UNIVERSITY OF MODENA AND REGGIO EMILIA

Department of Sciences and Methods for Engineering

Doctorate School in Industrial Innovation Engineering

XXXVII Cycle

Addressing Grand Societal Challenges through Data Sharing: Essays on Data Ecosystems in the Context of Integrated Mobility in Europe

---

DOCTORAL THESIS

Author: Giulia Renzi

Supervisor: Professor Paula Ungureanu

Co-Supervisor: Professor Mauro Dell'Amico

School Coordinator: Professor Franco Zambonelli

Academic year 2023/2024

Abstract

Data ecosystems are interconnected networks where stakeholders use digital technologies to leverage large amounts of data in order to create value and drive innovation. They represent an emerging and impactful phenomenon, characterized by multi-stakeholder data sharing collaborations, often addressing grand societal challenges like environmental sustainability and scientific advancement. Despite some success, many data ecosystems have not met expectations. Unlike other ecosystems, they face unique challenges related to data governance, privacy, trust, and regulatory frameworks, further complicated by the rapid evolution of data technologies. These factors highlight the need to study European mobility data ecosystems as unique phenomena, shaped by an intricate interplay of technological and organizational mechanisms. While their importance is widely recognized, research still lacks a deeper understanding of how technological aspects interact with organizational and managerial dimensions. This dissertation addresses this gap, contributing both theoretically and practically to the emerging field of data ecosystems.

The three empirical studies in this dissertation draw on the research context of the MobiDataLab project, a Horizon 2020 initiative aimed at developing a European data ecosystem for integrated mobility. Our data-driven approach combined in-depth semi-structured interviews and archival data, covering European data ecosystems, regulatory frameworks, participation models, and governance. The anonymized data was transcribed, imported into the QACDAS platform DEDOOSE, and analyzed using Grounded Theory to systematically organize the data, identify recurring themes, and develop new theoretical insights.

The results reveal that not everyone in a data ecosystem shares data in the same way. This dissertation identifies key factors and mechanisms driving the diversity of sharing practices and examines when this diversity presents challenges or opportunities for data

ecosystems. The studies highlight the role of data as boundary objects - adaptable enough to meet local needs while maintaining consistency across contexts - suggesting that different sharing practices depend on stakeholders' willingness to share and on the tension between standardization and trust. These insights underpin three distinct studies exploring critical aspects of data ecosystems.

The first study examines how stakeholders' interpretations of data and data sharing practices shape the evolution of data ecosystems via roles and sharing behaviors which either foster innovation or cause stagnation. The second study explores the motivational, cognitive, and relational factors influencing data-sharing practices. It theorizes willingness to share in an ecosystem as a multidimensional interplay between stakeholders' different social and utilitarian motivations, cognitive evaluations about risk-opportunity accounts, and relational dynamics. The third study focuses on the tension between standardization and trust, highlighting how efforts to reduce costs through standardization can increase the need for trust-based collaboration when rigid systems fail. Together, these studies provide valuable insights into the functioning, governance, and evolution of data ecosystems, adding to literatures on big data governance, digital and innovation ecosystems, and data sharing in complex organizational forms. As discussions on data privacy, governance, and sharing become increasingly relevant, this dissertation advances understanding of how data ecosystems can adapt to future challenges in a data-driven world.

Keywords: Data Ecosystems, Data Sharing, European Mobility, Trust, Standardization.

Contents

List of Tables

List of Figures

1. Introduction[1]

1.1. The Role of Data and Data Sharing in Addressing the Grand Challenges of Our Time

In recent years, humanity has faced a series of deep and interconnected challenges that jeopardize the sustainability of our societies, economies, and environment. Back in 2015, the United Nations launched the Sustainable Development Goals (SDGs) as a universal call to action to eradicate poverty, protect the planet, and ensure that by 2030, all people can live in peace and prosperity (United Nations Development Programme, 2024). As we approach the end of 2024, with only six years remaining, this ambitious vision appears increasingly difficult to achieve. Yet, the challenges we face must not serve as a reason to abandon efforts or cease striving for a better and more equitable future.

As Sidze et al. (2022) emphasize, addressing complex challenges requires holistic and inclusive strategies that not only prevent the exacerbation of existing inequalities but also deliver meaningful benefits to society as a whole. While new technologies often solve specific problems, they frequently create new ones in the process: technological advancements and global networks, in fact, have introduced both opportunities and risks (Demchenko et al., 2014; Wareham et al., 2014). Digitalization, for instance, has the potential to democratize access to essential services such as education and healthcare, thereby improving social conditions (Hansen & Pang, 2023). Numerous digitalization initiatives have been implemented to enhance transparency, reduce corruption in public governance, support law enforcement, facilitate smart urban planning, advance healthcare, and combat climate change (Davidson et al., 2023; Kitsios

[1] This chapter corresponds to a Working Paper to be submitted to the *European Journal of Innovation Management*. Authors: Giulia Renzi, first author (conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, validation, visualization, writing – original draft), Paula Ungureanu (conceptualization, investigation, methodology, supervision, validation), and Natalia Selini Hadjidimitriou (data curation, writing – review and editing).

et al., 2017). These innovations are increasingly viewed as critical tool for addressing systemic problems, fostering cooperation among stakeholders, and working towards a unified vision for societal progress (Susha, van den Broek, et al., 2023). However, digitalization – and innovation more broadly – also tends to amplify existing inequalities, as, for instance, those lacking digital access are often left even further behind (Helsper, 2021). This duality highlights the pressing need for strategies that balance the transformative potential of innovation with the principles of inclusivity and sustainability: addressing these grand challenges effectively requires solutions that not only leverage technological progress but also ensure that its benefits are equitably distributed across society (Bonina et al., 2021; Davidson et al., 2023; Susha, van den Broek, et al., 2023).

This thesis seeks to examine the multifaceted and complex dynamics of data ecosystems designed to address society's most pressing grand challenges. It critically analyzes the collaboration among diverse actors, exploring the motivations that drive individuals to engage in data ecosystems and share their data. The study delves into their willingness to share, the variety of data-sharing practices they adopt, the evolving roles they assume within these ecosystems, and the protocols developed to facilitate data exchange while finding a delicate balance between trust and standardization. The focus on data emerges from its transformative potential in tackling today's global challenges. In particular, data sharing has become a critical force, offering innovative pathways to address systemic issues. Although the practice of leveraging shared data for societal benefits is still in its early stages, its capacity to improve social outcomes is both significant and undeniable (Hansen & Pang, 2023). When effectively utilized, data sharing acts as a catalyst for innovation, enabling stakeholders to confront complex problems through collective intelligence and evidence-based decision-making (Hansen & Pang, 2023; Davidson et al., 2023; Kitsios et al., 2017). This thesis aligns with SDG 9: *Industry, Innovation, and Infrastructure*, which highlights the critical importance of

developing efficient and sustainable transport systems grounded in technological advancements and digital transformation. By addressing these priorities, we aim to contribute to the broader mission of creating resilient transport networks, fostering innovation, and leveraging digitalization to advance sustainable development.

Open data initiatives offer a compelling example of the transformative potential of data sharing. Globally, such programs have been implemented to improve transparency in public administration, combat corruption, and enhance accountability within governance structures: by making previously inaccessible information publicly available, these initiatives empower citizens, strengthen law enforcement, and support more effective urban planning (Davidson et al., 2023). In healthcare, shared data resources have driven breakthroughs in medical research, optimized the allocation of resources, and improved patient outcomes. Similarly, in the environmental sector, data sharing has informed strategies to combat climate change and promote sustainability (Kitsios et al., 2017). What sets data sharing apart is its unparalleled ability to break down traditional silos, fostering connections among stakeholders across diverse sectors and geographies (Demchenko et al., 2014; Granstrand & Holgersson, 2020; Lis & Otto, 2020). Modern data ecosystems are increasingly recognized as vital frameworks for addressing systemic challenges: they provide platforms where governments, private organizations, researchers, and civil society can pool their knowledge, resources and expertise, fostering cooperation and pave the way for a unified vision of societal progress (Susha, van den Broek, et al., 2023).

However, the increasing significance of data ecosystems underscores not only the vast opportunities they offer but also the critical need to establish trust, interoperability, and shared governance principles: without these foundational elements, the transformative potential of data sharing is at risk of being compromised by concerns surrounding privacy, inequality, and misuse (Aaen et al., 2022; Susha et al., 2019). While much of the discourse has centered on commercial

data ecosystems (Zillner et al., 2021), there is a growing recognition of their potential to address societal challenges and contribute to the public good across a variety of domains. From education (Cazier et al., 2015) and healthcare (Raghupathi & Raghupathi, 2014), to environmental sustainability (Newell & Marabelli, 2015) and public governance (G. H. Kim et al., 2014), data ecosystems serve as essential frameworks for fostering collaboration, driving innovation, and solving complex problems (Demchenko et al., 2014; Granstrand & Holgersson, 2020; Lis & Otto, 2020; Micheli et al., 2020). Collectively, these efforts illustrate the immense potential of data ecosystems to harness the power of data, creating systemic impacts that transcend traditional boundaries and support progress across multiple sectors.

The European Commission has been instrumental in supporting data-driven initiatives, providing funding and resources to projects in critical areas such as mobility, healthcare, and environmental sustainability (Petre et al., 2019). By fostering collaboration among diverse stakeholders, these initiatives aim to achieve outcomes that exceed the capacities of individual actors. However, despite their promise, data ecosystems often fail to meet their ambitious objectives. Effective implementation demands the ability to navigate diverse data types and reconciling conflicting stakeholder interests, both of which can pose significant challenges (Lis & Otto, 2020; Marcelo et al., 2019). The successes and failures of these ecosystems highlight their inherent complexity. In some cases, data ecosystems have revolutionized their respective fields: predictive policing systems, for instance, leverage advanced data analytics to prevent crime (Davidson et al., 2023; Iannacci et al., 2022; Lum et al., 2016; Schneider et al., 2020). Yet, concerns about data quality, interoperability, and stakeholder alignment frequently compromise their credibility (Iannacci et al., 2022; Lum et al., 2016). Similarly, environmental initiatives like Global Forest Watch, which unite governments, NGOs, and local communities to monitor deforestation in real time, continue to face persistent challenges in data integration and collaboration (Schneider et al., 2020). Even in healthcare, a field with significant stakes,

notable projects like Google Health and Denmark's General Practitioners Database (DAMD) have encountered major setbacks. Despite initial enthusiasm, DAMD, once heralded as a groundbreaking innovation, ultimately faltered due to issues related to privacy, governance, and technical complexity, leading to its discontinuation (Aaen et al., 2022; Spil & Klein, 2014).

These examples highlight the complex challenges that data ecosystems encounter, ranging from technical obstacles to issues of trust and regulatory concerns (Aaen et al., 2022; R. Agarwal & Dhar, 2014). However, the problem faced by data ecosystems extend far beyond these concerns. A particularly persistent issue is the lack of alignment among stakeholders: despite ongoing efforts to create common standards and frameworks, data ecosystems often reflect the conflict priorities of their participants (Aaen et al., 2022; Susha et al., 2019, Ungureanu, 2021). For instance, policymakers may prioritize balancing societal benefits with privacy considerations (Shin & Choi, 2015), while private organizations tend to focus on business opportunities and are often reluctant to adopt regulations they perceive as burdensome (Janhunen, 2019). Meanwhile, citizens frequently approach data sharing with caution, influenced by privacy concerns, trust issues, and varying levels of technological familiarity (Günther et al., 2017; Jäckle et al., 2019; Schudy & Utikal, 2017). Another significant challenge lies in the technical aspects of data management: issues such as data incompleteness, scalability, portability, interconnectivity, timeliness, and security can impede the (Bonina et al., 2021; Günther et al., 2017; Micheli et al., 2020). Data sharing, in particular, requires robust mechanisms for capturing, storing, searching, analyzing, and visualizing data to ensure its utility and reliability (Abraham et al., 2019; Alhassan et al., 2016; Kitsios et al., 2017). Data quality is an equally critical concern. Inconsistent or incomplete datasets erode trust in data-driven decisions, while the lack of standardized formats complicates the integration and analysis of information across systems (Han & Dong, 2015; N. Khan et al., 2014). Moreover, gaps in analytic skills among stakeholders further limit their ability to derive meaningful

insights from data, hindering efforts to address grand challenges (Han & Dong, 2015). Privacy and security represent additional persistent issues that demand urgent attention: as data sharing becomes more prevalent, the risk of breaches, misuse, and violations of individual rights increase (Günther et al., 2017; Micheli et al., 2020; Moreno et al., 2019). These threats are exacerbated by the absence of comprehensive data governance frameworks that clearly outline responsibilities, establish standards, and enforce mechanisms for accountability (Bonina et al., 2021; Cui et al., 2020). Without such frameworks, data ecosystems remain vulnerable to ethical breaches, privacy violations, and a decline in public trust.

These challenges associated with data sharing give rise to a thought-provoking paradox: is it worth creating new problems to solve existing ones? The complexities inherent in data sharing can sometimes lead to unforeseen obstacles that rival the issues they were designed to address: This paradox is particularly salient as the world navigate the ethical, technical, and societal implications of utilizing data for public good (Demchenko et al., 2014; Hanseth et al., 2006; Wareham et al., 2014). On one hand, the promise of data ecosystems is undeniable: by pooling resources, combining expertise, and employing advanced technologies, these ecosystems present innovative solutions to persistent global challenges (Davidson et al., 2023; Micheli et al., 2020; Susha, van den Broek, et al., 2023). On the other hand, their creation and management often introduce new hurdles that risk undermine their intended benefits (Bonina et al., 2021; Demchenko et al., 2014; Hanseth et al., 2006). This paradox raises critical questions about the trade-offs involved in pursuing data-driven innovation. Are the advantages of enhanced efficiency, transparency, and collaboration worth the potential drawbacks, such as heightened privacy concerns, increased inequality, and governance complexities? Can the challenges introduced by data ecosystems be justified if they yield only incremental progress in addressing pressing global issues?

The paradox of creating new problems to address existing ones is not exclusive to data ecosystems. Historically, technological and social advancements have often brought unintended consequences, ranging from the environmental degradation caused by industrialization to the ethical dilemmas posed by artificial intelligence. What distinguishes data ecosystems is the unprecedented speed and scale at which their impacts are felt, amplifying both their potential and their risks: ultimately, the true value of data ecosystems lies in their capacity to navigate these trade-offs effectively (Davidson et al., 2023; Günther et al., 2017; Micheli et al., 2020). By tackling the governance, ethical, and technical challenges they present, stakeholders can strive to maximize the benefits of data sharing while minimizing its drawbacks. Achieving this balance demands a collective effort to develop robust frameworks, foster trust among participants, and ensure that innovation is pursued in a way that prioritizes societal well-being.

1.2. Data Governance in Data Ecosystems: Balancing Structure, Collaboration and Trust

At the heart of addressing this paradox lies the concept of data governance, a cornerstone of modern data ecosystems. Data governance refers to a systematic framework for defining decision-making rights, responsibilities, and processes necessary for managing data as a critical organizational asset (Otto, 2011). It emphasizes the importance of treating data as a strategic resource that requires careful stewardship to unlock its full potential and from both academic and practical perspectives, data governance seeks to establish a universal approach to data accountability, covering every aspect of organizational data management (Demchenko et al., 2014; Günther et al., 2017; Micheli et al., 2020). Data governance also provides the necessary structure to align data practices with organizational goals, ensuring accessibility, reliability, and compliance with ethical and regulatory standards (Weber et al., 2009; Wende, 2007). Effective

data governance addresses technical dimensions such as data quality, security, and interoperability, alongside organizational processes like assigning roles, creating decision-making hierarchies, and establishing accountability measures. By integrating these elements, it creates a comprehensive framework that guides how data is collected, stored, shared, and utilized (Demchenko et al., 2014; Günther et al., 2017; Micheli et al., 2020). Beyond its technical functions, data governance plays a pivotal role in fostering trust by ensuring transparency in data management practices: this transparency is critical for enabling collaboration and supporting informed decision-making in increasingly interconnected data ecosystems (Demchenko et al., 2014; Günther et al., 2017; Micheli et al., 2020).

However, implementing data governance is not without its challenges. Organizations often struggle to balance competing priorities, such as ensuring data accessibility while protecting privacy, or fostering innovation while maintaining regulatory compliance (Bonina et al., 2021; Demchenko et al., 2014; Micheli et al., 2020). Additionally, the universality of data governance frameworks is frequently debated, as the diverse needs and contexts of organizations often demand tailored approaches rather than a one-size-fits-all solutions (Bonina et al., 2021; Demchenko et al., 2014; Micheli et al., 2020).

With data growing rapidly in both volume and complexity, robust data governance frameworks are more important than ever. These frameworks are key to tackling the challenges of data sharing, staying compliant with evolving regulations, and addressing ethical concerns in a fast-changing world. By building a foundation of accountability, transparency, and collaboration, data governance may unlock the full potential of data ecosystems and their transformative power (Demchenko et al., 2014; Günther et al., 2017; Micheli et al., 2020).

Efforts to tackle the complexities of data governance and ecosystems have largely concentrated on structural factors, such as participant characteristics, organizational

frameworks, and defined roles within these systems (Günther et al., 2017; Jarvenpaa & Essén, 2023; Lis & Otto, 2020; Micheli et al., 2020; van den Broek & van Veenstra, 2018). However, much of the existing literature mirrors organizational-level analyses, often neglecting the unique challenges posed by inter-organizational data governance, such as fluid roles, undefined responsibilities, and the lack of shared objectives (Davidson et al., 2023; Jarvenpaa & Essén, 2023). To fully grasp the dynamics of complex collaborations within data governance, we argue that it is crucial to move beyond structural factors and consider motivational and relational aspects. These encompass the interaction between individual and collective goals, the role of mutual trust, attitudes toward data sharing, and how these factors influence the effectiveness and sustainability of data ecosystems (Wang et al., 2021). Such considerations are particularly critical in scenarios where collaboration involves across diverse organizations with differing goals and operational cultures (Gregory et al., 2021).

In this thesis, particular attention will be given to these emergent and relational dimensions of data ecosystems. By examining the fluidity of roles and practices, it seeks to provide a deeper understanding of how governance frameworks can be designed to address the evolving needs of stakeholders while ensuring the sustainability of data ecosystems. Greater emphasis is placed on the processes of data governance, examining how governance dynamically unfolds within these ecosystems. This approach highlights not only the frameworks and rules that underpin data governance but also the actions, perceptions, and interactions that influence governance outcomes in practice. Governance, we argue, is not a static set of rules, but a dynamic process shaped by continuous negotiations, adaptations, and interactions among stakeholders. The complexities of data ecosystems – characterized by their multi-actor composition, the diverse priorities of participants, and the evolving nature of data itself - render these processes inherently unpredictable. Uncertainties often stem not only from technical challenges but also from differing interpretations of data's value, expectations

surrounding collaboration, and the trust dynamics among participants (Demchenko et al., 2014; Günther et al., 2017; Micheli et al., 2020).

1.3. The key drivers of data ecosystems at the center of this thesis: participation, sharing and the balance between trust and standardization

This thesis explores three critical dimensions of data ecosystems: the role of participation and collaboration, the willingness to share, and the interplay between trust and standardization. These aspects are central to understanding how data governance processes unfold within multi-stakeholder environments and influence the effectiveness and sustainability of data sharing practices.

Regarding the participation and collaboration, our research shifts the focus from the technical capabilities often highlighted in big data literature (Huynh et al., 2023; Sjödin et al., 2024) to the diverse interpretations individuals have of data and their sharing practices. While previous studies acknowledge the inherent complexity, ambiguity, and uncertainty of data, they frequently propose that these challenges can be mitigated through training and expertise-based approaches. In contrast, we document a plurality of approaches that coexist and interact within ecosystems composed of diverse stakeholders. Collaboration in such ecosystems is shaped not only by technical competencies but also by the motivations and attitudes of participants, which serve as key drivers of data-sharing behavior. By examining the variations in these motivations and attitudes, this thesis highlights the necessity for governance models that accommodate diversity while promoting effective collaboration.

Regarding the willingness to share (WTS), it is considered one of the pivotal factors in the success of data ecosystems, shaped by participants' goals, perceived benefits, and potential risks (Ahrens et al., 2011; N. Khan et al., 2014; Kouzes et al., 2009; Manyika et al., 2011). Motivations for WTS can include the pursuit of competitive advantage, social recognition, or contributions to collective goals (Hillebrand, 2021; Hillebrand et al., 2023; K. K. Kim et al., 2017; Shore et al., 2022; Struminskaya et al., 2021; Susha, van den Broek, et al., 2023; van den Broek & van Veenstra, 2018). However, these motivations are tempered by about data privacy, security, and potential misuse. We argue that the decision to share data is not merely an individual calculation, but it reflects a complex interplay of personal, relational, and contextual factors. Trust emerges as a critical enabler: stakeholders are more likely to share data when they have confidence in the processes and entities involved, as this alleviates fears of misuse (X. Lin & Wang, 2020). Relational dynamics, such as interactions and mutual confidence among stakeholders, significantly influence readiness to share data. These factors, in turn, affect the ecosystem's overall success by shaping the depth and quality of collaboration. Our analysis reveals that WTS in data ecosystems depends on stakeholders' motivational orientations (e.g., social vs. utilitarian), their cognitive evaluation of the risks and opportunities associated with data sharing, and their social assessment of the ecosystem's relational dynamics.

Finally, the interplay between trust and standardization represents another focal point of this thesis. Standardization facilitates data sharing by establishing consistent criteria for data formats and processes, enabling interoperability across systems and simplifying the integration of new participants (Mentel & Hajduk-Stelmachowicz, 2020; Pessanha Santos, 2023). However, the rigidity of standardized practices must be balanced with flexibility to adapt to the evolving needs of ecosystem (Zang et al., 2020). In this context, trust plays an essential role. It underpins stakeholders' willingness to adopt standardized practices, fostering security and cooperation even within rigid frameworks (Hou & Jansen, 2023). When trust is present,

stakeholders are more likely to perceive standardization as a collective benefit rather than a threat to their autonomy or interests. Conversely, a lack of trust can result in resistance to standardization due to fears of misalignment or the misuse of shared data (Steinbruch et al., 2021). Our research shows that while both standardization and trust are widely regarded as essential for governing data ecosystems, stakeholders often express a preference for transitioning from trust-based collaboration to a framework driven by standardization, where trust is replaced by consistent practices. However, our findings reveal a paradox: formal mechanisms designed to reduce reliance on trust frequently reinforce the need for trust-based relational strategies. This paradox highlights the interdependence of trust and standardization, demonstrating that standardization cannot function effectively without trust, yet trust alone is insufficient to sustain long-term engagement within a data ecosystem.

By focusing on participation, willingness to share, and the interplay between trust and standardization, this thesis seeks to answer a central research question: *how do data ecosystems address grand societal challenges through process-oriented and relational mechanisms? In particular, what are the practices that enable data ecosystems to navigate differences in data-sharing practices, calibrate the willingness to share, and find a balance between formal (i.e., standardization) and relational (i.e., trust) practices?*

By weaving together these threads, this thesis takes a relational and process-oriented perspective, moving beyond static analyses of roles and frameworks to capture the dynamic and evolving nature of data ecosystems. It emphasizes how negotiation, adaptation, and collaboration among stakeholders' shape governance practices and enhance ecosystems' capacity to address pressing societal challenges. Through this lens, the study seeks to unravel the complexities of data governance, offering insights into how ecosystems can be designed and managed to achieve both technical robustness and relational resilience. The findings

contribute to a deeper understanding of what drives the success of data ecosystems, providing actionable guidance for unlocking their full potential.

1.4. Context

This thesis builds its foundation upon the MobiDataLab initiative, a European project financed under the Horizon 2020 Research and Innovation Programme. The initiative serves as a compelling case study for examining the development of a novel data ecosystem aimed at transforming urban mobility across Europe. Through an in-depth exploration of this project, this thesis provides valuable insights into the mechanisms through which integrated data solutions can address challenges like sustainability, accessibility, and efficiency in contemporary transportation systems.

Active from January 2021 to January 2024, MobiDataLab aimed to harness the transformative power of data to enhance urban transportation networks. Its core objective was to establish an innovative framework that encouraged collaboration among diverse mobility stakeholders. This participatory approach targeted critical challenges of our time such as congestion, environmental impact, digitalization, and urban connectivity. The project began with a consortium of 10 partners from seven European countries, including academic institutions, ITS providers, policy experts, and transportation associations. Over time, its scope expanded, drawing participation from private enterprises, local communities, and policymakers eager to engage with the emerging mobility ecosystem.

The project's progress hinged on collecting and analyzing diverse mobility datasets from sources such as GPS devices, mobile apps, public transport systems, and traffic sensors. Initially managed by project partners, this data gradually attracted contributions from a broader network of collaborators, reflecting a collective commitment to improving mobility. This inclusive

framework placed a strong emphasis on aligning data-driven solutions with real-world needs, ensuring transparency, accountability, and trust among stakeholders. One of MobiDataLab's most notable contributions was the creation of the Open Knowledge Base (OKB). This resource hub offered a rich collection of materials designed to address obstacles to data sharing in the mobility sector. It included case studies, guidelines, and actionable recommendations for enhancing data usability, accessibility, and quality. The OKB served as a cornerstone for enabling the development of the Transport Cloud - a prototype platform designed to facilitate secure and seamless data exchanges between mobility stakeholders.

In parallel, the project implemented experimental methodologies through Living Labs and Virtual Labs. These innovative settings enable real-time testing and co-creation of solutions, involving local stakeholders such as residents, policymakers, and community organizations. Employing advanced techniques such as machine learning and predictive analytics, the labs transformed raw data into actionable insights that could inform urban mobility planning and optimization efforts.

Community engagement played a pivotal role in MobiDataLab's success. By collaborating with public institutions, transportation agencies, and residents, the project ensured that its outcomes were both technologically advanced and socially relevant. This dual emphasis highlighted the importance of addressing community needs and building trust in data-sharing practices as a foundation for sustainable urban development.

Given its pioneering approach and diverse stakeholder involvement, MobiDataLab consciously prioritized experimentation over immediate commercial outcomes. While commercialization studies were undertaken to explore future business models and market opportunities, the initial focus remained on fostering a culture of collaborative data sharing within a secure and inclusive environment.

Ultimately, MobiDataLab provides a unique perspective for examining the complexities of a data ecosystem. By exploring the motivations, expectations, and concerns of those involved, the initiative provides critical insights into the governance structures and relational dynamics that underpin effective data-sharing practices. These findings contribute significantly to the broader discourse on creating resilient, equitable, and scalable solutions for the future of urban mobility and digitalization.

1.5. Methodology

The methodological framework adopted for this thesis integrates qualitative data collection and grounded theory analysis to build a comprehensive understanding of data ecosystems and their governance. This combined approach facilitated an in-depth examination of how data-sharing practices evolve, particularly in the context of collaborative initiatives like MobiDataLab, while ensuring theoretical insights were rooted in empirical evidence.

The primary data for this thesis were gathered through semi-structured interviews conducted over an 18-month period, from May 2022 to December 2023. These interviews targeted 40 stakeholders representing 18 distinct categories, including academics, researchers, ITS professionals, policymakers, and non-profit organizations, alongside participants from the Living and Virtual Labs facilitated by the MobiDataLab project. Given the absence of predefined ecosystem boundaries, a phased strategy was employed to identify stakeholders. Initially, interviews focused on MobiDataLab's original partners, whose close involvement in the project's operational phases provided crucial entry points. Through snowball sampling, the study expanded to include other stakeholders who joined the ecosystem at later stages, ensuring a holistic understanding of its evolution.

The interviews explored several dimensions of data-sharing practices. Questions were structured to gather information about the organizations' motivations, goals, and current data sharing practices. Subsequent sections examined stakeholders' perceptions of data quality, collaboration dynamics, and the progression of data-sharing approaches within the ecosystem. All interviews were recorded, anonymized, and transcribed in full to ensure accuracy and support comprehensive analysis.

In addition to interview data, extensive archival materials were analyzed to contextualize the findings. These included key project documents such as the MobiDataLab Grant Agreement, which detailed the legal and operational parameters of the initiative, and 24 project deliverables that outlined the objectives, activities, and results achieved throughout the project's lifespan. To situate MobiDataLab within a broader context, the study also examined 12 European, national, and international projects related to data sharing. Publicly available resources such as reports, websites, and social media posts were examined to understand the goals, outputs, and governance structures of these initiatives. This comparative analysis allowed for a richer understanding of the regulatory and operational environment of data ecosystems, as well as variations in their funding mechanisms, governance models, and stakeholder participation.

The qualitative data collected through these multiple channels were systematically analyzed using grounded theory methodology. Interview transcripts and archival data were compiled into a database within the DEDOOSE platform, which provided a structured environment for organizing and coding the material. Open coding was the initial step in the analytical process, aimed at identifying recurring ideas and emergent themes across the data. Building on these initial codes, axial coding was employed to establish relationships between the emerging categories and subcategories. The final stage of analysis involved synthesizing the findings into core categories that encapsulated the central themes of the research. This

iterative process enabled the construction of a coherent narrative, integrating empirical evidence with theoretical insights while contributing to the development of new theoretical constructs.

1.6. Findings

To address the overarching research questions, this thesis draws on three studies conducted during the doctoral journey. The first study investigates the relationship between stakeholders' data-sharing practices and the evolution of data ecosystems. The second examines the interplay of motivational, cognitive, and relational factors that shape data sharing behaviors within these ecosystems. The third explores how stakeholders collaborate to establish data-sharing protocols, highlighting the challenges associated with trust, standardization, and data sharing practices.

The **first study** reveals how data acts as a boundary object, facilitating stakeholder collaboration while embodying the inherent complexity of data ecosystems. This dual role enables stakeholders to develop a shared vision of data sharing, though their perceptions, practices, and involvement vary significantly based on individual objectives. Within the ecosystem, stakeholders navigate a persistent dilemma: should they share data, and if so, under what conditions? This tension arises from three core characteristics of big data: heterogeneity, ambiguity, and uncertainty. These features simultaneously position mobility data as both an opportunity and a challenge, influencing how stakeholders interpret and act upon it. The study highlights that data heterogeneity, stemming from diverse types, collection methods, and standards, offers both benefits and challenges. While heterogeneous data fosters innovative, tailored solutions to societal issues, it complicates the creation of a cohesive and reliable ecosystem. Ambiguity further complicates matters, as mobility data is often interpreted

differently by industries, governments, and researchers, resulting in conflicting visions of its potential applications. Uncertainty exacerbates these challenges, as stakeholders grapple with questions about how, when, and by whom the data will be used, and the potential consequences of such use. The study identifies four distinct stakeholder roles within the data ecosystem: i) explorers, the most proactive actors, drive innovation and collaboration, often exemplifying best practices; ii) exploiters focus on immediate technical challenges, such as standardization, while avoiding long-term commitments to the ecosystem; iii) conservatives adopt a cautious approach, prioritizing risk avoidance over progress and refraining from significant investment; iv) idealists envision a utopian future for data sharing but often contribute minimally to immediate ecosystem development. These roles, along with the tension between innovation-driven and stagnation-prone practices, shape the ecosystem's evolution and underscore the dynamic nature of data sharing.

The **second study** delves deeper into how stakeholders' willingness to share data is shaped by motivational orientations and socio-cognitive evaluations of data sharing and the ecosystem itself. Stakeholders generally approach data sharing from two primary perspectives: some are driven by a desire to contribute to societal impact, while others view data as a utilitarian tool for achieving personal or organizational objectives. These motivations strongly influence their evaluations of data sharing, particularly regarding its perceived benefits, risks, and the collaborative social conditions within the data ecosystem. The research identifies three distinct types of data-sharing practices that reflect varying level of willingness to share and collaborate: i) driving practices are characterized by proactive, co-creative approaches, such as open data sharing and strong governance, often enacted by stakeholders motivated by social impact; ii) pulling practices, in contrast, are risk-averse and aim to align ecosystem activities with specific stakeholder needs and expectations (these are commonly employed by those driven by utilitarian motives); iii) scaffolding practices provide a middle ground, enabling

cautious contributions from stakeholders with differing motivational orientations, balancing perceived risks and benefits. This nuanced understanding of data-sharing practices underscores how ecosystems function as constellations of diverse motivations and collaborative strategies, shaped more by cognitive and social evaluations than by demographic or professional factors.

The **third study** explores the paradoxical relationship between trust and standardization in data ecosystems. While standardization is often pursued as a means to reduce reliance on trust-based interactions, this research finds that it simultaneously amplifies the need for trust-driven relational approaches. Stakeholders recognize the value of standardized practices in reducing uncertainty and improving interoperability, yet these efforts are frequently impeded by challenges such as data heterogeneity, misaligned expectations, and collaborative obstacles. This cyclical dynamic demonstrates the importance of harmonizing these governance mechanisms rather than prioritizing one over the other. The study identifies three forms of trust as crucial to the functioning of data ecosystems: i) interpersonal trust, which fosters smooth collaboration between individuals; ii) institutional trust, which provides confidence in organizational structures; iii) system-based trust, which ensures faith in technological infrastructure and governance mechanisms. Each form of trust complements procedural and technical standardization, highlighting their interdependence. While technological solutions may facilitate trust-free systems, the findings reveal that human trust remains an essential element of successful collaboration. Stakeholders express a strong desire to transition toward standardized practices, yet the research emphasizes that trust continuous to be a cornerstone of effective ecosystem governance. By focusing on harmonizing trust and standardization, this study provides a roadmap for overcoming inherent tensions and creating sustainable governance models, offering actionable insights into the complexities of data ecosystems.

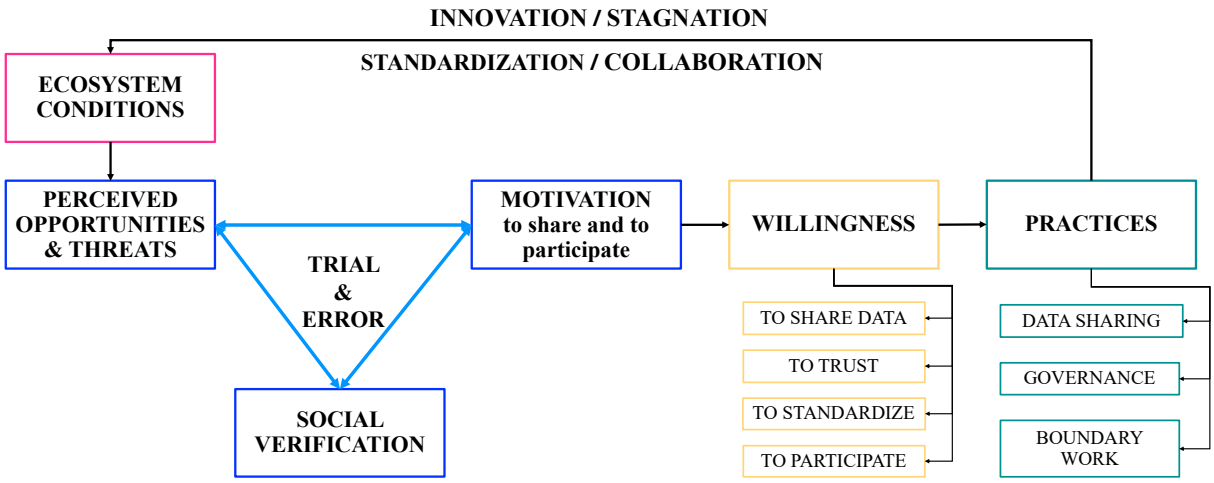

Figure 1: A grounded model of the ecosystem conditions and practices: the interplay between perceived opportunities and threats, the motivation and the social verification. Source: authors own work.

Building on our findings, we developed the model depicted in Figure 1, which provides a comprehensive summary of the conditions and practices within data ecosystems. The model emphasizes the iterative process - akin to trial and error - that occurs between perceived opportunities and threats, individual and collective motivation (to share and to participate), and social verification. This dynamic interplay is pivotal in shaping stakeholders' willingness to share data, trust partners, engage in the definition of standards, and actively participate in the data ecosystem.

Depending on the degree of willingness exhibited by stakeholders, various practices emerge, encompassing data-sharing behaviors, governance approaches, and boundary work practices. These practices are not neutral: their implementation by different actors significantly influences the trajectory of the data ecosystem. Such practices can either drive the ecosystem toward innovation and sustainable growth or lead to stagnation and fragmentation. Moreover, the adoption of specific practices may contribute to a data ecosystem grounded in standardization or, alternatively, one centered on collaboration and negotiation among stakeholders.

This model offers a visual synthesis of the key dynamics explored in this thesis, illustrating how the interplay of ecosystem elements unfolds within a complex socio-technical framework. By mapping these interactions, it provides a nuanced understanding of the factors that shape the evolution and outcomes of data ecosystems.

1.7. Contributions to the literature

The research presented in this thesis provides a multidimensional exploration of data ecosystems, offering insights into their theoretical foundations, practical implications, and the complex interplay among stakeholders, technologies, and governance mechanisms. Across three interrelated papers, the research examines data ecosystems, collaborative data-sharing practices, willingness to share, and the intricate dynamics of trust and standardization in multi-stakeholder environments. Each paper deepens our understanding of how data ecosystems function as socio-technical systems, highlighting both their potential to drive innovation and the challenges posed by their inherent complexities. Through an in-depth examination of stakeholder dynamics, governance frameworks, and data sharing practices, this thesis aim to advance theoretical and practical knowledge of these complex systems.

The first paper makes significant contributions to the literature on data ecosystems (Marcelo et al., 2019; Zuiderwijk et al., 2014), by advancing beyond traditional ecosystem studies (Thomas & Ritala, 2022). It emphasized the critical role of stakeholders' contributions, particularly their perceptions, attitudes, and behaviors toward data sharing (George et al., 2014). Rather than framing the challenges of big data ecosystems as solely related to governance or technical characteristics, as previous studies suggest (Marcelo et al., 2019; Zabel et al., 2023), this paper identifies uncertainties and ambiguities in stakeholders' perceptions and behaviors as central to shaping ecosystem dynamics, as highlighted by Kazantsev et al. (2023). Unlike

traditional research focused on enhancing data management capabilities (Huynh et al., 2023; Sjödin et al., 2024), this work shifts the focus to stakeholders' heterogeneous understanding of data and the plurality of practices within multi-stakeholder environments. The study documents the interactions of distinct stakeholder roles—explorers, exploiters, idealists, and conservatives—and demonstrates how these dynamics influence innovation or stagnation within ecosystems. It underscores the need to move beyond structural categorizations to explore the motivations and attitudes that drive participation. These findings add a critical dimension to the study of ecosystems as collective entities, particularly by enhancing the understanding of uncertainty and negotiation in ecosystem governance (Gomes et al., 2022). By bridging insights on collective action and uncertainty management, the paper portrays big data ecosystems as arenas for negotiation and collaboration (Bakici et al., 2023; Lu & Wang, 2018).

The second paper contributes to a practice-based perspective on data sharing within multi-stakeholder ecosystems. It enriches the literature by exploring how collaborative data practices are shaped by human intentions, with stakeholders engaging based on their specific goals and interests (Günther et al., 2017; Monteiro & Parmiggiani, 2019). This study shifts attention from static structural factors to the dynamic relational patterns that drive data-sharing behaviors. By uncovering the "backstage" processes of data-sharing, it extends the emergent literature on data ecosystems (Aaen et al., 2022; Gelhaar & Otto, 2020; Oliveira et al., 2019) and contributes to the field of information systems, particularly in the study of data work (Aaltonen & Tempini, 2014; Monteiro & Parmiggiani, 2019). The paper highlights the tensions between the widely accepted notion that data sharing generates societal and economic benefits (Raghupathi & Raghupathi, 2014) and the operational and organizational hurdles that create ambivalent cultures of sharing (Gelhaar & Otto, 2020; Klievink et al., 2018). These findings expand the understanding of how stakeholders' perceptions of risks and benefits influence governance decisions behavior (Davidson et al., 2023; Gelhaar & Otto, 2020; Klievink et al.,

2018; Shin & Choi, 2015). Additionally, the study underscores the interplay between stakeholders' evaluations of risks and opportunities and their assessment of the ecosystem's social and relational context. This nuanced perspective challenges static taxonomies of roles and motivations (Davidson et al., 2023; Jarvenpaa & Essén, 2023; Kitsios et al., 2017; Micheli et al., 2020), advocating instead for a dynamic understanding of why stakeholders exhibit diverse attitudes and behaviors toward data sharing. The findings emphasize the need for context-sensitive governance frameworks that address the complexities of decision-making in data ecosystems (Gelhaar & Otto, 2020; Guggenberger et al., 2020).

The third paper investigated the underexplored relationship between trust and standardization in data ecosystem, making important contributions to the literature. While trust and standardization are often treated as separate mechanisms (Backer, 2024; Z. Cao & Lumineau, 2015; Wareham et al., 2014), this study reveals their interconnected and, at times, paradoxical relationship. It addresses key gaps in understanding how trust dynamics and standardization practices influence governance and overall functionality of data ecosystems (Backer, 2024; Z. Cao & Lumineau, 2015). The paper challenges the notion that standardization can entirely replace trust, arguing instead that trust often serves as a prerequisite for the successful development and implementation of standards. This insight enriches studies documenting the ambivalent approaches to data sharing across organizations (Liu et al., 2022) by showing how trust and standardization can simultaneously support and hinder collaboration. The findings highlight the need for adaptive governance models that integrate relational and political practices, moving beyond purely technical solutions (Aaen et al., 2022; Davidson et al., 2023). Furthermore, the research contributes to the broader literature on data and digital ecosystems, which are increasingly recognized as dynamic environments where stakeholders, technologies, and governance interact to drive innovation (Backer, 2024; Z. Cao & Lumineau, 2015; Lioliou et al., 2014; Liu et al., 2022; Micheli et al., 2020; Oliveira et al., 2019; Wareham

et al., 2014). By aligning with calls for robust governance mechanisms that balance multi-stakeholder goals, the paper emphasizes flexibility, inclusivity, and the integration of diverse stakeholder interests. By highlighting the limitations rigid standardization and the fragility of trust, the study offers a renewed conceptualization of governance in data ecosystems. It underscores the importance of harmonizing these mechanisms to overcome tensions and foster sustainable collaboration.

Together, these papers provide a comprehensive view of data ecosystems as dynamic, multi-faceted environments. They emphasize the importance of stakeholder-centric perspectives, where motivations, perceptions, and relational dynamics are central. The research underscores the need for governance models that transcend static frameworks, treating ambiguity and complexity as drivers of innovation rather than obstacles. By challenging conventional approaches and introducing novel frameworks for understanding data sharing, trust, and governance, this thesis establishes a robust foundation for future research into the evolution of data ecosystems. In doing so, it positions these ecosystems as critical arenas for negotiation, collaboration, and adaptation within increasingly complex socio-technical landscapes.

1.8. Implication for Practice

The findings of this thesis carry significant practical implications for those involved in designing, managing, and sustaining data ecosystems. By focusing on the dynamics of data-sharing practices, the research provides a nuanced framework for practitioners to navigate the complexities of multi-stakeholder collaboration. A key insight is the recognition that data ecosystems cannot thrive through uniform approaches. Their success hinges on understanding and addressing the diverse motivations, behaviors, and relationships of stakeholders. For

practitioners, this means moving beyond idealized visions of seamless collaboration and confronting the reality of differing goals, perceptions, and practices. Embracing this diversity allows for the development of strategies that foster inclusivity, resolve conflicts, and build trust.

A central aspect of these implications lies in addressing motivational divergences within ecosystems. Understanding the factors that drive or inhibit stakeholders' willingness to share data is crucial. Policymakers and organizational leaders can leverage these insights to craft tailored interventions. For instance, establishing incentive structures that reward proactive data sharing can encourage participation, while educational initiatives can address uncertainties surrounding data governance, security, and quality. Additionally, collaborative frameworks that promote dialogue and negotiation can help stakeholders align their goals and foster mutual understanding.

This thesis also highlights the duality of trust and standardization as essential elements in ecosystem governance. Trust serves as a fundamental basis for collaboration, yet it is fragile, context-dependent, and often challenging to maintain. Standardization, on the other hand, provides predictability and structure but can lead to rigidity and exclusion. Practitioners must strike a balance between these elements by blending trust-based approaches with flexible mechanisms that adapt to the needs of different participants. Transparent decision-making processes and governance structures can enhance credibility and reduce skepticism, while involving stakeholders in the co-creation of standards fosters both buy-in and adaptability.

Policymakers play a particularly critical role in mitigating risks and maximizing the opportunities within data ecosystems. By breaking down the components of willingness to share, they can identify strategies to reduce dependence on fragile trust while fostering robust collaboration. Regulatory frameworks that establish baseline standards for data security, privacy, and interoperability can create a foundation of trust while allowing for contextual

customization. Public-private partnerships can capitalize on the strengths of various sectors to drive innovation while maintaining accountability and inclusivity. Additionally, conflict resolution mechanisms can address disputes and realign stakeholder goals, ensuring the ecosystem remains resilient and productive.

A key insight from this thesis is the evolutionary nature of data ecosystems. These ecosystems are not static; they mature, adapt, and face new challenges as they evolve. Practitioners must remain vigilant, employing strategies that allow for continuous improvement and resilience. This requires investing in long-term monitoring and evaluation systems to track ecosystem performance and anticipate emerging trends. It also involves recognizing that stakeholders' role and priorities may shift over time, necessitating adaptable governance approaches. Policies and frameworks must evolve alongside ecosystems to accommodate technological advancements and changing stakeholder needs.

