Beyond control over data

Conceptualizing data sovereignty from a social contract perspective

Abbas, Antragama Ewa; van Velzen, Thomas; Ofe, Hosea; van de Kaa, Geerten; Zuiderwijk-van Eijk, A.M.G.; de Reuver, Mark

DOI
10.1007/s12525-024-00695-2

Publication date
2024

Document Version
Final published version

Published in
Electronic Markets

Citation (APA)
Abbas, A. E., van Velzen, T., Ofe, H., van de Kaa, G., Zuiderwijk-van Eijk, A. M. G., & de Reuver, M. (2024). Beyond control over data: Conceptualizing data sovereignty from a social contract perspective. *Electronic Markets*, *34*(1), 1-21. Article 20. https://doi.org/10.1007/s12525-024-00695-2

Electronic Markets (2024) 34:20
https://doi.org/10.1007/s12525-024-00695-2

RESEARCH PAPER

Beyond control over data: Conceptualizing data sovereignty from a social contract perspective

Antragama Ewa Abbas[1] · Thomas van Velzen[1] · Hosea Ofe[2] · Geerten van de Kaa[1] · Anneke Zuiderwijk[1] · Mark de Reuver[1]

Received: 1 April 2023 / Accepted: 16 January 2024

Abstract

In the data economy, data sovereignty is often conceptualized as data providers' ability to control their shared data. While control is essential, the current literature overlooks how this facet interrelates with other sovereignty facets and contextual conditions. Drawing from social contract theory and insights from 31 expert interviews, we propose a data sovereignty conceptual framework encompassing protection, participation, and provision facets. The protection facets establish data sharing foundations by emphasizing baseline rights, such as *data ownership*. Building on this foundation, the participation facet, through *responsibility divisions*, steers the provision facets. Provision comprises facets such as *control*, *security*, and *compliance mechanisms*, thus ensuring that foundational rights are preserved during and after data sharing. Contextual conditions (data type, organizational size, and business data sharing setting) determine the level of difficulty in realizing sovereignty facets. For instance, if personal data is shared, *privacy* becomes a relevant protection facet, leading to challenges of ownership between data providers and data subjects, compliance demands, and control enforcement. Our novel conceptualization paves the way for coherent and comprehensive theory development concerning data sovereignty as a complex, multi-faceted construct.

Keywords Data economy · Data sharing · Data sovereignty · Data marketplaces · Meta-platforms

JEL Classification L86

Introduction

In today's digital era, numerous technologies generate vast amounts of data (Glennon et al., 2023). Yet, data itself has limited inherent value; its true significance emerges when transformed into highly contextualized insights to address business challenges (Aaltonen et al., 2021). However, data largely remain underutilized within many businesses (Gantz & Reinsel, 2012; Manyika et al., 2015). Unlocking data potential requires participation in the data economy, a global ecosystem driven by collecting, processing, and sharing data for economic and societal gains (Sestino et al., 2023). In the data economy, businesses increasingly share data with external parties (Richter & Slowinski, 2019). Data sharing is a process, irrespective of domains, where businesses (i.e., data providers) offer other businesses (i.e., data consumers) access to their data. Data consumers utilize these data to develop new applications and services. In return, data providers expect rewards, either monetary or other incentives, such as reciprocal data sharing (Jussen et al., 2023). For

Responsible Editor: Markus Bick

✉ Antragama Ewa Abbas
A.E.Abbas@tudelft.nl

Thomas van Velzen
Thomas.V.Velzen@gmail.com

Hosea Ofe
Hosea.Ofe@hh.se

Geerten van de Kaa
G.vandeKaa@tudelft.nl

Anneke Zuiderwijk
A.M.G.Zuiderwijk-vanEijk@tudelft.nl

Mark de Reuver
G.A.deReuver@tudelft.nl

[1] Faculty of Technology, Policy and Management, Delft University of Technology, Delft, The Netherlands

[2] School of Information Technology, Department of Intelligent Systems and Digital Design, Halmstad University, Halmstad, Sweden

Published online: 02 March 2024

instance, telecommunication operators could share aggregated call detail records with banks to help them develop more accurate models for assessing creditworthiness (e.g., Óskarsdóttir et al., 2019).

