforms.

#### **3.2 Case Selection and Context of Nursing Care**

We chose to focus on the care context because of the sheer need, limited resources (i.e., limited participation and time), and scarcity of available data, all of which highlight the potential of building a data platform. The high level of vulnerability of the data subjects (patients) and the limited and complex options for consent (e.g. dementia) further emphasise the importance of data governance in this context.

We therefore define the aggregation, processing and accessibility of data on care-related processes and outcomes that fall under the German Social Code, book XI or equivalent international standards or regulations on nursing and care services. We selected three cases: #1 fall prediction, #2 delirium diagnosis, and #3 secretion management. In each case, textual and tabular data consisting of clinical and/or inpatient care information (covered by Social Code, book XI) are collected from at least two data sources (i.e. clinical data and nursing home data) in an unstructured way, with the aim of processing and sharing the data in a structured, accessible way for secondary use in research. We identified key stakeholders in each case, including project coordinator (research institute i.e. university), industry partners, technical data processing partner, nursing specialist. All cases represent scientific projects started in 2022 and focused on building a nursing data platform. They represent influential cases (Gerring, 2006) in terms of their data governance configurations, as they were selected through a competitive process to represent the cutting edge of artificial intelligence (AI) in nursing data repositories/platforms. As such, these cases are diverse (Gerring, 2006) in their data governance configurations and the concrete strategies they pursue to build each platform. All projects operate under German and EU law, which provides a high level of comparable variables, as they have undergone a pre-selection of public funding with contextual influencing factors such as limited realisation periods (project duration between 24 and 36 months) and limited financial resources. In addition, the cases must not include investments for the launch of market-ready commercial products. As all projects integrate the use of AI, the process of building a data platform is strongly linked to AI modelling solutions, which are present in all cases. Finally, all are accessible for primary and secondary data collection in terms of interviewees and access<sup>1</sup> to project plans.

### **3.3 Data Collection and Analysis**

For each case, we started to collect data through semi-structured expert interviews and archival material (Meuser & Nagel, 2009). We collected and analysed archival data using internal project plans to derive strategies in the context of stakeholders and resource allocation. This was informed by a careful (hermeneutic) reading of the material (Mayring & Fenzl, 2014). Later, (2) we began to label and identify patterns with reference to managerial rationales for the design and evolution of specific platform arrangements (cf. Yoo et al., 2010), including the organizing vision and functionality of the platform (Fürstenau et al., 2021). In addition, (3) for our in-depth analysis, we conducted semi-structured expert interviews (Miles & Huberman, 1984) with key stakeholders (industry partner, project coordinator, nurses, technical partners), exploring each of the five construct characteristics of a data platform for underlying causes and actions that followed decisions. Our interview guide consisted of the following sections: general background, architectural choice for each data platform characteristic, decision making and stakeholder engagement regarding privacy and patient consent, anticipated outcomes and intentions for post-project phases to ensure sustainable implementation. To date, we have conducted seven interviews (3/7 for case #1, 2/7 for #2, 2/7 for #3) with an average length of 59 minutes in October 2022 and February 2023, following guidelines for semi-structured interviews (Spradley, 1979). In addition, we participated in a workshop on data protection and collaborative learning attended by representatives from all cases, followed by an on-site discussion of legal requirements and decision making. Our sample included 4 nursing scientists, 2 computer scientists, 1 doctor/medical data scientist and 1 software engineer. The selection of this sample indicates a diverse background of expertise and experience. All interviews were recorded and partially transcribed, we also took detailed notes for later discussion. This process was iterative, with the findings refined after each round of coding and discussion. Finally, (4)

---

<sup>1</sup> *Due to their scientific involvement in one of the cases, the authors possess a more comprehensive understanding of this particular case. To minimize the potential for bias and maintain the credibility of their research findings, the authors conducted interviews exclusively with individuals who are not affiliated with the author group. This approach aims to enhance objectivity in the research process, ultimately ensuring that the results of the study are reliable and trustworthy.*

following Mayring & Fenzl (2014), we focused on theory building around two themes: how to describe different data governance strategies for data platforms, and how to analyse them conceptually.