The intersection of data ecosystems with emerging technologies, particularly artificial intelligence and machine learning, represents a new frontier for exploration. These technologies rely heavily on shared data, and their integration into ecosystem practices introduces additional complexity. Addressing these challenges requires robust data governance mechanisms that ensure the equitable and responsible use of data while fostering innovation.

1.9. Limitations and future research

The findings of this thesis provide significant contributions to both the literature and practice, while also identifying several promising areas for future exploration, particularly as data ecosystems become increasingly complex and dynamic. Future research could enhance our understanding by examining the intricate relational dynamics among stakeholders. Specifically, investigating how archetypes such as explorers, exploiters, idealists, and conservatives interact,

negotiate, and influence one another within data ecosystems could yield richer insights. These archetypes represent diverse motivations, priorities, and approaches to data sharing, and their interactions may reveal critical tensions, compromises, and collaborative strategies that shape ecosystem's evolution. Understanding how these roles are balanced in practice could clarify the conditions that foster innovation and growth versus those that result in stagnation, conflict, or fragmentation.

The inherent diversity of data ecosystems presents both opportunities and limitations for future research. While this study focuses on ecosystems designed for research and social policy objectives, it does not encompass the full range of data ecosystems. Expanding future investigations to include ecosystems centered on data commercialization, operational data exchanges, or domain-specific applications such as healthcare, mobility, and environmental monitoring could uncover contrasting governance models, data-sharing practices, and stakeholder dynamics. Comparative research across these contexts could validate the findings presented here while uncovering new interaction patterns and collaborative frameworks. For example, commercial ecosystems driven by competitive pressures and profit-oriented goals might display distinct relational tensions and governance challenges compared to those focused on societal benefits.

Further exploration is also warranted regarding the relational dynamics of data-sharing practices identified in this study. Practices such as driving, scaffolding, and pulling reflect diverse strategies used by stakeholders to facilitate data sharing, yet their interactions remain underexplored. Future research could examine how these practices complement or conflict with one another in various ecosystem contexts. For instance, driving practices that promote innovation, and rapid development may clash with scaffolding practices that prioritize stability and incremental progress. Investigating how stakeholders navigate these tensions during their interactions could provide valuable insights into governance strategies that enhance

collaboration while effectively managing conflicts. Moreover, research could explore how these practices evolve over time, adapting to the shifting priorities and challenges of an ecosystem's lifecycle.

A long-term, longitudinal perspective is another critical avenue for research. Data ecosystems are inherently dynamic, with roles, relationships, and governance structures evolving as they mature. This study focuses on an ecosystem in its formative stages, highlighting early development and challenges. However, ecosystems are likely to encounter new dynamics, opportunities, and barriers as they grow and adapt to changing conditions (Alli & Alam, 2020; Hanseth et al., 2006; Lee et al., 2018; Oliveira et al., 2019; Rong & Luo, 2023; Shin & Choi, 2015). Longitudinal studies could track these changes over extended periods, offering a comprehensive understanding of how ecosystems navigate internal and external pressures. Such an approach would enable researchers to identify emerging patterns, anticipate challenges, and propose strategies for sustaining ecosystems over time.

Future research should also examine the broader socio-technical and regulatory landscapes in which data ecosystems operate. Factors such as data protection laws, advancements in data-sharing technologies, and evolving societal attitudes toward privacy, ethics, and data use could significantly influence ecosystem dynamics (Hanseth et al., 2006; Moreno et al., 2019; Van De Hoven et al., 2021). For instance, stricter data privacy regulations might affect stakeholders' willingness to share data, while advancements in interoperability technologies could enable more seamless collaboration (Hanseth et al., 2006; Moreno et al., 2019; Zang et al., 2020). Investigating the interplay between these external factors and internal ecosystem dynamics could lead to more robust and adaptive governance frameworks.

The integration of artificial intelligence (AI), machine learning (ML) and blockchain into data ecosystems represents yet another exciting frontier for research: they add layers of

complexity, as the quality, accessibility, and governance of shared data directly impact the accuracy, fairness, and ethical considerations of these technologies (Alli & Alam, 2020; Hanseth et al., 2006; Van De Hoven et al., 2021). Future studies could explore how ecosystems adapt to the unique demands of AI systems, addressing challenges such as data bias, algorithmic accountability, and the integration of automated decision-making into collaborative data-sharing practices. Understanding these dynamics is essential to ensuring that ecosystems not only foster innovation but also adhere to ethical and equitable principles. Additionally, research could examine how emerging technologies reshape the operational and governance models of data ecosystems. Exploring how these technologies interact with existing ecosystem dynamics could provide valuable insights into their potential to address current limitations and unlock new opportunities.

In summary, the future of research on data ecosystems lies in deepening our understanding of their relational, contextual, and temporal dimensions. By examining how stakeholder dynamics evolve, practices interact, and ecosystems respond to external socio-technical changes, researchers can build a more comprehensive and actionable body of knowledge. Such ongoing exploration will be essential for developing resilient, inclusive, and innovative data ecosystems capable of addressing the societal grand challenges of an increasingly interconnected and data-driven world.

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. https://doi.org/10.1080/0960085X.2021.1947753

Aaltonen, A., & Tempini, N. (2014). Everything counts in large amounts: A critical realist case study on data-based production. Journal of Information Technology, 29(1), 97–110. https://doi.org/10.1057/JIT.2013.29

Agarwal, R., & Dhar, V. (2014). Big data, data science, and analytics: The opportunity and challenge for IS research. Information Systems Research, 25(3), 443–448. https://doi.org/10.1287/ISRE.2014.0546

Ahrens, J., Hendrickson, B., Long, G., Miller, S., Ross, R., & Williams, D. (2011). Data-intensive science in the US DOE: case studies and future challenges. Computing in Science & Engineering, 14–24. https://ieeexplore.ieee.org/abstract/document/5999634/

Alli, A. A., & Alam, M. M. (2020). The fog cloud of things: A survey on concepts, architecture, standards, tools, and applications. In Internet of Things (Netherlands) (Vol. 9). Elsevier B.V. https://doi.org/10.1016/j.iot.2020.100177

Backer, L. C. (2024). Trust platforms: The digitalization of corporate governance and the transformation of trust in polycentric space. Regulation and Governance. https://doi.org/10.1111/rego.12614

Bakici, T., Nemeh, A., & Hazir, O. (2023). Big Data Adoption in Project Management: Insights From French Organizations. IEEE Transactions on Engineering Management, 70(10), 3358–3372. https://doi.org/10.1109/TEM.2021.3091661

Bonina, C., Koskinen, K., Eaton, | Ben, Gawer, | Annabelle, & School, S. B. (2021). Digital platforms for development: Foundations and research agenda. Wiley Online LibraryC Bonina, K Koskinen, B Eaton, A GawerInformation Systems Journal, 2021•Wiley Online Library, 31(6), 869–902. https://doi.org/10.1111/isj.12326

Cao, Z., & Lumineau, F. (2015). Revisiting the interplay between contractual and relational governance: A qualitative and meta-analytic investigation. Journal of Operations Management, 33–34, 15–42. https://doi.org/10.1016/j.jom.2014.09.009

Cazier, J., Cech, T., & Spaulding, T. (2015). Applying Business Analytic Methods To Improve Organizational Performance In The Public School System. Proceedings of the Twenty-First Americas Conference on Information Systems, Puerto Rico, 13–15.

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. https://doi.org/10.1016/J.INFOANDORG.2023.100454

Demchenko, Y., De Laat, C., & Membrey, P. (2014). Defining architecture components of the Big Data Ecosystem. 2014 International Conference on Collaboration Technologies and Systems, CTS 2014, 104–112. https://doi.org/10.1109/CTS.2014.6867550

de Vasconcelos Gomes, L. A., dos Santos, M. G., & Facin, A. L. F. (2022). Uncertainty management in global innovation ecosystems. Technological Forecasting and Social Change, 182, 121787.

Gelhaar, J., & Otto, B. (2020). Challenges in the Emergence of Data Ecosystems. PACIS, 175.

George, G., Haas, M. R., & Pentland, A. (2014). Big Data and Management, 57(2), 321–326. https://doi.org/10.5465/AMJ.2014.4002

Granstrand, O., & Holgersson, M. (2020). Innovation ecosystems: A conceptual review and a new definition. Technovation, 90–91, 102098. https://doi.org/10.1016/J.TECHNOVATION.2019.102098

Guggenberger, T., Schweizer, A., & Urbach, N. (2020). Improving Interorganizational Information Sharing for Vendor Managed Inventory: Toward a Decentralized Information Hub Using Blockchain Technology. IEEE Transactions on Engineering Management, 67(4), 1074–1085. https://doi.org/10.1109/TEM.2020.2978628

Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. The Journal of Strategic Information Systems, 26(3), 191–209. https://doi.org/10.1016/J.JSIS.2017.07.003

Hansen, J., & Pang, Y. S. (2023). Opening industry data: The private sector's role in addressing societal challenges. Data & Policy. https://doi.org/10.1017/dap.2023.15

Hanseth, O., Jacucci, E., Grisot, M., & Aanestad, M. (2006). Reflexive Standardization: Side Effects and Complexity in Standard Making. In Source: MIS Quarterly (Vol. 30).

Helsper, E. J. (2021). The digital disconnect: The social causes and consequences of digital inequalities. SAGE Publications.

Hillebrand, K. (2021). The Role of Fear and Trust When Disclosing Personal Data to Promote Public Health in a Pandemic Crisis. Lecture Notes in Information Systems and Organisation, 46, 247–262. https://doi.org/10.1007/978-3-030-86790-4_18

Hillebrand, K., Hornuf, L., Müller, B., & Vrankar, D. (2023). The social dilemma of big data: Donating personal data to promote social welfare. Information and Organization, 33(1), 100452. https://doi.org/10.1016/J.INFOANDORG.2023.100452

Hou, F., & Jansen, S. (2023). A systematic literature review on trust in the software ecosystem. Empirical Software Engineering, 28(1). https://doi.org/10.1007/s10664-022-10238-y

Huynh, M. T., Nippa, M., & Aichner, T. (2023). Big data analytics capabilities: Patchwork or progress? A systematic review of the status quo and implications for future research. Technological Forecasting and Social Change, 197, 122884.

Iannacci, F., Simeonova, B., & Kawalek, P. (2022). Investigating the determinants of inter-organizational information sharing within criminal justice: A context-mechanism-outcome approach. Journal of Information Technology, 37(2), 188-208.

Jarvenpaa, S. L., & Essén, A. (2023). Data sustainability: Data governance in data infrastructures across technological and human generations. Information and Organization, 33(1), 100449. https://doi.org/10.1016/J.INFOANDORG.2023.100449

Kazantsev, N., Islam, N., Zwiegelaar, J., Brown, A., & Maull, R. (2023). Data sharing for business model innovation in platform ecosystems: From private data to public good. Technological Forecasting and Social Change, 192, 122515. https://doi.org/10.1016/J.TECHFORE.2023.122515

Khan, N., Yaqoob, I., Hashem, I. A. T., Inayat, Z., Mahmoud Ali, W. K., Alam, M., ... & Gani, A. (2014). Big data: survey, technologies, opportunities, and challenges. The scientific world journal, 2014(1), 712826.

Kim, G. H., Trimi, S., & Chung, J. H. (2014). Big-data applications in the government sector. Communications of the ACM, 57(3), 78–85. https://doi.org/10.1145/2500873

Kim, K. K., Sankar, P., Wilson, M. D., & Haynes, S. C. (2017). Factors affecting willingness to share electronic health data among California consumers. BMC Medical Ethics, 18(1), 1–10. https://doi.org/10.1186/S12910-017-0185-X/TABLES/6

Kitsios, F., Papachristos, N., & Kamariotou, M. (2017). Business models for open data ecosystem: Challenges and motivations for entrepreneurship and innovation. IEEE 19th Conference on Business Informatics (CBI), 1, 398–407. https://ieeexplore.ieee.org/abstract/document/8010744/?casa_token=k9a7WXEvKCkAAAAA:2x_MHCN_u1JscIYKwTUcFb5aCFD5x8Feo3-zo7-JaDv0Q_NVnNYFL-yQt3uxzT4_Hl4U5sDx

Klievink, B., Van Der Voort, H., & Veeneman, W. (2018). Creating value through data collaboratives. Information Polity, 23(4), 379–397. https://content.iospress.com/articles/information-polity/ip180070

Kouzes, R. T., Anderson, G. A., Elbert, S. T., Gorton, I., & Gracio, D. K. (2009). The changing paradigm of data-intensive computing. Computer, 42(1), 26-34.

Lee, S. U., Zhu, L., & Jeffery, R. (2018). Designing Data Governance in Platform Ecosystems. Proceedings of the Annual Hawaii International Conference on System Sciences, 2018-January, 5014–5023. https://doi.org/10.24251/HICSS.2018.626

Lin, X., & Wang, X. (2020). Examining gender differences in people's information-sharing decisions on social networking sites. International Journal of Information Management, 50, 45–56.

Lioliou, E., Zimmermann, A., Willcocks, L., & Gao, L. (2014). Formal and relational governance in IT outsourcing: Substitution, complementarity and the role of the psychological contract. Information Systems Journal, 24(6), 503–535. https://doi.org/10.1111/isj.12038

Lis, D., & Otto, B. (2020). Data governance in data ecosystems–insights from organizations.

Liu, B., Shao, Y. F., Liu, G., & Ni, D. (2022). An Evolutionary Analysis of Relational Governance in an Innovation Ecosystem. SAGE Open, 12(2). https://doi.org/10.1177/21582440221093044

Lu, J., & Wang, W. (2018). Managerial conservatism, board independence and corporate innovation. Journal of Corporate Finance, 48, 1–16. https://doi.org/10.1016/J.JCORPFIN.2017.10.016

Lum, K., & Isaac, W. (2016). To predict and serve?. Significance, 13(5), 14-19. https://doi.org/10.1111/j.1740-9713.2016.00960.x

Manyika, J., Chui, M., Brown, B., Bughin, J., & Dobbs, R. (2011). Big data: The next frontier for innovation, competition, and productivity. http://dln.jaipuria.ac.in:8080/jspui/bitstream/123456789/14265/1/mgi_big_data_full_report.pdf

Marcelo, M. I., Barros Lima, G. de F., & Farias Lóscio, B. (2019). Investigations into Data Ecosystems: a systematic mapping study. Knowledge and Information Systems, 61(2), 589–630. https://doi.org/10.1007/S10115-018-1323-6

Mentel, U., & Hajduk-Stelmachowicz, M. (2020). Does standardization have an impact on innovation activity in different countries? Problems and Perspectives in Management, 18(4), 486–503. https://doi.org/10.21511/ppm.18(4).2020.39

Micheli, M., Ponti, M., Craglia, M., & Berti Suman, A. (2020). Emerging models of data governance in the age of datafication. Big Data and Society, 7(2). https://doi.org/10.1177/2053951720948087

Monteiro, E., & Parmiggiani, E. (2019). Synthetic Knowing: The Politics of the Internet of Things. https://doi.org/10.25300/MISQ/2019/13799

Moreno, J., Fernandez, E. B., Serrano, M. A., & Fernandez-Medina, E. (2019). Secure Development of Big Data Ecosystems. IEEE Access, 7, 96604–96619. https://doi.org/10.1109/ACCESS.2019.2929330

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. https://doi.org/10.1016/J.JSIS.2015.02.001

Oliveira, M. I., Barros Lima, G. D. F., & Faris Lóscio, B. (2019). Investigations into data ecosystems: a systematic mapping study. Knowledfe and Information Systems. https://idp.springer.com/authorize/casa?redirect_uri=https://link.springer.com/article/10.1007/s10115-018-1323-6&casa_token=UDAOg80S8HcAAAAA:1NbhNy9IaMq2LxoDsI6h3xj7D6MlXTVbjPkF8yWANqvgnw91_Qm2HWYq2bhoGFl0pOhUNEU8JqjgSJM

Pessanha Santos, N. (2023). The Expansion of Data Science: Dataset Standardization. Standards, 3(4), 400–410. https://doi.org/10.3390/standards3040028

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. Health Information Science and Systems, 2(1). https://doi.org/10.1186/2047-2501-2-3

Rong, K., & Luo, Y. (2023). Toward born sharing: The sharing economy evolution enabled by the digital ecosystems. Technological Forecasting and Social Change, 196, 122776.

Schneider, B., & Olman, L. (2020). The geopolitics of environmental global mapping services: An analysis of Global Forest watch. In A research agenda for environmental geopolitics (pp. 44-57). Edward Elgar Publishing.

Shin, D. H., & Choi, M. J. (2015). Ecological views of big data: Perspectives and issues. Telematics and Informatics, 32(2), 311–320. https://doi.org/10.1016/J.TELE.2014.09.006

Shore, A., Prena, K., & Cummings, J. J. (2022). To share or not to share: Extending Protection Motivation Theory to understand data sharing with the police. Computers in Human Behavior, 130, 107188. https://doi.org/10.1016/J.CHB.2022.107188

Sidze, E. M., Wekesah, F. M., Kisia, L., & Abajobir, A. (2022). Inequalities in Access and Utilization of Maternal, Newborn and Child Health Services in sub-Saharan Africa: A Special Focus on Urban Settings. Maternal and Child Health Journal, 26(2), 250–279. https://doi.org/10.1007/S10995-021-03250-Z/TABLES/2

Sjödin, D., Liljeborg, A., & Mutter, S. (2024). Conceptualizing ecosystem management capabilities: Managing the ecosystem-organization interface. Technological Forecasting and Social Change, 200, 123187.

Spil, T., & Klein, R. (2014). Personal health records success: Why Google health failed and what does that mean for Microsoft HealthVault? Ieeexplore.Ieee.Org, 2818–2827. https://ieeexplore.ieee.org/abstract/document/6758953/

Steinbruch, F. K., Nascimento, L. da S., & de Menezes, D. C. (2021). The role of trust in innovation ecosystems. Journal of Business and Industrial Marketing, 37(1), 195–208. https://doi.org/10.1108/JBIM-08-2020-0395

Struminskaya, B., Toepoel, V., Lugtig, P., Haan, M., Luiten, A., & Schouten, B. (2021). Understanding Willingness to Share Smartphone-Sensor Data. Public Opinion Quarterly, 84(3), 725–759. https://doi.org/10.1093/POQ/NFAA044

Susha, I., Grönlund, Å., & Van Tulder, R. (2019). Data driven social partnerships: Exploring an emergent trend in search of research challenges and questions. Government Information Quarterly, 36(1), 112–128. https://www.sciencedirect.com/science/article/pii/S0740624X17302708

Susha, I., van den Broek, T., van Veenstra, A. F., & Linåker, J. (2023). An ecosystem perspective on developing data collaboratives for addressing societal issues: The role of conveners. Government Information Quarterly, 40(1). https://www.sciencedirect.com/science/article/pii/S0740624X22000995

Thomas, L. D., Autio, E., & Gann, D. M. (2014). Architectural leverage: Putting platforms in context. Academy of management perspectives, 28(2), 198-219. https://doi.org/10.5465/amp.2011.0105

United Nations Development Programme. (2024). United Nations Development Programme. Https://Www.Undp.Org/Sustainable-Development-Goals.

Van De Hoven, J., Comandè, G., Ruggieri, S., Domingo-Ferrer, J., Musiani, F., Giannotti, F., Pratesi, F., & Stauch, M. (2021). Towards a Digital Ecosystem of Trust: Ethical, Legal and Societal Implications. In Opinio Juris In Comparatione. https://shs.hal.science/halshs-03510761

van den Broek, T., & van Veenstra, A. F. (2018). Governance of big data collaborations: How to balance regulatory compliance and disruptive innovation. Technological Forecasting and Social Change, 129, 330–338. https://doi.org/10.1016/J.TECHFORE.2017.09.040

Wareham, J., Fox, P. B., & Giner, J. L. C. (2014). Technology ecosystem governance. Organization Science, 25(4), 1195–1215. https://doi.org/10.1287/orsc.2014.0895

Zabel, C., O'Brien, D., & Social, J. N. (2023). Sensing the Metaverse: The microfoundations of complementor firms' dynamic sensing capabilities in emerging-technology ecosystems. Technological Forecasting and Social Change, 192. https://www.sciencedirect.com/science/article/pii/S0040162523002470

Zang, C. S., Buras, A., Esquivel-Muelbert, A., Jump, A. S., Rigling, A., & Rammig, A. (2020). Standardized drought indices in ecological research: Why one size does not fit all. In Global Change Biology (Vol. 26, Issue 2, pp. 322–324). Blackwell Publishing Ltd. https://doi.org/10.1111/gcb.14809

Zillner, S., Le Bars, L., de Lama, N., Scerri, S., García Robles, A., Tonna, M. C., Kenneally, J., Mayer, D., Hahn, T., Caj, S., Seidl, R., Carbonare, D. D., & Curry, E. (2021). A roadmap to drive adoption of data ecosystems. The Elements of Big Data Value: Foundations of the Research and Innovation Ecosystem, 41–62. https://doi.org/10.1007/978-3-030-68176-0_3/FIGURES/2

Zuiderwijk, A., Janssen, M., & Davis, C. (2014). Innovation with open data: Essential elements of open data ecosystems. Information Polity, 19(1–2), 17–33. https://doi.org/10.3233/IP-140329

2.

Unpacking Dynamics of Big Data Ecosystems in Mobility. Big Data as Boundary Objects and Data Sharing as Boundary Work[2]

Abstract

This study investigates big data ecosystems with the aim of uncovering their specific features and functioning mechanisms. Through qualitative research in a European multi-stakeholder ecosystem focused on data sharing for integrated mobility, we investigate the relationship between stakeholders' data sharing practices and the evolution of big data ecosystems. Drawing on sociological and organizational theories, our findings highlight big data as boundary objects and data sharing as boundary work. We emphasize how stakeholders in big data ecosystems create a collective vision that differentiates their data sharing practices, perceptions, and level of involvement. Specifically, we identify different roles in big data ecosystems based on stakeholders' perceptions of data sharing - how they balance perceived risks and advantages - and their degree of involvement in data sharing practices. We demonstrate how the interactions between different data sharing positions and practices can significantly influence the evolution of a big data ecosystem, leading either to innovation or stagnation. These findings uncover how big data ecosystems evolve and transform into successful initiatives, contributing to the literature on big data ecosystems, and on sharing and collaboration in ecosystems more broadly.

Keywords: big data, big data ecosystems, inter-organizational collaboration, boundary work, boundary objects, mobility.

[2] The paper presented in this chapter is currently under review in the journal *Information Technology and People.* Authors: Giulia Renzi, first author (conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, validation, visualization, writing – original draft), Paula Ungureanu (conceptualization, investigation, methodology, supervision, validation), and Natalia Selini Hadjidimitriou (data curation, writing – review and editing).

2.1. Introduction

It is widely acknowledged that we are living the big data era, which fosters a culture of data sharing. *Big data* refers to datasets so large, complex, and rapidly generated that traditional software cannot handle them (Gandomi and Haider, 2015). A *big data ecosystem* comprises components and collaborations enabling storage, processing, visualization, and delivery of data-driven insights across its lifecycle; serving as essential infrastructure that support various data processes, from acquisition to application (Demchenko et al., 2014).

While public attention often centers on commercial data ecosystems (Zillner et al., 2021), the use of big data for public good is growing across fields such as education (Cazier et al., 2015), healthcare (Raghupathi and Raghupathi, 2014), environmental sustainability (Newell and Marabelli, 2015), public safety and security (Newell and Marabelli, 2015) and government operations (G. H. Kim et al., 2014). Here, actors collaborate to address critical challenges of our times, by unlocking data's potential and shaping inter-organizational and cross-sectoral data sharing. Initiatives known as 'big data ecosystems' (Demchenko et al., 2014) or simply 'data ecosystems' (Lis and Otto, 2020; Oliveira et al., 2019) exemplify these efforts that include open data ecosystems (Yoon and Copeland, 2020; Zuiderwijk et al., 2014) and sector-specific data ecosystems that engage services and guide policy (Demchenko et al., 2014).

The European Commission plays a crucial role in supporting these ecosystems through funding initiatives in areas like mobility, healthcare, and environmental sustainability, which surpass the scope and capabilities of individual stakeholders (Petre et al., 2019). However, understanding the value of data sharing also requires considering stakeholders' perspectives and practices (Q. H. Cao et al., 2016; Mello et al., 2013). Successful ecosystems must navigate heterogeneous data and differing stakeholder goals to achieve value extraction (Lis and Otto, 2020; Marcelo et al., 2019).

Originally borrowed from biology, the term *ecosystem* now describes cross-sector actors co-creating, using, and regulating big data (Jacobides et al., 2018; Krasyuk et al., 2022). It has been argued that big data ecosystems warrant dedicated study due to their unique collaborative challenges in managing, analyzing, integrating, and standardizing big data (Cui et al., 2020; Huynh et al., 2023). However, studies regarding big data often focus on the quality and properties of the data itself (Vranopoulos et al., 2022) or on the capabilities necessary to manage big data (Helfat and Raubitschek, 2018; Huynh et al., 2023) without considering the attitudes, perceptions, and actual behaviors of those who participate in data sharing (Kazantsev et al., 2023). This can reduce big data to static assets, ignoring the dynamic interactions within ecosystems (Walker, 2014).

This study proposes a practice-based view of big data ecosystems, examining big data as boundary objects and data sharing as boundary work. Understanding these dynamics is crucial as the literature still lacks clarity on how big data ecosystems evolve and succeed. This perspective shift is critical, as successful ecosystems require navigating different data standards, interests, and sharing perspectives and practices towards standardization and common sharing practices – yet how this happens remains unclear (Bonina et al., 2021; Cui et al., 2020; Demchenko et al., 2014; Huynh et al., 2023).

By positioning big data as boundary objects and viewing data sharing as boundary work, this study brings these concepts into direct conversation, illustrating how the sociotechnical dynamics of big data shape and are shaped by evolving data ecosystems, illuminating how participants negotiate meaning, alignment, and collaboration.

This study addresses the following research question: *how do stakeholders in big data ecosystems navigate differences in data-sharing perspectives and practices, and what are the consequences for ecosystem organization and evolution?* This theoretical puzzle guides our inquiry into the dual role of big data as both a unifying and differentiating force. From a

boundary object perspective, big data bridge the diverse perceptions, goals, and interests of ecosystem stakeholders. Accordingly, data sharing become a boundary work practice, mediating between collective ecosystem goals and individual stakeholders' private activities (Granstrand and Holgersson, 2020; Kreiner et al., 2009).

Following a qualitative methodology in a big data ecosystem for integrated mobility solutions, results highlight the need to understand diverse data diverse data-sharing approaches to big data sharing. We identify four stakeholder types - idealists, explorers, exploiters, and conservatives – each with distinct data sharing practices shaped by perceived risks, advantages, and their involvement in the ecosystem. These practices influence whether the ecosystem progresses or stagnates through boundary work. Big data, as boundary objects, are found to be concrete enough to align actors with similar aspirations, yet flexible enough to support varied interpretations and uses. This finding suggests that ecosystem evolution is intricately linked to the boundary work undertaken by stakeholders.

This study advocates moving beyond idealized ecosystems models toward a realistic view that sees boundary data work as a strategy to manage risks and harness value across diverse data sharing practices. These findings contribute to the literature on big data ecosystems and broader ecosystems literature on sharing and collaboration, while offering practical guidance for policymakers and stakeholders to support effective ecosystem development.

2.2. Theoretical Framing

2.2.1. The promises of big data and the rise of big data ecosystems

Organizations collect data to create value, predict actions, understand behaviors, and forecast events (George et al., 2014; Rammer et al., 2023). Connecting and interpreting this data to address societal challenges requires complementary perspectives and strategies across organizational and sector boundaries. While giants like Apple and Google dominate attention

with profitable data ecosystems (Lu and Wang, 2018), less is known about data ecosystems promoting innovation and sharing across different domains, such as open data, research data and sector-specific data in healthcare, mobility, or sustainability (Petre et al., 2019; Demchenko et al., 2014).

However, these initiatives are gaining significant traction. The U.S Government, the World Bank, the United Nations, and the European Commission have founded projects to foster data sharing culture, standardization, and collaboration, driving innovation (Werbrouck et al., 2023). An underlying assumption of these projects is that big data can bring significant value to society, from firms to governments and citizens, through a common vision, shared culture, and accepted norms, practices, and standards for big data sharing (Olszak and Mach-Król, 2018). Yet, there is still little understanding of how these ecosystems should function to achieve these goals.

According to Thomas and Ritala (2022), ecosystems are:

> "communities of interdependent yet hierarchically independent heterogeneous participants who collectively generate an ecosystem value proposition" (p. 1).

These ecosystems involve an evolving set of actors, activities, artifacts, institutions, and relationships, including complementary and substitutive elements (Gomes et al., 2018; Granstrand and Holgersson, 2020). The literature on data ecosystems differentiates them based on the characteristics of big data (Rong and Luo, 2023; Shin and Choi, 2015) such as the 5Vs: volume, velocity, variety, value, and veracity of data (Demchenko et al., 2014; Tools et al., 2023). Studies suggest that successful big data ecosystems rely on data quality, standards, profiling, analysis capabilities, and the overall privacy and security (Bonina et al., 2021; Cui et al., 2020; Demchenko et al., 2014).

Despite some attention to technical and social factors, the human factor has received relatively limited attention. Aligning different stakeholders' interests is crucial for the success of big data ecosystems (Pesce et al., 2019; Sjödin et al., 2024). Managing big data requires a sustainable ecosystem where stakeholders participate in a common goal while finding their niches (Shum et al., 2012). Maaroof (2015) identifies different types of stakeholders in a big data ecosystem and reveals the challenge of overcoming different views and aligning interests regarding data sharing (Keser et al., 2012). Other studies highlighted difficulties in creating common expertise pools and visions among stakeholders (Bakici et al., 2021; Tools et al., 2023).

Thus, what is missing is an understanding of how social and technical factors interplay in creating big data ecosystems (Gomes et al., 2018). George et al., (2014) highlighted the need to unpack how perceptions, attitudes, and behavior patterns regarding ubiquitous data contribute to value creation and capture (Kazantsev et al., 2023). Referring to innovation ecosystems, Vasconcelos Gomes et al. (2018, 2022) and Granstrand and Holdersson (2020) suggest focusing on how activities and relationships in an ecosystem interplay and contribute to its co-evolution.

Building on Thomas and Ritala's (2022) conceptualizations of ecosystems, we propose moving beyond viewing big data as static entities. Instead, we embrace a broader social practice perspective where big data are collective, socially negotiated objects that change properties as stakeholders interact. To this end, we turn to the boundary objects literature.

2.2.2. Big data as boundary objects and data sharing as boundary work

Despite increasing research, the definition of big data is often assumed rather than clarified (De Mauro et al., 2015; Olszak and Mach-Król, 2018). We previously referred to big data as datasets so vast, complex, and rapidly produced that conventional data-processing software is inadequate for managing them (Gandomi and Haider., 2015). However, the lack of consensus

on the definition of big data leads to misunderstandings and varying visions, ranging from entities that can be hoarded, quantified, accumulated, and exchanged to processes of value extraction, institutional opportunities, and ethical dilemmas (De Mauro et al., 2015). Some authors suggest that this ambiguity is a primary obstacle to successful big data ecosystems: according to Singh *et al*., (2019) the open-source systems dominating big data have created a complex ecosystem of new frameworks, tools, and libraries, that are being released nearly daily, leading to confusion among technologists.

A broader social practice perspective on big data as boundary objects suggests that varying understandings of big data are natural social phenomena with different consequences depending on their social functions. To further explore this perspective, we turn to social studies on boundary work and on boundary objects.

Initially coined by sociologist Thomas Gieryn (1983), the term "boundary work" refers to strategies used by scientists to distinguish science from non-science, legitimizing their practices. Over time, this concept has been applied to multiple fields of human activity, encompassing efforts to influence the social, symbolic, material, and temporal boundaries affecting groups, occupations, and organizations (Lamont and Molnár, 2003). Numerous studies have documented the dynamics of inclusion and exclusion especially in large-scale projects requiring cross-boundary collaboration (Bucher et al., 2016; Mørk et al., 2012). Examples include functionally differentiated organizations, cross-sector collaboration projects, technological innovation ecosystems, and professional service firms serving global clients (Austin and Seitanidi, 2012; Barley et al., 2017; Kellogg et al., 2006; Lifshitz-Assaf, 2018; Suddaby et al., 2007; van Tulder et al., 2016). These situations pose unique challenges to organizational actors who need to collaborate and share knowledge with unfamiliar members from other organizations, often lacking the structure, time, or incentives to interact meaningfully (Anteby et al., 2016; Barley et al., 2017).

Boundary work strategies help members of different organizations balance between sharing and withholding knowledge, defending their identity, and building new identities simultaneously (Bechky, 2003; Carlile, 2004; Boland and Tenkasi, 1995). To enact boundary work, individuals often leverage boundary objects (Bourdieu, 2020; Henderson, 1991; Boland & Tenkasi, 1995; Star and Griesemer, 1989). These are:

> "objects which are both plastics enough to adapt to local needs and the constraints of the several parties employing them, yet robust enough to maintain a common identity across sites. (…) They have different meanings in different social worlds, but their structure is common enough to more than one world to make them recognizable, a means of translation. The creation and management of boundary objects is a key process in developing and maintaining coherence across intersecting social worlds." (Star and Griesemer, 1989, p.393).

Star and Griesemer (1989) study of the Museum of Vertebrate Zoology at the University of California, Berkley, showed how differentiation and cooperation coexist as stakeholders use the museum as a boundary object to negotiate knowledge and relationships. While studies have identified multiple boundary objects such as physical objects, digital artifacts, and cultural symbols, any social object can become a boundary object if used as a tool for boundary work (Clarke et al., 2016).

Given these considerations, it is evident that boundary work and boundary objects can be useful for studying the behavior of stakeholders in a big data ecosystem. Boundary objects align with de Granstrand and Holgersson's definition of ecosystems as an evolving set of actors, activities, artifacts, institutions, and relations (2020). Perceiving big data as boundary artifacts helps us to understand how stakeholders with diverse backgrounds, experiences, and visions about big data may collaborate for a common goal while pursuing personal goals within digital ecosystems. Viewing big data as boundary objects helps to overcome static and dichotomous views of big data as either shared or not shared and embraces a more nuanced understanding of big data as relational practices constituting sociotechnical systems.

2.3. Methods

2.3.1. Context: The MobiDataLab project

Our analysis focuses on the European project called MobiDataLab (Labs for prototyping future Mobility Data sharing cloud solutions), an EU-funded project under the Horizon 2020 Research and Innovation Program (H2020). Chosen for its exemplary collaborative big data environment, MobiDataLab unites diverse stakeholders to achieve common goals and drive innovation through effective data exchange and utilization. Running from January 2021 to January 2024, MobiDataLab aimed to foster data sharing practices in the mobility sector. It tackled urban mobility challenges by gathering, analyzing, and interpreting large datasets. The project aimed to enhance efficiency, sustainability, and accessibility of urban mobility. Data from GPS devices, mobile apps, public transit systems, and traffic sensors, were processed using advanced techniques like machine learning and statistical modeling to gain insights into travel patterns, congestion, and infrastructure usage. As part of Horizon 2020, MobiDataLab integrated research, dissemination, and commercialization efforts, engaging researchers studying mobility trends, private organizations identifying mobility network bottlenecks and opportunities for value creation, policymakers involved in urban planning and mobility policy, and local communities seeking to enhance mobility and living conditions.

MobiDataLab was structured around four key pillars to foster a dynamic big data ecosystem in the mobility sector: i) establishing an open knowledge base with resources on data sharing challenges, best practices, standardization efforts, real-world use cases analyses, and recommendations for enhancing data quality, accessibility, and usability; ii) developing the Transport Cloud, a prototype cloud-based platform for seamless mobility data sharing among stakeholders; iii) implementing Living and Virtual Labs for real-time data exchange and co-creation sessions; iv) conducting complementary activities to foster a data sharing culture within the project and the broader mobility sector.

The project supported sharing different types of data, from open research data to technical data on standards and cloud solutions, as well as market analyses of the mobility sector. While the project did not focus on data commercialization, stakeholders also explored innovative business models. Starting with 10 partners from various sectors across seven countries (further details in Table 1), MobiDataLab attracted additional stakeholders through its Living and Virtual Labs. These stakeholders benefited from information to address service and data sharing gaps, analyze challenges associated with new mobility services, and tackle obstacles to interoperability, connectivity, and standardization.

MobiDataLab provided a unique opportunity to study the establishment of a big data ecosystem, offering insights into stakeholder perceptions, concerns, and expectations regarding data sharing practices. Additionally, it allowed us to explore stakeholder perspectives on mobility data sharing, understanding how such perspectives shaped the formation of a big data ecosystem. Given its alignment with Sustainable Development Goals 11 (SDGs - https://sdgs.un.org/goals): *Make cities and human settlements inclusive, safe, resilient, and sustainable (including sustainable transport)*, MobiDataLab also offered a rich landscape for examining big data as boundary objects and data sharing as boundary work.

2.3.2. Data collection and Data Sources

To enhance theory construction, we gathered information from diverse sources, including semi-structured interviews and archival records.

Interviews

We interviewed key stakeholders in the MobiDataLab ecosystem to understand their perspectives, objectives, and readiness for data sharing in integrated mobility solutions. Initially, we engaged with MobiDataLab's original partners, including academics, researchers,

Information Technology specialists (IT), members of a nonprofit association, and policymakers. One authors' involvement in in the MobiDataLab project facilitated access to these stakeholders and helped establish connections with project partners and the project coordinator. This network allowed us to involve additional stakeholders who joined the ecosystem, expanding our sample to other pertinent stakeholders engaged in Living and Virtual Labs activities throughout the project.

Between May 2022 and December 2023, we conducted 40 semi-structured interviews covering 19 distinct stakeholder categories (see Table 1). All interviews were recorded, anonymized, and fully transcribed. The interview protocol consisted of four sections: 1) general information about interviewees and their organizations' motivations, objectives, and data sharing practices; 2) perceptions of data quality and its role in data sharing; 3) collaboration practices, including forms and types of data sharing activities; and 4) stakeholders' visions and practices concerning big data and data sharing within the ecosystem.

Archival data

We began by compiling archival data from the MobiDataLab Grant Agreement, which outlines the legal rights and responsibilities of both the European Commission and the project partners. Additionally, we gathered and analyzed the 24 public deliverables produced within MobiDataLab to understand the project's primary endeavors, objectives, and duties.

To provide a comprehensive context for MobiDataLab, we also examined 12 past and ongoing European, national, and international initiatives and projects related to big data and data sharing (see Table 1). By reviewing these projects' websites, social media, and reports, which outlined goals, outcomes, and accomplishments, we gained broader insights into big data ecosystems across Europe. This examination helped us to understand the regulatory framework, explore various forms of big data ecosystems, and analyze the articulation of diverse interests,

objectives, and practices within these ecosystems. We also delve into the mechanisms of participation and governance.

The selection of initiatives for this analysis was guided by a projects review conducted by the authors and supplemented by recommendations from interviewees. This approach facilitated a more thorough contextualization of the attributes and practices of MobiDataLab.

Table 1: Summary of the sources of data exploited to support the theory building process. Source: Authors own work.

Source	Amount and Description	Details
Semi-structured interviews	*Total number:* 40, fully recorded-fully transcribed with 19 categories of mobility data sharing stakeholders. Average length: 60 minutes. *Timespan:* May 2022 – December 2023.	Interviewed 8 partners of the MobiDataLab consortium (4 ITS providers, 2 academic and research institutions, 1 policy maker, 1 association of municipalities) spanning 7 countries. Expanded interviews to participants in MobiDataLab Living and Virtual Labs who took part in the mobility challenges and solutions identified by MobiDataLab and used snowball procedures from these stakeholders to further expand sample. The final sample included the following categories of actors: • Researchers & Academics (4). • Transport operators (4). • Citizens involved in data sharing projects (2). • Policymakers (3). • Municipalities (2). • Association of Municipalities (2). • Government (1). • Public Transport Authorities (1). • Rail Infrastructure Authorities (1). • Traffic Management Centre (1). • Trade Associations (2). • Logistics Operators (5). • Airlines (1). • Intelligent Transport Systems (ITS) Providers (5). • Software Providers (2). • Cloud Providers (1). • Car Manufacturers (2) • Emergency services (1)

Source	Amount and Description	Details
Archival data	MobiDataLab documentation Total number: 27 *Timespan:* 2021 – 2023	*Type of documents:* deliverables produced within the MobiDataLab project. A complete overview of the MobiDataLab deliverables can be found at the following link: https://mobidatalab.eu/publications/
	Related projects analysis: Total number: 12 *Documents and topics analyzed:* website, social media, reports on consortium composition, main objectives, main outcomes.	Related projects: • AEOLIX: https://aeolix.eu • SELIS: https://selisproject.eu • FENIX: https://fenix-network.eu • NAPCORE project: https://napcore.eu/ • Inspire initiatives: https://inspire.ec.europa.eu/ • Le Fabrique du mobilite: https://lafabriquedesmobilites.fr/ • Mobility as a service: https://www.maas4eu.eu/ • Convex: https://convex-project.de/index.html • Openstreetmap: https://www.openstreetmap.org/ • TALOS: https://cordis.europa.eu/project/id/218081/reporting • Transmodel Norm: https://www.transmodel-cen.eu/use-of-the-transmodel/ • W3C (World Wide Web Consortium): https://www.w3.org/

2.4. Data analysis

The interview recordings were transcribed verbatim, anonymized manually, and enrich with fields notes detailing our observations. They were imported into an integrated database stored on the QACDAS platform, DEDOOSE, a cross-platform application created for analyzing qualitative and mixed methods research (https://www.dedoose.com). Using grounded theory, we systematically classified the written material into identified categories of similar meanings, comparing empirical data with concepts. This method involved integrate empirical data with existing theories and generate new ones (Strauss and Corbin, 1998).

We started by open coding of recurring concepts, as established by grounded theory, identifying multiple codes. Moving from these initial codes to broader themes, we conducted axial coding to systematically organizing data by connecting categories and subcategories (Strauss and Corbin, 1998).

We discerned common underlying meanings behind respondents' use of different terms. For instance, we categorized advantages and risks of data sharing by meaning rather than specific words. From first-order codes, we proceeded to analyzing second-order themes, focusing on more abstract theoretical constructs, as advised by Strauss and Corbin (1998). Concepts such as trust, standardization, common vision, best practices in data sharing, and pioneers emergent prominently in our dataset, leading to the creation of second-order codes.

In the third level of analysis, we developed even more abstract categories that encompassed various meanings and relationships among second-order themes. This iterative process involved navigation between existing literature and our emergent empirical analysis, consistent with grounded theory methodology (Strauss and Corbin, 1998). The data structure of our analysis can be seen in Figure 2.

Our data analysis culminated in a grounded model depicting big data as boundary objects within big data ecosystems, depicted Figure 3 and elaborated in subsequent sections.

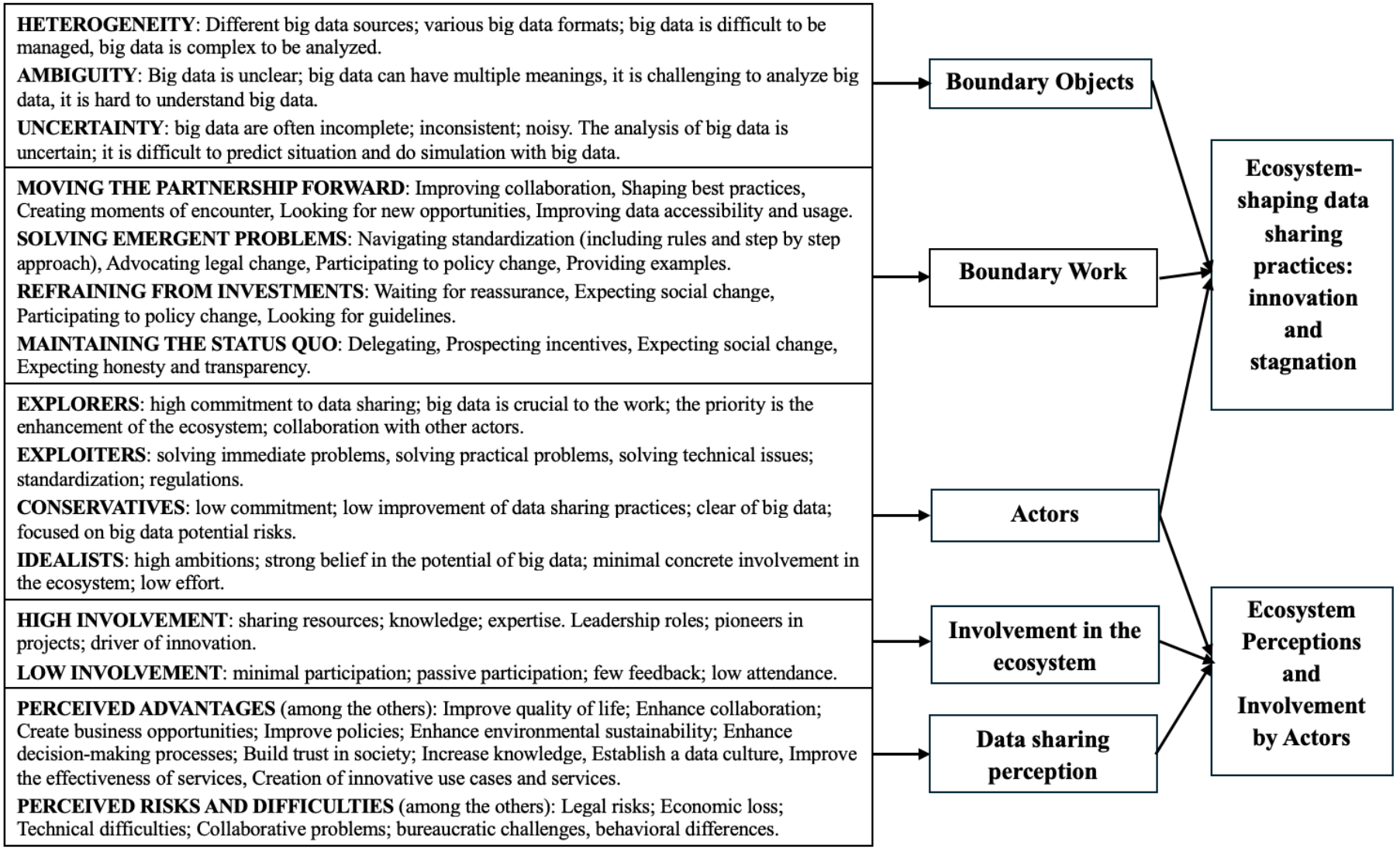

Figure 2: Data Structure. Source: Authors own work.