Data sovereignty is an essential requirement for data sharing (e.g., Scheider et al., 2023; Schweihoff et al., 2023). Existing literature interprets data sovereignty as the ability of data providers to control shared data (Hellmeier & von Scherenberg, 2023). While societal groups, from individuals to communities, might focus on personal autonomy or collective histories in understanding data sovereignty, this study delves into its implications for businesses. A lack of sovereignty harms the willingness to share data (Opriel et al., 2021), thereby hindering the overall growth of the data economy.

The predominant perspective in literature takes a control-centric view on data sovereignty, assuming that sole control over data is sufficient to attain sovereignty (see a review by Hellmeier & von Scherenberg, 2023). However, data sovereignty also relates to concepts beyond control, such as ownership and security (Hummel et al., 2021). The prevalent view that equates sovereignty to control might lead to developing partial solutions. For example, existing technical solutions often focus on access control to specify data access rights (e.g., Munoz-Arcentales et al., 2019) or data provenance to analyze the origin and history of the shared data (e.g., Olufowobi et al., 2017). While these solutions address control, they leave other facets of sovereignty unaddressed. Further, when other facets of sovereignty remain unobserved, the risk emerges that studies will produce contradictory findings due to unobserved variable bias. Preventing conceptual reductionism (i.e., an oversimplification of a complex phenomenon) is especially timely as the interest in data sovereignty is rapidly rising.

To move beyond reductionistic views of data sovereignty as sole control over data, we first must uncover and delineate additional key facets. This step establishes descriptive knowledge (Gregor, 2006), addressing the *what* aspect of data sovereignty. Subsequently, we must explore the *how* and *why* aspects of data sovereignty: the *how* involves relationships between facets, while the *why* uncovers causal mechanisms that explain these relationships (Dubin, 1978). Finally, we need to consider contextual conditions that challenge the realization of data sovereignty facets, such as business data sharing settings (e.g., van den Broek & van Veenstra, 2018). Despite these required steps, the data sharing literature seldom addresses the *what*, *how*, and *why* aspects of data sovereignty, nor its contextual conditions.

Overall, the neglect of treating data sovereignty as a multi-faceted and contextualized construct implies that a conceptual basis is missing for higher levels of theory development (i.e., explanatory and prediction theory) (see Gregor, 2006 about levels of theory development). Investigating facets and contextual conditions could further clarify inconsistencies in studies on data sovereignty, for instance, why control is sometimes considered less important (e.g., Shah et al., 2019) or crucial (e.g., van den Broek & van Veenstra, 2015).

This paper aims to develop a multi-faceted conceptual framework of data sovereignty. Specifically, we ask:

RQ1: What are the key facets of data sovereignty in data sharing by businesses?
RQ2: How are the key facets of data sovereignty interrelated?
RQ3: How do contextual conditions influence the difficulties of realizing the key facets of data sovereignty?

To develop a coherent conceptual framework, we build on Social Contract Theory (SCT), which addresses sovereignty within social systems. We use SCT to understand what sovereignty is and how it can be achieved. We interview 31 experts experienced in conducting or building solutions for data sovereignty to specify the substantive aspect of SCT in the data sharing context. To increase the chances of finding meaningful insights, we focus on a setting in which sovereignty is ultimately challenged: a meta-platform that interconnects heterogeneous data marketplaces, which may each have their own ways of safeguarding data provider interests.

This study primarily contributes to the Information Systems literature on data sharing by proposing a multi-faceted conceptual framework of data sovereignty, paving the way for further theory development. As data sovereignty is receiving increased scholarly attention, it is vital to avoid reductionistic views that overlook key facets of the concept. With our alternative conceptualization, researchers and practitioners can identify and evaluate alternative solutions for data sovereignty, which go beyond mere control measures. More specifically, our contributions are three-fold. First, we provide evidence that, next to control, data sovereignty has key facets of privacy, ownership, security, compliance, and responsibility. These key facets inform scholars on what to consider in their conceptualizations of data sovereignty. Second, we identify relationships between these key facets, which serve as foundations for explanatory theory on data sovereignty. Finally, we find contextual conditions for specifying boundary conditions that affect the difficulty in realizing sovereignty facets, which help researchers to identify and prioritize key sovereignty facets relevant to specific contexts, enhancing the applicability of our framework. Our findings provide insights for policymakers, highlighting that focusing excessively on one aspect of sovereignty can lead to unintended consequences in others, thereby potentially compromising the effectiveness of data economy-related policies.