## 4 Preliminary Results

Our preliminary results show that in each case data governance was strategised as a direct derivative of the target platform architecture. Therefore, in order to identify Data Governance Strategies (DGS), we first needed to understand the anticipated building approach. This chapter can therefore be divided into two parts to help us (1) identify different approaches to building a data platform along five construct characteristics and (2) identify the conflicting interests to be negotiated in each case. We summarise our findings in three distinct data governance strategies (DGS), which we have termed: *Moonshot*, *Umbrella* and *Run-of-the-Mill*.

We call the first strategy ‘**Moonshot**’ because it addresses conflicting interests and general debates about governance with an ambitious technical solution. This strategy involves a radical type of innovation, as it uses different sources of real patient data to generate representative synthetic data. In this way, the main challenges (“At the moment, our biggest challenge is still data protection law” - industry partner) are addressed in a pragmatic way. “Traceability of what happens to the data” (industry partner) is ensured through interdisciplinary committees. The second strategy, ‘**Umbrella**’, primarily responds to the high uncertainty of the regulatory environment with a medium-liberal governance platform strategy: “We face the problem that our solution has to withstand local and international interpretations; especially for the validation of AI algorithms, the network of norms is not yet viable, you still have to wait and see.” (project coordinator). By handing over both data and control to a parent organisation (e.g. a national or transnational research data centre), this strategy covers itself under the umbrella of an established parent for legal protection. Finally, the ‘**Run-of-the-Mill**’ strategy is a more conventional, standardised strategy of both building a data platform and data governance. The platform uses multiple data sources as input (e.g., private data donors, real patient datasets) and aims to incentivise shared data output (e.g., open source machine learning models), while governance is continuously negotiated among various stakeholders, particularly with the intended platform operator. “As a developer who wants to use the platform, it would be desirable for the subject of data protection to be taken over by the platform” (project coordinator). Table 1 provides a comprehensive overview of the main characteristics of each data platform (DP) approach.

	<b>Moonshot</b>	<b>Umbrella</b>	<b>Run-of-the-Mill</b>
<b>Organising Vision</b>	DP with representative synthetic data	Integration of dataset to parent DP	DP with various real patient data & data user sovereignty
<b>Openness</b>	Open access for data users	Restricted for data users	Open for data provider, Restricted for data user
<b>Distribution of Storage</b>	Decentral data sources & federated learning => Central DP	Decentral data sources & federated learning => Decentral DP	Decentral data sources & federated learning => Central DP
<b>Architecture of Participation</b>	Little to no participation	Parent’s choice	Incentives for data donation & data users
<b>Economic Model</b>	Operator is interdisciplinary research group	Operator is parent	Operator is to be negotiated

Table 1. Comparing anticipated data platforms along five construct characteristics.

In terms of the *normative* dimension of governance, the same legislation was mentioned as relevant in all cases (namely GDPR, hospital legislation, Social Code Book XI, medical products/licensing regulation). GDPR was the basic premise on which stakeholders negotiated, in practice, conflicting interests are answered in all cases with pragmatic solutions revolving around three issues: (1) informed

patient consent, (2) anonymisation of real patient data, and (3) data minimisation. As the current regulations in this area are met with a great deal of ambiguity and a wide scope for interpretation, the legal requirements are seen as a challenge to be worked around. In one illustrative example, informed patient consent could be circumvented with a different interpretation of the law: “The data subjects are the carers, not the patients...there are people in the ethics committee with whom such issues can be discussed constructively.” (project coordinator).