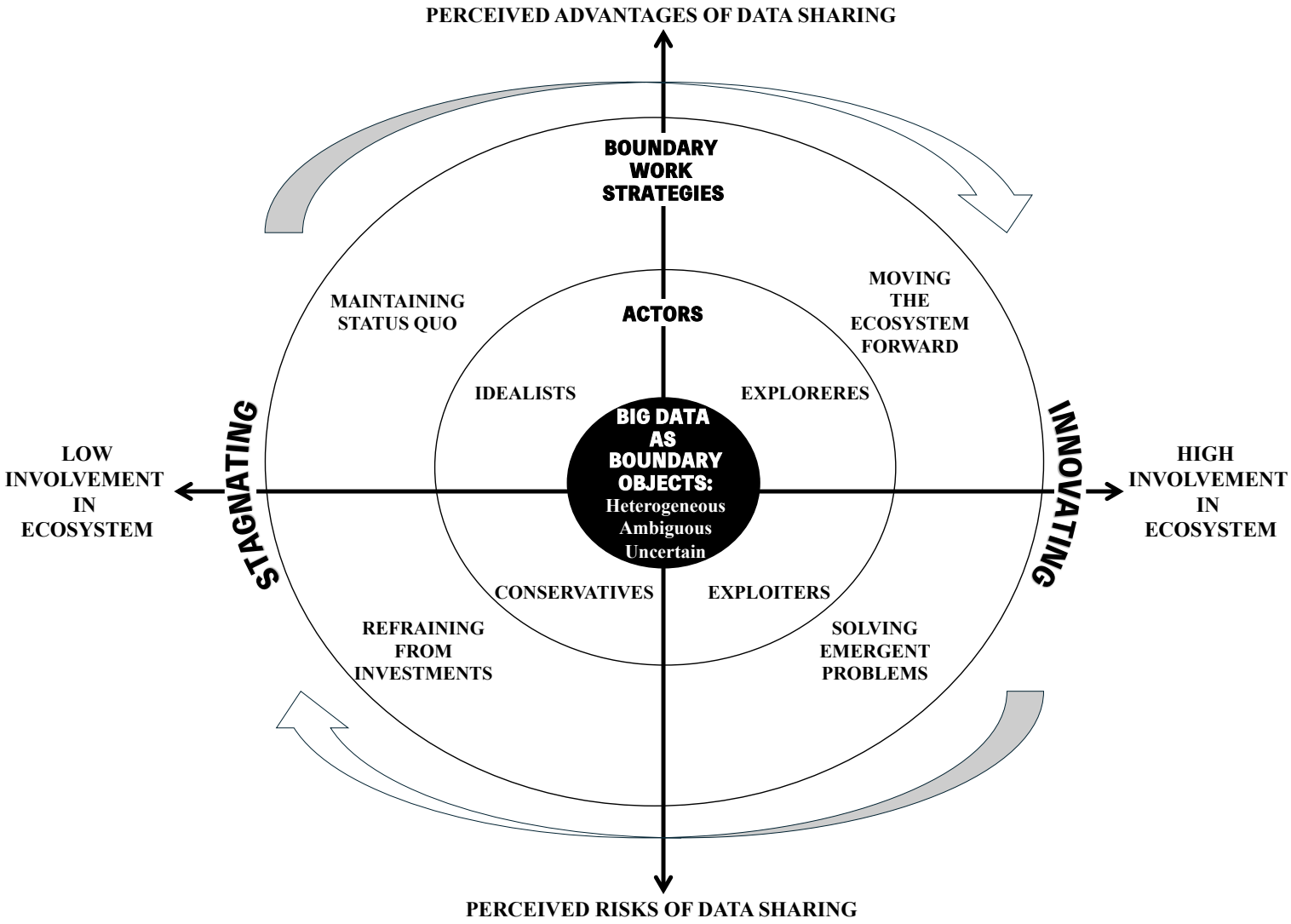

Figure 3: A grounded model of big data as boundary objects in data sharing ecosystems. Source: Authors own work.

2.5. Findings

Our findings reveal that big data can act as boundary objects and data sharing practices as boundary work. This enables stakeholders in big data ecosystems to develop a common vision of data sharing that unifies the ecosystem while significantly differentiating perceptions, involvement, and practices according to their specific objectives. Our analysis identified various types of actors within a big data ecosystem, each engaging in different data sharing practices based on their relationship with big data and data sharing. Stakeholders' positions within the ecosystem depended on their perception of the balance between risks and advantages and their degree of involvement in data sharing practices, as explained in the following sections. We demonstrate that interactions between different data sharing positions and practices can significantly influence the evolution of a big data ecosystem, either fostering generative innovation or leading to stagnation.

2.5.1. Big data as boundary objects

To understand Figure 3, it is necessary to start from the center. Our analyses revealed that the ecosystem's stakeholders had very different understandings of the value of their own and others' mobility data. They all faced the dilemma: '*Share or not share data?*', and '*Is it worth improving data sharing practices in the ecosystem?*'. From this standpoint, big data acted as a boundary object, guiding stakeholders' interpretations, and action.

Three main features of big data influenced the data sharing dilemma: heterogeneity, ambiguity and uncertainty regarding the usability and scalability of mobility data. These features qualified mobility data as a boundary object because stakeholders viewed them as both advantages and threats to data sharing. They also justified stakeholders' varying levels of involvement in the ecosystem and their different data sharing practices. Thus, big data acted as

a trigger for stakeholders' boundary work in the ecosystem. Table 2 below provides examples of the advantages and risks associated with each feature.

Table 2: Big data as boundary objects: advantages and risks associated with heterogeneity, ambiguity and uncertainty. Source: Authors own work.

Big data characteristics	**Perceived advantages**	**Perceived risks**
Heterogeneity	**Creating more interesting use cases and services**: "If you share the data and you know with who you share the data, if there's a contract, for me there's no risk since you know who's going to use the data and if he/she does use it in a bad way, you just cut him/her from the data source. So, for me, there's no real risk. It is more an opportunity because with data, it is sure that we'll be able to have more interesting use cases" (Consultant 1). **Improving the effectiveness of services, environmental sustainability, and people's quality of life**: "The opportunity is to give the end user the best travel experience. Imagine when you are in a town that you don't know. Good data sharing and also the application that this data use will give you the full information on how to move around this place that you don't know" (Trip planner 1). **Creating a data culture based on data value extraction**: "I think the opportunity is that the more data you share, the more value you realize (data) becomes an asset to be curated and to be exploited at least from a public health authority perspective, that's what we've seen in the last 10 years with this movement towards open data" (Project Manager 2 in an Association of municipalities).	**Legal, technological, and organizational challenges**: "We have already a lot of data shared by users about their mobility habits just using phone. The issue is that, in my experience, most of these data are not accurate, so they must be analyzed a lot before being used to find solution of mobility. And this is because many times there is no common standard on how this data should be collected." (Consultant 1). **Difficult Collaboration between different stakeholders:** "So, we have a quite heterogeneous set of stakeholders coming from a law department, from community building, public relations stakeholders, companies, research centers. Every stakeholder has a different view on the project. This was in the beginning, and this will be at the end as well and we will narrow these perspectives." (ITS provider 1).
Ambiguity	**Envisioning different opportunities in the data:** "The data can improve life in so many ways. It is clear that if I have a more massive, punctual, and efficient data exchange, and I can make my work more efficient, I can save on consumption and pollute less. I will have fewer running trucks waiting outside the terminal, I will make the cranes work less, because I have the data on the position of the containers, it is a cascade effect. The train, if I know it arrives earlier, I can arrange the means to work it, I can better organize	**Fear of misunderstandings:** "If we are not able to understand what has been shared, we cannot trust this this information" (Consultant, 3). **Fear of making mistakes:** "Here is the fear of being accountable for mistakes or a misunderstanding and bad use of personal data, for instance. So that can prevent services from going further into

Big data characteristics	Perceived advantages	Perceived risks
	people's shifts, fewer people who drive on the street with the car. A more efficient data transmission has a positive effect on sustainability and on social issues, less acoustic noise, and less traffic". (CEO 1, railway transport operator). **Clarifying common standards, best practices:** "Some advantages are not available immediately, but after a period of time when there is full implementation when it's learning from different practices. As a community, you always have to make investments for data sharing (…) and you might have to change your way of working" (project manager 4 in an airline company).	data sharing because, well, we are not sure that if what we are doing is legal, let's say so, that might be a big trigger for public actors" (Software provider, 2). **Lack of integration:** "I think actors are committed in the ecosystem. I think the problem at the moment is that each partner is working in its domain, and I think it still needs to come together. That's what I'm missing" (project manager 1 in an association of municipalities).
Uncertainty	**Creating large scale alignment:** "We still miss common standards and rules to be followed by everyone to share data in the same way. In our sector, (…) there are no rules that we are obliged to follow when we must collect and share specific kind of data and it is so challenging to find partners that are able to share the data you need (…) but if we manage to create an integrated system each of us will be able to extract exactly what we need, even if we don't know what exactly the others need" (Public transport operator, 2). **Seeing future possibilities:** "We really have the feeling that, (…) the companies really have this vision to create this digitized process and that they are really into this Community feeling to strive this common goal of achieving an improved, more efficient digitized process" (project manager 1 in an airline company).	**Uncertainty about data use and fear of giving up data**: "The big other problem is the fear of giving up one's data, because there is no awareness among the logistics nodes, and no one gives it to them, to then say how that data will be processed, and by whom above all, because as long as the relationship is from A to B, I am calm, the data goes from me to you and they are used so that you can transport (…). It is different to make data available to a community, to a system." (Rail Infrastructure Authority 1). **Uncertainty about ecosystem scalability:** "We are all working at a different speed with a different attention towards the subject of digitalization. And that makes it quite difficult to put in your effort and to collaborate on a national scale (…)" (Municipality representative, 3).

Heterogeneity

First, all MobiDataLab stakeholders highlighted the heterogeneous nature of big data in the mobility sector, arising from diverse data types, sources, standards, collection methods, and usage practices. While heterogeneous data offers opportunities to create tailored solutions for

various societal issues, it also poses challenges in creating a reliable data ecosystem. Key advantages mentioned include innovative use cases, improved services effectiveness, enhanced environmental sustainability, better quality of life, and a data culture focused on value extraction (see Table 2 for examples). However, stakeholders also noted significant challenges, such as legal, behavioral, collaborative, economic, and technical issues.

As explained by our informants, a major challenge of data heterogeneity is the variety of stakeholders in the ecosystem with differing visions and objectives. Most stakeholders stress the need for strategies to overcome heterogeneity through collaboration, trust, and legal, economic, and technical support systems. The ultimate goal is to create a "win-win situation" where each stakeholder benefits from heterogeneous data, as shown in Table 2.

Ambiguity

Data was then frequently described as ambiguous. Since mobility data was held by different stakeholders, it had different meanings for each. As a result, stakeholder envisioned and enacted different, often conflicting, opportunities around big data. For instance, industry stakeholders saw opportunities to understand markets and customer needs, governments saw ways to improve citizen welfare and justify investment, policymakers aimed to innovative policy-making practices and legitimize their operations, while researchers sought data quality and quantity to expand projects and databases for future research. Other stakeholders, like trip planners and software providers focused on improving knowledge, forecasting, and efficient connections in the fragmented mobility services industry.

Due to these varies understandings, the potential uses of big data in mobility were ambiguous and open to multiple interpretations. This ambiguity generated both enthusiasm and fear. A lack of standardization was seen as a root cause of this ambiguity (see Table 2). Stakeholders feared that differing interpretations would slow progress and prevent the

development of a common data sharing culture. Disambiguating data required significant coordination efforts, leading to concerns about increased data sharing costs, as shown in Table 2.

In sum, stakeholders acknowledged the ambiguity of mobility big data, viewing it as both an opportunity and a threat. Many anticipated a future where doubts about mobility big data would be clarified, and new alignments created. This future implied standardizing data sharing practices to clarify what data should be shared, how, and under what conditions.

Uncertainty

Lastly, stakeholders expressed uncertainty about how, why, and when the data would be used, by whom, and with what consequences. Citizens' representatives or government officials often doubted how private stakeholders, like mobility operators, would use the data or if their actions would jeopardize them (see Table 2).

Beyond usability concerns, there were both hype and fears about the ecosystem's scalability. Stakeholders worried that others might not take data sharing seriously, investing minimal effort and resources. This implied that standardization sharing practices could become a slow, complicated, and resource-intensive process. However, uncertainty also presented future possibilities. The idea of a mobility ecosystem with widespread participation and diverse data opened opportunities that stakeholders were beginning to see: the hope of a quickly scaling ecosystem was often mentioned as a core reason for participation (see Table 2).

Overall, big data as a boundary object - heterogeneous, ambiguous, and uncertain - contributed to a sharing dilemma, triggering stakeholders' boundary work within the ecosystem. While everyone envisioned benefits from enhanced, integrated mobility data, they also feared its limited usability, doubting the ecosystem's success. The heterogeneous, ambiguous, and

uncertain nature of mobility data contributed to this dilemma, leading stakeholders to develop varied perceptions and data sharing practices, as discussed in the following section.

2.5.2. Boundary work and the ecosystem stakeholder matrix

Based on their different views on the value of mobility big data, the stakeholders of the nascent MobiDataLab ecosystem engaged in various data sharing strategies to define, negotiate, and maintain boundaries between their activities. We called these strategies "data sharing boundary work". We identified four types of actors in the ecosystem: explorers, exploiters, conservatives, and idealists. These actors enacted boundary work practices, by moving the ecosystem forward, solving emerging problems, refraining from investment, and maintaining the status quo, respectively.

Our analyses found no correlation between these practices and the socio-demographic characteristics of stakeholders or their roles in the ecosystem. Instead, these actors deferred based on their boundary work strategy, which depended on their relationship with big data. Specifically, it hinged on how stakeholders perceived the balance between data sharing risks and advantages and their level of involvement in the ecosystem.

Explorers: moving the ecosystem forward

In our research, explorers are actors within the ecosystem who show high willingness and commitment to data sharing. They see data as central to their work and align their organizations' interests with the broader goal of enhancing the ecosystem. Explorers drive innovation, efficiency, and collaboration through continuous improvement. Their role as catalysts comes from a boundary work strategy that prioritizes active mitigation of risks and capitalization on opportunities. While aware of the risks and challenges of data sharing, explorers focus on

overcoming them, emphasizing benefits for themselves and the ecosystem. This approach fosters an optimistic and community-driven culture, as shown in Table 3.

Explorers prioritize collaboration to advance the MobiDataLab ecosystem. They are willing to shares heterogeneous, ambiguous, and uncertain mobility data, trusting others to engage in continuous collaboration. Consequently, they expect everyone to collaborate, minimizing selfish behaviors and fostering constant exchanges. As the most active participants in the data ecosystem, explorers act as pioneers setting positive examples. They initiate best practices, provide inspiration, and improve data accessibility and usability. During the early stages of ecosystem creation, explorers analyzed the mobility sector, identified stakeholders' needs, and systematized information into a common knowledge base accessible to all.

Additionally, explorers led the creation of a cloud-based platform prototype for searching, accessing, and integrating multimodal mobility data. This platform showcased tools to facilitate data access and exchange, support data standard conversions, and connect data and service providers through interoperable interfaces. Explorers also contributed to ecosystem scalability by organizing workshops, hackathons and living labs to expand the network, seek new business advantages, and share success and challenges. In particular, living labs, open innovation initiatives, played a key role in validating and improving the cloud-based prototype through real use sessions. These labs served to assess and measure the data sharing culture in the nascent ecosystem and generate scalable ways to evaluate and develop MobiDataLab tools and technologies. The following field notes serve to better understand the MobiDataLab goals, organizations, and structure, while in Table 3 field notes from actors labeled as explorers can be found.

"More specifically, the main activities (within the Living Labs) were: identify, develop and contribute to the creation of open tools; allow the discovery of data for operations or research, static and dynamic, real-time data and historized data (Findable data); provide centralized access to mobility data (Accessible data); ease the conversion from one standard to the other (…) (Interoperable data); trigger and provoke interoperable interfaces to connect data providers, service providers and mobility clouds;

prototype data processors for adding value to the data (…); demonstrate anonymization and privacy-preserving tools, in line with private data sharing principles" (MobiDataLab Grant agreement).

"Since MobiDataLab is concentrated on a holistic approach of data exchange that will involve actors ranging from municipalities to private entities and citizens, we wanted to examine whether a living Lab approach would be suitable for the scope of our project. Focus is put on the integration of potential users, where innovation processes should be accelerated, as a quicker step from an idea to a market-ready and demand-oriented product or service" (MobiDataLab Deliverable 3.1).

Exploiters: solving emerging problems

Exploiters are actors focused on solving immediate problems, particularly technical issues like standards and regulation. Their primary concern is standardization, allowing them to work with business partner without seeking commonly agreed long-term solutions. For instance, during MobiDataLab, these stakeholders contributed both to the creation of the knowledge base and to the setup of the cloud solution prototype, but they focused on technical aspects such as analyzing mobility data standards and developing technical specifications for the cloud-based prototype. They clarify doubts, inform stakeholders about standard's pros and cons, and address high risk associated with a lack of standardization.

The following excerpt exemplify how their technical and problem-oriented interventions stem from their view of big data as heterogeneous, ambiguous, and uncertain, and specifically from the high risks that they associate to lack of standardization:

"Given the heterogeneity and peculiarities of mobility data, and the various constraints related to their safe and trusted sharing, MobiDataLab envisions the usage of an open federated cloud architecture where complex and often contrasting requirements coming from FAIR (Findability, Accessibility, Interoperability, and Reusability) and privacy principles can be enforced easily and practically. (…) A federated architecture would be beneficial for MobiDataLab given its main objective of fostering the sharing of data amongst transport authorities, operators and other mobility stakeholders operating in Europe which in most of the cases want to maintain the governance of their data" (MobiDataLab, Deliverable D2.6).

Externally, exploiters participate in working groups on data sharing regulations and policies beneficial to their position and the ecosystem. Internally, they adopt a step-by-step approach, staying updated on news, regulations, policy, and funding opportunities. They participate in

events organized by explorers, sharing their wisdom and experience, but rarely leading shared and shareable solutions for the ecosystem. Exploiters maintain their own data sharing platforms, striving to promote their own standards for the entire ecosystem. The following field notes show the strategic importance of the standardization among the exploiters.

> "The aim is to provide suggestions about future standards and specifications to be adopted for improved Mobility data sharing and to show how these standards have been implemented in the different stages of the MobiDataLab project (…). To attain this objective, the methodology followed was based on the identification of standards and their context, their application in MobiDataLab, and their development and related support for implementation" (MobiDataLab, Deliverable D2.5).

> "When it comes to standardization, the first thing to do is to clearly define the use cases to be solved through their intermediary. It then becomes possible to explore what exists in terms of standards and specifications and select the one that will best meet the identified need" (MobiDataLab, Deliverable D2.5).

As suggested by the interview transcript provided, instead of actively contributing to a collaborative vision for developing comprehensive data sharing platforms, exploiters prefer to use different platforms for different purposes and push their platforms as standards for the entire ecosystem (see also other excerpts reported in Table 3).

Conservatives: refraining from investment

Conservatives in our research are stakeholders who refrain from investing in data ecosystems. They participate in big data ecosystems without significantly committing to improving sharing practices. While they acknowledge the advantages of big data sharing, they focus on potential risks, such as low data quality, thus its low usability, and the behavior of other stakeholders in the ecosystem, as illustrated in the excerpts in Table 3 and in the following one:

> "There might be some risks in case that data quality is not that high (…) that's why we want to share right and correct data: it is very critical for us, and it is a matter of trust. Further, I believe there are a lot of risks in publishing data, because everybody can evaluate on his/her own way the data, later claiming that the data were wrong" (Logistics operator 2).

Conservatives' cautious approach stagnates the ecosystem, as their boundary work strategy mostly relied on refraining from investments. For instance, as MobiDataLab engaged in the creation of the common knowledge base and prototyped the cloud solution, conservatives participated to standard overview, provided input for the common knowledge base, and participated in Living Labs and hackathons whenever asked by explorers or exploiters, but without leading or taking significant responsibility over any of these activities. Conservatives constantly seek reassurances, guidelines, recommendations, and best practices to follow, but they were often unwilling to share, cooperate or trust more committed partners. Instead, they tended to shift responsibilities to external environment, other stakeholders, and society. Paradoxically, these stakeholders show no intention to exit MobiDataLab. Instead, they continue participating in the ecosystem, waiting for conditions to improve and risks to decrease. Table 3 provides additional excerpts exemplifying conservatives' boundary work strategy based on refraining from investments.

Idealists: maintaining the status quo

Idealists in our study participate in the project with high ambitions but contribute little to the common cause. Idealists foster trust in the ecosystem and its stakeholders, assuming goodwill from all involved. However, their minimal involvement sets them apart from more active actors like explorers. Idealists made limited contribution to the knowledge base, the cloud prototype, and practical Living Labs, often urging others to initiate and lead by example (see excerpts in Table 3). As they delegated change to others, they prefer to wait:

"For the right moment in which social change will bring incentives, honesty, and transparency"
(informant's words)

envisioning a utopian society where data sharing solves significant societal changes effectively and safely. When the long-awaited moment arrives, idealists promise to enhance data sharing conditions, as shown in the following example:

"Participating in EU projects adds another layer of motivation for data-sharing. Beyond the funding received, the experience gained from collaborating with stakeholders from different countries, traveling, and visiting various places creates a sense of commitment. This commitment stems from the human aspect – the satisfaction of working on projects that bring value not just in the short term but contribute to the objectives of the project in the long term" (MobiDataLab, Deliverable D5.3).

While our data does not suggest that idealists create obstacles, waiting for the perfect moment implies sitting still without contributing to the improvement of the ecosystem. This behavior, like the conservatives' one, risks causing vicious circles of self-fulfilling prophecies.

"I think we are still yet to understand what the demand on the network really will be and what those patterns are likely to be. Post COVID we've certainly made some early observations in the use of the network and using public transport, but what that long term picture looks like I still don't think we fully understand that, and we'll probably need another year worth of data to understand and to get operative" (Trade association 1).

In Table 3 we summarized the ecosystem stakeholder matrix.

Table 3: The ecosystem Stakeholder Matrix. Source: Authors Own Work.

Explorers: moving the ecosystem forward	**Improving collaboration process, creating moments of encounter:** "Collaboration between partners is key and at least in the next 10 years cannot be substituted by data analytics and machine learning. People need to work together, need to understand the different positions to commonly agree on the way forward" (Coordinator 1 of a Transport Agency). **Shaping Best Practices:** "I think someone must do the first step. (…) We often do the first step, this is our credo, to be a pioneer in publishing mobility data: for us it doesn't have to be balanced. We like to give data, and we hope the others use our data and give their data too. The balance is that you have to publish data needed by others, and others will also share their data, this could be the secret" (Project manager 1 in an institute for logistics). **Improving Data Accessibility/usage, overcoming risks:** "For us there is no risk involved in using this data sets because the cloud is secure. The risk is that the data set might not be complete enough, but in any case, it can provide some answers to some questions, (…) better than not having answers at all" (Cloud provider 1).

Exploiters: solving emergent problems	**Navigating Standardization (pushing their strategies as the standard ones):** "We can integrate data and we can use it, but it would be better for everyone if they would use the existing standard; in this way we could improve our service for everyone and also (…) our customers could benefit from all the new features and evolutions we could do to our main products" (ITS provider 2). **Advocating Legal Change: "**Well, I can speak for the work we are doing, so there are a lot of legal challenges. (…) What datasets can be provided and how to integrate them? There are also interoperability issues. How to make sure that the different datasets work and communicate with each other to make the systems work? Then there is again legal issue about contracting on how to share the data." (Researcher 1). **Participating to Policy Change: "**Of course, you need data if you want to be able to understand what the impacts of policies are in these areas. If you're implementing a policy for job creation, if you're implementing a policy to encourage more women in the job market, then of course you need data too (…) you need to have access to it, and you may also need to share it" (representative of an association of municipalities).
Conservatives: refraining from investments	**Waiting for reassurance, focusing on potential risks:** "It's an ecosystem situation, meaning that on one side each organization has their own agenda, individual agenda, whether it's growing their business or providing services to citizens. But at the same time, pushing the standardization requires some efforts and costs and the balance is not always in favor" (Consultant 2). **Waiting for the social change, shifting responsibilities to the external environment:** "Today there are no strategies (for the data sharing). Data is shared according to the time axis with which it is to be used. With time, if things will change, also our involvement could be a different one" (CEO 1, railway transport operator). **Looking for the right partners: "**Choosing the right partner is a nuanced process, especially when considering various projects and activities. While the desire is to collaborate with all partners, practical limitations necessitate selecting those with the best ideas, a timely approach, and a profile suitable for the specific job at hand. The question of partner selection, therefore, becomes a delicate balance of fairness, encouraging active participation, and diversifying opportunities, even for those who may not have been highly active in previous collaborations" (D5.3).
Idealists: maintaining status quo	**Delegating, following other's good example:** "Well opportunities for sharing data, I would say there would not be if my company shared data, but more if a city did it, if people managing mobility did it, if highway managers did it, etc. If they could share more data (…) then of course that will be interesting because then all the use cases that we developed or that other companies like us which are based on data coming from these people would work much better, and these might be deployed at a larger scale" (ITS provider 4). **Wait for the "perfect" moment to act: "**If I had superpowers, I would infuse into the people and companies' mind the idea that sharing is believing in a new kind of global world, which is not necessarily globalization, but it's a global place for everybody. So, we share the data, and all the processes get better, including sustainability" (Researcher 1).

	Expecting honesty and transparency: "We really have the feeling that, even though the advantage is not immediate for some companies, they really have this vision to create this digitized process and that they are really into this Community feeling to strive this common goal of achieving an improved, more efficient digitized process (…) based on information partnerships" (project manager 1 in an airline company).

2.5.3. Ecosystem consequences: innovation and stagnation

Based on our study of the sharing dilemma, we identified two trends in the ecosystem: explorers and exploiters actively engage in boundary work practices that drive ecosystem innovation, while idealists and conservatives enact boundary work practices contributing to stagnation.

Our model, depicted in Figure 3, illustrates a circular movement where the four boundary work practices represent a continuum rather than binary positions. The horizontal axis measures involvement in the ecosystem (from high to low), and the vertical axis measures the perception of data sharing (from opportunity to risk). Exploiters, explorers, idealists, and conservatives each occupy unique positions within this matrix. Assuming that almost all respondents are aware of the main risks and advantages of data sharing, explorers and idealists predominantly see the advantages, while exploiters and conservatives focus more on the risks, leading to greater caution in their data sharing practices. We have shown that the promise of future advantages leads idealists to adopt utopian visions of big data ecosystems and explorers to adopt proactive behaviors for ecosystem improvement. Conversely, focus on data sharing risks leads conservatives to adopt a passive waiting behavior, and exploiters to work on external factors such as policymaking or standardization process to improve the ecosystem's performance and lower perceived risks.

Explorers and exploiters are deeply engaged in the ecosystem, striving for greater integration. Although they are aware of potential risks, they implement strategies to mitigate these risks in favor of the benefits. This is evident in the following excerpts, the first from an explorer and the second from an exploiter:

"I would say collaboration is more important than the risks out there. I mean, creating technology is something that one partner can do on their own, but the prerequisite for a new mobility ecosystem to work is for the partners to be able to work together. So, if that doesn't happen, even if you have the technology, if you don't have the collaboration, then that is not possible to happen" (Researcher 2).

"I think that key is being very, very clear upfront about just knowing what data it is you want to make available, what are the limitations and capabilities of that data, and how could that data be used or potentially misused and just to understand what those risks are for any dataset I think striking the balance just comes in. I suppose I'm just being very clear. What is that thing you want to make available, how you intend for it to be used, and: is there the potential for misuse?" (Researcher 3).

Conversely, idealists and conservatives adopt fewer active roles, remaining cautious and static in their engagement. A conservative often assumes marginal, static, and cautious roles in both relational and procedural aspects. While data sharing may not be their top priority, they remain connected to the big data ecosystem with a "just in case" approach.

"I would say that for a lot of organizations, the topic of data is not the number one topic on the list. And, when it climbs up the list, sharing is really not the by-default option (…) But things can certainly change (…) it is good to stick around just in case" (Citizen representative 1).

Since boundary work is a relational process dependent on how stakeholders perceive each other's commitment and contributions regarding data sharing risks and advantages, forces of innovation and stagnation manifest in constant dialogue. The model is circular to emphasize the possibility of transitioning between roles as perceptions of risks and advantages change or commitment shifts within the ecosystem. Stakeholders' perception of mobility big data as a heterogeneous, ambiguous, and uncertain boundary object allows for this flexible approach. As explained by the stakeholders we interviewed, advantages to learn from others and create opportunities for exchange and growth are fundamental for transitioning between different positions in a big data ecosystem:

"There are a lot of different attitudes and behaviors out there (…) some people will push their own standards, some will work hard for a common vision and others will not do much (…) I really think that having this community feeling helps a lot because creating this common vision is based on trust; data is the new gold, and you have the trust and you're giving actually away your data, and you have to trust that it is used in a good way and that you won't gain a disadvantage" (Airline project manager 1).

Overall, our findings suggest that big data sharing is an evolving process, where stakeholders - from pioneers to conservatives - collaborate to shape the future. Building community and fostering dialogue among these roles can mitigate stagnation, driving innovation in the ecosystem.

2.6. Discussion

The complexity of big data ecosystems, characterized by rapid expansion and dynamic interaction patterns (Demchenko et al., 2014; Rammer et al., 2023), has positioned them as standalone phenomena, particularly in addressing societal challenges that span organizational and sectoral boundaries and require complex information and large-scale collaboration (Agarwal et al., 2013; Jacobides et al., 2018; Zabel et al., 2023).

Our research contributes to the literature on big data ecosystems and related phenomena such as open data and domain-specific data ecosystems (e.g., healthcare, mobility, environmental accountability) (Marcelo et al., 2019; Zuiderwijk et al., 2014), advancing beyond the ecosystems literature (Thomas and Ritala., 2022) by highlighting the importance of stakeholders contributions, including their perceptions, attitudes, and behaviors towards data sharing (George et al., 2014). We adopted a boundary work perspective to understand how stakeholders' perspectives and practices shape the evolution of big data ecosystems.

In the context of big data ecosystems, our work expands the theoretical understanding of how actors behave, perceive, and interact around big data (Chen et al., 2016; Kazantsev et al., 2023; Sjödin et al., 2024). Drawing on a practice-based sociotechnical perspective, we highlight how big data act as boundary objects within complex socio-technical ecosystems.

These boundary objects exhibit plasticity, adapting to local needs while maintaining a common identity across the ecosystem. Our findings show that this plasticity arises from stakeholders' perceptions of heterogeneity, ambiguity, and uncertainty regarding the usability and scalability of mobility data. We propose that the challenges of studying big data ecosystems are linked not only to governance issues or the characteristics of data itself, as previous studies suggested (Marcelo et al., 2019; Zabel et al., 2023; Zuiderwijk et al., 2014), but also to the uncertainties and ambiguities associated with data sharing perceptions and actors' behaviors within the ecosystem (Kazantsev et al., 2023).

Our work shifts attention from data management capabilities, often focused on big data literature (Huynh et al., 2023; Sjödin, Liljeborg, Social, et al., 2024), to the heterogeneous understandings that individuals have of big data and their data sharing practices. While previous studies acknowledge the complexity, ambiguity, and uncertainty of big data, they suggest these challenges can be mitigated through training and expertise-based capabilities. Conversely, our study suggests that the uncertain, heterogeneous, and ambiguous nature of big data challenges the notion of one-size-fits-all capabilities and practices. We document a plurality of approaches that coexist and interact within a multi-stakeholder ecosystem. In such scenarios, we argue that motivations and attitudes become critical drivers of data sharing practices and warrant scrutiny.

We proposed a typology of contributions to big data ecosystems based on boundary work practices, moving beyond traditional views of ecosystems as merely constellations of structural roles related to big data management. The typology, organized in a matrix based on actors' ecosystem involvement and focus on advantages or risks, includes explorers, exploiters, idealists, and conservatives. This classification offers a nuanced understanding of ecosystems dynamics. Importantly, these categories are distinct from the structural roles like policymakers, data providers, data aggregators, and orchestrators, as there was no significant overlap between attitudes, practices, and structural roles. Thus, our findings emphasize the critical role of

participants' attitudes and practices, extending beyond the current focus on structural roles (Kazantsev et al., 2023).

Our investigation into the attitudes and practices of data sharing highlights the role of boundary work in organizing and developing big data ecosystems. Depending on how explorers, exploiters, idealists, and conservatives interact, a big data ecosystem can foster innovation or become stagnant. Our study emphasizes the necessity of moving beyond structural roles and adopting a dynamic approach that focuses on attitudes and practices in ecosystem participation. Big data ecosystems function as dynamic arenas where actors with diverse goals intersect in constructing big data. This perspective adds a new dimension to existing literature, portraying big data as plastic objects that create negotiation arenas for individuals and organizations with varied interests in data sharing (Bakici et al., 2021; Lu and Wang, 2018; Ramezani and Camarinha-Matos, 2020). It also contributes to the broader literature on ecosystems as collective entities.

While current studies acknowledge the interplay between shared and differentiated goals within ecosystem, they have yet to fully elucidate these dynamics' implications (Thomas et al., 2014; Thomas and Ritala, 2022). Ecosystems primarily rely on noncontractual mechanisms such as role definitions, super modular complementarity, and alignment facilitated by coordination structures like platforms, to balance generative and coherent outcomes (Jacobides et al., 2018). Collective action theory views groups of actors as causal agents and explorers how their interactions facilitate or constrain shared goals (Thomas and Ritala, 2022). However, while collective action perspectives emphasize participants interactions in pursuing shared value, little is known about how the interplay between similarity and distinctiveness without common coordination structures (Thomas and Ritala, 2022). Our study advances this conversation by proposing boundary objects and boundary work as mechanisms for collective action within ecosystems.

Our findings contribute to research on the role of uncertainty in ecosystem governance (Gomes et al., 2022). Previous research lacks insight into how focal firms orchestrate ecosystems to cope with uncertainty, leaving the relationship between uncertainty and ecosystem management understudied. Using boundary work and boundary objects, we bridge findings on collective action with insights on uncertainty management in ecosystems.

In the MobiDataLab system, we discovered an interplay between a shared vision on big data and fundamentally different perspectives and attitudes, influencing actors' behavior. While there was agreement on the advantages and risks of sharing big data, explorers, exploiters, idealists, and conservatives adopted distinct practices. These practices either drove or impeded the ecosystem, leading to innovation or stagnation. Strategies for managing uncertainty explain stakeholders' varied behavior within the ecosystem (Gomes et al., 2022).

Our results show that all stakeholders, even skeptics, desire to be part of big data ecosystems. Conservatives and idealists, despite perceiving more risks than advantages and making limited investments were reluctant to exit. In line with the current literature on ecosystems, we notice that factors such as affiliation with established organizations, future benefits, and motivations for social impact influence the decision to remain in an ecosystem (Davenport, 2014; Olszak and Mach-Król, 2018). Our study underscores the importance of further investigating motivations for participation and boundary work in multi-stakeholder ecosystems.

Ultimately, connecting the lens of boundary work with that of boundary objects allows us to see big data ecosystems not just as technical or managerial systems, but as evolving social arrangements in which data, practices, and roles co-constitute each other. This conceptual integration deepens our understanding of how big data ecosystems are formed, maintained, and transformed through actors' everyday practices.

2.7. Conclusions, limitations, and future research directions

The analytical work conducted for this research has certain limitations. Firstly, our typology is not intended to be exhaustive. Different types of digital ecosystems, such as innovation ecosystems or knowledge and research ecosystem, may exhibit different (Cobben et al., 2022). Analyzing additional big data ecosystems could validate the patterns identified in this research and uncover new ones. For example, while we emphasize risk-opportunity perceptions and involvement as primary factors, other attitudes, practices, and roles may also play significant roles in diverse types of big data ecosystems (see Kazantsev et al., 2023). Such studies would also contribute to the ongoing conversation on whether big data ecosystems differ significantly from other digital ecosystems like innovation and commercial ecosystems (Cobben et al., 2022) and should be studied as a self-standing category (Cui et al., 2020; Huynh et al., 2023).

Additionally, future research could delve deeper into the relational dynamics among stakeholders within these ecosystems. This would involve examining how explorers, exploiters, idealists, and conservatives mutually influence each other and interpret their differences during ecosystem interactions. Recognizing and negotiating these different positions in interactions could profoundly impact whether behaviors fostering innovation or stagnation prevail, as observed in our case.

On a practical level, by describing the practice-based view of big data ecosystems and emphasizing interactions among various actors, we offer insights for constructing, defining, and managing these ecosystems. Understanding different actors' positions and perspectives can help resolve conflicts and ensure high involvement, enhancing security and alleviating fears that often hinder progress.

Understanding the motivations behind exploratory, exploitative, idealistic, and conservative behaviors is crucial for leveraging these tendencies while fostering an environment based on shared visions and collaboration. In big data ecosystems, it is essential

to adopt a realistic perspective that acknowledges the risks and benefits of different data-sharing practices among actors.

Our study provides guidance for policymakers and stakeholders, highlighting both positive and negative dynamics in these sharing ecosystems. We introduce criteria for identifying various stakeholder types and influencing their contributions. Our findings offer theoretical support for understanding the interconnectedness of big data, ecosystems, diversity, and collaboration.

We demonstrate that adopting a sociotechnical perspective on big data ecosystems provides a more realistic understanding. Heterogeneity, ambiguity, and uncertainty are integral to these ecosystems. Embracing stakeholders' diverse approaches to data sharing is essential for advancing research, particularly in understanding the different standards, approaches, and practices within such environments (Jones, 2019). Balancing commonalities and differences is key to future advancements in society through new big data technologies and multi-stakeholder ecosystems.

References

Agarwal, A., Jain, V., & Karamchandani, A. (2013). State of the Low-Income Housing Market Encouraging Progress & Opportunity to Realize Dreams of Millions. New York City: Monitor Deloitte.

Anteby, M., Chan, C. K., & DiBenigno, J. (2016). Three Lenses on Occupations and Professions in Organizations: Becoming, Doing, and Relating. Https://Doi.Org/10.5465/19416520.2016.1120962, 10(1), 183–244.

Austin, J. E., & Seitanidi, M. M. (2012). Collaborative Value Creation: A Review of Partnering Between Nonprofits and Businesses. Part 2: Partnership Processes and Outcomes. Nonprofit and Voluntary Sector Quarterly, 41(6), 929–968. https://doi.org/10.1177/0899764012454685/ASSET/IMAGES/LARGE/10.1177_0899764012454685-FIG4.JPEG

Bakici, T., Nemeh, A., & Hazir, Ö. (2021). Big data adoption in project management: insights from french organizations. IEEE Transactions on Engineering Management, 70(10). https://ieeexplore.ieee.org/abstract/document/9502021/

Barley, S. R., Bechky, B. A., & Milliken, F. J. (2017). The Changing Nature of Work: Careers, Identities, and Work Lives in the 21st Century. Https://Doi.Org/10.5465/Amd.2017.0034, 3(2), 111–115.

Bechky, B. A. (2003). Object lessons: Workplace artifacts as representations of occupational jurisdiction. American Journal of Sociology, 109(3), 720-752. Https://Doi.Org/10.1086/379527, 109(3).

Bonina, C., Koskinen, K., Eaton, | Ben, Gawer, | Annabelle, & School, S. B. (2021). Digital platforms for development: Foundations and research agenda. Wiley Online LibraryC Bonina, K Koskinen, B Eaton, A GawerInformation Systems Journal, 2021•Wiley Online Library, 31(6), 869–902. https://doi.org/10.1111/isj.12326

Bourdieu, P. (2020). Outline of a Theory of Practice. The New Social Theory Reader, 80–86. https://doi.org/10.4324/9781003060963-11

Bucher, S. V., Chreim, S., Langley, A., & Reay, T. (2016). Contestation about Collaboration: Discursive Boundary Work among Professions. Http://Dx.Doi.Org/10.1177/0170840615622067, 37(4), 497–522.

Cao, Q. H., Khan, I., Farahbakhsh, R., Madhusudan, G., Lee, G. M., & Crespi, N. (2016). A trust model for data sharing in smart cities. 2016 IEEE International Conference on Communications, ICC 2016. https://doi.org/10.1109/ICC.2016.7510834

Carlile, P. R. (2004). Transferring, Translating, and Transforming: An Integrative Framework for Managing Knowledge Across Boundaries. Https://Doi.Org/10.1287/Orsc.1040.0094, 15(5), 555–568.

Cazier, J., Cech, T., & Spaulding, T. (2015). Applying Business Analytic Methods To Improve Organizational Performance In The Public School System. Proceedings of the Twenty-First Americas Conference on Information Systems, Puerto Rico, 13–15.

Chen, C., Ma, J., Susilo, Y., Liu, Y., & Wang, M. (2016). The promises of big data and small data for travel behavior (aka human mobility) analysis. Transportation Research Part C:

Emerging Technologies, 68. https://www.sciencedirect.com/science/article/pii/S0968090X16300092

Clarke, H. D., Kellner, P., Stewart, M. C., Twyman, J., & Whiteley, P. (2016). Austerity and political choice in Britain. Austerity and Political Choice in Britain, 1–177. https://doi.org/10.1057/9781137524935

Cobben, D., Ooms, W., Roijakkers, N., & Radziwon, A. (2022). Ecosystem types: A systematic review on boundaries and goals. Journal of Business Research, 142, 138–164. https://doi.org/10.1016/J.JBUSRES.2021.12.046

Cui, Y., Kara, S., & Chan, K. C. (2020). Manufacturing big data ecosystem: A systematic literature review. Robotics and Computer-Integrated Manufacturing, 62, 101861. https://doi.org/10.1016/J.RCIM.2019.101861

Davenport, T. (2014). Big data at work: dispelling the myths, uncovering the opportunities. Harvard Business Review Press.

Davenport, T. H., & Kudyba, S. (2016). Designing and developing analytics-based data products. MIT Sloan Management Review. https://sloanreview.mit.edu/article/designing-and-developing-analytics-based-data-products/

De Mauro, A., Greco, M., & Grimaldi, M. (2015). What is big data? A consensual definition and a review of key research topics. AIP Conference Proceedings, 1644(1), 97–104. https://doi.org/10.1063/1.4907823

de Vasconcelos Gomes, L. A., Facin, A. L. F., Salerno, M. S., & Ikenami, R. K. (2018). Unpacking the innovation ecosystem construct: Evolution, gaps and trends. Technological forecasting and social change, 136, 30-48.

de Vasconcelos Gomes, L. A., dos Santos, M. G., & Facin, A. L. F. (2022). Uncertainty management in global innovation ecosystems. Technological Forecasting and Social Change, 182, 121787.

Demchenko, Y., De Laat, C., & Membrey, P. (2014). Defining architecture components of the Big Data Ecosystem. 2014 International Conference on Collaboration Technologies and Systems, CTS 2014, 104–112. https://doi.org/10.1109/CTS.2014.6867550

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. International journal of information management, 35(2), 137-144.

George, G., Haas, M. R., & Pentland, A. (2014). Big Data and Management. Https://Doi.Org/10.5465/Amj.2014.4002, 57(2), 321–326.

Gieryn, T. F. (1983). Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists. American Sociological Review, 48(6), 781. https://doi.org/10.2307/2095325

Granstrand, O., & Holgersson, M. (2020). Innovation ecosystems: A conceptual review and a new definition. Technovation, 90–91, 102098. https://doi.org/10.1016/J.TECHNOVATION.2019.102098

Hargrave, T. J., & Van de Ven, A. H. (2006). A collective action model of institutional innovation. Academy of management review, 31(4), 864-888. https://doi.org/10.5465/AMR.2006.22527458

Helfat, C. E., & Raubitschek, R. S. (2018). Dynamic and integrative capabilities for profiting from innovation in digital platform-based ecosystems. Research Policy, 47(8), 1391–1399. https://doi.org/10.1016/J.RESPOL.2018.01.019

Henderson, K. (1991). Flexible Sketches and Inflexible Data Bases: Visual Communication, Conscription Devices, and Boundary Objects in Design Engineering. Http://Dx.Doi.Org/10.1177/016224399101600402, 16(4), 448–473.