Research background

The research aims to create a multi-faceted conceptual framework of data sovereignty, challenging the prevalent assumption that equates sovereignty solely with control over data. To achieve this, we need to delve into the key facets of data sovereignty, examine the interrelationships and causal mechanisms among these facets, and identify contextual conditions challenging the realization of these facets. However, the data sharing literature rarely discusses these topics. To explore these topics, selecting an appropriate theory as an analytical tool is essential. This requires revisiting the foundational concept of sovereignty in the political science field.

Sovereignty has historically been understood as the ultimate governing power over a political body (Hinsley, 1986). Bodin (1576) relates sovereignty closely to divine-fated monarchic rule for a country. However, the Enlightenment era of the seventeenth and eighteenth centuries shifted this understanding. This era, driven by empirical investigation and rational thought, led to new theories about the relationship between a country and its citizens, most notably social contract theory (SCT). This theory is central to our study and elaborated in "Section Social contract theory."

In the digital era, discussions on sovereignty have been refocused to encompass individuals, communities, and organizations' control over their data, primarily due to the rise of cloud computing (De Filippi & McCarthy, 2012) and the Snowden revelations on state-approved surveillance (Lyon, 2014). More recently, data sovereignty has become central to European Union measures to unlock the full potential of the data economy (e.g., see European Strategy for Data). This led to initiatives like the International Data Spaces Association and Gaia-X, which emphasize data sovereignty.

The previous discussion describes the historical context of sovereignty. The following sections delve into "Section Social contract theory," which examines SCT and provides the study's theoretical foundation, and "Section Contextualizing social contract theory to data sovereignty," which contextualizes SCT to data sovereignty.

Social contract theory

The Social Contract Theory (SCT) can be used to examine sovereignty, as it posits that individuals relinquish some freedoms to a governing entity for societal benefits (Friend, 2004). In the context of data sovereignty, this implies that data providers agree to certain compromises to gain benefits. For instance, they may adhere to predefined data sharing protocols. In return, they receive benefits offered by platform operators, such as the ability to control their shared data. Such notions of trade-offs between freedoms and benefits align with the principles of SCT. Hence, we consider SCT to be suitable for exploring data sovereignty facets and contextual conditions in our study.

While alternative theories, such as Transaction Cost Economics (TCE) and Social Exchange Theory (SET), may be used to explore data sovereignty, SCT is more appropriate for our purposes. TCE examines transactions from a cost perspective, aiming to identify organizational structures to minimize these costs (Williamson, 1989; Young, 2013). While TCE focuses on economic costs, SET provides a more comprehensive perspective by considering trade-offs between costs and benefits.

SET focuses on social interactions within exchange processes (Cook, 2015), asserting that if benefits outweigh costs, involved actors will likely engage in exchanges (Homans, 1958). SET may thus be used to consider the reasons why data providers would make their data available for transactions. In contrast, SCT takes a broader perspective, including transactional actors (e.g., sellers and buyers) and sovereign entities (e.g., countries) overseeing these transactions at a macro level. Such macro-level is essential to understanding data sovereignty, as data sharing is safeguarded both at the level of the transactional actors and through governmental oversight. SCT thus considers sovereignty at the levels of both individual and state-level actors.

Within SCT, diverging perspectives exist regarding the assumptions to be taken on how individuals and countries establish social contracts, a crucial element in interpreting sovereignty in data sharing. Therefore, exploring SCT aspects requires analyzing the perspectives of Hobbes (1651), Locke (1689), and Rousseau (1762) on individual behavior assumptions.

Hobbes (1651) takes a pessimistic view of human nature, suggesting that human life is "solitary, poor, nasty, brutish, and short" without regulation (Ch. XIII). In Hobbes' view, individuals should surrender their freedom entirely to a sovereign, usually a monarch, to prevent societal chaos driven by inherent self-interest. This sovereign entity, in return, would provide peace and security. For Hobbes, this powerful central rule is indispensable for stability. While Bodin (1576) believes monarchial power was divinely determined, Hobbes argues that it arises from human self-interest to prevent societal anarchy.