Each strategy took a different approach to the *organisational* dimension of governance. We found two overarching themes: (1) accountability and (2) oversight. With Moonshot, trust and transparency are addressed through the establishment of decision-making and oversight bodies (e.g. scientific advisory board). “Decision-making power should be the first priority: Make fairly quick, very pragmatic decisions” (industry partner). For the Umbrella strategy, governance issues are almost exclusively addressed on the organisational dimension (data integration into the parent structure). For the Run-of-the-Mill, on the other hand, accountability is considered to be “shared” (project coordinator) and therefore has to rely on subcontracted legal advice from third parties in its negotiation process.

In the *technological* dimension, architectural design choices differ in terms of storage (centralised/decentralised) of data sources, data processing and data usage (see Table 1). Data quality (DQ) emerged as another relevant governance metric within this dimension. The Moonshot relies heavily on this dimension of governance, with synthetic data as its main output on a cloud solution. “We have based our whole system on containers in the cloud, we use Kubernetes for orchestration like many others, these are the systems that give us the best possible management and centralisation of data, but also data security and backups. Data repositories should follow this approach” (industry partner). Umbrella addresses data quality with FHIR compatibility and the merging of different data sources: “they [parent] have decided to aggregate data on a central FHIR data repository, this FHIR data will be validated with profiles, we would also like to ... achieve distributed evaluations”. (project coordinator). The technical dimension in Run-of-the-Mill also focuses on data quality: an on-premises algorithm is to provide data quality alerts, as no pre-selection of data sources is intended, in order to incentivise the principle of getting as much data onto the platform as possible.

The following Table 2 highlights the different characteristics of each DGS, as well as the expected outcome acceptable to all stakeholders, in order to provide a source of orientation and summary along the different dimensions of governance.

	<b>Anticipated Outcome:</b>	<b>Moonshot:</b> stand alone data repository + proof of concept (synth data)	<b>Umbrella:</b> integrated data platform + API interface	<b>Run-of-the-Mill:</b> stand alone data platform + data quality assessment
<b>NORMATIVE</b>	<b>Consent</b>	broad consent	broad consent	earmarked, informed consent
	<b>Level of Anonymisation</b>	anonymized data	pseudonymised data	pseudonymised data
	<b>Data Minimisation</b>	data maximisation	data minimisation	data maximisation
<b>ORGANIZATIONAL</b>	<b>Accountability</b>	data provider (clinics)	parent	DP operator
	<b>Oversight</b>	tech partner	parent	data provider (patients, clinics)
<b>TECHNOLOGICAL</b>	<b>Data Quality</b>	to be negotiated	FAIR & FHIR	data quality assessment (machine learning based)

Table 2. Comparing DGS along dimensions of governance.



## **5 Discussion and Conclusion**

We used a multiple case study approach in the context of care to illustrate current data governance strategies for building a data platform. We deliberately chose to focus our research on data platforms, as opposed to data repositories, because platforms suggest a broader concept that includes, for example, economic models, implying a more integrated and thus sustainable view of how to co-create digital futures. Our study aims to contribute to the literature on platforms in three ways. First, we provide a definition of data platforms and characterise them along five primary dimensions. Second, we adopt a data governance framework from Grafenstein (2022), which we argue is well suited to the requirements of data platform governance strategies. Finally, we provide an overview of three implemented data governance strategies that exemplify how data governance is implemented in the context of data platform building, and that differ significantly in their architectural design, mode of integration, and handling of data sharing and privacy concerns. From a practical point of view, we show how data platforms can currently be designed in the care sector.