Huynh, M. T., Nippa, M., & Aichner, T. (2023). Big data analytics capabilities: Patchwork or progress? A systematic review of the status quo and implications for future research. Technological Forecasting and Social Change, 197, 122884.

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

Jones, M. (2019). What we talk about when we talk about (big) data. The Journal of Strategic Information Systems, 28(1), 3–16. https://doi.org/10.1016/J.JSIS.2018.10.005

Kazantsev, N., Islam, N., Zwiegelaar, J., Brown, A., & Maull, R. (2023). Data sharing for business model innovation in platform ecosystems: From private data to public good. Technological Forecasting and Social Change, 192. https://www.sciencedirect.com/science/article/pii/S0040162523002007

Kellogg, K. C., Orlikowski, W. J., & Yates, J. A. (2006). Life in the Trading Zone: Structuring Coordination Across Boundaries in Postbureaucratic Organizations. Https://Doi.Org/10.1287/Orsc.1050.0157, 17(1), 22–44.

Keser, C., Suleymanova, I., & Wey, C. (2012). Technology adoption in markets with network effects: Theory and experimental evidence. Information Economics and Policy, 24(3–4), 262–276. https://doi.org/10.1016/J.INFOECOPOL.2012.03.001

Kim, G. H., Trimi, S., & Chung, J. H. (2014). Big-data applications in the government sector. Communications of the ACM, 57(3), 78–85. https://doi.org/10.1145/2500873

Krasyuk, I. A., Kolgan, M. V., & Medvedeva, Y. (2022). Digital ecosystem structure formation depending on the archetype of distribution network participants. European Journal of Management and Business Economics, 31(2), 176–191. https://doi.org/10.1108/EJMBE-07-2021-0202/FULL/HTML

Kreiner, G., Hollensbe, E., & Sheep, M. (2009). Balancing Borders and Bridges: Negotiating the Work-Home Interface via Boundary Work Tactics. Https://Doi.Org/10.5465/Amj.2009.43669916, 52(4), 704–730.

Lamont, M., & Molnár, V. (2003). The Study of Boundaries in the Social Sciences. Https://Doi.Org/10.1146/Annurev.Soc.28.110601.141107, 28, 167–195.

Lifshitz-Assaf, H. (2018). Dismantling knowledge boundaries at NASA: The critical role of professional identity in open innovation. Administrative Science Quarterly, 746–782. https://journals.sagepub.com/doi/abs/10.1177/0001839217747876

Lis, D., & Otto, B. (2020). Data governance in data ecosystems–insights from organizations.

Lu, J., & Wang, W. (2018). Managerial conservatism, board independence and corporate innovation. Journal of Corporate Finance, 48, 1–16. https://doi.org/10.1016/J.JCORPFIN.2017.10.016

Maaroof, A. (2015). BIG DATA AND THE 2030 AGENDA FOR SUSTAINABLE DEVELOPMENT. Report for UN-ESCAP.

Marcelo, M. I., Barros Lima, G. de F., & Farias Lóscio, B. (2019). Investigations into Data Ecosystems: a systematic mapping study. Knowledge and Information Systems, 61(2), 589–630. https://doi.org/10.1007/S10115-018-1323-6

Mello, M. M., Francer, J. K., Wilenzick, M., Teden, P., Bierer, B. E., & Barnes, M. (2013). Preparing for responsible sharing of clinical trial data. New England Journal of Medicine, 369(17), 1651-1658.

Mørk, B. E., Hoholm, T., Maaninen-Olsson, E., & Aanestad, M. (2012). Changing practice through boundary organizing: A case from medical R&D. Https://Doi.Org/10.1177/0018726711429192, 65(2), 263–288.

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. https://doi.org/10.1016/J.JSIS.2015.02.001

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.

Olszak, C. M., & Mach-Król, M. (2018). A Conceptual Framework for Assessing an Organization's Readiness to Adopt Big Data. Sustainability 2018, Vol. 10, Page 3734, 10(10), 3734. https://doi.org/10.3390/SU10103734

Pesce, D., Neirotti, P., & Paolucci, E. (2019). When culture meets digital platforms: value creation and stakeholders' alignment in big data use. Current Issues in Tourism, 22(15), 1883-1903. https://doi.org/10.1080/13683500.2019.1591354

Petre, I., Boncea, R., Radulescu, C. Z., Zamfiroiu, A., & Sandu, I. (2019). A time-series database analysis based on a multi-attribute maturity model. Studies in Informatics and Control, 28(2), 177-188.

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. Health Information Science and Systems, 2(1). https://doi.org/10.1186/2047-2501-2-3

Ramezani, J., & Camarinha-Matos, L. M. (2020). Approaches for resilience and antifragility in collaborative business ecosystems. Technological Forecasting and Social Change, 151, 119846. https://doi.org/10.1016/J.TECHFORE.2019.119846

Rammer, C., & Es-Sadki, N. (2023). Using big data for generating firm-level innovation indicators-a literature review. Technological Forecasting and Social Change, 197, 122874.

Richard J. Boland, Jr., & Tenkasi, R. V. (1995). Perspective Making and Perspective Taking in Communities of Knowing. Https://Doi.Org/10.1287/Orsc.6.4.350, 6(4), 350–372. 350

Rong, K., & Luo, Y. (2023). Toward born sharing: The sharing economy evolution enabled by the digital ecosystems. Technological Forecasting and Social Change, 196, 122776.

Shin, D. H., & Choi, M. J. (2015). Ecological views of big data: Perspectives and issues. Telematics and Informatics, 32(2), 311–320. https://doi.org/10.1016/J.TELE.2014.09.006

Shum, S. B., Aberer, K., Schmidt, A., Bishop, S., Lukowicz, P., Anderson, S., Charalabidis, Y., Domingue, J., De Freitas, S., Dunwell, I., Edmonds, B., Grey, F., Haklay, M., Jelasity, M., Karpištšenko, A., Kohlhammer, J., Lewis, J., Pitt, J., Sumner, R., & Helbing, D. (2012).

Towards a global participatory platform: democratising open data, complexity science and collective intelligence. SpringerS Buckingham Shum, K Aberer, A Schmidt, S Bishop, P Lukowicz, S AndersonThe European Physical Journal Special Topics, 2012•Springer, 214(1), 109–152. https://doi.org/10.1140/epjst/e2012-01690-3

Singh, K. N., Behera, R. K., & Mantri, J. K. (2019). Big data ecosystem: Review on architectural evolution. Advances in Intelligent Systems and Computing, 813, 335–345. https://doi.org/10.1007/978-981-13-1498-8_30/COVER

Sjödin, D., Liljeborg, A., & Mutter, S. (2024). Conceptualizing ecosystem management capabilities: Managing the ecosystem-organization interface. Technological Forecasting and Social Change, 200, 123187.

Star, S. L., & Griesemer, J. R. (1989). Institutional Ecology, `Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. Http://Dx.Doi.Org/10.1177/030631289019003001, 19(3), 387–420.

Strauss, A., & Corbin, J. (1998). Basics of qualitative research techniques. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=18c7cfe7a46c7771b60dc384b1b4e350f65b13e2

Suddaby, R., Cooper, D. J., & Greenwood, R. (2007). Transnational regulation of professional services: Governance dynamics of field level organizational change. Accounting, Organizations and Society, 32(4–5), 333–362. https://doi.org/10.1016/J.AOS.2006.08.002

Sun, A. Y., & Scanlon, B. R. (2019). How can Big Data and machine learning benefit environment and water management: a survey of methods, applications, and future directions. Environmental Research Letters, 14(7), 073001. https://doi.org/10.1088/1748-9326/AB1B7D

Thomas, L. D., Autio, E., & Gann, D. M. (2014). Architectural leverage: Putting platforms in context. Academy of management perspectives, 28(2), 198-219. https://doi.org/10.5465/amp.2011.0105

Tools, M., Saeed, N., Malik, H., Naeem, A., & Bashir, U. (2023). Incorporating big data and IoT in intelligent ecosystems: state-of-the-arts, challenges and opportunities, and future directions. Multimedia Tools and Applications 2023, 1–43. https://doi.org/10.1007/S11042-023-16328-3

van Tulder, R., Seitanidi, M. M., Crane, A., & Brammer, S. (2016). Enhancing the Impact of Cross-Sector Partnerships: Four Impact Loops for Channeling Partnership Studies. Journal of Business Ethics, 135(1), 1–17. https://doi.org/10.1007/S10551-015-2756-4/TABLES/2

Vranopoulos, G., Clarke, N., & Atkinson, S. (2022). Addressing big data variety using an automated approach for data characterization. Journal of Big Data, 9(1), 1–28. https://doi.org/10.1186/S40537-021-00554-3/FIGURES/12

Walker, S. J. (2014). Big Data: A Revolution That Will Transform How We Live, Work, and Think. International Journal of Advertising, 33(1), 181–183. https://doi.org/10.2501/IJA-33-1-181-183

Werbrouck, J., Pauwels, P., Beetz, J., Verborgh, R., Mannens, E., & Aameri, B. (2023). ConSolid: A federated ecosystem for heterogeneous multi-stakeholder projects. Semantic Web, Preprint(Preprint), 1–32. https://doi.org/10.3233/SW-233396

Yoon, A., & Copeland, A. (2020). Toward community-inclusive data ecosystems: Challenges and opportunities of open data for community-based organizations. Journal of the Association

for Information Science and Technology, 71(12), 1439–1454. https://doi.org/10.1002/ASI.24346

Zabel, C., O'Brien, D., & Social, J. N. (2023). Sensing the Metaverse: The microfoundations of complementor firms' dynamic sensing capabilities in emerging-technology ecosystems. Technological Forecasting and Social Change, 192. https://www.sciencedirect.com/science/article/pii/S0040162523002470

Zillner, S., Le Bars, L., de Lama, N., Scerri, S., García Robles, A., Tonna, M. C., Kenneally, J., Mayer, D., Hahn, T., Caj, S., Seidl, R., Carbonare, D. D., & Curry, E. (2021). A roadmap to drive adoption of data ecosystems. The Elements of Big Data Value: Foundations of the Research and Innovation Ecosystem, 41–62. https://doi.org/10.1007/978-3-030-68176-0_3/FIGURES/2

Zuiderwijk, A., Janssen, M., & Davis, C. (2014). Innovation with open data: Essential elements of open data ecosystems. Information Polity, 19(1–2), 17–33. https://doi.org/10.3233/IP-140329

3. Not Everyone Shares the Same. Unpacking the Motivational, Cognitive, and Relational Factors of Willingness to Share Data in Data Ecosystems[3]

Highlights:

- Data ecosystems are an emerging and increasingly impactful phenomenon based on multi-stakeholder data sharing collaborations across sectors, often aimed at solving grand societal challenges, such as environmental sustainability and scientific advancement.
- Willingness to share data is an important process in explaining how stakeholders of a data ecosystem share data: this study investigates the relationship between willingness to share and various data sharing practices through a qualitative case study in a data ecosystem for integrated mobility.
- Willingness to share in a data ecosystem depends on stakeholders' motivational orientations (social vs. utilitarian), their cognitive evaluation of data sharing risks and opportunities, and their social assessment of the ecosystem's relational dynamics.
- Data ecosystems are constituted by a constellation of data sharing practices (driving, pulling, and scaffolding) depending on stakeholders' different levels of willingness to share.

[3] The paper presented in this chapter is currently under review in the journal *Information and Organization*. Authors: Giulia Renzi, first author (conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, validation, visualization, writing – original draft), Paula Ungureanu (conceptualization, investigation, methodology, supervision, validation), and Natalia Selini Hadjidimitriou (data curation, writing – review and editing).

- This study integrates motivational orientations, cognitive assessment, and socio-relational evaluation to enhance the understanding of willingness to share and data sharing practices within data ecosystems.

Abstract

Data ecosystems are fundamental to address relevant societal challenges, by leveraging multi-stakeholder participation. However, our understanding of what motivates or hinders individuals and organizations from sharing data within such ecosystems remains limited. While existing studies have emphasized many factors such as data quality, participants' motivations, and governance mechanisms, an integrated perspective is missing. This study focuses on the interplay between motivational, cognitive, and relational factors that influence data sharing behaviors in data ecosystems. By conducting a qualitative study in a multi-stakeholder data ecosystem, we adopt a practice-based approach to investigate how data sharing perceptions, concerns, and relational expectations shape the willingness to share data and contribute towards different sharing practices. Our findings document a constellation of data sharing practices within a data ecosystem – namely driving, pulling, and scaffolding. These practices depend on the social construction of willingness to share data resulting from stakeholders' motivational orientations (social vs. utilitarian), their cognitive evaluations of the risks and opportunities associated with data sharing, and their social evaluation of the ecosystem's conditions. Our practice-based view integrates motivational orientation with socio-cognitive evaluation and relational processes of data sharing, advancing the literature on the governance of data ecosystems and providing new insights into the factors shaping the willingness to share data.

Keywords: data ecosystems, data sharing, willingness to share, mobility, data governance, data partnerships.

3.1. Introduction

The world faces major social challenges like resources exploitation, unequal access to education and healthcare, and poor living conditions (among the others: Farquharson et al., 2022; Hardoy & Satterthwaite, 1991; Sidze et al., 2022). Although still emerging, data sharing shows strong potential to help to improve these conditions (Hansen & Pang, 2023). Open data initiatives, for example, aim to enhance transparency, fight corruption, support law enforcement, foster smart urban planning, promote healthcare, and address climate change (Davidson et al., 2023; Kitsios et al., 2017). Today, data ecosystems are seen as privileged paths to address these problems by encouraging stakeholder (Susha, van den Broek, et al., 2023).

However, while some data ecosystems succeed, many fall short (Aaen et al., 2022; Chaudhari & Sinha, 2021; Nguyen et al., 2017; Wang et al., 2021). For instance, predictive policing aims to prevent crime, but faces issued like inaccurate data, poor interoperability, and misaligned stakeholder interests (Iannacci et al., 2022; Lum et al., 2016). Similarly, environmental conservation projects like Global Forest Watch enables real-time forest monitoring but struggles with data integration and stakeholder coordination (Schneider et al., 2020). In healthcare, even high-profile projects like Google Health and the Danish General Practitioners Database (DAMD) failed despite early enthusiasm and investment (Aaen et al., 2022; Spil & Klein, 2014). For instance, the DAMD once hailed as revolutionary, collapsed due to disputes over privacy, compliance, and control (Aaen et al., 2022). These examples show how difficult it is to build and maintain successful data ecosystems (Aaen et al., 2022; R. Agarwal & Dhar, 2014; Anwar et al., 2021; Günther et al., 2017).

One common challenge is fragmentation: Despite efforts to create shared frameworks and standards, data ecosystems often remain disjointed (Aaen et al., 2022; Susha et al., 2019). This is partly because stakeholders - including public and private actors, policymakers, and civil society – bring differing views on risks and benefits, shaped by their roles and goals. For

instance, the European Commission promotes unified data-sharing frameworks balancing public good with privacy (Shin & Choi, 2015). In contrast, private organizations often prioritize business opportunities and may hesitate to adopt these frameworks if they are seen as burdensome or uncertain (Janhunen, 2019). Governments weigh collaboration benefits against regulatory constraints, often relying on existing guidelines (Aaen et al., 2022). Meanwhile, citizens often hesitate due to privacy concerns, tech disparities, and trust issues (Günther et al., 2017; Jäckle et al., 2019; Schudy & Utikal, 2017).

These diverging interests co-exist and influence how ecosystems evolve. This matters because data ecosystems cross boundaries between established rules, roles, and best practices within sectors (private, and public), and their success in addressing societal concerns may strongly depend on how these differences are navigated and managed at the inter-organizational level (Davidson et al., 2023; Susha et al., 2019). However, few studies explore data ecosystems as standalone phenomena (Jarvenpaa & Essén, 2023; Lee et al., 2018; Munoz-Arcentales et al., 2019; A. L. Oliveira, 2019; Sifah et al., 2021; van den Broek & van Veenstra, 2018) and even less is known about how stakeholder diversity shape data-sharing dynamics (Susha, Rukanova, et al., 2023; Susha, van den Broek, et al., 2023). Existing research on data sharing behaviors identifies some key factors behind different ways of sharing data, such as attitudes towards collaboration, financial incentives, power and control concerns, legal barriers, and self-promotion and legitimation interests (Devriendt et al., 2021; Y. Kim & Stanton, 2016; Oliveira et al., 2019; van den Broek & van Veenstra, 2018). These factors vary in nature, encompassing motivations, expectations, attitudes, and social and institutional practices. Despite these studies, an integrated perspective is lacking, as researchers are only beginning to explore how these factors interact when mobilized by different stakeholders within an ecosystem.

We argue that grasping data ecosystems development requires studying how perceptions, expectations, and practices interact. Current work on data governance - i.e., stakeholder roles

in managing and controlling data (Günther et al., 2017; Jarvenpaa & Essén, 2023; Lis & Otto, 2020; Micheli et al., 2020; van den Broek & van Veenstra, 2018) – focuses mainly on structural factors like organizational setup, disregarding the specific challenges of data governance at the inter-organizational level such as fluid roles, rules and norms, limited responsibility, and lack of common grounds for goal alignment (Davidson et al., 2023; Jarvenpaa & Essén, 2023). These characteristics, we argue, demand attention beyond structural factors, to also include motivational orientations (individual and collective) and relational factors, like trust and willingness to share, as well as the interplay between them (Wang et al., 2021).

To do this, we studied a data ecosystem focused on integrated mobility. Using a practice-based approach, we looked at how stakeholder views, concerns, and expectations affect their willingness to share. We identified three data sharing patterns - driving, pulling, and scaffolding - linked to different motivations (social vs. utilitarian), risk assessments, and views on ecosystem conditions. Our study combines motivational, cognitive, and relational factors to better understand data sharing. We aim to contribute new insights to the literature and help policymakers and ecosystem actors manage different data-sharing behaviors.

3.2. Theoretical Framing

3.2.1. Data ecosystems: from structural to motivational and relational perspectives

In response to the increasing volumes of heterogeneous data generated in an increasingly digitalized world, collaborative approaches to data sharing have emerged as a growing societal trend and a new area of research (Abraham et al., 2019a; Davidson et al., 2023; Jarvenpaa & Essén, 2023).

Although there is still no agreed-upon definition of what constitutes a data ecosystem (Arrieta-Ibarra et al., 2018; Prainsack, 2019; Micheli et al., 2020), the concept has been used to refer to complex arrangements of distributed actors, resources, and infrastructures, as well as to

the socio-technical relations that support the circulation of data (Oliveira, 2019). While some ecosystems are primarily built for commercial and competitive purposes (e.g., Amazon Web Services, Google Cloud), others aim to create collective value and serve the public interest (e.g., Global Forest Watch, Human Genome Project). Regardless of their goals and domains, all these ecosystems pose significant challenges related to the fragmentation of data, technological heterogeneity, quality issues, and the absence of clear roles and shared norms among participating stakeholders (Wang et al., 2021; Bonina et al., 2021; Shah et al., 2020; R. Taylor et al., 2020; Oliveira et al., 2019; Abraham et al., 2019; van den Broek & van Veenstra, 2018; Jenkins & Schaap, 2018; Günther et al., 2017; Kitsios et al., 2017; Alhassan et al., 2016; Q. H. Cao et al., 2016; Shin & Choi, 2015; Zwart, 2015; Demchenko et al., 2014).

Understanding the challenges of complex collaboration schemes for data governance (Gregory et al., 2021) requires attention to at least three interrelated aspects: the structure of collaborative data organization, motivations, and relational patterns (Grover & Kohli, 2012; Tiwana et al., 2010). So far, studies have focused on the structural analyses of data governance (De Prieelle et al., 2022; Jarvenpaa & Essén, 2023; Tenopir et al., 2015), distinguishing between centralized and decentralized data ecosystems. Centralized ecosystems are managed by an orchestrator or steward, such as a public or private organization, which holds authority and ensures the ecosystem's success. Decentralized ecosystems, on the other hand, are participatory, with stakeholders holding different roles (e.g., data providers, users, technology providers) and collaboratively making decisions to advance the ecosystem (Dhanaraj & Parkhe, 2006; Ford et al., 2011; Gupta et al., 2020; Jarvenpaa & Essén, 2023; Morabito, 2015; Pikkarainen et al., 2017; Susha, van den Broek, et al., 2023).

While these studies are important, they often present static views of data ecosystems and treat data as assets that stakeholders, within well-defined roles in well-defined ecosystems, possess, accumulate, quantify, and exchange (Cox & Ellsworth, 1997; De Mauro et al., 2015;

Demchenko et al., 2013; Günther et al., 2017; Parmiggiani et al., 2022). However, the lack of structured roles, rules, and practices, combined with the complexity of relations, make data ecosystems a highly uncertain and constantly evolving endeavor (De Prieelle et al., 2022; Jarvenpaa & Essén, 2023; Lee et al., 2018; Oliveira et al., 2019). In response, an increasing number of recent studies call for attention to emergent and fluid roles in data ecosystems, which are likely to influence key data governance dilemmas such as data ownership, security, sharing incentives, usability, and usefulness. From this perspective, instead of assigning ex ante specific data roles like owner, steward, or provider (Jarvenpaa & Essén, 2023), we may need to explore how these dynamic roles, rules and practices influence the governance and development of data ecosystems (Aaen et al., 2022; Lee et al., 2018; Lis & Otto, 2020; Markus et al., 2012; Susha et al., 2019).

In line with these observations, the Information System (IS) literature has increasingly argued that the value of data depends on the 'backrooms' of data production, where actors engage in 'data work' - emergent, ongoing and context-dependent practices of setting up, producing and curating data (Aaltonen & Tempini, 2014; Aversa et al., 2019; Baesens Leuven et al., 2014; Davidson et al., 2023; Gelhaar & Otto, 2020; Monteiro & Parmiggiani, 2019). This perspective is crucial for moving beyond static visions of data ecosystems but has produced limited insights into data sharing practices, their drivers, and interplay at the supra-organizational level (Baesens Leuven et al., 2014; Howe et al., 2018).

In this paper, we build on these insights to explore how data ecosystems are characterized by the interplay of different dispositions towards data sharing, influenced by stakeholders' motivations, cognitive evaluation and social dynamics. To better unpack this interplay, we now turn to the literature on data sharing and willingness to share.

3.2.2. Willingness to share: toward a practice-based relational perspective

It is widely acknowledged that much of what happens in data ecosystems depends on human intentions, as stakeholders contribute to data sharing based on specific goals and interests (Fosso Wamba et al., 2018; A. Gupta et al., 2020; Lavalle et al., 2020; Mcafee & Brynjolfsson, 2012). This brings to the fore the need to study data ecosystems from the perspective of their stakeholders' willingness to share.

Although, to our knowledge, there is no universally accepted definition of the concept of willingness to share (WTS), numerous studies describe it as an individual's, organization's, or community's readiness and openness to share information, data, knowledge, or resources with others (Ackermann et al., 2021; Bansal et al., 2016; Kayhan et al., 2016; Marwick et al., 2018; Roeber et al., 2015; J. F. Taylor et al., 2015). In the digital age, understanding the inclination to share data has become crucial, leading to extensive research.

Rather than being a simple trade-off, WTS involves a multi-dimensional evaluation of both benefits and concerns. On one hand, the vast volumes, diverse sources, low cost, abundance, and dynamism of data can significantly enhance aspects such as operational efficiency, strategic decision-making, customer service, and product innovation (N. Khan et al., 2014; Manyika et al., 2011). On the other hand, data sharing presents significant challenges regarding capture, storage, search, analysis, and visualization, as well as issues of inconsistency, incompleteness, scalability, timeliness, privacy, and security (Ahrens et al., 2011; Kouzes et al., 2009).

What remains less understood is how these benefits and drawbacks interact in shaping actual sharing behavior. Stakeholders in data ecosystems must balance their different goals and interests - such as operational efficiency, strategic advantages, and innovation - to address multiple and cross-cutting concerns like data inconsistency and lack of analytic skills (Han & Dong, 2015; N. Khan et al., 2014). This underscores the need to explore the motivational

orientations for data sharing. For instance, personal interests play a crucial role: individuals are more likely to share data if they perceive direct benefits, such as resource protection, competitive advantage, power gain, social reputation, or incentives (Jentzsch et al., 2013; Leon et al., 2013; Shore et al., 2022; Struminskaya et al., 2021; Susha, van den Broek, et al., 2023; van den Broek & van Veenstra, 2018; Ziefle et al., 2016). Conversely, the willingness to share data can also be driven by the desire to contribute to the greater good. Promoting social welfare, public health, and supporting medical research are powerful motivators (Hillebrand, 2021; Hillebrand et al., 2023; K. K. Kim et al., 2017). Additionally, contextual factors such as governance arrangements within an organization or adherence to legal prerequisites and institutional conditions can significantly influence the willingness to share data (Ackermann et al., 2021; De Reuver et al., 2017; Polanin & Terzian, 2019).

In sum, the willingness to share data is shaped by a complex interplay of personal interests, perceived benefits, privacy concerns, and contextual factors. Yet, focusing only on individual or organizational motivations may overlook the social and relational dynamics that emerge in data ecosystems, where diverse actors must coordinate, negotiate, and co-evolve.

In an ecosystem where stakeholders' pursuit of different intrinsic values is interdependent, we propose viewing the willingness to share data not merely from an individual standpoint but as a collaborative endeavor with a relational dimension. From this perspective, by pursuing their own goals, actors can also create extrinsic value for others and contribute to collective purposes, such as solving grand societal challenges and fostering resilient, scalable, and durable ecosystems (Iansiti & Richards, 2006; Jacobides et al., 2018; Tiwana et al., 2010).

Despite growing interest in data sharing, relational aspects of ecosystem governance remain underexplored. Yet, many reasons associated with the failure of data ecosystems point to the relationship between stakeholders, from skepticism about others' ability to manage and derive value from data (Bakici et al., 2023; Tools et al., 2023), to lack of trust regarding others'

possible use of data (Parmiggiani et al., 2022) and difficulty in co-creating common visions of data ecosystems among different stakeholders (Tools et al., 2023).

While it is intuitive to think that the performance of an ecosystem and of the actors in the ecosystem depend on each other, the nature, and mechanisms of such part–whole interdependence remain to be explored. Therefore, despite the importance of prior studies, an integrated perspective on willingness to share data is still lacking (Kitchin et al., 2015).

Some studies suggest that a more process-based view can help account for the relational aspects of data sharing, such as how actors in an ecosystem relate to data and to each other: for instance, citizens may boycott or exhibit passive sharing behaviors if they perceive more risks to their privacy than opportunities from third-party data usage, but trust in those in possession of their data (e.g., governments, private companies) may play a mitigating role (Hillebrand et al., 2023). Similarly, even if private organizations expect commercial or strategic benefits from participation in data ecosystems, their willingness to share can be influenced by their relationships with other stakeholders, such as governments involving them in public administration agendas, and policymakers creating incentives schemes. Collaborations and relations in data ecosystems can thus impact sharing behaviors, sometimes even at the cost of potential individual advantages (Du et al., 2012).

Therefore, to better understand what facilitates or impedes data sharing, we adopt a practice-based lens that considers how perceptions, motivations, and inter-organizational relations interact over time. To contribute to a more comprehensive understanding of data ecosystems, we thus ask the following research question: *How do different factors (i.e., perceptions, motivations, and relations) associated with willingness to share shape data sharing practices in data ecosystems*?

3.3. Methods

3.3.1. Context

The foundation of our analysis rests upon the Labs for prototyping future Mobility Data sharing cloud solutions (MobiDataLab) project, a European initiative funded by the Horizon 2020 Research and Innovation Programme (H2020). This project was selected for its exemplary role in the development of a new data ecosystem for integrated mobility at the European level.

MobiDataLab operated from January 2021 to January 2024, seeking to address challenges in urban mobility and transportation through the utilization of data analytics. The primary goals of the project were to improve integration, efficiency, sustainability, and accessibility within European urban transportation systems. To achieve these objectives, MobiDataLab promoted data sharing practices among a wide range of stakeholders of the mobility sector. Throughout its three-year duration, MobiDataLab evolved into a data ecosystem for integrated mobility, aiming to foster trust, transparency, and accountability in urban mobility initiatives, ultimately striving to create more livable, accessible, and environmentally friendly cities. Initially composed of 10 partners from various sectors - including ITS providers, academics, research institutions, policy makers and transportation associations spanning across 7 countries (further details can be found in Table 4) – the project gradually attracted interest and contributions from a wide variety of stakeholders in the mobility sector. These included researchers interested in mobility trends, private organizations seeking to identify bottlenecks and create new value in transportation networks, policymakers and government agencies involved in urban planning and transportation policy, and local communities looking to improve their mobility and living experiences.

The collection, analysis, and interpretation of large-scale transportation datasets served as the foundation for developing the integrated mobility data ecosystem. Data sourcing and elaboration from various channels, such as GPS devices, mobile applications, public

transportation systems, and traffic sensors, were initially performed by the project's partners. Over time, groups of mobility stakeholders began contributing to the nascent ecosystem. MobiDataLab also emphasized public engagement, collaborating closely with government agencies, transportation authorities, community groups, and residents to ensure that research findings aligned with the needs and preferences of the local population.

Key activities of the MobiDataLab project included the creation of an Open Knowledge Base (OKB), a repository of resources on data sharing challenges and best practices in the mobility sector. This included real world use-cases and recommendations for improving data quality, accessibility, and usability. The OKB facilitated the development of the Transport Cloud, a prototype cloud-based platform designed to enable seamless sharing of transport data among mobility stakeholders. Additionally, Living and Virtual Labs were conducted - real-time experiments involving data exchanges and co-creation among local stakeholders - using advanced data analysis, including machine learning algorithms and statistical modeling, to derive actionable insights. These activities aimed to help researchers and policymakers to better understand mobility trends, identify bottlenecks, and optimize transportation systems.

The project aimed to foster a data sharing culture related to urban mobility by providing a collaborative and safe environment for multi-stakeholder experimentation with data sharing. Due to its novelty, the diversity of stakeholders involved, and its institutional embeddedness, the project did not initially focus on data commercialization, even if market studies on commercialization opportunities and potential business models are ongoing.

Given the diversity of stakeholders and practices involved, MobiDataLab provided a unique opportunity to study the creation of a data ecosystem. It offered insights into the perceptions, concerns, and expectations surrounding data sharing practices and contributed to a deeper understanding of the structural, motivational, and relational aspects of data ecosystem governance.

3.3.2. Data Collection

We collected data from various sources to support the theory-building process, including semi-structured interviews and archival data.

Interviews

To understand perceptions, visions, and willingness towards data sharing for integrated mobility solutions, we interviewed all the main categories of stakeholders in the MobiDataLab ecosystem. Since there were no clearly pre-defined boundaries for the categories of actors, we started by interviewing MobiDataLab's original partners, including academics, researchers, ITS experts, representatives of a non-profit association of municipalities, and policymakers. One author's involvement in some operational phases of the MobiDataLab project facilitated access to these stakeholders, helping to establish relationships with the project's partners and coordinators. Through these initial contacts, we reached other stakeholders who joined the ecosystem at later stages. Using a snowball sampling procedure, we extended our interviews to other relevant stakeholders involved in the Living and Virtual Labs activities during the project.

We conducted 40 semi-structured interviews between May 2022 and December 2023, covering 18 different categories of actors, as detailed in Table 4. All interviews were recorded, anonymized, and fully transcribed. The interview protocol was structured into four sections: 1) general information about the interviewees and their organizations' motivations, goals, and practices related to data sharing; 2) understandings of data and the role of data quality in data sharing; 3) collaboration practices related to data sharing; 4) visions and practices of data sharing within the ecosystem, and their evolution over time.

Archival Data

We began collecting archival data from the MobiDataLab Grant Agreement, which outlined the legal rights and obligations between the European Commission and the project partners. Additionally, we gathered and analyzed 24 deliverables produced within MobiDataLab to understand its main activities, goals, and responsibilities.

To better contextualize the MobiDataLab project and its dynamics, we also examined 12 previous and ongoing European, national, and international projects and initiatives related to data sharing. By studying these projects' websites, social media, and reports, which detailed their objectives, outputs, and achievements, we gained a broader view of data ecosystems in Europe. This analysis helped us understand the regulatory framework and context of data ecosystems, allowing us to compare our case study with different types of data ecosystems (such as research-driven, commercialization-driven, hybrid, externally vs. internally funded). We also delved deeper into participation schemes and governance arrangements. Initiatives for this analysis were selected through a project review led by the authors and supplemented with suggestions from interviewees.

Overall, this approach helped us better contextualize our findings regarding the characteristics and practices of the MobiDataLab project within the broader landscape of data ecosystems in Europe. Table 4 summarizes the multiple data sources exploited to support the theory-building process.

Table 4: Summary of the sources of data exploited to support the theory building process. Source: authors own work.

Source	Amount and Description	Details
Semi-structured interviews	*Total number:* 40, fully recorded-fully transcribed with 18 categories of mobility data sharing stakeholders. *Average length:* 60 minutes. *Timespan:* 2022 –2023.	Interviewed 8 partners of the MobiDataLab consortium spanning 7 countries. Expanded interviews to participants in MobiDataLab Living and Virtual Labs who took part in the mobility challenges and solutions identified by MobiDataLab and used snowball procedures from these stakeholders to further expand sample. The final sample included the following categories of actors: Researchers & Academics (4); Transport operators (4); Citizens involved in data sharing projects (2); policymakers (3); Municipalities (2); Association of Municipalities (2); Government (1); Public Transport Authorities (1); Rail Infrastructure Authorities (1); Traffic Management Centre (1); Trade Associations (2); Logistics Operators (5); Airlines (1); Intelligent Transport Systems (ITS) Providers (5); Software Providers (2); Cloud Providers (1); Car Manufacturers (2); Emergency services (1).
Archival data	*MobiDataLab documentation:* *Total number:* 27. *Timespan:* 2021 – 2023.	*Type of documents:* deliverables produced within the MobiDataLab project (report, technical diagram brochure, state of the art analysis, legal regulatory overview, data governance assessment, use cases definition, actors and stakeholders' definition, market analysis, data sharing business models, communication strategies, virtual lab description and analysis, events with the stakeholders' report, cooperation activities report). A complete overview of the MobiDataLab deliverables can be found at the following link: https://mobidatalab.eu/publications/
	Related projects analysis: *Total number*: 12. *Documents and topics analyzed:* website, social media, reports on consortium composition, main objectives, main outcomes.	Related projects, reports and docs suggested by informants and/or retrieved by researchers: AEOLIX: https://aeolix.eu; SELIS: https://selisproject.eu; FENIX: https://fenix-network.eu; NAPCORE project: https://napcore.eu/; Inspire initiatives: https://inspire.ec.europa.eu/; Le Fabrique du mobilite: https://lafabriquedesmobilites.fr/; Mobility as a service: https://www.maas4eu.eu/; Convex: https://convex-project.de/index.html; Openstreetmap: https://www.openstreetmap.org/; TALOS: https://cordis.europa.eu/project/id/218081/reporting; Transmodel Norm: https://www.transmodel-cen.eu/use-of-the-transmodel/; W3C (World Wide Web Consortium): https://www.w3.org/

3.3.3. Data Analysis Using the Grounded Theory Approach

After transcribing and anonymizing the interviews, we used archival data to clarify references, allusions, and connections within the interviews. The anonymized data was then imported into an integrated database on the QACDAS platform, DEDOOSE. Guided by the grounded theory

approach, we classified the written material to identify recurring themes and categories of similar meanings, integrating empirical data with existing theoretical concepts, and new theory development (A. Strauss & Corbin, 1998; A. L. Strauss & Corbin, 1997).

In line with grounded theory methodology, we began the analysis by open coding to identify frequent concepts and recurrent themes. We then progressed from first-order codes to second-order themes though axial coding, systematically linking categories and subcategories identified during the initial open coding process (A. Strauss & Corbin, 1998). Open coding revealed various advantages and benefits associated with data sharing, as well as concerns about potential issues and risks, the need for changes to encourage data sharing, criteria for defining data quality, and prerequisites for more collaborative data sharing practices.

During axial coding, we identified a recurring distinction between motivations for and against collaborative data sharing, which we labelled as *perceived opportunities of data sharing* and *perceived risks of data sharing*. We also noticed that informants frequently referred to two main types of motivations for engaging in data ecosystems: social impact goals (e.g., environmental sustainability, contributing to next generation mobility services and smart cities, fostering partnerships) and utilitarian goals (e.g., efficiency, efficacy, legitimation). Analyzing the 12 data sharing projects selected for background context revealed similar patterns of opportunities and risks related to data sharing. Additionally, during axial coding, we observed that informants often referred to conducting 'checks' on other stakeholders to assess alignment with their own motivations for data sharing and to determine whether these stakeholders could potentially mitigate risks and/or enhance expected advantages. We termed these practices as *social evaluation of the ecosystem's conditions for data sharing*. Finally, we identified a set of data sharing practices reported by our informants, which we compared with practices from other projects and grouped them into three main categories based on meaning assimilation: *driving*, *scaffolding*, and *pulling*.

Following axial coding, we used selective coding to identify core categories encapsulating the main theme and storyline of the data, developing a comprehensive and coherent narrative explaining the phenomena observed – as recommended by Strauss & Corbin, (1997). Specifically, by comparing themes from axial coding, we found that utilitarian and social impact goals related to motivational orientations for data sharing, perceived risks and opportunities referred to cognitive aspects, and social evaluation of the ecosystem's conditions for data sharing echoed relational aspects. We thus focused on identifying the relationships between motivational, cognitive, and rational aspects that lead to different data sharing practices within an ecosystem. We ultimately labeled the interplay between these aspects as '*willingness to share'*.

Figure 4 provides the data structure behind our grounded theory development. Figure 5 presents the model, based on grounded theory, explaining the factors influencing data sharing practices in data ecosystems. Tables 2, 3 and 4 in the manuscript provide examples of the coding process and data structure described in this section.

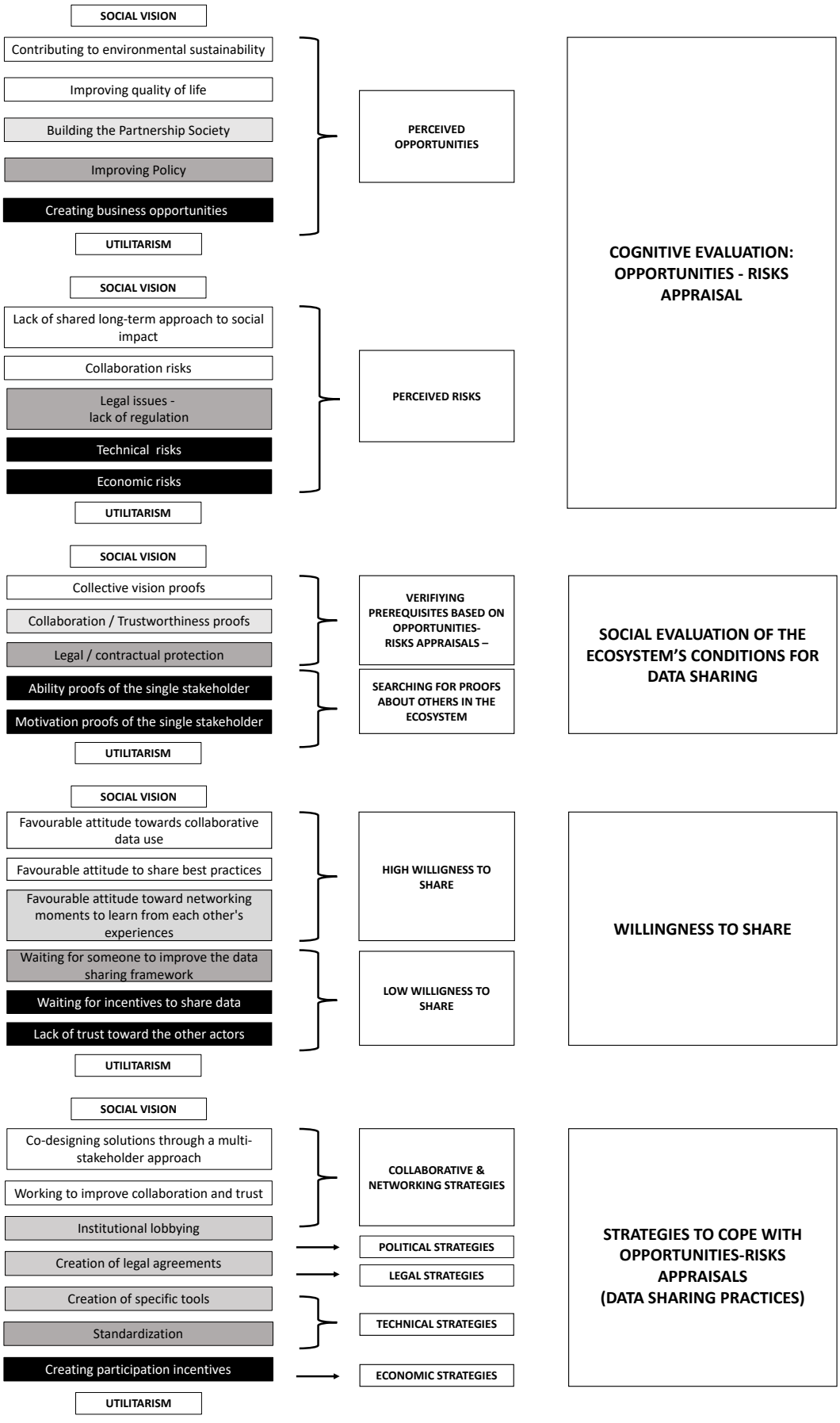

Figure 4: Data structure behind the grounded theory development. Source: authors own work.

3.4. Results

Our study reveals that data sharing practices within a data ecosystem are influenced by stakeholders' willingness to share data, shaped by their motivational orientations as well as their cognitive and social evaluations of the ecosystem's data sharing conditions. We identified two main motivational orientations: some stakeholders engage in data sharing to contribute to social impact and change, while others view it as a utilitarian tool for achieving personal goals. Cognitive evaluation involves how stakeholders assess the perceived advantages and risks of data sharing, while social evaluation involves assessing the social conditions within the partnership, including the stakeholders' interests, practices, and the nature of their collaboration. Interestingly, we found no correlation between sharing practices and stakeholders' socio-demographic characteristics and background (e.g., sector or industry) or their roles in the ecosystem (e.g., data provider, orchestrator, policymaker). Instead, data sharing practices were primarily influenced by the interplay between cognitive and social evaluations of ecosystem conditions and each stakeholder's motivational orientation.

Figure 5 illustrates the grounded model that explains the factors influencing data sharing practices in data ecosystems, highlighting that the willingness to share stems from the interplay between motivational orientations and cognitive and social evaluations of data sharing.

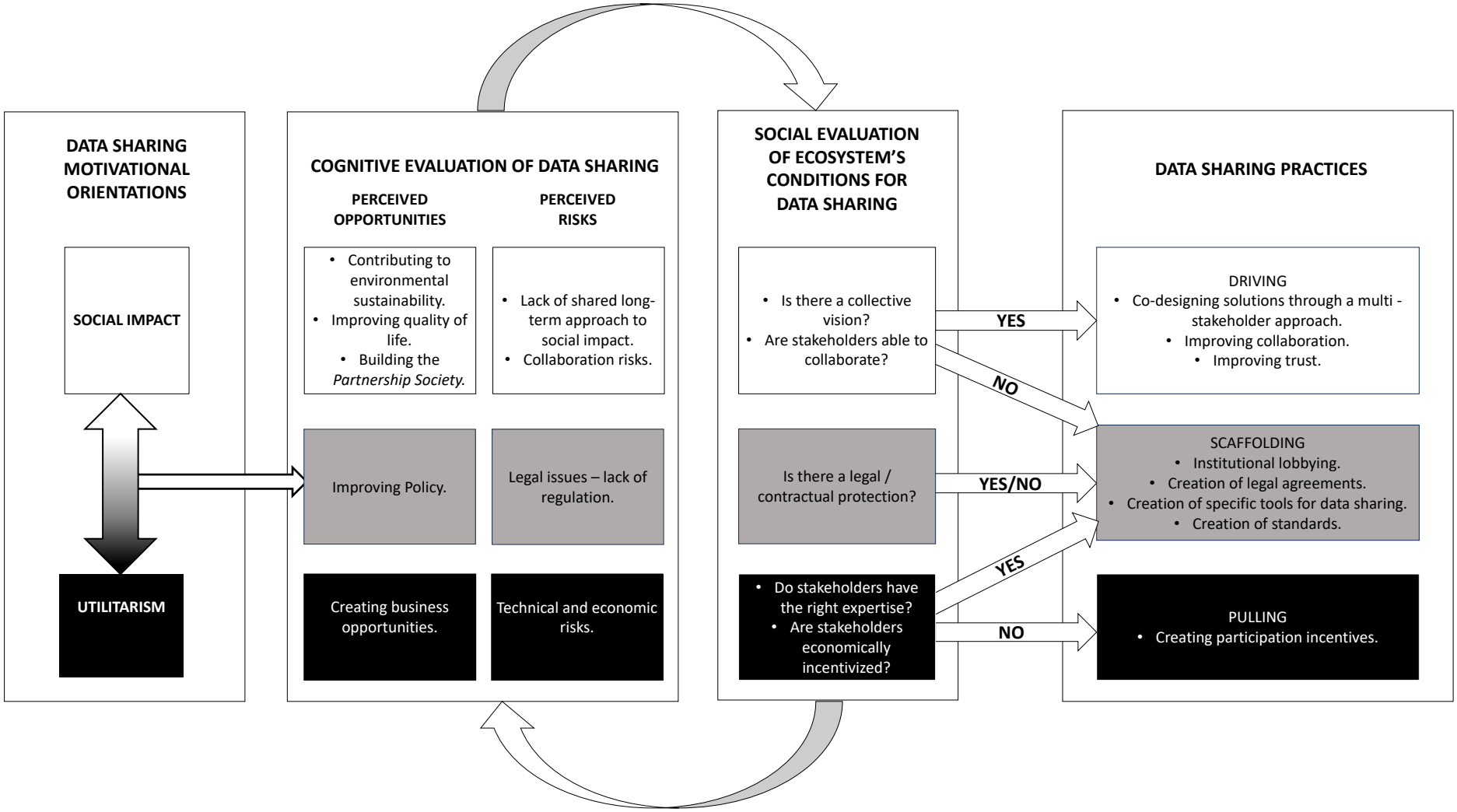

Figure 5: A grounded model of relation between willingness to share and sharing practices in data ecosystems: the interplay of motivational orientations, cognitive and social evaluation. Source: authors own work.