Locke (1689), on the other hand, has a more optimistic view of human nature. He sees humans as naturally social and rational, respecting others' rights. However, Locke believes that the lack of an unbiased authority to settle disputes in a pre-governed state necessitates forming a government. His vision of the social contract does not involve individuals giving up all their rights but only some. This

government is obligated to serve the individuals' best interests. Critically, Locke emphasizes that governmental power is not absolute but derived from the people's consent. People hold the right to reelect the government if it fails in its responsibilities. Locke's view represents representative democracy, wherein elected representatives act in the populace's best interest.

Rousseau (1762), while sharing some views with Locke, took a different stance on the essence of the social contract. For him, humans in their natural state were peaceful and lived solitary lives, free from the corruption and evils of society. However, with the introduction of private property, humans agreed to a social contract to protect their property rights, consequently forming a government. Rousseau emphasized the notion of *general will*, the collective will of the individuals, which should guide a country. For Rousseau, sovereignty lies with the people, and the government merely administers the people's will. Rousseau's theory resonates strongly with the principles of direct democracy, where citizens have a direct say in their government's decisions. In this paper, the complementary views of Hobbes, Locke, and Rousseau on SCT will serve as a lens to interpret data sovereignty.

To apply Social Contract Theory (SCT) as an analytical tool, we must consider its key aspects: spatial, temporal, and substantive (Loewe et al., 2021). The *spatial* aspect of SCT specifies *who* participates and *where* their influence applies in a societal contract. The *who* includes varied parties like governments, societal groups, and individuals. The *where* implies the territorial extent of the agreement, which could span sub-national, national, transnational, or supranational levels (Loewe et al., 2021).

The *temporal* aspect, concerning the *when*, explores the dynamic of social contracts over time. Social contracts can differ significantly in their duration and the timing of their renegotiations. While social contracts aim to bring stability to state-society relations, they often require renegotiation and adaptation due to changes in power distribution or the perceived failure of countries to meet their obligations (Loewe et al., 2021).

The *substantive* aspect of SCT describes vertical arrangements between a nation and societal groups. These are known as the three Ps: protection, provision, and participation. The three Ps explain the *what* of social contracts (Loewe et al., 2021). Protection focuses on recognizing and acknowledging inherent rights that need safeguards (Ellis, 2006; Hickey, 2011). Provision encompasses the various services and resources a country provides to society, including healthcare, education, and infrastructure (e.g., Sobhy, 2021). Participation involves citizens actively engaging in public affairs and interacting with government processes (Loewe et al., 2021). The following section discusses how this study applies SCT in the context of data sovereignty.

Contextualizing social contract theory to data sovereignty

For contextualizing Social Contract Theory (SCT) to data sovereignty, we must define the spatial, temporal, and substantive aspects that shape social contracts. To do so, understanding the setting in which data sharing occurs is essential. Business data sharing operates in three primary modes: hierarchy, network, and market (van den Broek & van Veenstra, 2015). In the hierarchy mode (e.g., supply chains), focal partners orchestrate data sharing through formalized, centralized control. The network mode is characterized by lateral relationships between data ecosystem members, emphasizing social agreements and collaborative approaches (Otto & Jarke, 2019). The market mode has recently gained traction, where data is shared as a commercial product through formal contracts via data marketplaces (Spiekermann, 2019). Such marketplaces are a subtype of digital platforms that create value by connecting data providers with consumers, facilitating smooth data sharing, and maintaining a modular infrastructure for third-party providers to add additional offerings and services (Abbas et al., 2021; Fruhwirth et al., 2020; Spiekermann, 2019). We focus on the market mode, which is the most complicated setting for sovereignty issues. Within this setting, precisely defining data sharing terms and conditions is difficult (Virkar et al., 2019), increasing the risk of misinterpreting data ownership and usage rights.

Within the market mode, one specifically challenging setting is when meta-platforms federate and interconnect heterogeneous data marketplaces. A meta-platform is a platform designed to operate atop two or more existing platforms, connecting their respective ecosystems (Floetgen et al., 2021; Zhang & Williamson, 2021). They do so by providing (1) technology architecture and integration standards for interoperability and (2) offering a central hub to connect platform actors (Rossmannek & Chen, 2023). For data marketplaces specifically, the meta-platform setting (e.g., i-3 Market[1]) presents a highly complicated setting for examining data sovereignty because it interconnects multiple marketplaces, each with its distinct spatial territory. Therefore, the challenges of realizing and aligning the three Ps are enormous, making it a worthwhile context to study.