While other studies have focused on the legal (Shabani, 2022), organizational (Winter & Davidson, 2017), or system design (Otto et al., 2022) dimensions of data governance, we interweave these dimensions to highlight the importance of normative, organizational, and technical aspects of data governance strategies. For our selected cases, it has proved particularly fruitful to argue how normative conceptions of data platform construction are operationalised or mitigated – to varying degrees – by means located in the organisational or technical dimensions. In all cases, the proper handling of data protection law was identified as one of the main obstacles to be overcome. It has already been established in the literature that the scarce public availability of databases and the restrictions on access to existing datasets due to regional data protection laws constitute barriers to data-driven research with a high need for data protection (Seibert et al., 2021). The prevailing governance strategies in the cases at hand provide either organisational or technology-centred approaches to address the issue of consent in the processing of patient data. Technology-centred methods include anonymisation techniques, such as the creation of synthetic data with deep neural networks (Beaulieu-Jones et al., 2019) or by means of differential privacy (Ficek et al., 2021) - both current hot topics in the medical and computer science literature. Examples of organisational approaches include obtaining (broad) consent from patients as a sub-contractual agreement during treatment, or ceding control to established public data platform initiatives.

Finally, we add to the discussion on the formation of platform models. In classic platform models, the value creation is distributed but value capture is fairly centralized (Gawer, 2022). In the context of data platforms, actors can access and further reuse data – the main offering of a data platform – while maintaining sovereignty. In particular, the value of voluntary horizontal knowledge spillovers is emphasized more than in traditional business models. In contrast to vertical knowledge spillovers in user-producer dyads, horizontal knowledge spillovers tend to occur in learning and production ecosystem-type clusters that foster the formation of an innovation culture (Autio et al., 2018). In this paper, we outline governance strategies to mitigate data sharing problems, thereby potentially enabling collaborating efforts more easily than in traditional platform ecosystems.

## **6 Outlook**

For future research, we envision extending the research in three directions. First, as an initial set of interviews was conducted, additional interviews will be conducted to explore more broadly, as research in width, how to extend the strategy proposals. Second, a research in depth will address more topics within the established frameworks, i.e. discussions around medical device regulation, a highly relevant topic mentioned in all interviews. Finally, we want to expand our research in distance by comparing the initially intended strategies covered so far with actual deliberate and emergent strategies (Mintzberg and Waters, 1990). To this end, an additional round of qualitative data collection is planned in 12–24 months.

## References

- Abraham, R., Schneider, J., and vom Brocke, J. (2019). “Data governance: A conceptual framework, structured review, and research agenda”. *International Journal of Information Management* 49, 424–438.
- Alvesson, M. and Kärreman, D. (2007). “Constructing mystery: Empirical matters in theory development”. *Academy of Management Review* 32(4), 1265–1281.
- Autio, E., Nambisan, S., Thomas, L. D. W., and Wright, M. (2018). “Digital affordances, spatial affordances, and the genesis of entrepreneurial ecosystems”. *Strategic Entrepreneurship Journal* 12(1), 72–95.
- Bardhan, I., Chen, H., and Karahanna, E. (2020). “Connecting systems, data, and people: A multidisciplinary research roadmap for chronic disease management”. *MIS Quarterly: Management Information Systems* 55(1), 185–200.
- Bartelheimer, C., zur Heiden, P., Lüttenberg, H., and Beverungen, D. (2022). “Systematizing the lexicon of platforms in information systems: A data-driven study”. *Electronic Markets* 32(1), 375–396.
- Beaulieu-Jones, B. K., Wu, Z. S., Williams, C., Lee, R., Bhavnani, S. P., Byrd, J. B., and Greene, C. S. (2019). “Privacy-Preserving Generative Deep Neural Networks Support Clinical Data Sharing”. *Circulation: Cardiovascular Quality and Outcomes* 12(7).
- Beverungen, D., Hess, T., Köster, A., and Lehrer, C. (2022). “From private digital platforms to public data spaces: Implications for the digital transformation”. *Electronic Markets* 32(2), 493–501.
- Beynon-Davies, P., and Wang, Y. (2019). “Deconstructing Information Sharing”. *Journal of the Association for Information Systems* 20(04), 476–498.
- Brous, P., Janssen, M., and Vilminko-Heikkinen, R. (2016). “Coordinating Decision-Making in Data Management Activities: A Systematic Review of Data Governance Principles”, in: H. J. Scholl, O. Glassey, M. Janssen, B. Kliavink, I. Lindgren, P. Parycek, E. Tambouris, M. A. Wimmer, T. Janowski, & D. Sá Soares (eds.), *Electronic Government* 9820, 115–125.
- Chandler, A. D. (1990). *Strategy and structure: Chapters in the history of the industrial enterprise*. M.I.T. Press.
- Constantinides, P., Henfridsson, O., and Parker, G.G. (2018). “Introduction—Platforms and infrastructures in the digital age”. *Information Systems Research* 29(2), 381–400.
- Davidson, E., Winter, J., Wessel, L., and Winter, S. (2021). “Data Governance, Digital Innovation and Grand Challenges”. *Call for Papers: Information and Organization*.
- Davidson, E., Wessel, L., Winter, J., and Winter, S. (2023). “Future directions for scholarship on data governance, digital innovation, and grand challenges”. *Information and Organization* 33(1), 100454.
- de Reuver, M., Sørensen, C., and Basole, R. C. (2018). “The Digital Platform: A Research Agenda”. *Journal of Information Technology* 33(2), 124–135.
- Eisenhardt, K. M. (1989). “Building Theories from Case Study Research”. *The Academy of Management Review* 14(4), 532.
- Ficek, J., Wang, W., Chen, H., Dagne, G., and Daley, E. (2021). “Differential privacy in health research: A scoping review”. *Journal of the American Medical Informatics Association* 28(10), 2269–2276.
- Fürstenau, D., Klein, S., Vogel, A., and Auschra, C. (2021). “Multi-sided platform and data-driven care research: A longitudinal case study on business model innovation for improving care in complex neurological diseases”. *Electronic Markets* 31(4), 811–828.