As depicted in Figure 5, we observed three types of data sharing practices within the ecosystem, each reflecting different degrees of willingness to share: *driving, pulling, and scaffolding*. Driving refers to highly proactive, collaborative practices, like co-design, open data sharing, and strong governance. These practices are enacted by stakeholders motivated by social impact, who find social proofs that the ecosystem can help them achieve their goals. Pulling practices are common among utility-driven stakeholders who prioritize minimizing risks when these outweigh the perceived benefits, thereby incentivizing other stakeholders to align with their expectations. Scaffolding practices serve as a compromise between immediate collaboration and future commitment, allowing stakeholders with different motivational orientations to contribute cautiously yet progressively as they balance perceived risks and benefits of data sharing.

In sum, we conceptualize data ecosystems as constellations of data sharing practices that are influenced by stakeholders' motivational orientations and socio-cognitive evaluations of the ecosystem conditions.

The following subsections of the findings (4.1 to 4.3) provide a detailed description of the main categories in our grounded theory model.

3.4.1. Data sharing motivational orientations: social and utilitarian impact

In the analyzed ecosystem, we observed two predominant motivational orientations towards data sharing. Some stakeholders prioritized the social impact of data sharing, while others viewed it as a utilitarian tool to achieve personal goals.

Participants with a social impact orientation often articulated a vision of data sharing's transformative potential, viewing it as a catalyst for significant positive global change. By actively engaging in the data ecosystem, these stakeholders aimed to address pressing global challenges, particularly in relation to environmental sustainability. They believed that data

sharing could enhance societal well-being, improving the quality of life for current and future generations. This perspective is reflected in the following excerpts:

> "If we can imagine a system where every kind of mobility data is shared, then we could benefit from having all this information shared between all the platforms, so maybe we can find opportunities to collect more goods, reduce the overall pollution and be more sustainable. If I had superpowers, I would infuse into the people's and companies' mind the idea that sharing is believing in a new kind of global world, which is not necessarily globalization, but it's a global place for everybody. So, we share the data, and all the processes get better, including sustainability" (Cloud provider 1).

Conversely, stakeholders with a utilitarian mindset prioritized their organizational needs over broader societal benefits, focusing on maximizing individual or corporate gains. They adopted a strategic approach to data sharing, aiming to optimize internal processes, resource allocation, and strategic decision-making to strengthen their competitive position and financial stability. Common goals included improving efficiency and effectiveness, accessing new markets opportunities, and achieving economies of scale and scope. These stakeholders tended to favor outcomes that aligned with their immediate interests, sometimes at the expense of wider societal or ethical considerations. The following excerpts illustrate this utilitarian perspective:

> "They're in business, so they have to make money. I just think that's just the nature of the private sector and the public sector as well. I mean, public sector is there really to try and improve our quality of life and deliver public services and the private sector? That's not strictly true, but when it comes to data, of course they're gathering data because they want to sell it. That's their business (…) Obviously, we are also focused on meeting our own goals, that's a primary requisite, I think we have different utilities in mind (…)" (Association of municipalities, 1).

As suggested in the excerpts above, both utilitarian and social impact motivations are common among stakeholders, regardless of whether they are public or private organizations, and irrespective of their roles within the partnership. It is important to highlight that social impact, and utilitarian motivations exist on a continuum rather than as a strict dichotomy, with many stakeholders falling somewhere in between (see the "grey zone" in Figure 5). While numerous stakeholders mentioned both type of motivations, they often leaned more heavily toward one

aspect when discussing their reasons for engaging in data sharing. However, mixed motivations were also frequently observed, where stakeholders sought to contribute to the common good while simultaneously pursuing strategic market or financial goals. These mixed motivations were often reconciled through a stronger commitment to data quality and improving end-user experiences. The following excerpt illustrates how generating high-quality data and using it to enhance user experience was perceived by stakeholders as a means of achieving utilitarian goals while also contributing to the greater good:

> "Data sharing and quality of data is always at the center of our daily activities. So, improving the data on user experience as well as providing stakeholders with the best data and also connected among different processes. We can look at the data in public transport from a different point of view. We can look at the sales, monetary part of the data. We can share with the transport company the profile of this user, whether a student or a worker. Where does he move from to? Does he go from school to home or from home to the workplace? Does the user use the transport network during the weekend or after school or working hours? So, the improvement of the quality of the data that we share is constantly part of our job, and it can reconcile different goals out there, offering some advantages for everybody involved in data sharing (…)" (Municipality, 1).

Our findings suggest that motivational orientation plays a pivotal role in shaping an ecosystem's data sharing practices, including both stakeholders' cognitive and social evaluation of the ecosystem's data sharing conditions. The following section (4.2) delve into these aspects in greater detail.

3.4.2. Socio-cognitive evaluation of data sharing: an account of opportunities and risks

In addition to motivational orientation, stakeholders' willingness to share data within a data ecosystem depended on their socio-cognitive assessment of the opportunities and risks associated with both the data and the ecosystem conditions for data sharing. Specifically, we found that all stakeholders interviewed in the MobiDataLab ecosystem conducted cognitive

evaluations of the potential risks and opportunities of data sharing, as well as social assessments of the supporting evidence, to help them establish their data sharing practices.

Cognitive evaluation risk-opportunities account

We found that perceived opportunities of data sharing were often accompanied by perceived risks, leading stakeholders to engage in what we termed as *accounts of opportunities and risks*. While all interviewed stakeholders agreed that data sharing involved multifaceted risks and opportunities, these perceptions ranged along a continuum from social impact to individual value creation. Stakeholders focused on social impact predominantly emphasized on social impact opportunities, whereas those with prevalent utilitarian motivations focused on technical and economic concerns.

Regarding social impact opportunities, stakeholders frequently highlighted environmental sustainability as a key issue. They envisioned data sharing as a tool to optimize supply chains, reduce pollution, and promote sustainable transportation, ultimately improving the quality of life for transportation end-users. Additionally, many referred to the opportunity of creating a *'Partnership Society'*, founded on a culture of data sharing that could enhance operational efficiency and drive continuous innovation across industries. However, significant risks were also noted. Aligning goals among stakeholders - especially those focused on social impact - was identified as a major risk due to the lack of a shared long-term approach to societal challenges. This misalignment could hinder the establishment of a common agenda for data sharing, limiting collective efforts to address complex issues like environmental sustainability. Excerpts in Table 5 below exemplify these themes identified in the data.

On the utilitarian side, data sharing presented both technical and economic opportunities and risks. A wide range of MobiDataLab stakeholders, from startups to established companies and public organizations, noted that participating in data ecosystems could create new business

opportunities, drive innovation, and provide competitive advantages. However, they also expressed concerns about technical challenges and economic risks. On the technical side, while engaging in the construction of the Common Data Knowledge (CDK), stakeholders highlighted inaccuracies in user-shared data, the presence of multiple standards across databases, and the lack of common standards. These challenges required extensive analysis before effective utilization, complicating data sharing and hindering the efficiency of shared information in the mobility sector. Economic concerns also weighed heavily on stakeholders. For instance, they worried about the lack of return on investments, the escalation of costs due to technical issues, and the potential loss of competitive advantage. These concerns are illustrated by the excerpts in Table 5, which highlight the significant obstacles to adopting data sharing practices.

Furthermore, intermediate positions that combined multiple advantages and disadvantages were a common pattern. For instance, many stakeholders we interviewed oscillated between social impact and business opportunities. We found that stakeholders with a mixed motivational orientation (see the "grey area" of Figure 5) were more inclined to shift between social impact and utilitarian concerns, focusing on mid-range opportunities and risks such as contributing to policymaking and navigating with legal complexities. These stakeholders viewed data sharing projects as opportunities for institutional change and improved policy effectiveness in areas like transportation, job creation and workforce diversification. However, they also expressed fears regarding legal and regulatory complexities. The lack of clear regulatory frameworks complicates the possibility to define permissible datasets, ensure interoperability, and foster clear contracts for data sharing, while contributions to policymaking can be highly resource-intensive, as explained by informants in the excerpts in Table 5.

In summary, our analysis shows that motivational orientation significantly shapes how stakeholders evaluate the risks and opportunities of data sharing, and how they develop

accounts of risk and opportunity. Stakeholders with a social impact orientation mainly focused on social impact opportunities and risks, while those with utilitarian motivations highlighted opportunities and risks related to individual advantages such as business strategy, technological advancements, and economic factors. However, although initial motivation influences the cognitive evaluation process, opportunities and risks are evaluated transversally and often represent a continuum, with many evaluations falling in between. The key difference lies in the priority and weight assigned to each type of opportunity and risk, depending on the motivational orientation.

Table 5 below synthetizes the main themes regarding the cognitive evaluation of opportunities and risks of data sharing within the MobiDataLab data ecosystem, providing exemplary data excerpts for each theme.

Table 5: Cognitive evaluation: opportunities and risks of data sharing within data ecosystem. Source: authors own work.

View of Data Sharing	Opportunities	Risks
Environment	**Contributing to Environmental Sustainability:** "The main opportunity is to contribute with digital information that is shared to a more climate-friendly environment and to a better and more effective mobility system. If you know what the current information is about what's going on in the traffic, you'll not have delays in public transport, nor traffic jams. The data exchange can contribute to effectiveness and to more sustainability" (ITS provider – 2).	No risks have been detected regarding the relationship between data sharing and environmental sustainability.
Relational	**Improving Quality of Life:** "The data can improve life in so many ways. It is clear that if I have a more massive, punctual, and efficient data exchange, and I can make my work more efficient, I can save on consumption and pollute less.	**Collaboration problems:** "So, we have a quite heterogeneous set of stakeholders coming from a law department, from community building, public relations stakeholders, companies, research centers. Every stakeholder has a different view on the project.

View of Data Sharing	Opportunities	Risks
	I will have fewer running trucks waiting outside the terminal, I will make the cranes work a lot less, because I have the data on the position of the containers, it is a cascade effect. The train, if I know it arrives earlier, I can better organize people's shifts, fewer people who drive on the street with the car. A more efficient data transmission has a positive effect on sustainability and on social issues, less acoustic noise, and less traffic" (CEO 1, railway transport operator).	This was in the beginning, and this will be at the end as well if we don't narrow these perspectives" (ITS provider 1). "I don't believe in fairy tales, so I don't think that everybody is transparent, honest and works as an actor for the global wealth" (Policymaker 3).
Social	**Building the Partnership Society:** "We really have the feeling that, even though the advantage is not immediate for some companies, they really have this vision to create this digitized process and that they are really into this Community feeling to strive this common goal of achieving an improved, more efficient digitized process (…) based on information partnerships" (Project manager 1 in an airline company). "Collaboration for data sharing is key and at least in the next 10 years cannot be substituted by data analytics and machine learning. People need to work together, need to understand the different positions to commonly agree on the way forward" (Coordinator 1 of a Transport Agency).	**Lack of shared long-term approach to social impact (wicked problems):** "We are all working at a different speed with a different attention towards the subject of digitalization. And that makes it quite difficult to put in your effort and to collaborate on a national scale (…) I think it largely depends on how much each of us will commit to data sharing" (Municipality representative, 3). "Today there are no strategies (for the data sharing). Data is shared according to the time axis with which it is to be used. With time, if things will change, also our involvement could be a different one" (CEO 1, railway transport operator).
Legal	**Improving Policy:** "Of course, you need data if you want to be able to understand what the impacts of policies are in these areas. If you're implementing a policy for job creation, if you're implementing a policy to encourage more women in the job market, then of course you need data too (…) you need to have access to it, and you may also need to share it" (Representative of an association of municipalities).	**Legal issues - lack of regulation:** "Well, I can speak for the work we are doing, so there are a lot of legal challenges. (…) What datasets can be provided and how to integrate them? There are also interoperability issues. How to make sure that the different datasets work and communicate with each other to make the systems work? Then there is again legal issue about contracting on how to share the data" (Researcher 1).

View of Data Sharing	Opportunities	Risks
Creating Business Opportunities	**Creating Business Opportunities:** "The main goal for business (through data sharing) is to make money. Environmental aspects to me seems to be more a way to make somebody look prettier (Policymaker 2). "If we look at who is pushing for sharing to see what are the, the reasons (..), um, sometimes some organizations maybe smaller, maybe the new entrance in an industry, or (..) data sharing becomes a way to position in and to differentiate against competition" (ITS provider 2).	**Economic Problems:** "Let's be honest, every organization is afraid of sharing data because they will lose something" (Municipality representative 1).
Technical	**Technical Opportunities: "**It's not only always a matter of revenues, but let's say that's a lot of companies have an interest pushing their own standards, (…) At the same time, pushing the standardization through data sharing requires some efforts and as a cost and the balance is not always in favor" (Citizen, 2).	**Technical Problems:** "We have already a lot of data shared by users about their mobility habits just using phone. The issue is that, in my experience, most of these data are not accurate, so they must be analyzed a lot before being used to find solution of mobility. And this is because many times there is no common standard on how this data should be collected" (Consultant 1). "The main challenges will be to be able to integrate all datasets from different parts of the world, of Europe" (Trip Planner1). "We still miss common standards and rules to be followed by everyone to share data in the same way. In our sector, specifically in the freight transport sector, there are no rules that we are obliged to follow when we must collect and share specific kind of data and it is so challenging to find partners that are able to share the data you need (…)" (Transport operator, 2).

Social evaluation: social proofs regarding the account of opportunities and risks

In addition to cognitive evaluation of the risks and opportunities of data sharing, MobiDataLab stakeholders also engaged in social evaluation of the ecosystem's conditions. Social proofs acted as heuristics, enabling stakeholders to assess whether the social conditions in the ecosystem could mitigate perceived risks, thereby positively shifting the risk-opportunity

balance, or whether they constituted further obstacles, increasing perceived risks. Thus, social evaluation was thus pivotal in shaping stakeholders' willingness to share data, as it allowed them to contextualize perceived advantages and risks within the ecosystem's social fabric.

We found that MobiDataLab's stakeholders used social evaluations to seek social proofs regarding the ecosystem's collective vision, collaboration practices, legal protection, and the motivation and incentives of individual stakeholders engaged in data sharing initiatives. Alongside cognitive evaluation of risks and opportunities, motivational orientation significantly influenced the search for social proofs. Stakeholders focused on social impact frequently sought evidence of common visions and collaborative practices within the ecosystem. Conversely, those with utilitarian goals primarily looked for social proofs related to the motivation, technical abilities, and economic incentives of individual stakeholders. Stakeholders in the 'grey area' sought proofs of legal and institutional protections for data sharing.

Stakeholders motivated by social impact sought social proof of a shared vision within the ecosystem, viewing it as essential for aligning diverse goals and ensuring effective collaboration. As explained by some, a collective vision helps mitigate risks such as standard fragmentation, opportunism, improper data use, and inefficiency, while promoting transparency, knowledge sharing, and resource optimization. Conversely, a lack of shared vision - indicated by incoherence between goals and actions, lack of common language, and evidence of greenwashing - signals potential failure in achieving desired outcomes.

Additionally, stakeholders sought effective collaboration practices within the ecosystem. Elements such as effective communication, coordination, negotiation, structured collaboration roles, consolidated best practices, and clear milestones and communication channels were considered proxies for successful data ecosystems. In contrast, the absence of these elements served as a warning sign that reinforced perceived risks associated with data sharing, as illustrated by the following excerpt and others retrievable in Table 6.

> “I would say collaboration is more important than the risks out there. I mean creating technology is something that one partner can do on their own, but the prerequisite for a new mobility ecosystem to work, is for the partners to be able to work together. So, if that doesn't happen, even if you have the technology, if you don't have the collaboration, then that is not possible to happen” (Researcher 2).

As indicated in the excerpts above, certain behaviors in the ecosystem were perceived by stakeholders with social impact orientations as warning signs. These included inconsistencies between declared goals and actions, a lack of common language in documents and communications, evidence of greenwashing, multiple short-term goals disconnected from broader social transformation, and futuristic goals about social transformation.

Differently, stakeholders with a utilitarian motivational orientation paid less attention to the ecosystem’s overall social conditions and focused on social proofs from individual stakeholders they interacted with. In such cases, social evaluation centered on these stakeholders’ technical expertise, motivations, and incentives to contribute to data sharing. Proof of technical ability was considered essential reassurance against the technical risks of complex data sharing solutions. Additionally, due to the intricate nature of technical, legal, institutional, and economic challenges in data ecosystems, stakeholders prioritize establishing economic incentives to encourage virtuous behaviors and discourage opportunistic, incorrect, or negligent behaviors that could undermine their interests and position within the ecosystem. This type of social evaluation is exemplified in the following excerpt containing the opinion of a Logistic Operator involved in mobility data sharing activities:

> “Private actors want the data only to make more money, so that’s maybe why also the public doesn’t want to share that data so easily. I’ve heard that some private companies get the data for free from the public, transform it a little bit and sell it for a higher price (…) Before data sharing, it is pivotal for us to assess that these operators are incentivized to act in a correct way, and sanctioned in case they are negligent or misuse data” (Logistic Operator 1).

Lastly, stakeholders with hybrid motivational orientations tended to focus on the contractual aspects of data sharing within the ecosystem. Legal protection mechanisms were viewed as crucial safeguards that could increase willingness to share by alleviating fears of improper or unethical data usage in the ecosystem.

Table 6 summarizes the evidence the interviewees sought as social proof to assess the opportunities and risks of participating in a data ecosystem, along with illustrative data excerpts for each theme.

Table 6: Social evaluation: social proofs regarding opportunities – risks account. Source: authors own work.

Social Proofs	
Is there a collective vision?	"Without community building, without exchanges between all the actors of the ecosystem, you could not go far (in such an ecosystem)" (Software Provider 1). "(When you set up data sharing systems) having a collective vision is not a prerequisite, but I think this collaboration and this knowledge sharing is super helpful because you generate transparency, you learn from each other. (…) It is super helpful to have a fruitful and efficient collaboration, and therefore I think the initiatives like MobiDataLab are super helpful for that (…). In my view (this) is what distinguishes successful cases of sharing from not so successful ones" (ITS provider 1).
Are stakeholders able to collaborate?	"I would say collaboration is more important than the risks out there. I mean creating technology is something that one partner can do on their own, but the prerequisite for a new mobility ecosystem to work, is for the partners to be able to work together. So, if that doesn't happen, even if you have the technology, if you don't have the collaboration, then that is not possible to happen" (Researcher 2).
Is there a contractual protection?	"So even if data is being shared, there must be rules and contract so that people, which are able of getting the data, use it only for good reasons and, of course, that those people who use it are known or knowable. It's not possible to share this data without knowing who is going to use it and in what way: this is a risk of having this data shared" (Logistics operator 1). "I think that comes down to contractual agreements. Of course, trust is one thing, when we have a new party on boarding our data sharing platform, there's always this contractual agreement that their data will only be used for the purposes that they intend, we call it the *data sharing rules*" (Airline project manager 3).

Social Proofs	
Do stakeholders have the right expertise? **Are stakeholders economically incentivized?**	"We do want to share data with municipalities, but they don't have a decent data platform on which to land this data; then it's going to be quite difficult to go for data sharing" (Municipality 1). "Choosing the right partner is a nuanced process, especially when considering various projects and activities. While the desire is to collaborate with all partners, practical limitations necessitate selecting those with the best ideas, a timely approach, and a profile suitable for the specific job at hand. The question of partner selection, therefore, becomes a delicate balance of fairness, encouraging active participation, and diversifying opportunities, even for those who may not have been highly active in previous collaborations" (Deliverable 5.3). "The companies know how to make data available, maybe they want the money to develop such an option or maybe they don't really want to share those data because they fear sharing it" (ITS provider 3). "I think there are times when different stakeholders try to find solution together and identify common standard to be applied to solve issue. But it always depends on the goals and abilities that each stakeholder has. So, if there is a common motivation it is reasonable to think there is an interest to investigate for new solutions but if there is no a common interest, it's difficult to solve problem (…) Are there at least technical competencies? That can be problematic for a public organization like ours and for private companies as well probably for everyone". (Public Transport Operator 1).

Our grounded model includes two circular arrows between cognitive and social evaluations to emphasize the potential feedback loop between these processes. This loop signifies that both social and cognitive evaluations contribute to forming risk-opportunities accounts, with aspects of one potentially leading to a reconsideration of aspects of the other. However, we have shown that social evaluation constitutes an additional step beyond cognitive evaluation, as it provides contextualization and solutions to risk-opportunity dilemmas within the ecosystem's social conditions. Consequently, social evaluations play a crucial role in influencing data sharing practices, as will be explained in the following sub-section (3.4.3).

3.4.3. From willingness to share to types of data sharing practices

We identified evidence of three distinct types of data sharing practices in the MobiDataLab ecosystem: driving, pulling, and scaffolding. When stakeholders with social impact orientations find social proofs supporting their prerequisites, they adopt mobilizing and participative data

sharing practices, which we refer to as ‘driving’. Conversely, when stakeholders with utilitarian perspectives do not find confirmation for their prerequisites, they engage in practices to incentivize other stakeholders to align with their expectations of value extraction. We refer to these practices as ‘pulling’. Lastly, we identified a set of preparatory practices aimed at creating structured frameworks to support future data sharing, which we termed ‘scaffolding’. These practices were enacted by different types of stakeholders, such as socially oriented individuals lacking sufficient social support, utilitarian whose prerequisites are socially confirmed, and stakeholders in the ‘grey area’.

The following excerpt illustrates the link between motivational orientation, cognitive and social evaluation, and data sharing practices. A project manager from one of the Municipalities involved in the MobiDataLab ecosystem highlights her/his organization’s focus on social impact and explains that their willingness to drive collaborative data sharing practices depends on the extent to which other actors in MobiDataLab demonstrate a shared commitment to a common cause, providing social proofs of their dedication “to the cause” and willingness to “work together”:

> “Provided we see that others are also dedicated to the cause, and everyone is working together, we always want to share all the data collected and we want to work collectively on spreading them out towards our communities or local governments so they can make a good use of it in their decision making (…) so they can give good services towards our population” (Project manager in municipality 1).

The following sub-sections describe these three practices, which are later summarized in Table 7, along with excerpts to clarify the descriptions.

Driving

Driving practices refer to proactive, collaborative behaviors such as co-design, open data sharing, and the implementation of data governance mechanisms that foster shared solutions

and relationships within a data ecosystem. These practices typically involve actors motivated by the desire to achieve social impact through data sharing. When these actors find cognitive and social reassurances about the risks of data sharing (e.g., perceiving social benefits overweight risks, and prerequisites like a common visions and effective collaboration are met), they develop positive risk-opportunities assessments. This leads them to engage in proactive practices of "*diving and driving data sharing*" as informants termed it. For instance, during the trials and errors of building the CKB, stakeholders driven by social impact concerns frequently mentioned transparency, accessibility, and commitment to open collaboration as key motivations for engaging in data sharing initiatives. When they also found social proofs that the ecosystem had a common vision and was governed through collaborative practices, they adopted a set of open data strategies, deliberately sharing information with other stakeholders in the ecosystem and ensuring that data was accessible to all.

> "We believe there are the right prerequisites for collaborative data sharing here therefore we are committed to push this forward (…) We want to strengthen and enhance data sharing on the basis of an open data approach, wherever we think that data should belong to the public, like information about current transport alternatives or delays in the traffic, but we are also aware of the risks out there so we want to share and strengthen data protection where individual data, for example, is concerned, and data sovereignty for the people taking part in the mobility system so that they can decide on their own what happens to their data which has connection to personal activities (…)" (ITS provider 4).

The main driving practice identified in the MobiDataLab ecosystem was co-designing solutions through a multi-stakeholder approach. This implied leading the setup of the CKB repository, identifying real world use-cases, offering recommendations to improve data quality, accessibility and usability, and collaborating on prototyping the Transport Cloud – a cloud-based platform designed to facilitate seamless sharing of transport data among mobility stakeholders. Driving practices also included organizing and moderating Living and Virtual Labs, which took the form of experimental co-creation sessions among local mobility stakeholders.

"More specifically, the main activities (within the Living Labs) were: identify, develop and contribute to the creation of open tools; allow the discovery of data for operations or research, static and dynamic, real-time data and historized data (Findable data); provide centralized access to mobility data (Accessible data); ease the conversion from one standard to the other (…) (Interoperable data); trigger and provoke interoperable interfaces to connect data providers, service providers and mobility clouds; prototype data processors for adding value to the data (…); demonstrate anonymization and privacy-preserving tools, in line with private data sharing principles" (MobiDataLab Grant Agreement).

As many informants explained, these experimental moments of collective work and co-creation played a crucial role in valuing the diverse experiences and expertise of stakeholders in the mobility sector. They also improved collaboration and trust among stakeholders, with one participant stating:

"Without community building, without exchanges between all the actors of the ecosystem, you could not go far (…) in setting up an ecosystem" (Software Provider 1).

This proactive and collaborative attitude distinguishes driving practices from the other data sharing approaches discussed in the following sections.

Pulling

The pulling practice involves creating incentives within the ecosystem to influence the behavior of others, aiming to maximize value extraction while minimizing personal risks. Predominantly exhibited by utilitarian stakeholders, this approach emerges after a thorough cognitive and social evaluation of data sharing opportunities and risks. Utilitarian stakeholders specifically seek evidence of technical expertise and economic incentivization from those with whom they have established business relationships. If social evaluations reveal that technical and economic risks outweigh the benefits, or if others' technical abilities and economic incentives are insufficient to provide reassurance, stakeholders adopt pulling data sharing practices. A

common pulling strategy involves offering economic incentives to encourage participation in data sharing while reducing personal costs.

Unlike driving practices, which promote proactive and open data sharing, pulling practices suggest a limited commitment to data sharing, with stakeholders often waiting for others to make progress before making significant moves. For instance, while pulling practices helped incentivize MobiDataLab's stakeholders to contribute to the creation of the CKB, they played a limited role in co-creation and experimentation activities, such as the cloud solution prototype and the organization of Living Labs. By adopting pulling practices, stakeholders took a cautious approach with limited contributions to ecosystem innovation. This strategy focused on safeguarding risks and maintaining a secure position within the ecosystem.

In the following excerpt, Municipality 3 expresses skepticism about businesses prioritizing profit over genuine environmental efforts, viewing sustainability as often superficial. It calls for economic incentives and stricter controls, like preventing "data cooking," to turn these statements into real action and ensure meaningful progress.:

> "The main goal for business is to make money. Environmental aspects to me seems more like a way to make somebody's speech look prettier (…) Since we do have an environmental agenda, we'll need to figure out how to turn socially desirable statements into real facts (…) (…) Until they (policymakers) do not economically incentivize this, and we don't control for irresponsible behaviors such as data cooking, we will not be able to make this come true (…)" (Municipality, 3).

Municipality 3 expresses skepticism that businesses prioritize profit over genuine environmental goals, seeing sustainability as mere rhetoric. This lack of trust in others' commitment leads them to adopt a passive approach to data sharing, focusing on monitoring behaviors and pushing for economic incentives to mitigate risks from opportunistic actions.

Scaffolding

Scaffolding practices were the most common data sharing practices in our dataset, characterized by a moderate and conservative data sharing behaviors along with preparatory actions to increase future engagement. Firstly, stakeholders with social impact orientation resort to scaffolding when lacking social reassurance against perceived risks. For example, if a collective vision or strong collaboration is absent, they remain committed but proceed more cautiously than those with positive risk-opportunity assessments (e.g., stakeholders who adopt driving practices). This is evident in the following excerpt from ITS provider 1, who expresses doubts about the ecosystem's ability to align diverse perspectives but remains hopeful about future contributions:

> "So, we have a quite heterogeneous set of stakeholders coming from a law department, from community building, public relations stakeholders, companies, research centers. Every stakeholder has a different view on the project. This was in the beginning, and this will be at the end as well if we don't narrow these perspectives. The first step is to find common goals where we do have overlap, where we are complementary, and I think the successful initiation of this project shows that partners can combine their different perspectives and can contribute to a common goal (…) We're definitely not there yet but we'll see, for now we'll just stay in the loop" (ITS provider 1).

Similarly, Airline 1 is cautious about engaging in current data sharing practices but remains optimistic about future developments. These kinds of stakeholders adopt a strategy of staying in the loop and preparing for what may come, balancing hope for future potential with a reserved approach for now.

> "I really think that the social community feeling will help a lot in this ecosystem (…): data is the new gold, and you have to believe that it is used for the right reason:(…) some advantages are not available immediately, but after a period of time when there will be an advantage for all, in the meantime we need to properly prepare for it" (Airline 1).

The main scaffolding practices adopted by these stakeholders focused on creating ecosystem infrastructure to facilitate future data sharing, including tools, platforms, and applications. A municipality representative, concerned about the lack of best practices for collaboration in the

mobility sector, emphasized that these platforms are essential for managing transport networks and aligning ecosystem interests with policy goals. Although reluctant to share data until the ecosystem becomes more trustworthy, they expressed interest in improving the ecosystem infrastructure by contributing to the CKB, and the cloud-based platform, as well as providing input for Application Programming Interfaces (APIs) (see excerpts in Table 7).

Secondly, stakeholders with utilitarian motivations engage in scaffolding when they perceive more risks than advantages or when the abilities of other stakeholders do not meet their technical or economic expectations. These stakeholders advocate for standardization to address technical and economic issues and to simplify relational challenges in the ecosystem, such as divergent objectives and difficulty to collaborate.

> "The aim is to provide suggestions about future standards and specifications to be adopted for improved Mobility data sharing and to show how these standards have been implemented in the different stages of the MobiDataLab project (…). To attain this objective, the methodology followed was based on the identification of standards and their context, their application in MobiDataLab, and their development and related support for implementation" (MobiDataLab, Deliverable D2.5).

For instance, a logistics operator emphasized the need for common standards to simplify data sharing in the mobility sector. However, since they found the current state of standardization in the CKB unsatisfactory, they adopted a cautious, fence-sitting approach, motivated by the fact that the CKB was still in its infancy:

> "We still miss common standards and rules to be followed by everyone to share data in the same way. In our sector, specifically in the freight transport sector, there are no rules that we are obliged to follow when we must collect and share specific kind of data and it is so challenging to find partners that are able to share the data you need (…) but if we manage to create an integrated system of standards each of us will be able to extract exactly what he or she needs, even if we don't know what exactly the others need (…) but there is still a long way to go, eventually we'll get there, I hope" (Transport operator, 2).

Notably, each stakeholder sought to promote their practices as standards within the entire ecosystem, as articulated by a logistics operator in the following excerpt:

"We can integrate (a new standard), and we can use it, but it would be better for everyone if they used our existing standard" (Logistics Operator, 1. See Table 7 for the complete excerpt).

Lastly, stakeholders in the 'grey area', motivated by both social impact and utilitarian concerns, engaged in scaffolding by developing legal and institutional frameworks for the ecosystem. These actors lobbied policymakers and regulators to influence the external context. For instance, an ITS provider emphasized their role in implementing new EU regulations for supply chain sustainability, demonstrating a proactive approach to define and executing legal standards:

"Starting 2023, the EU has a new regulation for the sustainability of supply chains, and we are going to be the first one to implement it" (ITS provider, 5).

In summary, scaffolding practices exhibit common elements that reflect a compromise between social impact and utilitarian motivations, aiming to build structured frameworks that address collaboration challenges. By contributing to infrastructure, standardization, and institutional lobbying, scaffolding practices serve as bridges between present and future conditions, facilitating a gradual move toward more robust data sharing. Stakeholders adopt a cautious approach in the present, envisioning a future of more efficient and risk-free data sharing. This strategy allows them to engage actively within a controlled environment, minimizing exposure while demonstrating moderate commitment to current ecosystem conditions. Notably, despite the cautious and /or critical attitudes, no stakeholder expressed an intention to exit the data ecosystem, even when they perceived more risks than opportunities and lacked reassuring social proofs. For these reasons, scaffolding practices constitute a convergence of various motivations and evaluation processes within the MobiDataLab data ecosystem.

Table 7 provides additional examples of driving, scaffolding and pulling practices, along with their connections to other categories identified in the data.

Table 7: Data sharing practices in a data ecosystem: driving, pulling, scaffolding. Source: Authors own work.

	Description	**Relation with willingness to share (motivational orientation and socio-cognitive evaluation of data sharing)**	**Sub-themes (second and first order) and exemplary excerpts**
DRIVING	Embodies a proactive and collaborative attitude among stakeholders, driven by a desire to co-design solutions and foster trust and collaboration within the ecosystem. Actors emphasize community building and exchange to drive progress collectively. Prioritizes collaboration over individual gains, recognizing that successful data sharing relies on mutual understanding and cooperation.	Actors typically have social impact motivational orientations. Engage in cognitive evaluation of data sharing process focusing on social impact like environmental sustainability and collaboration. Go through social evaluation of ecosystem's conditions, finding reassurance in collective visions, existence of collaborative practices and/or predisposition to collaborate.	<u>Co-Designing solutions through a multi-stakeholder' approach</u> "Well, we know that some urban contexts and municipalities are like pilot sites. So, they have some experience that some other cities do not have. Even the most advanced cities have different experiences with data sharing because they use different technologies, different approaches (…). Without a shared vision there is no way to make this work, and that is what we are committed to (…)" (Software provider 1). <u>Improving collaboration and trust</u> "Collaboration between partners is key and at least in the next 10 years cannot be substituted by data analytics and machine learning. People need to work together, need to understand the different positions to commonly agree on the way forward" (Coordinator 1 of a transport Agency).
PULLING	Focuses on creating incentives, particularly economic ones, to encourage participation in data sharing activities. Actors primarily driven by utilitarian motivations lead this approach, recognizing the power of financial rewards.	Actors with utilitarian motivational orientations engage in cognitive and social evaluation but don't find enough reassurance against perceived risks.	<u>Creating data sharing incentives, especially economic ones, for data sharing participation and influence behavior within the ecosystem through strategic incentive offerings.</u> "Well, the strongest incentives are always financial, I guess and, what our company for example also has done was to take care of all the financial cost dedicated to the innovation for startup that focus on data sharing" (Airport manager, 2).

	Description	Relation with willingness to share (motivational orientation and socio-cognitive evaluation of data sharing)	Sub-themes (second and first order) and exemplary excerpts
	Strategically offers incentives to stimulate participation while protecting stakeholders' interests, thereby influencing the behavior of others within the ecosystem.	Committed to creating and pursuing participation incentives offering strong economic advantages unrelated to data sharing directly, but to the participation in data haring activities and ecosystem.	"Everybody loves incentives, so if there is money involved, this goes directly into the heart of our company and so people will get crazy to get that money. So yes, economic incentives will help" (ITS provider, 4). "For the public no because they already wanted to do it, it's complex to do it. For the private, I don't see what kind of incentives you can give them, except money. They will do it if they are paid to do it. So, they won't do it for free in any case (…) until incentives are in place we will not be able to contribute too much with our own data and abilities" (Logistics Operator, 2).
SCAFFOLDING	Involves the creation of structured frameworks to support data sharing, such as legal agreements and standardization. Engages in activities like institutional lobbying and the development of tools and standards to facilitate data sharing while mitigating risks and ensuring compliance with regulations.	Actors with social impact data sharing motivational orientations may adopt this approach if they realize a lack of collective vision or strong collaboration among stakeholders. Actors with motivational orientation between social and utilitarian impact of data sharing. Actors with utilitarian motivational orientation may adopt this approach after cognitive evaluation, finding reassurance in social evaluation that stakeholders have the expertise and are economically incentivized.	<u>Infrastructuring: setting up technological tools and frameworks for data sharing (e.g., applications, platforms).</u> "So, in acquiring data they want to be able to manage the transport network in a better fashion and according to their policy objectives, and I'm not sure if those who are generating the data and selling the data not necessarily, they're doing it to improve quality of life. They're in business, so they have to make money. I just think that's just the nature of the private sector and the public sector. I mean, public sector is there really to try and improve our quality of life and deliver public services and the private sector? That's not strictly true, but when it comes to data, of course they're gathering data because they want to sell it. That's their business (…) And that's why we need to be careful about what we do in terms of data sharing, just take it easy, one step at a time (…) platforms can help increase transparency and figure out where opportunities and dangers are" (Municipality 2). <u>Standardizing: developing standards to facilitate data sharing</u>

	Description	Relation with willingness to share (motivational orientation and socio-cognitive evaluation of data sharing)	Sub-themes (second and first order) and exemplary excerpts
			"We can integrate (a new standard), and we can use it, but it would be better for everyone if they used our existing standard; in this way we could improve our service for everyone and also, by following the standards, our customers could benefit from all the new features and evolutions we could do to our main products, so basically we can improve the quality by working more on the existing standards" (Logistics Operator, 1). <u>Legal and institutional lobbying: strategically engage in institutional and legal advocacy activities</u> "We share ideas, and we work with apparatus like European Commission, with the European parliament, to create a legislation of data sharing" (Vehicle manufacturer, 1). "If there is a regulation to share, it's very easy: then you have to share and you are not worried about the kind of data out there because there is a legal agreement that can protect you" (City representative, 1). "The main area that I'm working on, and that's more related to European policy and new rules that we might see in the future" (Association of municipality, 1).

3.5. Discussions

Data ecosystems present a novel approach to addressing societal challenges through collaborative participation from multiple stakeholders. Existing literature highlights the importance of understanding the various factors that shape data sharing practices within these ecosystems, including social dynamics, perceptions of sharing, and motivations (Alhassan et al., 2016; Davidson et al., 2023; De Prieelle et al., 2022; Jarvenpaa & Essén, 2023; Lis & Otto, 2020; Morabito, 2015).

Our study aimed to address this gap by exploring the factors that contribute to the co-existence of different data sharing practices within a data ecosystem. Through qualitative research in a data ecosystem, we found that data sharing practices are influenced by a recursive relationship between motivation, cognitive evaluation, and social evaluation. Our contribution enhances the emerging literature on data ecosystems by examining data sharing governance practices within complex multi-stakeholder collaborations and providing a reconceptualization of willingness to share data in these contexts.

3.5.1. Advancing the understanding of governance mechanisms in data ecosystems: from structural arrangements to motivational, cognitive, and relational practices

Understanding how data sharing practices coexist and interact is fundamental for advancing our comprehension of complex multi-stakeholder data ecosystems.

We contribute to a practice-based view of data sharing within these ecosystems, demonstrating that those data sharing practices heavily rely on human intentions, with stakeholders engaging based on their specific objectives and interests (Günther et al., 2017; Monteiro & Parmiggiani, 2019; Pesce et al., 2019; van den Broek & van Veenstra, 2018). Our perspective sheds light on ‘backrooms’ of collaborative data practices in multi-stakeholder

environments, enriching both the emergent literature on data ecosystems (among the others: Aaen et al., 2022; Gelhaar & Otto, 2020; Lis & Otto, 2020; Oliveira et al., 2019) and the information system literature on data work (Aaltonen & Tempini, 2014; Aversa et al., 2019; Baesens Leuven et al., 2014; Monteiro & Parmiggiani, 2019; Parmiggiani et al., 2022).

Our findings unpack the role of social, cognitive, and motivational factors in data sharing practices. We argue that different motivational orientations shape perceptions of data sharing opportunities and risks, influencing social evaluations of the ecosystem's relational context. We refer to this process as the social construction of willingness to share, demonstrating its constitutive effect on the constellation of data sharing practices that form an emerging data ecosystem.

These findings are significant for several reasons. First, we complement existing literature, which often provides a static view of data ecosystems by focusing on structural factors rather than the motivations and relational patterns driving data sharing (Kitsios et al., 2017; Wang et al., 2021). While considerable attention has been given to data-related challenges - such as data quality, standards, and analytical capabilities - as distinctive features of these ecosystems (Bonina et al., 2021; Shin & Choi, 2015), much less is known about the roles of the actors managing the data. Our work fills this gap and contributes towards a practice-based view of data ecosystems.

Second, while some studies have shown that data are not merely neutral objects reflecting an independent reality (Burton-Jones et al., 2021; Jones, 2019), but shaped by practices of production, dissemination, and use (Bailey et al., 2011; Davidson et al., 2023; Monteiro & Parmiggiani, 2019; Parmiggiani et al., 2022), we still know relatively little about how data work unfolds in complex multi-stakeholder settings such as data ecosystems. Our study thus aims to extend the conversation around data work beyond organizational boundaries to include the dynamics of data governance at the ecosystem level.

Third, our practice-based approach contributes to studies that document ambivalent cultures of data sharing within and across organizations. These studies highlight the tension between the consensus that data sharing can generate positive impact (Günther et al., 2017; Raghupathi & Raghupathi, 2014), and the lack of understanding regarding the costs and benefits associated with data use (or misuse) in such ecosystems (Gelhaar & Otto, 2020; Margetts et al., 2013; Oliveira et al., 2019). This results in an ambivalent data-sharing culture, where the potential to address significant economic, environmental, and social challenges is tempered by operational and organizational hurdles, necessitating a balance between controlling owned data and trusting partner organizations (Eurich et al., 2010; M. Gupta & George, 2016; Klievink et al., 2018; Zrenner et al., 2019). Our study provides a nuanced understanding of how the manifold risks and opportunities of data sharing impact data governance, particularly shaping stakeholders' data sharing behaviors.

Importantly, our study contributes to a renewed conceptualization of data governance in data ecosystems. We align with Aaen et al.'s (2022) calls for pragmatic and political views on data ecosystems. Accordingly, while emerging research has focused on data value creation (Azkan, Iggena, et al., 2020; De Prieelle et al., 2022; Gelhaar & Otto, 2020; Oliveira et al., 2019), we argue that an ecosystem perspective also allows for a critical understanding of the politics that emerge as data are shared, repurposed, and reused across organizational contexts. From this standpoint, we argue that data collaborations at the ecosystem level may require dramatically new approaches to data governance within and across organizations, incorporating sophisticated relational and political practices (Davidson et al., 2023; Donge et al., 2022; Gregory et al., 2021).

In these contexts, each actor can perform multiple fluid roles, leading to complex interdependencies where actors simultaneously cooperate and compete (Gelhaar & Otto, 2020; Guggenberger et al., 2020; Nalebuff & Brandenburger, 1997). To contribute to this vision, we

unpacked the pragmatic and relational processes of data sharing within the MobiDataLab ecosystem. By analyzing the interaction of motivational orientations, cognitive assessments, and social evaluations, we offer new insights into the governance of multi-stakeholder data ecosystems, viewing it as a dynamic interplay between different sharing practices rather than a static collection of roles, rules and responsibilities. We found that sharing practices depend heavily on the interplay between pragmatics - perceptions of risks and opportunities associated with data ecosystems - and politics - assessment of the ecosystem's social and relational context.

We thus highlight the complexity of decision-making processes within data ecosystems, where stakeholders balance various motivations, evaluations, and strategies to advance divergent data sharing goals. In light of these considerations, we emphasize the need to move beyond taxonomies of roles and motivations in data ecosystems to understand why stakeholders display different attitudes towards data sharing and engage in various sharing practices.

3.5.2. Data ecosystems as constellations of sharing practices and the social construction of willingness to share

We have demonstrated that the emergence of a data ecosystem can be conceptualized as a constellation of data sharing practices. This highlights that not everyone shares data in the same way or for the same reasons. Different organizations may engage in collaborative data sharing to accrue resources, support innovation processes, improve data quality, enhance security or architecture, evolve data management culture, gain legitimacy, achieve social impact, or pursue any combination of these goals (in line with Bresciani et al., 2021; Devriendt et al., 2021; Y. Kim & Stanton, 2016). Thus, data sharing practices depend on how these motivations interplay and are negotiated among multiple actors. We refer to this process as the social construction of willingness to share. Our findings shed light on this ongoing exercise of sensemaking, driven

by individuals' motivations and pragmatic and political evaluations of the ecosystem's conditions.

So far, willingness to share has received limited attention in data ecosystem research (M. Khan et al., 2016). Our findings emphasize the need for a renewed and integrated perspective on willingness to share in data ecosystems. Since not everyone participates in data sharing for the same reasons or perceives the same advantages and disadvantages, commonly cited challenges of data ecosystems - such as data quality or standardization - may depend on how goals and interests co-exist and shape constellations of data sharing practices.