Having selected the meta-platform for data marketplace as a business data sharing setting, we can now contextualize SCT to data sovereignty by considering its spatial, temporal, and substantive aspects (refer to Table 1).

Regarding the spatial aspect of SCT, we need to identify the key actors and determine the territorial settlement of social contracts. The main actors participating in business

[1] https://www.i3-market.eu/, accessed on November 26, 2023.

Table 1 The mapping between social contract theory and data sovereignty in this study

Social contract theory		Data sovereignty
Spatial aspect	Actors (who)	(Meta-platform and data marketplace) operators, data providers, data subjects, and data consumers
	Territorial settlement (where)	Meta-platform ecosystems, including data marketplaces they federate
Temporal aspect	Duration and timing (when)	Focusing on post-2018 as this time shows significant regulatory changes in the data economy
Substantive aspect	Vertical arrangements (the three Ps: protection, provision, and participation)	*Underexplored*

data sharing via meta-platforms include sovereign entities, such as (meta-platform and data marketplace) operators, and societal groups, such as data providers, data subjects, and data consumers (cf. Azcoitia & Laoutaris, 2022). In this context, operators handle platforms as business ventures. Data providers are companies that offer data in meta-platforms. When these data include personal data, data subjects—identifiable natural persons from whom the data originates—emerge as another important actor. Data consumers leverage these shared data for several tasks, like analytics and strategic planning. Regarding territorial settlement, the social contract of data sovereignty is applicable within the scope of the meta-platform ecosystem, covering the data marketplaces they federate.

Considering the temporal aspect of SCT, we direct attention to the period post-2018. Within the European Union, which is the sociopolitical area for our study, this time signifies a crucial regulatory evolution in the data economy. Notable legislation arose, such as the Data Governance Act, setting a vision for a Single European Data Market. Concurrently, these rules emphasized the importance of data sovereignty. This period is, thus, appropriate for studying the effects of regulatory influences on social contracts within the meta-platform territory.

The substantive aspect of SCT is represented by vertical arrangements, the three Ps (protection, provision, and participation). However, the existing understanding of the three Ps, initially explored within the context of the relationship between citizens and countries, cannot be directly transferred to the context of data sovereignty in business data sharing. This is due to the fundamentally different characteristics of these contexts. In the former case, the relationship revolves around well-established societal structures and tangible resources; in the latter case, data sovereignty generally resides within the more abstract and fluid scope of data sharing. Hence, it is unclear how the three Ps manifest in data sovereignty.

To explore the three Ps, we utilize notions that correlate with data sovereignty provided by Hummel et al. (2021). As a starting point, we focus on the most potentially suitable notions in business data sharing: *control, ownership, privacy*, *security*, and *responsibility* (Hummel et al., 2021). In addition, we investigate *compliance* as a facet, given its recent legal prominence in contexts such as the European Data Governance Act (Duisberg, 2022). Here, control refers to the capability to influence and direct information flows. Ownership refers to data property rights, indicating the privileges over data resources. Privacy encapsulates the protection of personal data. Security, on the other hand, focuses on preventing potential threats and risk mitigation concerning data. Additionally, responsibility delineates roles and expectations, while compliance represents the adherence to relevant legal and regulatory frameworks.

In summary, Fig. 1 illustrates how this study contextualizes SCT to data sovereignty, highlighting the three Ps as the primary focus of the empirical investigation. This figure is adapted from Furness and Trautner (2020), who apply SCT to the relationship between societal groups and countries. We tailor this figure by specifying SCT aspects to the data sovereignty context, as shown in Table 1. This figure is used as a frame of reference in our subsequent analysis.

Research approach

Given our focus on empirically exploring data sovereignty with the Social Contract Theory (SCT), we employed an exploratory qualitative approach. This approach excels in studies that need contextualization and interpretation (Glesne, 2016). We gathered data through semi-structured interviews, which allowed for a combination of structured queries and the ability to adapt the interviews according to responses (Edwards & Holland, 2013). We required a flexible approach as the application of the substantive aspect of SCT to the data sovereignty context is not yet clear, allowing us to delve into follow-up inquiries.