- Gawer, A. (2022). "Digital platforms and ecosystems: Remarks on the dominant organizational forms of the digital age". *Innovation* 24(1), 110–124.
- Gerring, J. (2006). *Case study research: Principles and practices*. Cambridge University Press.
- Ghazawneh, A. and Henfridsson, O. (2013). "Balancing platform control and external contribution in third-party development: The boundary resources model: Control and contribution in third-party development". *Information Systems Journal* 23(2), 173–192.
- Hagiu, A., and Wright, J. (2015). "Multi-sided platforms". *International Journal of Industrial Organization* 43, 162–174.
- Harris, S. G. and Sutton, R. I. (1986). "Functions of parting ceremonies in dying organizations". *Academy of Management Journal* 29(1), 5–30.
- Hofmann, J., Katzenbach, C., and Gollatz, K. (2017). "Between coordination and regulation: Finding the governance in Internet governance". *New Media & Society* 19(9), 1406–1423.
- Kretschmer, T., Leiponen, A., Schilling, M., and Vasudeva, G. (2022). "Platform Ecosystems as Meta- Organizations: Implications for Platform Strategies". *Strategic Management Journal* 43(3), 405–424.
- Locke, K., Golden-Biddle, K., and Feldman, M. S. (2008). "Perspective—Making Doubt Generative: Rethinking the Role of Doubt in the Research Process". *Organization Science* 19(6), 907–918.
- Madison, M. J. (2020). "Tools for Data Governance". *Technology and Regulation* 2, 29–43.
- Malik, P. (2013). "Governing Big Data: Principles and practices". *IBM Journal of Research and Development* 57(3/4), 1:1-1:13.
- Meuser, M. and Nagel, U. (2009). "The Expert Interview and Changes in Knowledge Production", in: A. Bogner, B. Littig, & W. Menz (eds.), *Interviewing Experts* (17–42). Palgrave Macmillan UK.
- 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).
- MiData (2023). *Meine Daten, unsere Gesundheit*. URL: <https://www.midata.coop/>
- Miles, M. B. and Huberman, A. M. (1984). *Qualitative data analysis: A sourcebook of new methods*. Sage Publications.
- Mintzberg, H. and Waters, J. A. (1990). "Tracking Strategy in an Entrepreneurial Firm". *Family Business Review* 3(3), 285–315.
- O'Reilly, T. (2004). *Architecture of Participation*. URL: [https://wiki.p2pfoundation.net/Architecture\\_of\\_Participation](https://wiki.p2pfoundation.net/Architecture_of_Participation) (visited on November 14, 2022).
- Otto, B. (2011). "Organizing Data Governance: Findings from the Telecommunications Industry and Consequences for Large Service Providers". *Communications of the Association for Information Systems* 29.
- Otto, B. and Jarke, M. (2019). "Designing a multi-sided data platform: Findings from the International Data Spaces case". *Electronic Markets* 29(4), 561–580.
- Otto, B., ten Hompel, M., and Wrobel, S. (2022). *Designing Data Spaces: The Ecosystem Approach to Competitive Advantage*. Springer International Publishing.
- Rosenbaum, S. (2010). "Data Governance and Stewardship: Designing Data Stewardship Entities and Advancing Data Access: Data Governance and Stewardship". *Health Services Research*, 45(5(2)), 1442–1455.
- Sabatello, M., Martschenko, D. O., Cho, M. K., and Brothers, K. B. (2022). "Data sharing and community-engaged research". *Science* 378(6616), 141–143.
- Schrieck, M., Wiesche, M., and Kremer, H. (2022). "Governing innovation platforms in multi-business organisations". *European Journal of Information Systems*, 1–22.