Previous studies have uncovered the various motivational determinants of willingness to share, including personal interests, incentives, and familiarity (Shore et al., 2022; Struminskaya et al., 2021; Susha, Rukanova, et al., 2023; Susha, van den Broek, et al., 2023; van den Broek & van Veenstra, 2018), as well as interests in contributing to the greater good (Hillebrand, 2021; Hillebrand et al., 2023; K. K. Kim et al., 2017). Other researchers have even suggested that roles in a data ecosystem (e.g., data collector, provider, steward) may depend on these underlying motivations (Azkan, Möller, et al., 2020; Gelhaar & Otto, 2020; Oliveira et al., 2019). A recent article by Susha et al. (2023) explores how organizations from various sectors can share data in partnerships to address societal challenges, finding that the partnership model type (resource-based, social issue, or societal) influences the drivers and challenges to data sharing. They conclude that, in the face of common risks and challenges, the willingness to engage with broader societal issues without immediate direct gains varies by partnership models and depends on partners' motivations. Our study adds nuance to Susha et al.'s (2023) findings by further exploring the motivational forces behind interorganizational data sharing. To our knowledge, ours is one of the first studies to document different data sharing dispositions within a motivational continuum ranging from social impact to utilitarian concerns.

However, in contrast to previous studies, we suggest that motivational orientations are just one building block of willingness to share data, alongside pragmatic and political evaluations enacted within the context of a data ecosystem. We have shown that stakeholders with social impact orientations often seek social proof regarding collective data sharing visions and collaborative practices. Depending on whether these prerequisites are met, they adopt either highly participative data sharing practices (*driving*) or more cautious practices (*scaffolding*). In contrast, stakeholders with utilitarian visions look for individual proofs of technical ability and financial incentives rather than for common visions and collaboration at the ecosystem level. Depending on whether these conditions are met, they engage in scaffolding practices or in more passive practices that incentivize others to align with their expectations (*pulling*).

These findings suggest that while initial motivations play an important role, the constellations of sharing practices also depend on cognitive and social evaluations of opportunities and risks within the ecosystem. Based on these findings, we call for studies of willingness to share data not only as a motivational process but also as a process of socio-cognitive assessment process. This perspective moves beyond the sharing-non sharing dichotomy that often characterizes debates about data sharing behaviors, especially in complex multi-stakeholder collaborations. The prevalence of scaffolding practices in our findings - i.e., preparatory practices aimed at creating structured frameworks to support future data sharing - emphasizes the need to further study mid-range sharing practices that allow stakeholders to navigate multiple motivations and socio-cognitive assessments. This aligns with Gunther et al. (2017) claim that many organizations are attempting to find a middle ground in the sharing-not sharing debate. For example, some organizations may collect data for their own purposes but refrain from selling it or allow individuals to 'opt out' of services that heavily rely on personal data (Greenaway et al., 2015). As Gunther et al. (2017) suggest, however, finding this middle ground is particularly challenging, especially when organizations aim to create social value. For

instance, a public organization may adopt data-driven surveillance to enhance public security, while simultaneously compromising other types of social value such as citizens' privacy rights (Newell & Marabelli, 2015). Our study documented the dilemmas between utilitarian and social impact motivations, illustrating the importance of scaffolding as a middle-ground strategic approach with political and pragmatic implications.

Interestingly, we found that stakeholders did not intend to leave the ecosystems, as even the most skeptical and frustrated participants in our study reported plans to pursue scaffolding practices, such as creating technological tools, engaging in legal and institutional lobbying, or advocating for standardization. Given the high occurrence of scaffolding practices in our case study, we invite future research to explore the role of middle-ground (i.e., the 'grey area' in our model) strategies in the evolution of multi-stakeholder data ecosystems. Future studies could investigate the conditions under which scaffolding becomes a dominant practice potentially able to stabilize a data ecosystem in the long-term perspective. Researchers should also examine potential counterstrategies enacted by stakeholders who adopt different data sharing practices - such as driving and pulling - following the path suggested by George (2014).

3.6. Conclusion

The analytical work conducted for this research is not devoid of limitations. Firstly, our examination of data sharing practices is not intended to be exhaustive. Investigating other data ecosystems could validate some of the patterns identified in this study while revealing additional ones. Our case study focuses on open data ecosystems for research and social policy purposes, rather than ecosystems centered on data commercialization, which may display different patterns.

Additionally, future research could further explore the relational dynamics among data sharing practices discussed in this study, examining how driving, scaffolding, and pulling

practices interact and influence each other within the ecosystem context. This includes analyzing how stakeholders interpret and respond to distinctions between these practices during their interactions.

Furthermore, while we identify certain risks and opportunities within the cognitive evaluation and specific social proofs related to the social evaluation of the ecosystem conditions, other factors may also play significant roles in shaping attitudes toward data sharing (Kazantsev et al., 2023).

Our study does not aim to provide a comprehensive taxonomy of data sharing practices, and we acknowledge that other practices may exist in different ecosystems. Although we documented a motivational continuum between social impact and utilitarianism, we do not exclude the possibility that in other contexts, motivational orientations could manifest along different, more relevant dimensions. An additional limitation is that our findings are strongly anchored in informants' reports of their perceptions, motivations, and practices, due to limited access to direct observation of their interactions. Finally, future studies should consider the dynamic nature of data ecosystems. We studied an ecosystem in its initial stages and followed its development, but new dynamics that are not covered in this study may emerge as MobiDataLab evolves and changes.

In terms of implication for practice, we provide insights for potential participants in data ecosystems, including private and public organizations, citizens, and policymakers. Our primary goal was to offer a realistic account of the possibilities and pitfalls of data ecosystems, beyond current idealizations and criticisms regarding the roles of data in our society. We thus argue for the need to move beyond idealized notions of successful data ecosystems characterized by unified visions, practices, and intents. We highlight both the risks, and the benefits associated with diverse data sharing practices and invite in-depth analyses of how

different motivations, perceptions and data sharing practices may shift the course of data ecosystems and impact their ability to reach expected outcomes.

Our study also aims to offer valuable guidance for policymakers, shedding light on both positive and negative dynamics that may arise in data ecosystems. By explaining the building blocks of willingness to share, we highlight potential actions to resolve motivational divergences – ranging from acknowledging diversity in data sharing perceptions to incentivizing desired behaviors and discouraging passive or hesitant practices to avoid ecosystem stagnation. Recognizing the motivation of various actors within the ecosystem is a crucial step toward resolving conflicts and ensuring high levels of engagement, thereby mitigating concerns that hinder progress.

Additionally, we emphasize the theoretical and practical importance of scaffolding practices in ecosystems characterized by heterogeneous, complex, and uncertain collaboration schemes. Future studies should investigate when and how the interaction between different data sharing practices, such as driving, scaffolding, and pulling, can propel an ecosystem forward, keep it in stasis, or lead it adrift.

In conclusion, this study has shed light on the factors influencing data sharing practices within mobility data ecosystems and invites future contributions to expand our understanding of these phenomena in other contexts and sectors, ideally by exploring related practices and mechanisms of data sharing.

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. https://doi.org/10.1080/0960085X.2021.1947753

Aaltonen, A., & Tempini, N. (2014). Everything counts in large amounts: A critical realist case study on data-based production. Journal of Information Technology, 29(1), 97–110. https://doi.org/10.1057/JIT.2013.29

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. https://doi.org/10.1016/J.IJINFOMGT.2019.07.008

Ackermann, K. A., Burkhalter, L., Thoralf Mildenberger, |, Frey, | Martin, & Bearth, A. (2021). Willingness to share data: Contextual determinants of consumers' decisions to share private data with companies. Wiley Online LibraryKA Ackermann, L Burkhalter, T Mildenberger, M Frey, A BearthJournal of Consumer Behaviour, 2022•Wiley Online Library, 21(2), 375–386. https://doi.org/10.1002/cb.2012

Agarwal, R., & Dhar, V. (2014). Big data, data science, and analytics: The opportunity and challenge for IS research. Information Systems Research, 25(3), 443–448. https://doi.org/10.1287/ISRE.2014.0546

Ahrens, J., Hendrickson, B., Long, G., Miller, S., Ross, R., & Williams, D. (2011). Data-intensive science in the US DOE: case studies and future challenges. Computing in Science & Engineering, 14–24. https://ieeexplore.ieee.org/abstract/document/5999634/

Alhassan, I., Sammon, D., & Daly, M. (2016). Data governance activities: an analysis of the literature. Journal of Decision Systems, 25(s1), 64–75. https://doi.org/10.1080/12460125.2016.1187397

Anwar, M. J., Gill, A. Q., Hussain, F. K., & Imran, M. (2021). Secure big data ecosystem architecture: challenges and solutions. Eurasip Journal on Wireless Communications and Networking, 2021(1), 1–30. https://doi.org/10.1186/S13638-021-01996-2/TABLES/13

Arrieta-Ibarra, I., Goff, L., Jiménez-Hernández, D., Lanier, J., & Weyl, E. G. (2018). Should We Treat Data as Labor? Moving beyond "Free" AEA Papers and Proceedings, 108, 38–42. https://doi.org/10.1257/PANDP.20181003

Aversa, P., Hervas-Drane, A., & Evenou, M. (2019). Business model responses to digital piracy. California Management Review, 61(2), 30–58. https://doi.org/10.1177/0008125618818841

Azkan, C., Iggena, L., Gür, I., Möller, F., & Otto, B. (2020). A Taxonomy for Data-Driven Services in Manufacturing Industries. In PACIS (Vol. 184).

Azkan, C., Möller, F., Meisel, L., & Otto, B. (2020, June). Service dominant Logic Perspective on Data Ecosystems-a Case Study based Morphology. In ECIS.

Baesens Leuven, B. K., Vanthienen Leuven, J. K., Baesens, B., Bapna, R., Marsden, J. R., JanVanthienen, B., & Leon Zhao, kuleuvenbe J. (2014). Transformational issues of big data and analytics in networked business. JSTOR. https://doi.org/10.25300/MISQ/2016/40:4.03

Bailey, D. E., Leonardi, P. M., & Barley, S. R. (2011). The Lure of the Virtual. Https://Doi.Org/10.1287/Orsc.1110.0703, 23(5), 1485–1504.

Bakici, T., Nemeh, A., & Hazir, O. (2023). Big Data Adoption in Project Management: Insights From French Organizations. IEEE Transactions on Engineering Management, 70(10), 3358–3372. https://doi.org/10.1109/TEM.2021.3091661

Bansal, G., Zahedi, F. M., & Gefen, D. (2016). Do context and personality matter? Trust and privacy concerns in disclosing private information online. Information and Management, 53(1), 1–21. https://doi.org/10.1016/J.IM.2015.08.001

Bonina, C., Koskinen, K., Eaton, | Ben, Gawer, | Annabelle, & School, S. B. (2021). Digital platforms for development: Foundations and research agenda. Wiley Online LibraryC Bonina, K Koskinen, B Eaton, A GawerInformation Systems Journal, 2021•Wiley Online Library, 31(6), 869–902. https://doi.org/10.1111/isj.12326

Bresciani, S., Ciampi, F., Meli, F., & Ferraris, A. (2021). Using big data for co-innovation processes: Mapping the field of data-driven innovation, proposing theoretical developments and providing a research agenda. International Journal of Information Management, 60, 102347. https://doi.org/10.1016/J.IJINFOMGT.2021.102347

Burton-Jones, A., Butler, B., Scott, S., & Xu, S. X. (2021). Next-Generation Information Systems Theorizing: A Call to Action. Management Information Systems Quarterly, 45(1). https://aisel.aisnet.org/misq/vol45/iss1/11

Cao, Q. H., Khan, I., Farahbakhsh, R., Madhusudan, G., Lee, G. M., & Crespi, N. (2016). A trust model for data sharing in smart cities. 2016 IEEE International Conference on Communications, ICC 2016. https://doi.org/10.1109/ICC.2016.7510834

Chaudhari, S. L., & Sinha, M. (2021). A study on emerging trends in Indian startup ecosystem: big data, crowd funding, shared economy. International Journal of Innovation Science, 13(1), 1–16. https://doi.org/10.1108/IJIS-09-2020-0156/FULL/HTML

Cox, M., & Ellsworth, D. (1997). Managing big data for scientific visualization. ACM Siggraph, 97, N1, 21–38. https://www.researchgate.net/publication/238704525

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. https://doi.org/10.1016/J.INFOANDORG.2023.100454

De Mauro, A., Greco, M., & Grimaldi, M. (2015). What is big data? A consensual definition and a review of key research topics. AIP Conference Proceedings, 1644(1), 97–104. https://doi.org/10.1063/1.4907823

De Prieelle, F., De Reuver, M., & Rezaei, J. (2022). The Role of Ecosystem Data Governance in Adoption of Data Platforms by Internet-of-Things Data Providers: Case of Dutch Horticulture Industry. IEEE Transactions on Engineering Management, 69(4), 940–950. https://doi.org/10.1109/TEM.2020.2966024

De Reuver, M., Sørensen, C., & Basole, R. C. (2017). The digital platform: a research agenda. Journals.Sagepub.ComM De Reuver, C Sørensen, RC BasoleJournal of Information Technology, 2018•journals.Sagepub.Com. https://doi.org/10.1057/s41265

Demchenko, Y., De Laat, C., & Membrey, P. (2014). Defining architecture components of the Big Data Ecosystem. 2014 International Conference on Collaboration Technologies and Systems, CTS 2014, 104–112. https://doi.org/10.1109/CTS.2014.6867550

Demchenko, Y., Grosso, P., De Laat, C., & Membrey, P. (2013). Addressing big data issues in scientific data infrastructure. 2013 International Conference on Collaboration Technologies and Systems (CTS), 48–55. https://ieeexplore.ieee.org/abstract/document/6567203/

Devriendt, T., Borry, P., & Shabani, M. (2021). Factors that influence data sharing through data sharing platforms: A qualitative study on the views and experiences of cohort holders and platform developers. PLOS ONE, 16(7), e0254202. https://doi.org/10.1371/JOURNAL.PONE.0254202

Dhanaraj, C., & Parkhe, A. (2006). Orchestrating Innovation Networks. Https://Doi.Org/10.5465/Amr.2006.21318923, 31(3), 659–669.

Donge, W., Bharosa, N., & Janssen, M. F. W. H. A. (2022). Data-driven government: Cross-case comparison of data stewardship in data ecosystems. Government Information Quarterly, 39(2), 101642.

Du, T., Lai, V., & Cheung, W. (2012). Willingness to share information in a supply chain: A partnership-data-process perspective. Information & Management, 49(2), 89–98. https://www.sciencedirect.com/science/article/pii/S0378720611000917?casa_token=h4Qym2uTtLAAAAAA:OuQXWlfyasYYeekAkl2yRWtAnzUPkUPI6-JhMb_BHAl02DlCOWXUGJhsbhSlzfT3QSbdth1I

Eurich, M., Oertel, N., & Boutellier, R. (2010). The impact of perceived privacy risks on organizations' willingness to share item-level event data across the supply chain. Electronic Commerce Research, 10(3), 423–440. https://doi.org/10.1007/S10660-010-9062-0

Farquharson, C., Mcnally, S., & Tahir, I. (2022). Education inequalities. IFS Deaton Review of Inequalities.

Ford, C. R., Laseter, S. H., Swank, W. T., & Vose, J. M. (2011). Can forest management be used to sustain water-based ecosystem services in the face of climate change? Ecological Applications, 21(6), 2049–2067. https://doi.org/10.1890/10-2246.1

Fosso Wamba, S., Gunasekaran, A., Dubey, R., & Ngai, E. W. T. (2018). Big data analytics in operations and supply chain management. Annals of Operations Research, 270(1–2), 1–4. https://doi.org/10.1007/S10479-018-3024-7/METRICS

Gelhaar, J., & Otto, B. (2020). Challenges in the Emergence of Data Ecosystems. PACIS, 175.

George, G., Haas, M. R., & Pentland, A. (2014). Big Data and Management. Https://Doi.Org/10.5465/Amj.2014.4002, 57(2), 321–326.

Greenaway, K. H., Alexander Haslam, S., Cruwys, T., Branscombe, N. R., Ysseldyk, R., & Heldreth, C. (2015). From "we" to "me": Group identification enhances perceived personal control with consequences for health and well-being. Journal of Personality and Social Psychology, 109(1), 53–74. https://doi.org/10.1037/PSPI0000019

Gregory, R. W., Henfridsson, O., Kaganer, E., & Kyriakoi, S. H. (2021). The role of artificial intelligence and data network effects for creating user value. Academy of Management Review, 46(3), 534–551. https://doi.org/10.5465/AMR.2019.0178

Grover, & Kohli. (2012). Cocreating IT Value: New Capabilities and Metrics for Multifirm Environments. MIS Quarterly, 36(1), 225. https://doi.org/10.2307/41410415

Guggenberger, T., Schweizer, A., & Urbach, N. (2020). Improving Interorganizational Information Sharing for Vendor Managed Inventory: Toward a Decentralized Information Hub

Using Blockchain Technology. IEEE Transactions on Engineering Management, 67(4), 1074–1085. https://doi.org/10.1109/TEM.2020.2978628

Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. The Journal of Strategic Information Systems, 26(3), 191–209. https://doi.org/10.1016/J.JSIS.2017.07.003

Gupta, A., Panagiotopoulos, P., & Bowen, F. (2020). An orchestration approach to smart city data ecosystems. Technological Forecasting and Social Change, 153, 119929. https://doi.org/10.1016/J.TECHFORE.2020.119929

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. Information & Management, 53(8), 1049–1064. https://www.sciencedirect.com/science/article/pii/S0378720616300787

Han, G., & Dong, M. (2015). Trust-embedded coordination in supply chain information sharing. International Journal of Production Research, 53(18), 5624–5639. https://doi.org/10.1080/00207543.2015.1038367

Hansen, J., & Pang, Y. S. (2023). Opening industry data: The private sector's role in addressing societal challenges. Data & Policy. https://doi.org/10.1017/dap.2023.15

Hardoy, J. E., & Satterthwaite, D. (1991). Environmental problems of third world cities: A global issue ignored? Public Administration and Development, 11(4), 341–361. https://doi.org/10.1002/PAD.4230110405

Hillebrand, K. (2021). The Role of Fear and Trust When Disclosing Personal Data to Promote Public Health in a Pandemic Crisis. Lecture Notes in Information Systems and Organisation, 46, 247–262. https://doi.org/10.1007/978-3-030-86790-4_18

Hillebrand, K., Hornuf, L., Müller, B., & Vrankar, D. (2023). The social dilemma of big data: Donating personal data to promote social welfare. Information and Organization, 33(1), 100452. https://doi.org/10.1016/J.INFOANDORG.2023.100452

Howe, N., Giles, E., Newbury-Birch, D., & McColl, E. (2018). Systematic review of participants' attitudes towards data sharing: A thematic synthesis. Journal of Health Services Research and Policy, 23(2), 123–133. https://doi.org/10.1177/1355819617751555

Iannacci, F., Simeonova, B., & Kawalek, P. (2022). Investigating the determinants of inter-organizational information sharing within criminal justice: A context-mechanism-outcome approach. Journal of Information Technology, 37(2), 188-208.

Iansiti, M., & Richards, G. L. (2006). The Information Technology Ecosystem: Structure, Health, and Performance. Antitrust Bulletin, 51(1), 77–110. https://doi.org/10.1177/0003603X0605100104

Jäckle, A., Burton, J., Couper, M. P., & Lessof, C. (2019). Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel: coverage and participation rates and biases. https://doi.org/10.18148/SRM/2019.V1I1.7297

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

Janhunen, T. (2019). Why do companies fail to create data driven B2B ecosystems? : case study on a data sharing ecosystem initiative. https://lutpub.lut.fi/handle/10024/160314

Jarvenpaa, S. L., & Essén, A. (2023). Data sustainability: Data governance in data infrastructures across technological and human generations. Information and Organization, 33(1), 100449. https://doi.org/10.1016/J.INFOANDORG.2023.100449

Jenkins, M., & Schaap, B. (2018). Forest ecosystem services. Background Analytical Study, 1.

Jentzsch, N., Sapi, G., & Suleymanova, I. (2013). Targeted pricing and customer data sharing among rivals. International Journal of Industrial Organization, 31(2), 131–144. https://www.sciencedirect.com/science/article/pii/S0167718712001191?casa_token=TvD1QILbF9gAAAAA:JJszFq4IcGBN0PuzlcM8GgfvSJEj8-n79AQXWNajpymyp7hxcXUtZ6ZcKqAtmTkhc1hNKvFb

Jones, M. (2019). What we talk about when we talk about (big) data. The Journal of Strategic Information Systems, 28(1), 3–16. https://doi.org/10.1016/J.JSIS.2018.10.005

Kayhan, V. O., & Davis, C. J. (2016). Situational privacy concerns and antecedent factors. Journal of Computer Information Systems, 56(3), 228-237. https://doi.org/10.1080/08874417.2016.1153913

Kazantsev, N., Islam, N., Zwiegelaar, J., Brown, A., & Maull, R. (2023). Data sharing for business model innovation in platform ecosystems: From private data to public good. Technological Forecasting and Social Change, 192, 122515. https://doi.org/10.1016/J.TECHFORE.2023.122515

Khan, M., Hussain, M., & Saber, H. M. (2016). Information sharing in a sustainable supply chain. International Journal of Production Economics, 181, 208–214. https://www.sciencedirect.com/science/article/pii/S0925527316300354

Khan, N., Yaqoob, I., Hashem, I. A. T., Inayat, Z., Mahmoud Ali, W. K., Alam, M., ... & Gani, A. (2014). Big data: survey, technologies, opportunities, and challenges. The scientific world journal, 2014(1), 712826.

Kim, K. K., Sankar, P., Wilson, M. D., & Haynes, S. C. (2017). Factors affecting willingness to share electronic health data among California consumers. BMC Medical Ethics, 18(1), 1–10. https://doi.org/10.1186/S12910-017-0185-X/TABLES/6

Kim, Y., & Stanton, J. M. (2016). Institutional and individual factors affecting scientists' data-sharing behaviors: A multilevel analysis. Journal of the Association for Information Science and Technology, 67(4), 776–799. https://doi.org/10.1002/ASI.23424

Kitchin, R., & Lauriault, T. P. (2015). Small data in the era of big data. GeoJournal, 80, 463-475. https://doi.org/10.1007/s10708-014-9601-7

Kitsios, F., Papachristos, N., & Kamariotou, M. (2017). Business models for open data ecosystem: Challenges and motivations for entrepreneurship and innovation. IEEE 19th Conference on Business Informatics (CBI), 1, 398–407. https://ieeexplore.ieee.org/abstract/document/8010744/?casa_token=k9a7WXEvKCkAAAAA:2x_MHCN_u1JscIYKwTUcFb5aCFD5x8Feo3-zo7-JaDv0Q_NVnNYFL-yQt3uxzT4_Hl4U5sDx

Klievink, B., Van Der Voort, H., & Veeneman, W. (2018). Creating value through data collaboratives. Information Polity, 23(4), 379–397. https://content.iospress.com/articles/information-polity/ip180070

Kouzes, R. T., Anderson, G. A., Elbert, S. T., Gorton, I., & Gracio, D. K. (2009). The changing paradigm of data-intensive computing. Computer, 42(1), 26-34.

Lavalle, A., Teruel, M. A., Maté, A., & Trujillo, J. (2020). Improving sustainability of smart cities through visualization techniques for big data from IoT devices. Sustainability, 12(14). https://doi.org/10.3390/su12145595

Lee, S. U., Zhu, L., & Jeffery, R. (2018). Designing Data Governance in Platform Ecosystems. Proceedings of the Annual Hawaii International Conference on System Sciences, 2018-January, 5014–5023. https://doi.org/10.24251/HICSS.2018.626

Leon, P. G., Ur, B., Wang, Y., Sleeper, M., Balebako, R., Shay, R., Bauer, L., Christodorescu, M., & Cranor, L. F. (2013). What matters to users? Factors that affect users' willingness to share information with online advertisers. SOUPS 2013 - Proceedings of the 9th Symposium on Usable Privacy and Security. https://doi.org/10.1145/2501604.2501611

Lis, D., & Otto, B. (2020). Data governance in data ecosystems–insights from organizations.

Lum, K., & Isaac, W. (2016). To predict and serve?. Significance, 13(5), 14-19. https://doi.org/10.1111/j.1740-9713.2016.00960.x

Manyika, J., Chui, M., Brown, B., Bughin, J., & Dobbs, R. (2011). Big data: The next frontier for innovation, competition, and productivity. http://dln.jaipuria.ac.in:8080/jspui/bitstream/123456789/14265/1/mgi_big_data_full_report.pdf

Margetts, H., & Sutcliffe, D. (2013). Addressing the policy challenges and opportunities of "Big data". Policy & Internet, 5(2), 139-146. https://doi.org/10.1002/1944-2866.POI326

Markus, M. L., Bui, Q. N., Lynne, M., Quang, M., Neo, ", & Bui, ". (2012). Going concerns: The governance of interorganizational coordination hubs. Taylor & FrancisML Markus, QN BuiJournal of Management Information Systems, 2012•Taylor & Francis, 28(4), 163–198. https://doi.org/10.2753/MIS0742-1222280407

Marwick, A., & Hargittai, E. (2019). Nothing to hide, nothing to lose? Incentives and disincentives to sharing information with institutions online. Information, Communication & Society, 22(12), 1697-1713. https://doi.org/10.1080/1369118X.2018.1450432

Mcafee, A., & Brynjolfsson, E. (2012). HBR.ORG Spotlight on Big Data Big Data: The Management Revolution.

Micheli, M., Ponti, M., Craglia, M., & Berti Suman, A. (2020). Emerging models of data governance in the age of datafication. Big Data and Society, 7(2). https://doi.org/10.1177/2053951720948087

Monteiro, E., & Parmiggiani, E. (2019). Synthetic Knowing: The Politics of the Internet of Things. https://doi.org/10.25300/MISQ/2019/13799

Morabito, V. (2015). Big data and analytics: Strategic and organizational impacts. Big Data and Analytics: Strategic and Organizational Impacts, 1–183. https://doi.org/10.1007/978-3-319-10665-6/COVER

Munoz-Arcentales, A., López-Pernas, S., Pozo, A., Alonso, A., Salvachua, J., & Huecas, G. (2019). An Architecture for Providing Data Usage and Access Control in Data Sharing Ecosystems. Procedia Computer Science, 160, 590–597. https://doi.org/10.1016/J.PROCS.2019.11.042

Nalebuff, B. J., & Brandenburger, A. M. (1997). Co opetition: Competitive and cooperative business strategies for the digital economy. Strategy & Leadership, 25(6), 28–33. https://doi.org/10.1108/EB054655/FULL/HTML

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. https://doi.org/10.1016/J.JSIS.2015.02.001

Nguyen, T., Tagett, R., Diaz, D., & Draghici, S. (2017). A novel approach for data integration and disease subtyping. Genome research, 27(12), 2025-2039. https://doi.org/10.1101/gr.215129.116

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.

Parmiggiani, E., Østerlie, T., & Almklov, P. G. (2022). In the Backrooms of Data Science. Journal of the Association for Information Systems, 23(1), 139–164. https://doi.org/10.17705/1jais.00718

Pesce, D., Neirotti, P., & Paolucci, E. (2019). When culture meets digital platforms: value creation and stakeholders' alignment in big data use. Current Issues in Tourism, 22(15), 1883-1903. https://doi.org/10.1080/13683500.2019.1591354

Pikkarainen, M., Ervasti, M., Hurmelinna-Laukkanen, P., & Nätti, S. (2017). Orchestration roles to facilitate networked innovation in a healthcare ecosystem. Jultika.Oulu.Fi, 7(9), 30–43. https://doi.org/10.22215/TIMREVIEW/1104

Polanin, J. R., & Terzian, M. (2019). A data-sharing agreement helps to increase researchers' willingness to share primary data: results from a randomized controlled trial. Journal of Clinical Epidemiology, 106, 60–69. https://doi.org/10.1016/J.JCLINEPI.2018.10.006

Prainsack, B. (2019). Logged out: Ownership, exclusion and public value in the digital data and information commons. Https://Doi.Org/10.1177/2053951719829773, 6(1)

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. Health Information Science and Systems, 2(1). https://doi.org/10.1186/2047-2501-2-3

Roeber, B., Rehse, O., Knorrek, R., & Thomsen, B. (2015). Personal data: how context shapes consumers' data sharing with organizations from various sectors. Electronic Markets, 25(2), 95–108. https://doi.org/10.1007/S12525-015-0183-0

Schneider, B., & Olman, L. (2020). The geopolitics of environmental global mapping services: An analysis of Global Forest watch. In A research agenda for environmental geopolitics (pp. 44-57). Edward Elgar Publishing.

Schudy, S., & Utikal, V. (2017). 'You must not know about me'—On the willingness to share personal data. Journal of Economic Behavior & Organization, 141, 1–13. https://doi.org/10.1016/J.JEBO.2017.05.023

Shah, S. I. H., Peristeras, V., & Magnisalis, I. (2020). Government (big) data ecosystem: definition, classification of actors, and their roles. International Journal of Computer and Information Engineering, 14(4), 102–112.

Shin, D. H., & Choi, M. J. (2015). Ecological views of big data: Perspectives and issues. Telematics and Informatics, 32(2), 311–320. https://doi.org/10.1016/J.TELE.2014.09.006

Shore, A., Prena, K., & Cummings, J. J. (2022). To share or not to share: Extending Protection Motivation Theory to understand data sharing with the police. Computers in Human Behavior, 130, 107188. https://doi.org/10.1016/J.CHB.2022.107188

Sidze, E. M., Wekesah, F. M., Kisia, L., & Abajobir, A. (2022). Inequalities in Access and Utilization of Maternal, Newborn and Child Health Services in sub-Saharan Africa: A Special Focus on Urban Settings. Maternal and Child Health Journal, 26(2), 250–279. https://doi.org/10.1007/S10995-021-03250-Z/TABLES/2

Sifah, E. B., Xia, Q., Agyekum, K. O. B. O., Xia, H., Smahi, A., & Gao, J. (2021). A blockchain approach to ensuring provenance to outsourced cloud data in a sharing ecosystem. IEEE Systems Journal, 16(1), 1673–1684. https://ieeexplore.ieee.org/abstract/document/9405789/

Spil, T., & Klein, R. (2014). Personal health records success: Why Google health failed and what does that mean for Microsoft HealthVault? Ieeexplore.Ieee.Org, 2818–2827. https://ieeexplore.ieee.org/abstract/document/6758953/

Strauss, A., & Corbin, J. (1998). Basics of qualitative research techniques. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=18c7cfe7a46c7771b60dc384b1b4e350f65b13e2

Strauss, A. L., & Corbin, J. M. (1997). Grounded Theory in Practice. Sage. https://books.google.it/books?hl=it&lr=&id=TtRMolAapBYC&oi=fnd&pg=PP9&dq=Strauss,+A.,+%26+Corbin,+J.+M.+(1997).+Grounded+theory+in+practice.+Sage.+&ots=DBhoexpDWk&sig=oMm_gdGbQKKNboV2guS1-E_SiSU&redir_esc=y#v=onepage&q=Strauss%2C%20A.%2C%20%26%20Corbin%2C%20J.%20M.%20(1997).%20Grounded%20theory%20in%20practice.%20Sage.&f=false

Struminskaya, B., Toepoel, V., Lugtig, P., Haan, M., Luiten, A., & Schouten, B. (2021). Understanding Willingness to Share Smartphone-Sensor Data. Public Opinion Quarterly, 84(3), 725–759. https://doi.org/10.1093/POQ/NFAA044

Susha, I., Grönlund, Å., & Van Tulder, R. (2019). Data driven social partnerships: Exploring an emergent trend in search of research challenges and questions. Government Information Quarterly, 36(1), 112–128. https://www.sciencedirect.com/science/article/pii/S0740624X17302708

Susha, I., Rukanova, B., Zuiderwijk, A., Gil-Garcia, J. R., & Gasco Hernandez, M. (2023). Achieving voluntary data sharing in cross sector partnerships: Three partnership models. Information and Organization, 33(1), 100448. https://doi.org/10.1016/J.INFOANDORG.2023.100448

Susha, I., van den Broek, T., van Veenstra, A. F., & Linåker, J. (2023). An ecosystem perspective on developing data collaboratives for addressing societal issues: The role of conveners. Government Information Quarterly, 40(1). https://www.sciencedirect.com/science/article/pii/S0740624X22000995

Taylor, J. F., Ferguson, J., & Ellen, P. S. (2015). From trait to state: understanding privacy concerns. Journal of Consumer Marketing, 32(2), 99–112. https://doi.org/10.1108/JCM-07-2014-1078/FULL/HTML

Taylor, R., Davis, C., Brandt, J., Parker, M., Stäuble, T., & Said, Z. (2020). The rise of big data and supporting technologies in keeping watch on the world's forests. International Forestry Review, 22(1), 129–141.

Tenopir, C., Dalton, E. D., Allard, S., Frame, M., Pjesivac, I., Birch, B., Pollock, D., & Dorsett, K. (2015). Changes in data sharing and data reuse practices and perceptions among scientists worldwide. PLoS ONE, 10(8). https://doi.org/10.1371/JOURNAL.PONE.0134826

Tiwana, A., Konsynski, B., & Bush, A. A. (2010). Research Commentary—Platform Evolution: Coevolution of Platform Architecture, Governance, and Environmental Dynamics. Https://Doi.Org/10.1287/Isre.1100.0323, 21(4), 675–687.

Tools, M., Saeed, N., Malik, H., Naeem, A., & Bashir, U. (2023). Incorporating big data and IoT in intelligent ecosystems: state-of-the-arts, challenges and opportunities, and future directions. Multimedia Tools and Applications 2023, 1–43. https://doi.org/10.1007/S11042-023-16328-3

van den Broek, T., & van Veenstra, A. F. (2018). Governance of big data collaborations: How to balance regulatory compliance and disruptive innovation. Technological Forecasting and Social Change, 129, 330–338. https://doi.org/10.1016/J.TECHFORE.2017.09.040

Wang, Z., Zheng, Z., Jiang, W., & Tang, S. (2021). Blockchain-Enabled Data Sharing in Supply Chains: Model, Operationalization, and Tutorial. Production and Operations Management, 30(7), 1965–1985. https://doi.org/10.1111/POMS.13356

Ziefle, M., Halbey, J., & Kowalewski, S. (2016). Users' Willingness to Share Data on the Internet: Perceived Benefits and Caveats. Human-Computer Interaction Center. https://doi.org/10.5220/0005897402550265

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

Zwart, H. (2015). Human Genome Project: History and Assessment. In International Encyclopedia of the Social & Behavioral Sciences (pp. 311–317). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.82036-X

4. Getting Where We (Do Not) Want To: Understanding Standardization-Trust Loops in Data Ecosystems[4]

Abstract

The rapid growth and evolving dynamics of data ecosystems highlight their crucial role in addressing the intricate relationship between trust and standardization. While both are recognized as fundamental to governance in data sharing ecosystems, their interaction and the tensions they create remain underexplored. This study investigates how data ecosystems navigate this interplay, examining how formal mechanisms designed to reduce reliance on trust can paradoxically intensify the need for trust-driven relational strategies, and vice versa. Through qualitative analysis, we explore how stakeholders in a multi-stakeholder data ecosystem collaborate to establish data-sharing protocols while addressing challenges related to trust and standardization. Our findings reveal a paradoxical relationship between standardization and trust. Stakeholders attempting to replace trust with standards faced significant challenges, while efforts to build relational trust without digital standardization were equally problematic. This dynamic creates a self-reinforcing loop where neither eliminating trust through standardization nor relying solely on trust can effectively drive progress and innovation. This study emphasizes the need to harmonize trust and standardization through adaptive governance models that balance consistency with flexibility. By advancing the understanding of their interaction, this research provides valuable insights into how the interplay between standardization and trust shapes the formation and evolution of multi-

[4] The paper presented in this chapter is currently under review in the journal *Organization Studies*. Authors: Giulia Renzi, first author (conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, validation, visualization, writing – original draft), Paula Ungureanu (conceptualization, investigation, methodology, supervision, validation, visualization), and Natalia Selini Hadjidimitriou (data curation, writing – review and editing).

stakeholder data ecosystems, contributing to the literature on data ecosystems, standardization and trust. Furthermore, we provide practical guidance for policymakers and stakeholders to address tensions, foster collaboration, and build resilient, scalable, and adaptable data ecosystems.

Keywords: data ecosystems, data sharing, standardization, trust, ecosystems, mobility, data governance, data partnerships.

4.1. Introduction

Data and digital ecosystems are dynamic, interconnected environments where stakeholders, organizations, technologies, and data interact to generate innovation (Micheli et al., 2020; Oliveira et al., 2019). Far from being static frameworks, they are complex, multi-sided phenomena characterized by evolving interactions and data practices. Scholars emphasize the need of understanding how data is collected, produced, consumed, and governed (Davidson et al., 2023; Oliveira et al., 2019). Fueled by rapid growth of heterogeneous data, these ecosystems hold transformative potential that requires a deeper exploration of their intricate dynamics (Davidson et al., 2023; Oliveira et al., 2019).

Unlike traditional ecosystems, data ecosystems involve shifting roles, rules, and practices adapting to emerging data challenges and requirements (Lee et al., 2018; Oliveira et al., 2019; Rong & Luo, 2023; Shin & Choi, 2015), particularly the “5Vs”: volume, velocity, variety, value, and veracity (Demchenko et al., 2014; Tools et al., 2023). Effective data management is crucial, encompassing quality, privacy, and security concerns (Bonina et al., 2021; Cui et al., 2020; Demchenko et al., 2014). However, challenges stem not only from the data itself but also from the diversity and interdependence of actors within these ecosystems

(Demchenko et al., 2014; Tools et al., 2023; Oliveira et al., 2019; Shin & Choi, 2015; Wang et al., 2021). Addressing these challenges requires robust data governance mechanisms that balance multi-stakeholders' goals, data quality, privacy, scalability, and interconnectivity (Abraham et al., 2019; Günther et al., 2017). Among these mechanisms, standardization and trust are often seen as critical (Abraham et al., 2019; Günther et al., 2017; Woolthuis et al., 2005).

Standards are often described as inter-organizational boundary resources with key roles in the governance of innovation ecosystems and platforms (Arnold & Loconto, 2020; Brunsson et al., 2012; Cepa et al., 2022; Cepa & Schildt, 2019; Constantinides & Barrett, 2014; Grillo et al., 2024; Hinings et al., 2018; Manning & Reinecke, 2016; Sandholtz, 2012; Timmermans & Epstein, 2010). Standardization aims to reduce complexity and ensure consistency by defining common protocols, formats, and frameworks that facilitate interoperability and reliability among stakeholders, while also enabling efficient communication and potentially reducing frictions (Pessanha Santos, 2023). Yet, excessive standardization may create rigidity and hinder adaptability to changing technological or stakeholder needs (R. Cao & Iansiti, 2023; Cusumano et al., 2015; Teece, 2018).

Trust, in contrast, addresses the human and relational dimensions of collaboration, reducing concerns about opportunistic behavior (Hou & Jansen, 2023; Lumineau et al., 2023; Ven et al., 2006). However, trust is fragile and takes time to build (Cvitanovic et al., 2021; Swärd, 2016), relying on informal mechanisms that may clash with formal standards (Cobben & Roijakkers, 2019; T. K. Das & Teng, 1998, 2002).

Despite the growing body of literature and their recognized importance, the interplay between trust and standardization remains underexplored, particularly in data ecosystems (Backer, 2024; Z. Cao & Lumineau, 2015; Lioliou et al., 2014; Wareham et al., 2014; Woolthuis

et al., 2005). While some highlight tensions, others note complementarity between standardization and trust. Standardization affords predictable and consistent data systems that reduce reliance on trust, but establishing standards frequently requires trust, and formalizing informal practices can create tensions and disruption (Backer, 2024; Z. Cao & Lumineau, 2015; Lioliou et al., 2014). Standardization supports scalable and interoperable data ecosystems, while trust addresses relational and ethical dimensions that formal mechanisms miss. These arguments are not new: research on strategic alliances has discussed trust and control as complementary or substitutive (Z. Cao & Lumineau, 2015; T. K. Das & Teng, 1998; Ven et al., 2006). However, understanding trust - standardization dynamics in digital contexts like data ecosystems requires further investigation.

Data ecosystems present challenges due to their scale, complexity, and reliance on data and technology-driven governance (Hanisch et al., 2023; He et al., 2020; Lumineau et al., 2023). Unlike traditional alliances, they rely on automation and significant standardization. Digitalization can exacerbate standardization but risks rigidity, sidelining tacit knowledge vital for coordination (Cepa et al., 2022; Cepa & Schildt, 2019; A. Das, 2020; Hanisch et al., 2023). Consequently, relational governance emerges to counter over-standardization, often reshaping how trust is enacted in digital settings (Cepa et al., 2022; Cepa & Schildt, 2019; Dattee et al., 2018; Lumineau et al., 2023). Balancing these forces is arguably harder than in prior organizational forms. In addition, as data ecosystems rapidly evolve with increasing data volume, variety, and velocity, trust-standardization tensions may intensify with stakeholders' heterogeneous attempts to navigate these two forces (Liu et al., 2022; Wareham et al., 2014). Despite the emphasis on hybrid governance (Grillo et al., 2024; Hanisch et al., 2023; He et al., 2020), challenges in balancing standardization and trust remain underexplored.

To address these gaps, our study explored the following research questions: *how is the tension between standardization and trust navigated in data ecosystems? How do formal*

mechanisms intended to reduce reliance on trust may ultimately increase the need for trust-based relational mechanisms among ecosystem stakeholders - and vice versa? We conducted a qualitative study within MobiDataLab, a data ecosystem focused on integrated mobility solutions in the European Union. Using a practice-based approach, we examined how stakeholders develop data-sharing protocols and navigate trust-standardization challenges.

In line with the emerging literature on governance in digital ecosystems (Backer, 2024; Z. Cao & Lumineau, 2015; Wareham et al., 2014), we document a dynamic relationship between trust and standardization but also emphasize a self-depleting loop between standard-based control practices and trust-based relational practices. We observed that neither full reliance on formal standards nor exclusive dependence on trust enabled effective coordination. Instead, participants had to navigate ongoing dilemmas, often realizing that both elements were indispensable, though difficult to reconcile.

We aim to contribute to the literature on data ecosystems, standardization, and trust, offering practical guidance for policymakers and stakeholders on addressing tensions between standardization and trust.

4.2. Theoretical Framing

To better understand data ecosystems and their complexities related to standardization and trust-based governance, we examine literatures on these topics as interdependent elements shaping data governance, collaboration, and the functioning of data-sharing practices.

4.2.1. Standardization Challenges in Data Ecosystems: From Centralized to Relational Governance

Standards are essential for ecosystems (Bogers et al., 2019; Shipilov & Gawer, 2020), the platform economy (Jacobides et al., 2024; Teece, 2018) and systemic change toward a more sustainable society (Geels, 2004; van Tulder & van Mil, 2022). Yet, while there are many references to standards in these contexts, they often remain vague on how standards contribute to transformations in business and society (Grillo et al., 2024). In data ecosystems, standards are crucial for establishing uniform criteria for data formats, definitions, and structures, ensuring consistency and compatibility across systems and stakeholders (R. Cao & Iansiti, 2023; Cepa et al., 2022). By harmonizing data structures and exchange protocols, standardization allows stakeholders to "speak the same language", facilitating efficient communication and reducing frictions among data stakeholders (Pessanha Santos, 2023). Additionally, it enhances data quality by reducing inconsistencies, ensuring reliable analytics and decision-making in complex ecosystems (Gelhaar & Otto, 2020). As ecosystems grow, standardization enhances scalability and reusability by defining common data formats and simplifying the integration of new actors or data sources (Mentel & Hajduk-Stelmachowicz, 2020). It also facilitates compliance with regulation like GDPR, aligning practices with privacy and security standards, and mitigating non-compliance risks (Moreno et al., 2019).

Despite its benefits, achieving data standardization is challenging. Lindgren et al. (2021) describe standardization as a complex, dialectic process marked by paradoxical tensions. These tensions arise from the diversity of interests, perspectives, and existing data standards, making alignment within a commonly agreed framework difficult (Brunsson et al., 2012; Haack et al., 2012; Reinecke et al., 2012). Reconciling diverse interests, technical capabilities, and pre-existing standards is particularly difficult in ecosystems involving both government and private entities that must harmonize their practices (Garud et al., 2002; Lnenicka et al., 2024; Oliveira

et al., 2019). The failure to establish universally accepted standards often leads to fragmented efforts, especially in data ecosystems with actors from varied sectors, or cultural context (Zang et al., 2020). Privacy, security and ethical concerns further complicate data standardization, particularly when sensitive data is involved (Moreno et al., 2019). Balancing data accessibility with privacy, security, and ethical standards is a significant challenge in data ecosystems where information flows across multiple stakeholders with differing priorities, requiring advanced governance mechanisms (Moreno et al., 2019).

A key issue in standardization practices in the governance structures employed. Governance encompasses actions to organize, regulate, and safeguard data flows through policies, standards, and procedures that ensure interoperability and data quality (Jarvenpaa & Essén, 2023; Micheli et al., 2020). Research emphasizes the need for governance models that combine structured frameworks with adaptable mechanisms, balancing compliance with standards and flexibility to accommodate evolving stakeholder needs (Hanseth et al., 2006; Zang et al., 2020). However, achieving this balance is inherently challenging: while existing research provides insights into the structural aspects of data ecosystems, it often presents static views on data as fixed assets exchanged between actors in predictable ways (Cox & Ellsworth, 1997; De Mauro et al., 2015).

The dynamic nature of data ecosystems represents one of the greatest challenges to standardization. These environments are constantly evolving, with new participants, data types, and use cases emerging over time. Maintaining standardized formats while adapting to change and innovation requires ongoing management, often introducing complexity to data structure and storage (Alli & Alam, 2020). Emerging technologies amplify this complexity by introducing diverse data types, standards, and evolving tools that challenge the adaptability of existing framework (Alli & Alam, 2020; Hanseth et al., 2006).

Centralized ecosystems, typically managed by orchestrators like governments, private companies, or brokers, provide solutions to challenges requiring consistency, efficiency and control (Dhanaraj & Parkhe, 2006; Ford et al., 2011; Jarvenpaa & Essén, 2023; Micheli et al., 2020). Centralized data governance promotes uniform standards, enabling efficient data management, streamlined decision-making, and stronger security. It also enhances accountability and compliance, which is vital in regulated industries (Dhanaraj & Parkhe, 2006; Ford et al., 2011; Moreno et al., 2019).

However, this structure reduces flexibility and ecosystem adaptability. Once standards are established, they prioritize consistency over adaptability, which can conflict with the dynamic needs of rapidly evolving fields like big data and technology (Dattee et al., 2018; Gawer & Cusumano, 2014). This rigidity may cause friction as standardized practices fail to meet diverse stakeholder needs or adapt quickly to changing conditions (Brunsson et al., 2012). Resistance to standardization is common both during adoption - where organizations hesitate to embrace change (Garud et al., 2002; Hanseth et al., 2006; Moreno et al., 2019; Pessanha Santos, 2023; Zang et al., 2020) – and post adoption, where organizations are reluctant to modify established standards due to perceived costs and unclear long-term benefits (Pessanha Santos, 2023).

As Hanseth et al. (2006) observed, standardization can lead to disorder when it fails to address diverse stakeholders' evolving needs, causing fragmentation as stakeholders create competing standards (Reinecke et al., 2012). This resistance can push stakeholders to work in silos, developing incompatible data practices that undermine standardization objectives (Brunsson et al., 2012; Hanseth et al., 2006; Manning & Reinecke, 2016).

To address these challenges, data ecosystems may turn to trust-based relational governance.

4.2.2. Trust Challenges in Data Ecosystems

Trust is a psychological state where one party accepts vulnerability based on positive expectations of another's actions or intentions (Cvitanovic et al., 2021; Mayer et al., 1995; Rousseau et al., 1998). In innovations ecosystems such as data ecosystems, trust is crucial for collaboration, value co-creation, and achieving shared goals. It improves performance and sustainability while mitigating risks tied to value roles, responsibilities, and contract terms (Hou & Jansen, 2023; Woolthuis et al., 2005). Trust allows stakeholders to collaborate despite differing objectives, reducing reliance on rigid controls and lowering monitoring costs (Hou & Jansen, 2023; Cvitanovic et al., 2021). It also supports risk-sharing, which enhances innovation and adaptability (Backer, 2024; Swärd, 2016; Woolthuis et al., 2005).

However, building trust in data ecosystems is challenging. These environments involve diverse actors with conflicting goals, values, and risk perceptions, making trust hard to establish (Cobben & Roijakkers, 2019). The complexity of data exchanges – ensuring integrity, privacy, and proper use – adds further difficulty (Van De Hoven et al., 2021). Opportunistic behavior can erode trust, particularly when individual gains are prioritized over collective outcomes (Cobben & Roijakkers, 2019). Trust is fragile, slow to build and easily broken by misconduct or misaligned expectations (Lumineau, 2014; Lumineau et al., 2022, 2023; Mayer et al., 1995; Rousseau et al., 1998). This underscores the need for robust governance mechanisms to support and sustain trust within these ecosystems (Cvitanovic et al., 2021).

These challenges can destabilize standardization process. Standardization often involves tension between the need for cooperation and conflicting stakeholder interests (Hanseth et al., 2006; Lindgren et al., 2021). Trust-based governance also evolves over time, requiring continuous adjustments (Z. Cao & Lumineau, 2015; T. K. Das & Teng, 1998). As such, efforts to impose order can inadvertently introduce new forms of instability (Hanseth et al., 2006).

4.2.3. Understanding the Interplay between Standardization and Trust in the Digital World

Standardization and trust are recognized as key to governing digital ecosystems, yet their interaction – especially in data ecosystems – remains underexplored. The core challenge lies in balancing trust-based collaboration with the automation and efficiency demands of digitalization. Emerging technologies require standardized protocols to ensure interoperability and scale, while trust is essential for managing uncertainty, enabling collaboration, and compensating for the rigidity of algorithmic or rule-based systems (Cepa et al., 2022; Hanisch et al., 2023; He et al., 2020).

These unique challenges highlight the intricate interplay between trust and standardization in evolving governance models. Yet, data ecosystems and digital ecosystems at large must find ways to traverse this dilemma (A. Das, 2020; Huber et al., 2017), even if these ways take the form of trials and errors (Rudmark et al., 2024).

Standardization reduces uncertainty and enhances collaboration by ensuring data quality, consistency, and interoperability, minimizing reliance on relational trust by lowering uncertainty and fostering collaboration (Backer, 2024). However, relational governance mechanisms remain essential for managing risks, particularly in complex ecosystems (Liu et al., 2022), where standards alone cannot prevent opportunism or adapt to evolving needs. (Z. Cao & Lumineau, 2015). Yet, attempts to standardize trust-based practices can backfire, as they may be perceived as signals of distrust. Stakeholders with long-standing relationships may resist standardization, fearing biases or unequal benefits, which can intensify tensions as ecosystems evolve (Z. Cao & Lumineau, 2015; Wareham et al., 2014).

Paradoxically, rigid standards may increase the need for trust-based mechanisms to address unforeseen challenges, while over-reliance on trust and reciprocity may lead to greater complexity, skepticism, or opportunistic behavior, straying ecosystems from their objectives.

In sum, while standardization and trust can ideally work in tandem to strengthen governance, overemphasis on one at the expense of the other, perceptions of tradeoffs or alternating practices, can potentially cause disruption in data ecosystems. As a consequence, we claim for increased attention to the conditions and processes leading to either trust-standardization trade-offs or complementarities.

4.3. Methods

4.3.1. Context

Our analysis centers on the MobiDataLab project, a European initiative funded by the Horizon 2020 Research and Innovation Programme (H2020) to advance a data ecosystem for integrated mobility across Europe. The project was funded for a three-year period, starting from January 2021, with the purpose of fostering data culture, common data governance and standardization across Europe's diverse mobility sectors. Since inception, MobiDataLab addressed urban mobility and transportation challenges through data analytics, with core objectives to improve integration, efficiency, sustainability, and accessibility across European urban transportation systems by promoting data-sharing practices among diverse mobility stakeholders. A key component of this endeavor was the development of a cloud-based prototype platform, known as the Transport Cloud, designed to facilitate seamless data exchange among mobility stakeholders. For this purpose, the ecosystem invested significant resources in studying existing standards across mobility systems and mobility stakeholders and provided occasions for discussion and negotiation of shared standards to be prototyped, tested and implemented on the

ecosystem's Transport Cloud. Throughout its three-years of funding, the project evolved into an integrated mobility data ecosystem where multiple stakeholders came together to create a large pool of mobility data with the aim of making their cities more livable, accessible, and environmentally sustainable. Initially composed of 10 partners, including ITS providers, academia, research institutions, policymakers, and transport associations from seven countries (see Table 8 for details), it eventually engaged a broader audience, including researchers analyzing mobility patterns, private companies identifying transport bottlenecks, policymakers involved in urban planning, and local communities aiming to improve mobility and quality of life.

Large-scale transportation data collection, analysis, and interpretation were central to building this ecosystem. Early data, sourced from GPS, mobile apps, public transit systems, and traffic sensors, was initially managed by project partners, with stakeholder contributions increasing as the project progressed. MobiDataLab emphasized public engagement, collaborating with government agencies, transportation authorities, community groups, and residents to ensure the research met local needs and preferences. Efforts were made to standardize data formats and protocols, enabling seamless data sharing and integration across different stakeholders and systems. Additionally, multiple governance structures and processes were discussed in the attempt to oversee the complex aspects of the data ecosystem, such as data access, privacy, ethical use, and long-term sustainability.

The key practices that defined the activity of MobiDataLab included developing an Open Knowledge Base (OKB) to address data-sharing challenges and best practices in the ecosystem, with real-world use cases and recommendations for improving data quality, accessibility, and usability. The OKB supported the creation of the Transport Cloud, a prototype platform enabling seamless transport data exchange among stakeholders. Additionally, Living and Virtual Labs provided real-time, collaborative experiments in data sharing, utilizing

advanced data analysis techniques like machine learning and statistical modeling to generate actionable insights. These efforts enabled researchers and policymakers to better understand mobility patterns, address bottlenecks, and optimize transport systems. While commercialization was not a primary focus, the project also explored potential business models and commercialization opportunities though ongoing market studies.

Through its broad engagement and innovative approach, MobiDataLab offered valuable insights into data ecosystem formation, highlighting stakeholder perceptions, concerns, and expectations around data-sharing, and revealing the structural, motivational, and relational dynamics of data ecosystem governance. Specifically, since one of MobiDataLab's core practices ('OKB') entailed data standardization efforts, it represented a valuable opportunity to study standardization efforts within a data ecosystem, as well as relational governance dynamics.

Regarding the latter, MobiDataLab aimed to cultivate a data-sharing culture in urban mobility by creating a collaborative and secure environment for multi-stakeholder experimentation. On the one hand, this approach was underpinned by efforts to standardize data formats, protocols, and interoperability frameworks within the OKB, ensuring seamless collaboration among diverse stakeholders. On the other hand, relational governance mechanisms such as need for trust and collaboration were also largely discussed as key to regulate data access, maintain security, and address ethical considerations in the ecosystem.

4.3.2. Data collection

We gathered information from multiple sources to aid in developing the theory, using semi-structured interviews and archival records.

Interviews

To explore the relation between trust and standardization in a data ecosystem, we conducted interviews with key stakeholder groups within the MobiDataLab ecosystem. Due to the fluid boundaries between stakeholder categories, we began by interviewing MobiDataLab's founding partners. One author's involvement in some of the operational phases of the project facilitated access to these stakeholders, enabling the establishment of relationships with partners and coordinators. These initial connections allowed us to reach additional stakeholders who joined the ecosystem at a later stage. Using a snowball sampling method, we expanded our interviews to other relevant actors involved in the Living and Virtual Labs activities throughout the project.

Between May 2022 and December 2023, we conducted 40 semi-structured interviews across 18 different stakeholder categories, as detailed in Table 8. All interviews were recorded, anonymized, and fully transcribed. The interview protocol comprised four main sections: 1) background information on the interviewees and their organizations, including motivations, objectives, and data-sharing practices; 2) understanding of data and the role of data quality in sharing processes; 3) collaboration practices related to data sharing; and 4) visions and evolving practices of data sharing within the ecosystem over time.

Archival Data

We collected archival data from the MobiDataLab Grant Agreement, which outlined the legal rights and obligations between the European Commission and project partners. Additionally, we analyzed all public deliverables produced during the project to understand its core activities, objectives, and responsibilities.

To further contextualize the MobiDataLab project, we examined 12 other European, national, and international initiatives related to data sharing - both past and ongoing - that were

selected based on a review conducted by the authors and supplemented with recommendations from interviewees. By reviewing these projects' websites, social media content, and reports detailing their goals, outputs, and accomplishments, we gained a broader perspective on European data ecosystems. We also investigated different participation models and governance structures.

This comprehensive approach helped us to better situate our findings on the MobiDataLab's unique characteristics and practices within the broader European data ecosystem landscape. Table 8 provides a summary of the diverse data sources that supported the theory-building process.

Table 8: Overview of data sources used to support the theory development process. Source: authors own work.

	Semi-structured interviews	**Archival data**	Relate Projects Analysis
Total Number	40 fully recorded and transcribed interviews *Categories of Stakeholders:* 18	27 documents	12 European, national and international projects.
Timespan	May 2022 – December 2023	2021 - 2023	2018 - 2023
Description	*Average Length:* 60 minutes	All the public deliverable of the MobiDataLab project can be found here: https://mobidatalab.eu/publications/	*Documents and topics analyzed:* Websites, social media, and reports on consortium composition, main objectives, and outcomes.
Details	*Interviewed Stakeholders:* 8 partners of the MobiDataLab consortium, divided as follow: 4 ITS providers, 2 academic/research institutions, 1 policymaker, 1 association of municipalities. *Extended interviews* with participants from MobiDataLab Living and Virtual Labs, focusing on mobility challenges and solutions. *Final sample* expanded using snowball sampling, including: Researchers & Academics (4), Transport Operators (4), Citizens (2), Policymakers (3), Municipalities (2), Association of Municipalities (2), Government (1), Public Transport Authorities (1), Rail Infrastructure Authorities (1), Traffic Management Centres (1), Trade Associations (2), Logistics Operators (5), Airlines (1), ITS Providers (5), Software Providers (2), Cloud Providers (1), Car Manufacturers (2), Emergency Services (1).	All the public deliverable of the MobiDataLab project can be found here: https://mobidatalab.eu/publications/ *Deliverables produced within the MobiDataLab project,* including reports, technical diagrams, brochures, state-of-the-art analysis, legal and regulatory overviews, data governance assessments, use case definitions, actors and stakeholder definitions, market analyses, data sharing business models, communication strategies, virtual lab descriptions and analyses, stakeholder event reports, and cooperation activity reports.	Related projects, reports, and documents recommended by informants or gathered by researchers: • AEOLIX: https://aeolix.eu • SELIS: https://selisproject.eu • FENIX: https://fenix-network.eu • NAPCORE project: https://napcore.eu/ • Inspire initiatives: https://inspire.ec.europa.eu/ • Le Fabrique du mobilite: https://lafabriquedesmobilites.fr/ • Mobility as a service: https://www.maas4eu.eu/ • Convex: https://convex-project.de/index.html • Openstreetmap: https://www.openstreetmap.org/ • TALOS: https://cordis.europa.eu/project/id/218081/reporting • Transmodel Norm: https://www.transmodel-cen.eu/use-of-the-transmodel/ • W3C (World Wide Web Consortium): https://www.w3.org/

4.3.3. Data analysis Using the Grounded Theory Approach

After transcribing and anonymizing the interviews, we used archival data to clarify references, allusions, and connections made during the discussions. The anonymized data was then imported into a comprehensive database on the QACDAS platform, DEDOOSE. Following the grounded theory approach, we organized the textual data to identify recurring themes and categories, integrating empirical evidence with theoretical frameworks to contribute to the development of new theories (A. Strauss & Corbin, 1998; A. L. Strauss & Corbin, 1997).

In alignment with grounded theory methodology, we began with open coding to identify frequent concepts and themes. This was followed by axial coding, where we connected the categories and subcategories identified in the open coding phase (A. Strauss & Corbin, 1998; A. L. Strauss & Corbin, 1997). Open coding revealed several challenges unique to data ecosystems, arising from conditions that differentiate these ecosystems from others – such as exchanges among multiple stakeholders, ambiguous legal agreements, differing technical capabilities, and varied objectives and economic incentives – as well as from the lack of a shared long-term approach, characterized by an absence of collective purpose, misalignment, and insufficient commitment. We labeled all these conditions and needs as "*Facing the Data Ecosystem Challenges*".

In axial coding, our analysis identified two key practices that interplayed continuously at the ecosystem level: standardization efforts and efforts to develop relational trust. Accordingly, stakeholders were working towards the setup of data standards while acknowledging that, given Europe's current data-sharing landscape and the nature of the data being exchanged, trust and strong relationships between actors remained vital for building a successful data ecosystem. We thus decided to further inquire into the conditions and dynamics of the trust-standardization interplay in MobiDataLab. We observed that participants frequently emphasized the need to establish a highly structured data sharing process at the ecosystem level.

This implied aiming to create data sharing standards at the ecosystem level in order to transition from a trust-based data ecosystem to a standard-based trustless governance process. This involved three key objectives: first, creating data-sharing standards at the ecosystem level to transition from a trust-based data ecosystem to a governance process based on standards without relying on trust; second, aligning on common data sharing standards – such as incorporating technical reviewers, improving and integrating standards, and reaching consensus on shared standards; third collaborating to achieve standardization through joint initiatives managed with a collaborative approach, and networking opportunities for sharing perspectives. We labeled this process as "*Setting up an Ecosystem Data Sharing Process: From Standardization to Trust*".

We further observed the central role of trust in the analyzed data ecosystem, and we started to open code for types of trust, targets and processes. We identified different trust practices that relied on different types of trust. These practices were: i) trusting standardization, which implied 'procedural trust' in how the standards were developed and would have allowed the ecosystem to function and the single stakeholders to draw value from the data through collaboration and a collective vision, ii) trusting the system which implied 'system-based trust' which relies on the ecosystem's functioning as a whole (internal capabilities and availabilities) and external conditions like existing norms, regulation, and procedures; iii) trusting the others -i.e., 'interpersonal trust' based on stakeholders' evaluations of each other's motivation, integrity, and benevolence. While comparing these patterns, we found that they differed based on whether they referred to the technical aspects of big data (i.e., 'procedural trust') or to the underlying relational conditions (i.e., 'interpersonal trust'). We thus labeled these categories as "*Trust Dynamics Related to Data Sharing: From Technical to Relational*".

Finally, we identified several challenges stakeholders faced in their activities within the data ecosystems as they dealt with the standardization-trust conundrum. These included

standardization challenges related to data, such as data heterogeneity, variable data quality, and multiple parameters and standards; alignment challenges, including unclear regulatory framework, and ambiguous practices and procedures; and collaboration challenges, such as stakeholder heterogeneity, unmet expectations, lack of transparency and difficulty in monitoring processes and verifying stakeholders' efforts. We labeled all these challenges *"Facing Data Sharing Challenges: From Standardization to Collaboration*".

To build a grounded model, we conducted further axial coding to identify relationships among second-order themes and abstract them into higher-level categories. For example, we analyzed how standardization efforts intersected with trust-building practices, revealing that procedural trust and system-based trust were essential for addressing alignment and collaboration challenges. Additionally, we mapped how each challenge influenced and was influenced by trust practices. For instance, unclear regulatory frameworks were linked to gaps in procedural trust, while stakeholder heterogeneity emphasized the importance of interpersonal trust to navigate differing priorities.

This iterative process of selective coding allowed us to integrate fragmented themes into a cohesive theoretical model that explains the dynamic interplay between trust and standardization in data-sharing ecosystems. By synthesizing these findings, we provide insights into the mechanisms and conditions necessary for developing resilient and effective data-sharing ecosystems capable of addressing key challenges.

Figure 6 illustrates the data structure underpinning our grounded theory development. Figure 7 presents the grounded theory-based model that explains the interplay between standardization and trust.

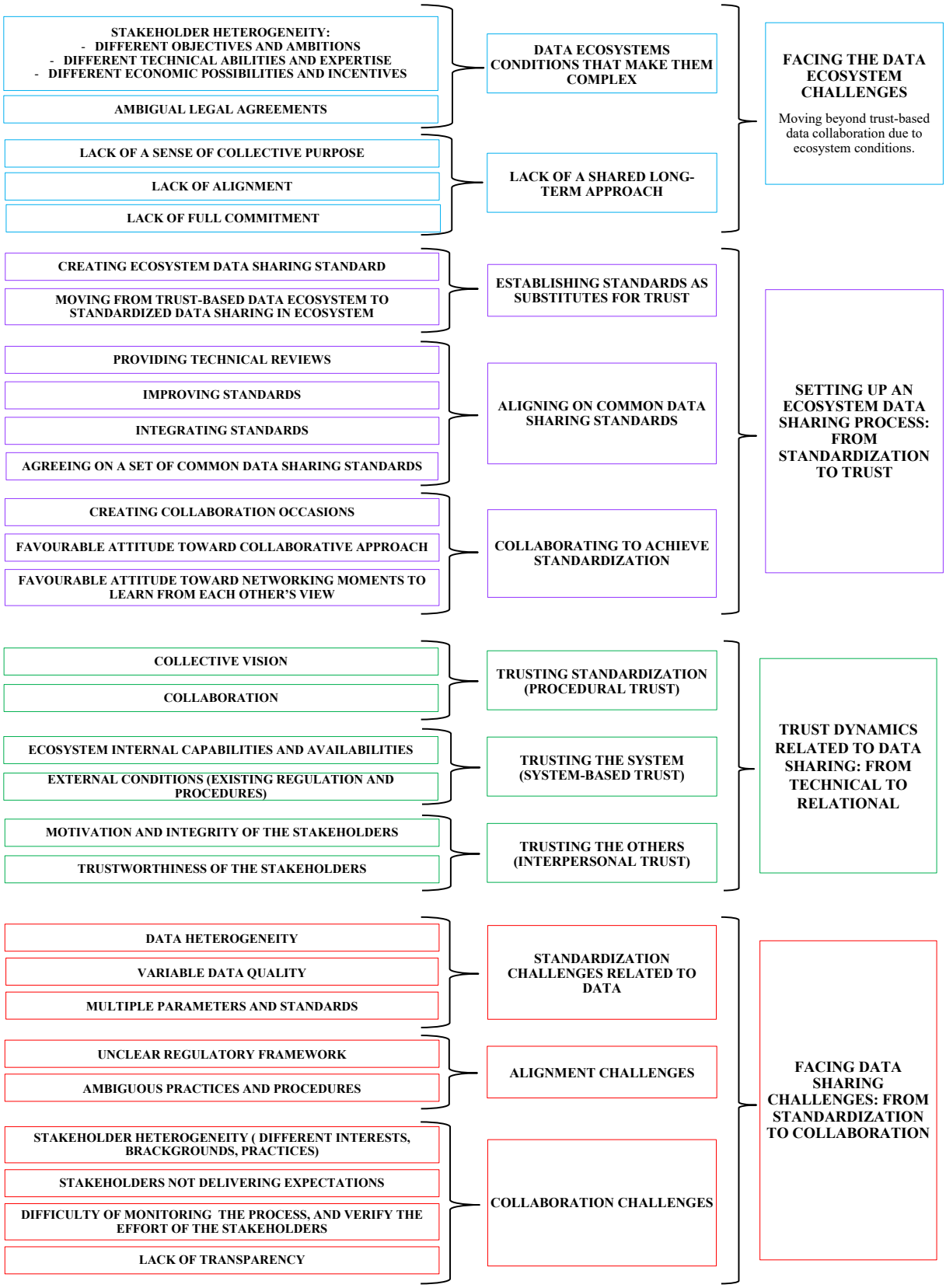

Figure 6: Data framework supporting the development of grounded theory. Source: authors own work.

4.4. Findings

Our study highlights that, while both standardization and trust are widely recognized as crucial elements of governance in data ecosystems, stakeholders in the MobiDataLab project ideally desire to move beyond trust-based data sharing towards a model governed by standardization – one where trust can be replaced by standards, and relational practices by automation. The conditions within the ecosystem that make relational mechanisms costly, complex, and uncertain drive the pursuit of this ideal, creating a loop of standardization-trust tensions that has stagnating consequences for the ecosystem. In exploring how the tension between standardization and trust is managed in the data ecosystem, we found that formal mechanisms intended to develop and automate standards in order to reduce reliance on trust ultimately increase the need for stakeholders to adopt trust-driven relational approaches, which in turn contributes to the perception of ecosystem complexity and exacerbates the perceived need for standard-driven simplification. This creates a paradoxical relationship between standardization and trust, as standardization cannot be pushed forward without trust, and trust alone is insufficient to sustain stakeholder exchanges within a data ecosystem. Notably, we did not find a resolution to this paradox. Instead, we observed that the data ecosystem was caught in a loop where governance progressed without choosing between standardization and trust.

Figure 7 presents the grounded model that explains the factors shaping the paradoxical relationship between standardization and trust-building practices in a data ecosystem. We identified a common pathway that stakeholders followed in their attempts to establish standards as a substitute for trust. We observe that overcoming standardization challenges often requires some level of trust, and conversely, addressing trust-related challenges frequently leads to standardization attempts. Importantly, we found that every standardization attempt faced the need for a different form of trust (procedural, system-based, interpersonal).

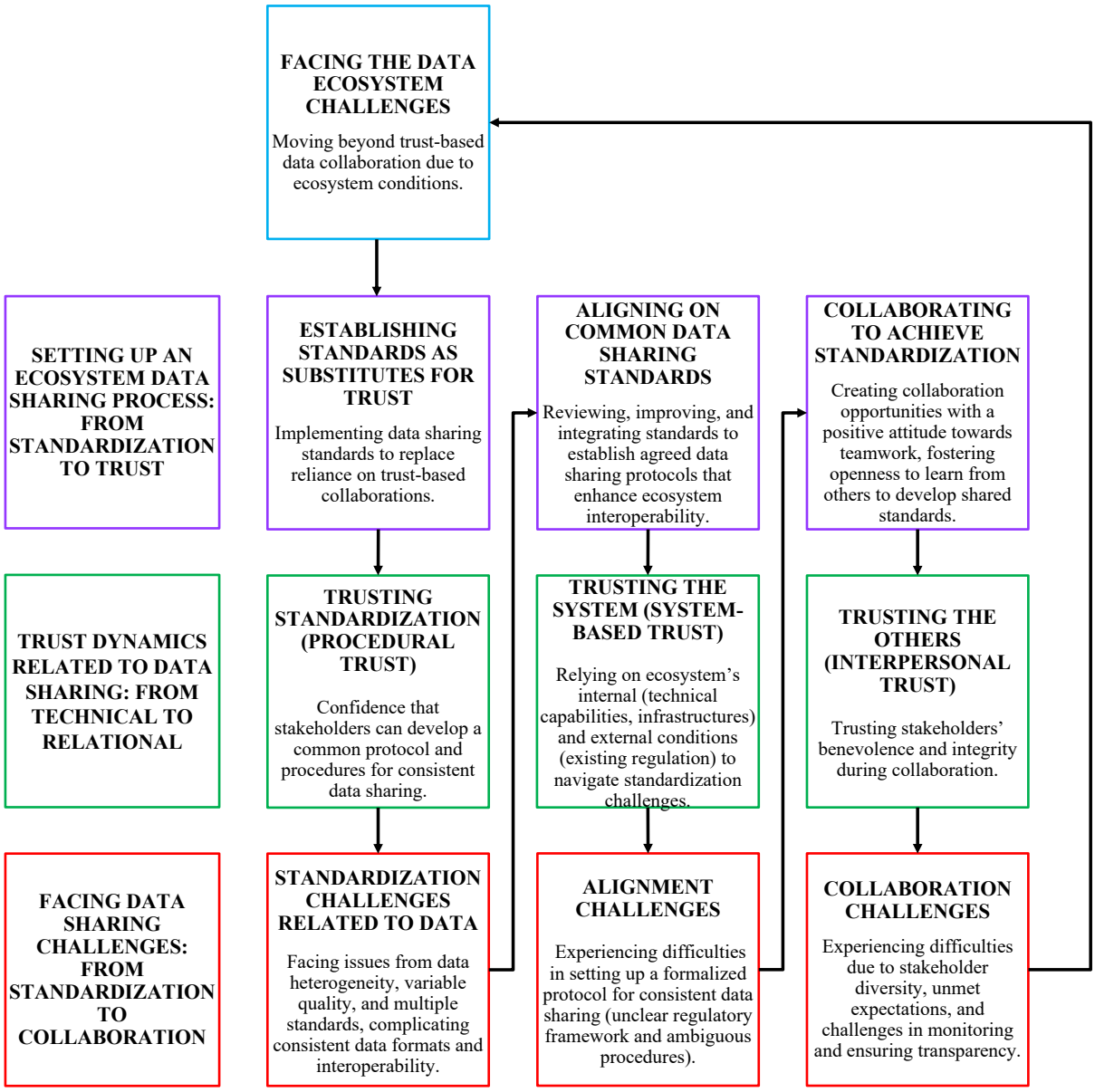

Figure 7: Setting up a protocol for data sharing in a data ecosystem: the paradoxical relation between standardization and trust-building practices. Source: authors own work.

As shown in Figure 7, the initial impetus is driven by the desire to address data ecosystem challenges, prompting stakeholders to move beyond trust-based data collaboration due to ecosystem conditions. This initial push sets off a process with three key practices (represented by horizontal lines): 1) setting up a data sharing process; 2) recognizing trust dynamics related to data sharing; and 3) addressing data sharing challenges. Following the arrows in the model (vertically), it becomes evident that although there is a need to establish a data ecosystem based

on standardization, every attempt to replace trust with standardization ultimately necessitates some form of trust (whether procedural, system-based, or interpersonal).

The following subsections of the findings (4.4.1 – 4.4.10) provide a detailed analysis of the main categories within our grounded theory model.

4.4.1. Facing the data ecosystem challenges

In the analyzed ecosystem, we observed an initial impetus towards standardization driven by the desire to address the inherent challenges of data ecosystems. While most stakeholders had prior experience with data sharing partnerships, they acknowledged the data ecosystem posed unique challenges, requiring a move beyond trust-based collaborations. Stakeholders recognized that several challenges - mainly related to ecosystem conditions and the need to go beyond trust-based data collaboration - could be better managed through standardization rather than relying mainly on relationships between parties, as noted by a public transport authority in the following excerpt:

> "Well, my organization publishes to the public a lot of usable information. I think we don't publish as much data as we could, but probably this is not entirely our fault (…). At the end of the day, it's not our data; this data doesn't belong to us, (…) We could achieve that (an effective data sharing in the ecosystem) but there is the need for standardization" (Public Transport Authority, 1).

As MobiDataLab evolved into a data ecosystem, the limitations of trust-based collaboration become clear: while trust is valuable, it often proves insufficient for managing the complexities of multi-stakeholder ecosystems. The main challenges underscoring the need to move beyond trust are linked to the intrinsic characteristics of data ecosystems - such as stakeholder

heterogeneity, the absence of a shared long-term approach and the lack of clear legal agreements.

The heterogeneity of stakeholders, each with distinct interests and objectives, was seen as a challenge to sustaining the trust needed for collaboration. Differences in interpretations of transparency, accountability, and data security often led to misunderstandings and potential conflicts. As an ITS provider noted, this diversity necessitated moving beyond trust-based relationships:

> "So, we have a quite heterogeneous set of stakeholders coming from a law department, from community building, public relations stakeholders, companies, research centers. Every stakeholder has a different view on the project. This was in the beginning, and this will be at the end as well if we don't narrow these perspectives (…) and provide a common ground" (ITS provider, 1).

Economic incentives further complicated collaboration. Some stakeholders were deterred by concerns about uncertain returns, risks to competitive advantage, and differing goals, which threatened full participation and diminished the ecosystem's value:

> "Private companies have a strong incentive to grow their business and revenue, which drives them, but it is not the same for all kind of companies. In other cases, companies may find sharing data interesting if it helps their growth, but sharing isn't always natural as it often prevents them from collecting revenue out of the data (…) this mentality becomes a problem for the ecosystem and its ability to generate value overall" (ITS provider, 1).

Technical capabilities were another concern. Stakeholders worried that uneven skills and resources could result in inconsistent data quality, incompatible standards, and divergent practices, creating bottlenecks and threatening the ecosystem's effectiveness.

"If we don't have common motivation and common vision: are there at least common technical competencies? The lack of that can be problematic for a public organization like ours as well as for private companies, as well for everyone, probably" (Public Transport Operator, 1).

As explained in the excerpts above, stakeholders feared that differing objectives, capabilities, and economic incentives could hinder the development of a unified vision, which they considered essential for advancing the ecosystem. Without this shared vision, they worried that data-sharing efforts might become fragmented, with some stakeholders moving forward while others hesitated. Additionally, the lack of alignment was seen as a potential source of risks such as opportunism, data misuse, and inefficiencies. These concerns led some stakeholders to express reluctance about fully committing to the ecosystem under such uncertain conditions, as stated by the CEO of a railway transport operator in the following excerpt:

"Today there are no strategies (for the data sharing). Data is shared according to the time axis with which it is to be used. With time, if things will change, also our involvement could be a different one" (CEO 1, Railway Transport Operator).

Last, stakeholders raised concerns about legal and regulatory inconsistencies. The absence of a unified framework created uncertainty over which datasets could be shared, how they should be integrated, and under what conditions. As reported by a researcher in the following excerpt, negotiating contracts that ensure both compliance and effective collaboration becomes challenging within ecosystems, often requiring the pursuit of alternative solutions:

"Well, I can speak for the work we are doing, so there are a lot of legal challenges. (…) What datasets can be provided and how to integrate them? There are also interoperability issues. How to make sure that the different datasets work and communicate with each other to make the systems

> work? Then there is again legal issue about contracting on how to share the data (…) so much regulatory uncertainty (…) new problems and new tools call for new solutions" (Researcher, 1).

In sum, stakeholders emphasized that tackling the inherent heterogeneity of data ecosystems and the lack of a unified long-term strategy is essential for establishing a resilient data-sharing framework that meets the diverse needs and expectations of all participants. As explained next, they identified the development of standards as a key solution to minimize these problems while also avoiding dependence on trust.

4.4.2. Establishing standards as substitutes for trust

To address challenges typical of data ecosystems, MobiDataLab's stakeholders emphasized standardization as a solution to move beyond trust-based relationships. In particular, they agreed that as data ecosystems expand, trust alone becomes insufficient to manage the complexities of multi-stakeholder collaboration. By contrast, adopting clear and consistent standards fosters system scalability, improves data interoperability, and reduces reliance on interpersonal trust.

A central effort was the development of a cloud-based prototype platform, the Transport Cloud, designed to facilitate seamless data exchanges among mobility stakeholders. Supporting this platform was the Open Knowledge Base (OKB), a web-based repository of best practices, guidelines, and technical documentation on mobility standards. This tool enabled stakeholders to access information and reflect on existing standards and methodologies, including legal frameworks, governance, data privacy, and technical protocols. Researchers conducted studies on mobility data-sharing standards, identified best practices, and shared findings with the ecosystem through reports, meetings, and workshops.

In addition to featuring the OKB, the Transport Cloud also served as a testing ground for standardization initiatives, allowing stakeholders to pilot and refine standardized protocols. Through workshops, hackathons, and technical reviews, stakeholders co-created protocols for secure and efficient data handling. These activities focused on improving data quality, addressing inconsistencies, and reducing interpretation errors to ensure that the Cloud was both a technological and a standardization tool:

"Data quality, for example, quality of maps is excellent, but for traffic information, description of data, identification, latitude, and longitude location, they're often wrong because there is no standard to update this, so no one is taking care of it" (Logistics Operator, 3).

"The aim is to provide suggestions about future standards and specifications to be adopted for improved Mobility data sharing and to show how these standards have been implemented in the different stages of the MobiDataLab project (…). To attain this objective, the methodology followed was based on the identification of standards and their context, their application in MobiDataLab, and their development and related support for implementation" (MobiDataLab, Deliverable D2.5).

"By mapping existing standards, protocols, and frameworks, the OKB facilitates understanding of what is already in place and identifies areas where new standards are needed. The OKB entails documents discussing potential standards for legal and technical data governance, as well as for privacy, and security" (MobiDataLab Grant Agreement).

Noteworthy, the Transport Cloud and the OKB were seen as ways to reduce the costs of trust-based data-sharing and facilitate standardization. According to most opinions, fostering a stable environment through standardization enabled more seamless and efficient exchanges while tackling challenges arising from diverse stakeholders and heterogeneous data types. The following excerpt highlights the high agreement in the ecosystem about the need for standardization as a trust-replacing mechanism:

"Data ecosystems are not supposed to be ran like startups (…) You can't survive vast amounts of data and dozens of companies with one-to-one agreements (…) What I expect is that they (the data ecosystems) will move forward with the standardization and with a common approach" (Researcher, 1).

A driving motivation behind this shift was the need to align multi-stakeholder data governance by using the OKB as a common framework to bridge differences and ensure consistency. Stakeholders believed shared protocols could foster mutual understanding, reduce misunderstandings, and establish a reliable foundation for data exchanges.

As the excerpts above suggest, standardization is a highly felt need across the ecosystem. While many agree that it may not resolve every challenge - such as proprietary data feeds – the expectation among the stakeholders was that it would minimize discrepancies, smoothing data exchange and enhancing predictability. Standardized protocols play a critical role in maintaining data quality, ensuring stakeholders can rely on shared information to support their operations effectively. Unlike trust-based models, which rely on individual relationships, interviewees see standardized protocols as a uniform set of rules for all participants, making data sharing scalable while reducing the need for personalized trust-building. In this way, new participants can more easily integrate by adhering to established standards, and all involved parties can be reassured that everyone will follow the same guidelines. However, the attempts to develop common standards through the Transport Cloud and the OKB also implied trusting the very process of standardization, as will be explained in the following section.

4.4.3. Trusting standardization (procedural trust)

Our informants in the MobiDataLab ecosystem believed that as standardized practices became widespread, the need for informal trust-building between actors would diminish. However, they also emphasized the importance of ensuring that all actors in the data ecosystem trusted both

the proposed standards and each other's ability to develop common protocols for consistent data sharing. This highlights the critical role of procedural trust in fostering confidence in the system.

Our findings suggest that as MobiDataLab's stakeholders got involved in the complex activities of the ecosystem, they began to see that effective collaboration was not easily achievable through technical compatibility alone; it also relied on procedural trust - confidence in the processes and protocols governing data sharing. Specifically, most of our informants anticipated that widespread adoption of standardized practices would reduce the reliance on informal trust-building. However, they stressed the necessity of ensuring that all participants trusted not only the proposed standards but also each other's ability to develop common protocols for consistent data sharing. This underscores the pivotal role of procedural trust in building confidence in the ecosystem.

"As data ecosystems grow, relying on trust in standards rather than individual trust relationships fosters a more inclusive and adaptive environment that meets the evolving needs of all stakeholders (…) If we can have this transparency and confidence, I am sure we can support both innovation and growth" (ITS Provider, 3).

"Standards without trust in the underlying standardization mechanisms are useless (…) only if people trust the process will they agree to use existing standards and invest in advancing them when the time comes" (ITS Provider, 2).

Procedural trust assures stakeholders that consistent, transparent procedures are followed to enable interoperability, aligning efforts and reducing the need for individualized trust-building. Workshops, seminars and hackathons were organized to raise awareness about existing standards and generate collective discussions about their utility and feasibility. These initiatives aimed at creating a "level playing field" where all participants understood and adhered to the

same rules. Clarity and confidence about how standards are set was considered vital for encouraging stakeholders to contribute to the ecosystem without concerns, minimizing the need for constant verification.

> "The main problem is following common standards. If we have common standards on the way to collect, analyze, and use data, we can trust because we know what the boundaries of this data are. If we are not able to understand what has been shared, we cannot trust this information" (Public Transport Operator, 1).

In the shift from relying on trust to embracing standardization, MobiDataLab stakeholders recognized that procedural trust was essential for sustainable data sharing. However, fostering procedural trust also came with challenges: those related to the data to be standardized.

4.4.4. Standardization challenges related to data

Although standardization was deemed essential, informants highlight several difficulties in trusting standards. Data heterogeneity, inconsistent data quality, and multiple, often conflicting, standards, complicated data consistency and interoperability, making trust in standardization highly challenging:

> "The problem is that there is so much noise in the data (…) everybody speaks their own language (…) I argue that there should be some sort of common language, for instance, traffic management data, [...] there needs to be more communication and perhaps a lot more collaboration, from both the private entities that are concerned but also public bodies to come up with this common language" (Municipality, 2).

A perceived barrier to standardization is data heterogeneity, referring to the varied types, formats, and management systems of data collected by different stakeholders. As explained

below, despite the consolidation of the OKB and the workshops, seminars and hackathons organized to prototype the Transport Cloud, achieving consensus was difficult as participants relied on diverse practices and technical frameworks. Such diversity was often seen as a challenge for standardization, as it hindered the creation of a unified framework, as explained in the following excerpt:

> "At the moment, my impression is there's a lot of fragmentation and a lot of different bodies that are discussing standards [...] I think it's hard to have a standard for everyone. This is also complicated by the need to cater to the unique needs and technical capabilities of each participant, making it difficult to develop a single standard that meets everyone's requirements" (Municipality, 2).

Inconsistent data quality posed significant challenges, as data from varied sources often lacked uniform accuracy and reliability, creating barriers to standardization:

> "Without common data quality standards, stakeholders are forced to spend significant time and resources cleaning and analyzing data before it can be effectively used, resulting in delays and reduced efficiency" (Transport Operator, 3).

A clear example of this challenge arose when transportation agencies attempted to standardize geolocation data for real-time traffic management. While some data providers used accurate mapping systems, others relied on outdated or incomplete datasets, causing discrepancies in location accuracy and delaying the deployment of traffic optimization solutions.

> "We have already a lot of data shared by users about their mobility habits just using phone. The issue is that, in my experience, most of these data are not accurate, so they must be analyzed a lot before being used to find solutions for mobility (…) this surely curtails our propensity to trust the standards" (Consultant, 1).

Finally, the presence of multiple parameters and standards further complicated interoperability. As noted by an interviewed researcher, stakeholders used very different methods and technical standards for data management, creating a fragmented landscape. This lack of a unified approach also introduced legal and contractual hurdles, leading to delays in navigating data-sharing agreements.

Together, these challenges highlight the complexities of implementing standardization in a data ecosystem and the difficulty of trusting standards amid heterogeneous data and concurrent practices. Addressing these challenges was deemed an essential step toward a more cohesive and efficient standardization process and required alignment on common data sharing standards.

4.4.5. Aligning on common data sharing standards

As MobiDataLab's stakeholders encountered standardization challenges related to data – such as data heterogeneity, variable quality, and multiple standards – they recognized the need to establish agreed-upon data sharing protocols, as illustrated in the following excerpt:

> "We still miss common standards and rules to be followed by everyone to share data in the same way. In our sector, (…) there are no rules that we are obliged to follow when we must collect and share specific kinds of data and it is so challenging to find partners that are able to share the data you need (…) but if we manage to create an integrated system each of us will be able to extract exactly what we need, even if we don't know what exactly the others need" (Public Transport Operator, 2).

As explained in the excerpt above, the stakeholders of MobiDataLab saw alignment on common data-sharing standards as essential for their standardization project. The alignment envisioned involved a continuous process of reviewing, improving, integrating, and reaching consensus on standards, ultimately creating a unified framework for data exchange.

“When you want to build a sustainable collaboration, trust is the main factor. If there is no trust, you cannot collaborate at all, so you need to have transparency at some part. For example, in data usage because what you need to keep in mind is you're enlarging the cake. So, every single player is not able to cover the whole thing, so you need your partners here, and simply having standards in place will not do the job, you will need alignment” (ITS Provider, 4).

However, they also acknowledged that alignment posed significant challenges. As explained in the excerpt above, in sectors without binding standards for data collection and sharing, such as mobility, stakeholders struggle to find compatible partners and face inefficiencies in adapting to various data formats.

MobiDataLab strongly advocated for technical reviews as a core part of the OKB in the hope of fostering consistency in data-sharing approaches. The continuous improvement and integration of standards was reported as essential for maintaining a healthy ecosystem, allowing stakeholders to assess current practices collectively, identify gaps, and establish standardized protocols. Lack of shared guidelines, on the contrary, posed a concrete risk to misinterpretations and inconsistencies, undermining trust and hindering collaboration. However, in their efforts to align common data-sharing standards to enhance ecosystem interoperability, stakeholders once again faced a trust challenge, this time that of system-based trust.

"Sure, they (technical reviews) are important. With every discussion we are sharing knowledge, with every meeting we are getting more experience. We are getting into assets that our partners are using. Work package members are getting in touch with the groups that we are sharing data with to figure out how to review and integrate their data in MobiDataLab" (Researcher, 1).

4.4.6. Trusting the system (system-based trust)

While working toward common data-sharing standards, MobiDataLab stakeholders emphasized the importance of systemic trust - confidence in the systems, structures, and

technologies underpinning the ecosystem. As data-sharing networks grew more complex, stakeholders relied on the ecosystem's infrastructure, which combined internal capabilities with external supports such as regulations, contractual frameworks, and industry standards. These elements collectively fostered a secure environment, enabling stakeholders to share data with confidence, knowing that robust systemic safeguards were in place.

Reliable technical infrastructure, data management protocols, and cybersecurity measures instilled confidence in the ecosystem's ability to achieve standardization. These factors also motivated stakeholders to enhance their own internal capabilities, aiming to become more attractive partners and more integral to the system, as highlighted in the following excerpts:

> "We do have specialist individuals in terms of data management [...] we've got a cybersecurity, especially as we've got a data management specialist within our team (…) having internal capabilities from multiple sources is reassuring, it's a proof that the system is reliable" (Transport Operator, 2).

External regulations further contributed to system trust. As often reported to us, clear regulatory frameworks would have enabled stakeholders to engage in data sharing without fear of misuse:

> "If there is a regulation to data sharing, it's very easy: then you have to share and you are not worried about the kind of data out there because there is a legal agreement that can protect you" (City representative, 1).

Last, informants also highlighted contracts and procedural agreements enhanced predictability in the system by formalizing commitments and establishing clear terms for data use, as reported in the following excerpt:

"I think that trust in the ecosystem essentially comes down to contractual agreements. Of course, trust is one thing, when we have a new party on board in our data sharing platform, there's always this contractual agreement that their data will only be used for the purposes that they intend; we call it the data-sharing rules (…) it's not really about trusting the other as it is about trusting the system as a whole (…) that it's reliable and works as expected (…)" (Airline Project Manager, 1).