- Seibert, K., Domhoff, D., Bruch, D., Schulte-Althoff, M., Fürstenau, D., Biessmann, F., and Wolf-Ostermann, K. (2021). “Application Scenarios for Artificial Intelligence in Nursing Care: Rapid Review”. *Journal of Medical Internet Research* 23(11), e26522.
- Shabani, M. (2022). “Will the European Health Data Space change data sharing rules?” *Science* 375(6587), 1357–1359.
- Spradley, J. P. (1979). *The ethnographic interview*. Holt, Rinehart and Winston.
- von Grafenstein, Max. (2022). *Reconciling Conflicting Interests in Data through Data Governance. An Analytical Framework (and a Brief Discussion of the Data Governance Act Draft, the Data Act Draft, the AI Regulation Draft, as well as the GDPR)*.
- Wang, M., Li, S., Zheng, T., Li, N., Shi, Q., Zhuo, X., Ding, R., and Huang, Y. (2022). “Big Data Health Care Platform With Multisource Heterogeneous Data Integration and Massive High-Dimensional Data Governance for Large Hospitals: Design, Development, and Application”. *JMIR Medical Informatics*, 10(4).
- Winter, J. S., & Davidson, E. (2017). “Investigating values in personal health data governance models”. 23rd Americas Conference on Information Systems (AMCIS), Boston, USA.
- Winter, J. S., & Davidson, E. (2019). “Big data governance of personal health information and challenges to contextual integrity”. *The Information Society* 35(1), 36–51.
- Yin, R. K. (1984). *Case study research: Design and methods*. Sage Publications.
- Yoo, Y., Henfridsson, O., & Lyytinen, K. (2010). “The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research”. *Information Systems Research* 21(4), 724–735.
- Zhang, Q., Sun, X., and Zhang, M. (2022). “Data Matters: A Strategic Action Framework for Data Governance”. *Information & Management* 59(4), 103642.
- Zittrain, J. (2008). *The Future of the Internet and How to Stop It*. Yale University Press.
- Zuboff, S. (2015). “Big other: Surveillance Capitalism and the Prospects of an Information Civilization”. *Journal of Information Technology* 30(1), 75–89.
- Zuboff, S. (2019). “Surveillance Capitalism and the Challenge of Collective Action”. *New Labor Forum* 28(1), 10–29.