In sum, as explained in the excerpts above, systemic trust relies on a blend of internal and external infrastructure which provides the checks and balances needed for ecosystem standards alignment. However, this process is not without limitations, as MobiDataLab participants often encountered alignment challenges along the way.

4.4.7. Alignment challenges

In the attempt to create a formalized protocol for consistent data sharing which leveraged both internal and external ecosystem infrastructure, stakeholders in MobiDataLab encountered several alignment challenges. A frequently reported challenge was the lack of a clear regulatory framework for data sharing. Ambiguities and inconsistencies in regulations made it difficult for stakeholders to understand and meet legal requirements, leading to delays and reluctance to share data. Additionally, varying rules across jurisdictions or sectors risked creating further complications, resulting in confusion and restricting data flows. Ultimately, these concerns threatened the development of system-based trust.

"Well, I can speak for the work we are doing, so there are a lot of legal challenges. (…) What datasets can be provided and how to integrate them? There are also interoperability issues. How to make sure that the different datasets work and communicate with each other to make the systems work? Then there is again the legal issue about contracting on how to share the data" (Researcher, 2).

Internal conditions such as ambiguities in data collection and sharing practices further complicated trust and alignment possibilities in the ecosystem. For instance, stakeholders in the MobiDataLab acknowledged the lack of a strong ecosystem infrastructure and feared inconsistent data handling methods and consequent inefficiencies, misunderstandings and obstacles to interoperability.

To address alignment challenges, collaboration became essential.

4.4.8. Collaborating to achieve standardization

As informants dealt with alignment challenges, they became increasingly aware that successful standardization depended on each other's ability to effectively collaborate.

As reported in the following excerpt, a key step in collaborating to achieve standardization was creating collaborative opportunities (as the Living and Virtual Lab) where mobility stakeholders could discuss requirements and align on shared standards for the ecosystem. From such standpoint, effective standardization required a shift from a competitive to a collaborative approach, fostering a shared framework that benefited the entire ecosystem.

"Collaboration between partners is key and at least in the next 10 years cannot be substituted by data analytics and machine learning. People need to work together, need to understand the different positions to commonly agree on the way forward" (Coordinator 1 of a Transport Agency).

The shift to collaboration was pursued in the meetings, workshops and hackathons that aimed at discussing different standards and data sharing protocols to be featured in the Transport Cloud. These interactions aimed at helping stakeholders see the benefits of standardization, while exploring best practices together.

However, to achieve collaboration for standardization, stakeholders confronted with the fundamental need for mutual interpersonal trust.

4.4.9. Trusting the others (interpersonal trust)

In the attempt to set up effective collaboration schemes in support of standardization, MobiDataLab's stakeholders acknowledged the importance of interpersonal trust. Specifically, trust in other stakeholders' benevolence and integrity was deemed essential for creating an environment where stakeholders confidently collaborated for the development of shared data and resources. For instance, as many of our informants explained, although procedural and systemic trust were very necessary, human trust remained fundamental.

Accordingly, key factors in building trust included shared motivation and vision and integrity among stakeholders. When stakeholders find the others committed to these values, their trust deepens, as well as their availability to engage in collaborative efforts. We found that stakeholders constantly searched for proofs about each other's motivation and explained how this was particularly valuable during conflicts or challenges, allowing stakeholders to address issues without concerns about hidden agendas. Specifically, trust in the benevolence and integrity of other stakeholders was deemed essential for creating an environment where participants confidently collaborated on the development of shared data and resources, as well as the construction of the OKB and the Transport Cloud. Building the OKB and the Transport Cloud required extensive trust among stakeholders to align on common goals, data standards, and interoperable systems to foster an environment where participants were willing to contribute openly to shared resources and infrastructure.

In the same way, they sought for proofs of integrity. For example, stakeholders explained that demonstrating integrity proved commitment to the social goal of more efficient and sustainable mobility solutions for the society. This is evident in the following excerpt:

"I think there are times when different stakeholders try to find solutions together and identify common standard to be applied to solve issue. But it always depends on the goals and abilities that each stakeholder has. So, if there is a common motivation it is reasonable to think there is an interest to investigate for new solutions, but if there is not a common interest, it's difficult to solve problems" (Public Transport Operator, 1).

4.4.10. Collaboration Challenges

While trust was considered crucial for effective collaboration in standardization, it often presented challenges. During meetings, workshops, and hackathons, conflicts arose from competing interests, divergent practices, and resource disparities. A recurring issue was the perception that influential stakeholders were prioritizing proprietary standards over collaboration:

"Yes, in general, I mean, it depends on the standards. If you take a look at the public transport sector, there we have standards that are developed by the industry. But is that really a question of collaboration? (…) we have powerful single players who sometimes develop their own standard" (Traffic Management Center, 1).

"It's not only always a matter of revenues, but let's say that's a lot of those companies, they have an interest pushing their own standards, because if their standards become the industry standards, they have a kind of interest because they, they, they manage. So that's something that, uh, is, um, that's been observed in, uh, in many fields. Um, but at the same time, pushing the standardization requires some efforts and, uh, as a cost and the balance is not always in, in, in favor. So, I would say that there is a tendency, but at the same time, um, there, there is a cost associated and not all companies are willing to support that cost" (Researcher, 3).

Overcoming these obstacles was considered essential to establishing a sustainable ecosystem where data can be effectively shared and utilized. Yet, navigating the balance between trust and standardization remains fraught with tensions:

> "We are all working at a different speed with a different attention towards the subject of digitalization. And that makes it quite difficult to put in your effort and to collaborate on a national scale (…) I think it largely depends on how much each of us will commit to data sharing" (Municipality representative 3).

As the above quote reveals, aligning stakeholders with diverse priorities and capacities remains a persistent challenge. The very effort to define clear roles and responsibilities requires mechanisms for monitoring participation, which introduces additional layers of collaboration challenges. Transparency is crucial, as it builds trust, aligns efforts, and ensures all participants understand their roles. Without open communication and clear data-sharing practices, stakeholders may hesitate to fully engage due to concerns over data misuse or hidden agendas.

> "(When you set up data sharing systems) having a collective vision is not a prerequisite, but I think this collaboration and this knowledge sharing is super helpful because you generate transparency, you learn from each other. (…) It is super helpful to have a fruitful and efficient collaboration, and therefore I think the initiatives like MobiDataLab are super helpful for that (…). In my view (this) is what distinguishes successful cases of sharing from not so successful ones" (ITS Provider, 1).

MobiDataLab's activities, such as the development of the Open Knowledge Base (OKB) and the Transport Cloud, further exemplify the delicate interplay between collaboration and trust. While designed to facilitate standardization, their development depended on collaboration. Stakeholders stressed the need to accommodate diverse needs, ensure regular monitoring, and

promote open practices to achieve shared goals. Strengthened collaboration provides the foundation for effective data sharing and innovation across the ecosystem:

> "I would say collaboration is more important than the risks out there. I mean creating technology is something that one partner can do on their own, but the prerequisite for a new mobility ecosystem to work, is for the partners to be able to work together. So, if that doesn't happen, even if you have the technology, if you don't have the collaboration, then that is not possible to happen" (Researcher, 2).

However, collaboration is not always straightforward. Stakeholders expressed concerns about data ownership, control, and the implications of sharing data more broadly:

> "The big other problem is the fear of giving up one's data, because there is no awareness among the logistics nodes, and no one gives it to them, to then say how that data will be processed, and by whom above all, because as long as the relationship is from A to B, I am calm, the data goes from me to you and they are used so that you can transport (…). It is different to make data available to a community, to a system" (CEO 1, Railway Transport Operator).

This reluctance underscores the paradox: stakeholders must trust both the system and each other to build a standardized, trust-free ecosystem. Tools like the OKB and Transport Cloud brought this tension into focus, revealing that standardization itself requires trust - not just in the standards but in the collaborative process of creating, implementing, and maintaining them.

At this point, the need to establish a data sharing protocol in a data ecosystem resurfaces, pushing stakeholders to adopt standards as substitutes for trust. This highlights the paradox of standardization and trust-building practices. On one hand, stakeholders desire a system where trust is no longer necessary for data sharing and ecosystem management. On the other, they find themselves relying on trust to build this trust-free system, realizing that trust cannot be entirely excluded from the process. Even with established standards, the ongoing need for sharing,

updating, and alignment will always require relational engagement among actors, where trust remains a crucial element.

4.5. Discussion

Data ecosystems offer an innovative approach to reconciling the interplay between standardization and trust through collaborative participation among multiple stakeholders. Despite the expanding literature, the relationship between trust and data standardization within these ecosystems remains underexplored, leaving significant gaps in our understanding of their dynamics and implications (Backer, 2024; Z. Cao & Lumineau, 2015; Lioliou et al., 2014; Liu et al., 2022; Wareham et al., 2014). Existing studies highlight the importance of examining the various factors that influence the tension between trust and standardization, noting that excessive standardization can suppress innovation, hinder collaboration, and reduce adaptability - all of which typically rely on trust (Liu et al., 2022; Wareham et al., 2014). Furthermore, while standardization can enhance trust by promoting transparency and accountability, trust often serves as a prerequisite for the successful adoption of standards, especially in ecosystems characterized by heterogeneous actors (Backer, 2024; Z. Cao & Lumineau, 2015; Wareham et al., 2014). Our study aims to explore the paradoxical relationship between standardization and trust, analyzing the processes that emerge from the necessity felt by the MobiDataLab actors to move beyond trust-based data collaboration.

Through qualitative research within a data ecosystem, we found that overcoming challenges related to standardization often requires some form of trust – whether in the standards themselves, the ecosystem, or the people behind them. Conversely, addressing issues related to trust frequently leads to aspirations towards standardization. We observed that data ecosystems are caught in a cyclical process where governance advances without forcing a

choice between standardization and trust, by constantly shifting between searches for the best standards, the most efficient infrastructures and the most advanced collaboration skills.

4.5.1. Advancing the understanding of the interplay between trust and standardization

In data ecosystems, both standardization and trust play pivotal roles in shaping interactions and collaborations among participants. Although their relationship is complex, their combined influence is essential for creating an environment conducive to effective data sharing and utilization. The relationship between standardization and trust is thus complementary as standardization can foster trust by providing transparency and accountability, while trust can be considered as a prerequisite for the successful adoption of standards (Backer, 2024). The trade-off lies in excessive standardization potentially stifling flexibility and innovation, while insufficient trust risks undermining stakeholder alignment and cooperation - leaving governance systems vulnerable to inefficiency or fragmentation. Our findings suggest that balancing the challenges and opportunities within standardization requires achieving an equilibrium between flexibility and rigidity – in line with Zang et al. (2020) -, but also that this equilibrium may often be a mere chimera.

Specifically, standards need to be rigid enough to ensure consistency while remaining flexible enough to accommodate new data types, stakeholders, and their relationship (Hanseth et al., 2006; Zang et al., 2020). This underscores the importance of directing future efforts toward developing widely accepted, context-sensitive standards and creating governance frameworks that facilitate collaboration among diverse participants while addressing the unique needs of evolving data ecosystems. To accomplish this, both trust and standardization are crucial in fostering effective data ecosystems.

Our research indicates that while standardization and trust are recognized as vital components of governance in data ecosystems, stakeholders show a strong inclination to transition away from trust-based collaborations toward models governed by standardized practices. However, in examining how the tension between standardization and trust is navigated, we found that formal mechanisms intended to reduce reliance on trust paradoxically increase the necessity for stakeholders to engage in trust-driven relational approaches. This dynamic becomes even more complex with the increasing integration of automation and digital technologies in data ecosystems. Automation, often perceived as a tool for eliminating human bias and enhancing efficiency, introduces new challenges in governance. While these technologies can streamline processes and establish predefined rules, they also shift some of the burden of trust from interpersonal or institutional relationships to the reliability of the technological infrastructure itself. Automated systems depend on accurate programming, unbiased algorithms, and transparent decision-making processes to function effectively, yet they cannot operate independently of human oversight and validation.

Stakeholders may perceive automation as a way to reduce the need for trust by creating consistent, repeatable processes; however, the deployment of these technologies often requires a foundational level of trust in their design, operation, and outcomes.

Furthermore, our findings reveal that ecosystem trust encompasses various forms, including interpersonal trust (trust between individuals), institutional trust (trust in institutions), and system-based trust (trust in the technological infrastructure). These forms of trusts are essential in data ecosystems, serving as a cornerstone of innovation by enabling the free exchange of ideas and resources without fear of misuse or misappropriation. For example, procedural trust - the confidence that stakeholders can develop common protocols and procedures for consistent data sharing - is fundamental: without it, stakeholders may hesitate to share data, limiting the ecosystem's ability to generate insights and foster innovation (in line

with Lin et al., 2020). Additionally, system-based trust - confidence in the ecosystem's internal and external conditions - assures participants of the capabilities to navigate standardization challenges (in line with Byabazaire et al., 2020; Hou & Jansen, 2023). Interpersonal trust - trust in stakeholders' benevolence and integrity during collaboration - enables the data ecosystem to function without continuous monitoring, ensuring reliable data processing and exchange, which are critical for both immediate operations and long-term relationships in line with the interpersonal trust studies proposed by Fachrunnisa & Hussain (2013).

These three types of trust not only operate individually but also interact dynamically, shaping the relationship between standardization and trust in complex ways. Procedural trust often serves as the starting point, as shared protocols and standards provide a foundation for collaboration by creating a predictable framework for interaction. However, the successful adoption of these protocols requires system-based trust, as stakeholders must believe in the technological infrastructure's ability to enforce and uphold the agreed-upon standards. Similarly, interpersonal trust can complement both procedural and system-based trust by fostering goodwill and reducing resistance to compliance, particularly in situations where rules or systems are ambiguous or evolving.

The relationship between these types of trust is not necessarily sequential or exclusive. While in some cases, procedural trust may precede system-based and interpersonal trust, this order is not fixed. For example, in our case study, we observed that interpersonal trust often emerged first during early stages of collaboration, enabling stakeholders to agree on initial procedural frameworks. Over time, as protocols were formalized, system-based trust became more prominent, reinforcing stakeholders' confidence in the ecosystem's scalability and reliability. In other contexts, such as those involving high levels of technological uncertainty, system-based trust might precede interpersonal trust, as participants first assess the robustness of the infrastructure before committing to relational investments.

Rather than operating in isolation, these forms of trust frequently work in tandem, creating feedback loops that enhance or weaken the overall trust ecosystem. For instance, a well-functioning system can strengthen procedural trust, which in turn fosters interpersonal trust as participants observe the system's fairness and consistency in action. Conversely, a breakdown in one type of trust - for example, a system failure - can undermine both procedural and interpersonal trust, causing disruptions in collaboration.

On the other hand, stakeholders express a strong desire to establish a data sharing process that advances toward standardization, aiming to implement standardized practices and reduce reliance on trust-based collaborations. This involves implementing data sharing standards, reviewing, enhancing, and integrating these standards to develop mutually agreed-upon protocols that improve ecosystem interoperability. However, the standardization process is challenged not only by the pursuit of trust but also by issues related to data heterogeneity and variable quality, alignment difficulties, and collaborative obstacles, in line with the current literature (Garud et al., 2002; Lingens et al., 2021; Lnenicka et al., 2024; Moreno et al., 2019; Oliveira et al., 2019; Zang et al., 2020).

Our study underscores the ongoing struggle between trust and standardization, where neither can fully replace the other. While innovative technology is often considered ideally a trust-free system, our results demonstrate that human trust remains essential at critical interaction points such as data input and decision-making. This aligns with Hanseth (2006) who argues that technology can enhance trust but cannot entirely replace the need for human trust. Our findings suggest that the focus should be on harmonizing standardization and trust rather than attempting to prioritize one over the other.

In contrast to prior literature, which often emphasizes the potential for automation and digitalization to reduce reliance on trust by creating predictable, standardized processes (Hanseth et al., 2006; Zang et al., 2020), our findings suggest a more complex reality. While

automation and widespread digitalization have undoubtedly introduced tools and infrastructures that enhance transparency, consistency, and efficiency, they also bring new challenges that deepen the interdependence between trust and standardization. Automated systems are only as reliable as their design, programming, and the data they process, which means that stakeholders must place significant trust in these systems' integrity, fairness, and robustness.

Moreover, the increasing reliance on automation reshapes the dynamics of trust by redistributing it across different levels. Stakeholders not only need to trust the technological infrastructure but also the human and institutional actors responsible for designing, implementing, and maintaining these systems. This layered trust requirement explains why formal mechanisms of standardization often fail to entirely replace relational trust.

4.5.2. Implication for existing literature and for practice

Our findings are significant for several reasons. First, we complement existing literature, which often provides a static view of standardization, trust and their interplay. The relationship between trust and data standardization in data ecosystems remains underexplored, leaving gaps in understanding their dynamics and implications (Backer, 2024; Z. Cao & Lumineau, 2015; Lioliou et al., 2014; Liu et al., 2022; Wareham et al., 2014). We argue that understanding the process by which one factor is prioritized over another in establishing data sharing protocols can shed light not only on the stakeholders involved but also on their relationships and the dynamic nature of the ecosystem.

Second, our practice-based approach contributes significantly to the growing body of research that documents the nuanced and often ambivalent approaches to data sharing within and across organizations (Z. Cao & Lumineau, 2015; Liu et al., 2022). While previous studies have identified the coexistence of conflicting priorities, our findings provide deeper insights into the inherent tension between the drive to establish standards as a means of reducing reliance

on trust and the persistent need for trust dynamics to enable effective collaboration. This tension arises because, while standardization can create clear rules and frameworks to facilitate interoperability and reduce uncertainty, it does not fully eliminate the relational elements required for successful data sharing. Trust dynamics remain pivotal in aligning stakeholder motivations, resolving conflicts, and ensuring adherence to shared standards. Furthermore, our approach highlights how these competing forces shape decision-making and collaboration in data ecosystems.

Third, our study adds to the literature on data and digital ecosystems, increasingly recognized as interconnected environments where stakeholders, organizations, technologies, and data interact dynamically to create value and drive innovation. We address a critical gap in the literature by emphasizing the need to study data ecosystems as distinct phenomena (Oliveira et al., 2019). Furthermore, we align with scholarship advocating for more robust data governance mechanisms that balance multi-stakeholder goals, data quality, privacy, scalability, and interconnectivity - elements where standardization and trust are often highlighted as key (Abraham et al., 2019; Günther et al., 2017).

Importantly, our study contributes to a renewed conceptualization of data governance in data ecosystems. We align with Aaen et al.'s (2022) call for pragmatic and politically aware approaches to data ecosystems. While emerging research has focused on data value creation (Azkan, Möller, et al., 2020; De Prieelle et al., 2022; Gelhaar & Otto, 2020; Oliveira et al., 2019), we argue that data collaborations at the ecosystem level require dramatically new approaches to data governance. These approaches must incorporate sophisticated relational and political practices, as standardization alone cannot provide a comprehensive solution (Davidson et al., 2023; Donge et al., 2022; Gregory et al., 2021).

In terms of practical implications, our study offers valuable insights for potential participants in data ecosystems, such as private and public organizations, citizens, and policymakers. Our primary aim is to present a realistic perspective on the possibilities and challenges of data ecosystems, moving beyond idealized notions and common criticisms regarding the roles of data, standardization, and trust in society.

We advocate for moving away from oversimplified and idealized notions of trust, which often assume unified visions, practices, and intentions among participants. Instead, we emphasize understanding the diverse and sometimes conflicting dynamics within data ecosystems. Both trust and standardization carry inherent risks and benefits, and our findings highlight the importance of adopting a nuanced approach. A deeper analysis of their interaction is essential to understanding how these factors influence data ecosystems and impact their potential to achieve desired outcomes.

Our study also offers guidance for policymakers by shedding light on the positive and negative dynamics that arise from the interplay between standardization and trust. Breaking down the components of each, we highlight potential pathways to reduce reliance on trust-based data sharing. On one hand, moving beyond purely trust-based collaborations can address motivational divergences among stakeholders, mitigating issues caused by misaligned goals and fostering more sustainable and effective collaboration within data ecosystems. On the other hand, the paradoxical process of building standardization through trust underscores the need for trust among participants to align practices and objectives.

The dual nature of standardization and trust reveals a complex dynamic where both elements can be simultaneously supportive and challenging. Standardization provides consistency and reduces uncertainty but can lead to tensions when perceived as inflexible or

misaligned with specific needs. Trust fosters cooperation and reduces resistance but it is inherently fragile and can be undermined by conflicting interests or past experiences.

Ultimately, we call for further exploration of these dynamics within different types of data ecosystems and their implications for achieving desired outcomes. Policymakers and practitioners must recognize these nuanced relationships and adapt strategies to balance the benefits of standardization with the subtleties of building and maintaining trust among diverse stakeholders. Such awareness is critical for fostering resilient data ecosystems capable of evolving to meet current and future challenges.

4.6. Conclusion, Limitations and Future Research Directions

The rapid growth and dynamic interactions within data ecosystems (Demchenko et al., 2014; Rammer et al., 2023) underscore their uniqueness, particularly in addressing the interplay between standardization and trust (Backer, 2024; Z. Cao & Lumineau, 2015; Liu et al., 2022; Wareham et al., 2014). This research explores these complexities through the lens of the MobiDataLab project - a Horizon 2020 funded initiative that evolved into an emerging mobility data ecosystem in Europe, centered on data sharing.

The findings highlight the critical need to transition from trust-based to standardized practices in data sharing. However, this shift revealed an inherent interdependence: while standardization ensures consistency and interoperability, trust remains essential for managing uncertainties, aligning stakeholder objectives, and fostering collaboration. The study identifies three forms of trust - procedural, system-based, and interpersonal - that underpin data-sharing practices. Paradoxically, efforts to reduce reliance on trust through formal mechanisms often increase the need for relational trust, reinforcing the interdependence between these two elements.

From a governance perspective, the research emphasizes the necessity of harmonizing trust and standardization to effectively manage the complexities of multi-stakeholder ecosystems. Consistent with existing literature (Abraham et al., 2019; Günther et al., 2017), the findings advocate for adaptive governance models that integrate the strengths of trust and standardization, promoting flexibility and scalability while maintaining shared protocols.

The study also provides actionable insights for organizations, policymakers, and citizens. It challenges idealized notions of trust and standardization, highlighting their nuanced roles in overcoming ecosystem challenges. While trust is vital for collaboration, it hinges on goal alignment - a difficult task in diverse ecosystems. Standardization, on the other hand, offers structural consistency but risks creating rigidity. A balanced approach is crucial for fostering sustainable collaboration and addressing stakeholder misalignments.

For policymakers specifically, the research outlines strategies to reduce over-reliance on trust-based systems while addressing misalignments among stakeholders. It also underscores the paradox of building trust through standardization, where formal mechanisms depend on relational trust to align practices. Striking this balance is essential to enhance cooperation and ensure the long-term success of data ecosystems.

However, the study has its limitations. Its focus on MobiDataLab constrains the generalizability of the findings, as other ecosystems, such as those centered on data commercialization, may exhibit different dynamics. The reliance on qualitative methods, particularly interviews, reflects subjective perceptions, while limited opportunities for direct observation may introduce bias. Additionally, capturing the ecosystem in its formative stage means that future developments could alter the observed relationships. Furthermore, the study does not aim to provide a comprehensive view of the trust-standardization relationship, or the

processes stakeholders follow to establish data-sharing protocols. Other factors may significantly influence collaboration, standardization, and trust in different ecosystems.

Future research should investigate the trust-standardization dynamic across varied contexts, including ecosystems driven by emerging technologies like blockchain and artificial intelligence. Comparative studies of ecosystems with different governance models could yield deeper insights into how contextual factors shape the trust-standardization interplay. Exploring socio-political dimensions, such as power dynamics and incentives, could further enhance understanding of how trust and standardization are negotiated. Additionally, drawing on paradox literature, future studies could use data ecosystems as empirical contexts to explore tensions such as those between knowledge sharing and protection, transparency and privacy, or openness and control - moving beyond purely theoretical framings to examine how such paradoxes play out in practice (Boders, 2011; Ahmed et al., 2014; Graham et al., 2016; Chen et al., 2021)

This study advances the understanding of governance in data ecosystems by illuminating the intricate relationship between trust and standardization. It demonstrates how formal mechanisms often reinforce the necessity of trust, and harmonizing these elements can foster resilient, collaborative environments that drive sustainable innovation. The findings highlight the importance of adaptive governance models that balance structure with flexibility, enabling stakeholders to navigate evolving challenges effectively.

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. https://doi.org/10.1080/0960085X.2021.1947753

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. https://doi.org/10.1016/J.IJINFOMGT.2019.07.008

Ahmed, G., Ragsdell, G., & Olphert, W. (2014). Knowledge sharing and information security: a paradox?.

Alli, A. A., & Alam, M. M. (2020). The fog cloud of things: A survey on concepts, architecture, standards, tools, and applications. In Internet of Things (Netherlands) (Vol. 9). Elsevier B.V. https://doi.org/10.1016/j.iot.2020.100177

Arnold, N., & Loconto, A. (2020). Serving Magically Perfect Fruit Globally: Local nesting in translating multiple standards. Https://Doi.Org/10.1177/0170840620935858, 42(2), 327–349.

Azkan, C., Möller, F., Meisel, L., & Otto, B. (2020, June). Service dominant Logic Perspective on Data Ecosystems-a Case Study based Morphology. In ECIS.

Backer, L. C. (2024). Trust platforms: The digitalization of corporate governance and the transformation of trust in polycentric space. Regulation and Governance. https://doi.org/10.1111/rego.12614

Bogers, M. (2011). The open innovation paradox: knowledge sharing and protection in R&D collaborations. European journal of innovation Management, 14(1), 93-117.

Bogers, M., Sims, J., & West, J. (2019). What Is an Ecosystem? Incorporating 25 Years of Ecosystem Research. SSRN Electronic Journal. https://doi.org/10.2139/SSRN.3437014

Bonina, C., Koskinen, K., Eaton, | Ben, Gawer, | Annabelle, & School, S. B. (2021). Digital platforms for development: Foundations and research agenda. Wiley Online LibraryC Bonina, K Koskinen, B Eaton, A GawerInformation Systems Journal, 2021•Wiley Online Library, 31(6), 869–902. https://doi.org/10.1111/isj.12326

Brunsson, N., Rasche, A., & Seidl, D. (2012). The Dynamics of Standardization: Three Perspectives on Standards in Organization Studies. Organization Studies, 33(5–6), 613–632. https://doi.org/10.1177/0170840612450120

Byabazaire, J., O'hare, G., & Delaney, D. (2020). Data quality and trust: Review of challenges and opportunities for data sharing in IoT. Electronics (Switzerland), 9(12), 1–22. https://doi.org/10.3390/electronics9122083

Cao, R., & Iansiti, M. (2023). Association for Information Systems Association for Information Systems AIS Electronic Library (AISeL) AIS Electronic Library (AISeL) 14th Scandinavian Conference on Information Systems Scandinavian Conference on Information Systems. https://aisel.aisnet.org/scis2023/2

Cao, Z., & Lumineau, F. (2015). Revisiting the interplay between contractual and relational governance: A qualitative and meta-analytic investigation. Journal of Operations Management, 33, 15–42. https://www.sciencedirect.com/science/article/pii/S0272696314000680

Cepa, K., & Schildt, H. (2019). Technological embeddedness of inter-organizational collaboration processes. Research in the Sociology of Organizations, 64, 91–115. https://doi.org/10.1108/S0733-558X20190000064007

Cepa, K., & Schildt, H. (2023). Data-induced rationality and unitary spaces in interfirm collaboration. Organization Science, 34(1), 129-155. https://doi.org/10.1287/orsc.2021.1566

Chen, L., Huang, Y., Ouyang, S., & Xiong, W. (2021). The data privacy paradox and digital demand (No. w28854). National Bureau of Economic Research.

Cobben, D., & Roijakkers, N. (2019). The Dynamics of Trust and Control in Innovation Ecosystems. International Journal of Innovation, 7(1), 01–25. https://doi.org/10.5585/iji.v7i1.341

Constantinides, P., & Barrett, M. (2014). Information Infrastructure Development and Governance as Collective Action. Https://Doi.Org/10.1287/Isre.2014.0542, 26(1), 40–56.

Cox, M., & Ellsworth, D. (1997). Managing big data for scientific visualization. ACM Siggraph, 97, N1, 21–38. https://www.researchgate.net/publication/238704525

Cui, Y., Kara, S., & Chan, K. C. (2020). Manufacturing big data ecosystem: A systematic literature review. Robotics and Computer-Integrated Manufacturing, 62, 101861. https://doi.org/10.1016/J.RCIM.2019.101861

Cusumano, M. A., Kahl, S. J., & Suarez, F. F. (2015). Services, industry evolution, and the competitive strategies of product firms. Strategic Management Journal, 36(4), 559–575. https://doi.org/10.1002/SMJ.2235

Cvitanovic, C., Shellock, R. J., Mackay, M., van Putten, E. I., Karcher, D. B., Dickey-Collas, M., & Ballesteros, M. (2021). Strategies for building and managing 'trust' to enable knowledge exchange at the interface of environmental science and policy. Environmental Science and Policy, 123, 179–189. https://doi.org/10.1016/j.envsci.2021.05.020

Das, A. (2020). Trust in "Trust-free" Digital Networks: How Inter-firm Algorithmic Relationships Embed the Cardinal Principles of Value Co-creation. AIS Transactions on Human-Computer Interaction, 12(4), 228–252. https://doi.org/10.17705/1thci.00137

Das, T. K., & Teng, B. S. (1998). Between trust and control: Developing confidence in partner cooperation in alliances. Academy of Management Review, 23(3), 491–512. https://doi.org/10.5465/AMR.1998.926623

Das, T. K., & Teng, B. S. (2002). Alliance constellations: A social exchange perspective. Academy of Management Review, 27(3), 445–456. https://doi.org/10.5465/AMR.2002.7389937

Dattee, B., Alexy, O., & Autio, E. (2018). Maneuvering in poor visibility: How firms play the ecosystem game when uncertainty is high. Academy of Management Journal, 61(2), 466–498. https://doi.org/10.5465/AMJ.2015.0869

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. https://doi.org/10.1016/J.INFOANDORG.2023.100454

De Mauro, A., Greco, M., & Grimaldi, M. (2015). What is big data? A consensual definition and a review of key research topics. AIP Conference Proceedings, 1644(1), 97–104. https://doi.org/10.1063/1.4907823

De Prieelle, F., De Reuver, M., & Rezaei, J. (2022). The Role of Ecosystem Data Governance in Adoption of Data Platforms by Internet-of-Things Data Providers: Case of Dutch Horticulture Industry. IEEE Transactions on Engineering Management, 69(4), 940–950. https://doi.org/10.1109/TEM.2020.2966024

Demchenko, Y., De Laat, C., & Membrey, P. (2014). Defining architecture components of the Big Data Ecosystem. 2014 International Conference on Collaboration Technologies and Systems, CTS 2014, 104–112. https://doi.org/10.1109/CTS.2014.6867550

Dhanaraj, C., & Parkhe, A. (2006). Orchestrating Innovation Networks. Https://Doi.Org/10.5465/Amr.2006.21318923, 31(3), 659–669.

Donge, W., Bharosa, N., & Janssen, M. F. W. H. A. (2022). Data-driven government: Cross-case comparison of data stewardship in data ecosystems. Government Information Quarterly, 39(2), 101642.

Fachrunnisa, O., & Hussain, F. K. (2013). A methodology for maintaining trust in industrial digital ecosystems. IEEE Transactions on Industrial Electronics, 60(3), 1042–1058. https://doi.org/10.1109/TIE.2011.2173890

Ford, C. R., Laseter, S. H., Swank, W. T., & Vose, J. M. (2011). Can forest management be used to sustain water-based ecosystem services in the face of climate change? Ecological Applications, 21(6), 2049–2067. https://doi.org/10.1890/10-2246.1

Garud, R., Jain, S., & Kumaraswamy, A. (2002). Institutional entrepreneurship in the sponsorship of common technological standards: The case of Sun Microsystems and Java. Academy of Management Journal, 45(1), 196–214. https://doi.org/10.2307/3069292

Gawer, A., & Cusumano, M. A. (2014). Industry platforms and ecosystem innovation. Journal of Product Innovation Management, 31(3), 417–433. https://onlinelibrary.wiley.com/doi/abs/10.1111/jpim.12105

Geels, F. W. (2004). From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory. Research Policy, 33(6–7), 897–920. https://doi.org/10.1016/J.RESPOL.2004.01.015

Gelhaar, J., & Otto, B. (2020). Challenges in the Emergence of Data Ecosystems. PACIS, 175.

Graham, F. S., Gooden, S. T., & Martin, K. J. (2016). Navigating the transparency–privacy paradox in public sector data sharing. The American Review of Public Administration, 46(5), 569-591.

Gregory, R. W., Henfridsson, O., Kaganer, E., & Kyriakoi, S. H. (2021). The role of artificial intelligence and data network effects for creating user value. Academy of Management Review, 46(3), 534–551. https://doi.org/10.5465/AMR.2019.0178

Grillo, F., Wiegmann, P. M., DE VRIES, H. J., Bekkers, R., Tasselli, S., Yousefi, A., & VAN DE KAA, G. (2024). Standardization: Research Trends, Current Debates, and Interdisciplinarity. Https://Doi.Org/10.5465/Annals.2023.0072, 18(2), 788–830.

Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. The Journal of Strategic Information Systems, 26(3), 191–209. https://doi.org/10.1016/J.JSIS.2017.07.003

Haack, P., Schoeneborn, D., & Wickert, C. (2012). Talking the Talk, Moral Entrapment, Creeping Commitment? Exploring Narrative Dynamics in Corporate Responsibility

Standardization. Organization Studies, 33(5–6), 815–845. https://doi.org/10.1177/0170840612443630

Hanisch, M., Goldsby, C. M., Fabian, N. E., & Oehmichen, J. (2023). Digital governance: A conceptual framework and research agenda. Journal of Business Research, 162. https://www.sciencedirect.com/science/article/pii/S0148296323001352

Hanseth, O., Jacucci, E., Grisot, M., & Aanestad, M. (2006). Reflexive Standardization: Side Effects and Complexity in Standard Making. In Source: MIS Quarterly (Vol. 30).

He, Q., Meadows, M., Angwin, D., Gomes, E., & Child, J. (2020). Strategic alliance research in the era of digital transformation: Perspectives on future research. Wiley Online LibraryQ He, M Meadows, D Angwin, E Gomes, J ChildBritish Journal of Management, 2020•Wiley Online Library, 31(3), 589–617. https://doi.org/10.1111/1467-8551.12406

Hinings, B., Gegenhuber, T., & Greenwood, R. (2018). Digital innovation and transformation: An institutional perspective. Information and Organization, 28(1), 52–61. https://www.sciencedirect.com/science/article/pii/S1471772718300265

Hou, F., & Jansen, S. (2023). A systematic literature review on trust in the software ecosystem. Empirical Software Engineering, 28(1). https://doi.org/10.1007/s10664-022-10238-y

Huber, T. L., Kude, T., & Dibbern, J. (2017). Governance Practices in Platform Ecosystems: Navigating Tensions Between Cocreated Value and Governance Costs. Https://Doi.Org/10.1287/Isre.2017.0701, 28(3), 563–584.

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. https://doi.org/10.1016/J.RESPOL.2023.104906

Jarvenpaa, S. L., & Essén, A. (2023). Data sustainability: Data governance in data infrastructures across technological and human generations. Information and Organization, 33(1), 100449. https://doi.org/10.1016/J.INFOANDORG.2023.100449

Lee, S. U., Zhu, L., & Jeffery, R. (2018). Designing Data Governance in Platform Ecosystems. Proceedings of the Annual Hawaii International Conference on System Sciences, 2018-January, 5014–5023. https://doi.org/10.24251/HICSS.2018.626

Lin, D., Crabtree, J., Dillo, I., Downs, R. R., Edmunds, R., Giaretta, D., De Giusti, M., L'Hours, H., Hugo, W., Jenkyns, R., Khodiyar, V., Martone, M. E., Mokrane, M., Navale, V., Petters, J., Sierman, B., Sokolova, D. V., Stockhause, M., & Westbrook, J. (2020). The TRUST Principles for digital repositories. In Scientific Data (Vol. 7, Issue 1). Nature Research. https://doi.org/10.1038/s41597-020-0486-7

Lindgren, R., Mathiassen, L., & Schultze, U. (2021). The dialectics of technology standardization. MIS Quarterly: Management Information Systems, 45(3), 1187–1212. https://doi.org/10.25300/MISQ/2021/12860

Lingens, B., Miehé, L., & Gassmann, O. (2021). The ecosystem blueprint: How firms shape the design of an ecosystem according to the surrounding conditions. Long Range Planning, 54(2), 102043. https://doi.org/10.1016/J.LRP.2020.102043

Lioliou, E., Zimmermann, A., Willcocks, L., & Gao, L. (2014). Formal and relational governance in IT outsourcing: Substitution, complementarity and the role of the psychological contract. Information Systems Journal, 24(6), 503–535. https://doi.org/10.1111/isj.12038

Liu, B., Shao, Y. F., Liu, G., & Ni, D. (2022). An Evolutionary Analysis of Relational Governance in an Innovation Ecosystem. SAGE Open, 12(2). https://doi.org/10.1177/21582440221093044/ASSET/IMAGES/LARGE/10.1177_21582440221093044-FIG12.JPEG

Lnenicka, M., Nikiforova, A., Luterek, M., Milic, P., Rudmark, D., Neumaier, S., Kević, K., Zuiderwijk, A., & Rodríguez Bolívar, M. P. (2024). Understanding the development of public data ecosystems: From a conceptual model to a six-generation model of the evolution of public data ecosystems. Telematics and Informatics, 94, 102190. https://doi.org/10.1016/j.tele.2024.102190

Lumineau, F. (2014). How Contracts Influence Trust and Distrust. Https://Doi.Org/10.1177/0149206314556656, 43(5), 1553–1577.

Lumineau, F., Long, C., Sitkin, S. B., Argyres, N., & Markman, G. (2023). Rethinking Control and Trust Dynamics in and between Organizations. Journal of Management Studies, 60(8), 1937–1961. https://doi.org/10.1111/JOMS.12999

Lumineau, F., Schilke, O., & Wang, W. (2022). Organizational Trust in the Age of the Fourth Industrial Revolution: Shifts in the Form, Production, and Targets of Trust. Https://Doi.Org/10.1177/10564926221127852, 32(1), 21–34.

Manning, S., & Reinecke, J. (2016). A modular governance architecture in-the-making: How transnational standard-setters govern sustainability transitions. Research Policy, 45(3), 618–633. https://www.sciencedirect.com/science/article/pii/S0048733315001821

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model Of Organizational Trust. Academy of Management Review, 20(3), 709–734. https://doi.org/10.5465/AMR.1995.9508080335

Mentel, U., & Hajduk-Stelmachowicz, M. (2020). Does standardization have an impact on innovation activity in different countries? Problems and Perspectives in Management, 18(4), 486–503. https://doi.org/10.21511/ppm.18(4).2020.39

Micheli, M., Ponti, M., Craglia, M., & Berti Suman, A. (2020). Emerging models of data governance in the age of datafication. Big Data and Society, 7(2). https://doi.org/10.1177/2053951720948087

Moreno, J., Fernandez, E. B., Serrano, M. A., & Fernandez-Medina, E. (2019). Secure Development of Big Data Ecosystems. IEEE Access, 7, 96604–96619. https://doi.org/10.1109/ACCESS.2019.2929330

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.

Pessanha Santos, N. (2023). The Expansion of Data Science: Dataset Standardization. Standards, 3(4), 400–410. https://doi.org/10.3390/standards3040028

Rammer, C., & Es-Sadki, N. (2023). Using big data for generating firm-level innovation indicators-a literature review. Technological Forecasting and Social Change, 197, 122874.

Reinecke, J., Manning, S., & von Hagen, O. (2012). The Emergence of a Standards Market: Multiplicity of Sustainability Standards in the Global Coffee Industry. Organization Studies, 33(5–6), 791–814. https://doi.org/10.1177/0170840612443629

Rong, K., & Luo, Y. (2023). Toward born sharing: The sharing economy evolution enabled by the digital ecosystems. Technological Forecasting and Social Change, 196, 122776.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. Academy of Management Review, 23(3), 393–404. https://doi.org/10.5465/AMR.1998.926617

Rudmark, D., Lindgren, R., & Schultze, U. (2024). Open data platforms: Design principles for embracing outlaw innovators. The Journal of Strategic Information Systems, 33(3), 101850. https://doi.org/10.1016/J.JSIS.2024.101850

Sandholtz, K. W. (2012). Making Standards Stick: A Theory of Coupled vs. Decoupled Compliance. Organization Studies, 33(5–6), 655–679. https://doi.org/10.1177/0170840612443623

Shin, D. H., & Choi, M. J. (2015). Ecological views of big data: Perspectives and issues. Telematics and Informatics, 32(2), 311–320. https://doi.org/10.1016/J.TELE.2014.09.006

Shipilov, A., & Gawer, A. (2020). Integrating research on interorganizational networks and ecosystems. Academy of Management Annals, 14(1), 92–121. https://doi.org/10.5465/ANNALS.2018.0121

Strauss, A., & Corbin, J. (1998). Basics of qualitative research techniques. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=18c7cfe7a46c7771b60dc384b1b4e350f65b13e2

Strauss, A. L., & Corbin, J. M. (1997). Grounded Theory in Practice. Sage. https://books.google.it/books?hl=it&lr=&id=TtRMolAapBYC&oi=fnd&pg=PP9&dq=Strauss,+A.,+%26+Corbin,+J.+M.+(1997).+Grounded+theory+in+practice.+Sage.+&ots=DBhoexpDWk&sig=oMm_gdGbQKKNboV2guS1-E_SiSU&redir_esc=y#v=onepage&q=Strauss%2C%20A.%2C%20%26%20Corbin%2C%20J.%20M.%20(1997).%20Grounded%20theory%20in%20practice.%20Sage.&f=false

Swärd, A. (2016). Trust, Reciprocity, and Actions: The Development of Trust in Temporary Inter-organizational Relations. Organization Studies, 37(12), 1841–1860. https://doi.org/10.1177/0170840616655488

Teece, D. J. (2018). Profiting from innovation in the digital economy: Enabling technologies, standards, and licensing models in the wireless world. Research Policy, 47 (8), 1367–1387. https://www.sciencedirect.com/science/article/pii/S0048733318300763

Timmermans, S., & Epstein, S. (2010). A world of standards but not a standard world: Toward a sociology of standards and standardization. Annual Review of Sociology, 36(Volume 36, 2010), 69–89. https://doi.org/10.1146/ANNUREV.SOC.012809.102629/CITE/REFWORKS

Tools, M., Saeed, N., Malik, H., Naeem, A., & Bashir, U. (2023). Incorporating big data and IoT in intelligent ecosystems: state-of-the-arts, challenges and opportunities, and future directions. Multimedia Tools and Applications 2023, 1–43. https://doi.org/10.1007/S11042-023-16328-3

Van De Hoven, J., Comandè, G., Ruggieri, S., Domingo-Ferrer, J., Musiani, F., Giannotti, F., Pratesi, F., & Stauch, M. (2021). Towards a Digital Ecosystem of Trust: Ethical, Legal and Societal Implications. In Opinio Juris In Comparatione. https://shs.hal.science/halshs-03510761

van Tulder, R., & van Mil, E. (2022). PRINCIPLES OF SUSTAINABLE BUSINESS: Frameworks for Corporate Action on the SDGs. Principles of Sustainable Business: Frameworks for Corporate Action on the Sdgs, 1–1019. https://doi.org/10.4324/9781003098355/PRINCIPLES-SUSTAINABLE-BUSINESS-ROB-VAN-TULDER-EVELINE-VAN-MIL/ACCESSIBILITY-INFORMATION

Ven, V. de., Andrew, H., & Ring, P. S. (2006). Relying on trust in cooperative inter-organizational relationships. Handbook of Trust Research, 144–164. https://www.elgaronline.com/downloadpdf/edcollchap/9781843767541.00015.pdf

Wang, Z., Zheng, Z., Jiang, W., & Tang, S. (2021). Blockchain-Enabled Data Sharing in Supply Chains: Model, Operationalization, and Tutorial. Production and Operations Management, 30(7), 1965–1985. https://doi.org/10.1111/POMS.13356

Wareham, J., Fox, P. B., & Giner, J. L. C. (2014). Technology ecosystem governance. Organization Science, 25(4), 1195–1215. https://doi.org/10.1287/orsc.2014.0895

Woolthuis, R. K., Hillebrand, B., & Nooteboom, B. (2005). Trust, contract and relationship development. Organization studies, 26(6), 813-840. https://doi.org/10.1177/0170840605054594

Zang, C. S., Buras, A., Esquivel-Muelbert, A., Jump, A. S., Rigling, A., & Rammig, A. (2020). Standardized drought indices in ecological research: Why one size does not fit all. In Global Change Biology (Vol. 26, Issue 2, pp. 322–324). Blackwell Publishing Ltd. https://doi.org/10.1111/gcb.14809

Acknowledgments

This work studies MobiDataLab which has received funding from the European Union under the Horizon 2020 Research and Innovation program (Grant Agreement N° 101006879). The content reflects only the authors' views, and the European Commission is not responsible for any use that may be made of the information it contains.