Participant selection criteria

We utilized the *purposive sampling strategy*, specifically *judgment sampling*, to select our interview participants by considering their expertise (Sekaran & Bougie, 2016).

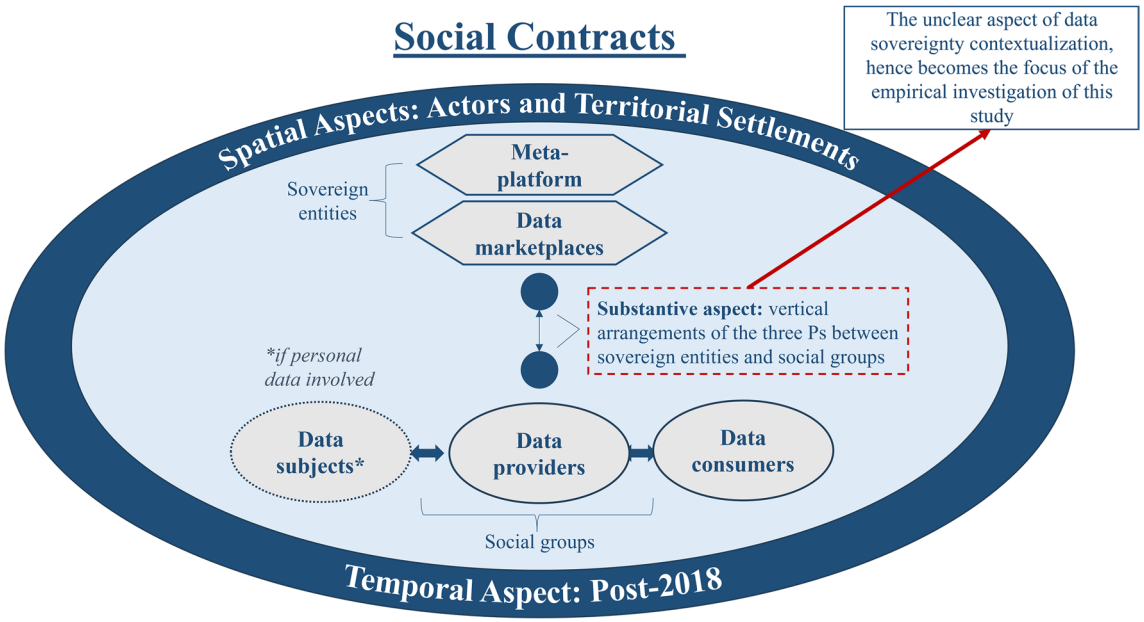

Fig. 1 Contextualizing the social contract theory to data sovereignty (adapted from Furness & Trautner, 2020)

This sampling strategy was appropriate as we investigated a novel phenomenon that only a few individuals were knowledgeable about; therefore, we selected individuals who were ideally situated and most capable of offering the necessary insights (Etikan et al., 2016). We focused on experts who had experience with and had a pivotal role in conducting or building solutions for sovereign data sharing in different settings (e.g., data marketplaces, data ecosystems). Furthermore, the interview participants should possess experience in decision-making processes related to sovereign data sharing, as this reflects their expertise. Proficiency in English was also a key selection criterion. We verified participants' expertise by looking at their public professional profiles (e.g., via LinkedIn) and asked them to share their experience for sovereign data sharing during the interview. Leveraging our networks in EU data sharing projects, we identified potential interviewees and stopped once code saturation was achieved. From July 2021 to June 2022, we conducted 31 online interviews via Microsoft Teams, averaging 45 min each.

Appendix 1 lists our interviewees, who predominantly have mid to senior management roles, with an average work experience of 15 years. Our sample contains a varied and balanced range of expertise, including strategic roles (e.g., director of innovation, chief data officer, commercial director), technical positions (e.g., technical researcher, information technology architect), project management (e.g., project managers), and legal roles (e.g., data protection specialist, risk manager). Most of our interviewees are professionals from either the telecommunication or financial sectors.

Interview protocol

We developed an initial interview protocol focusing on three key concepts central to our study: data marketplaces, meta-platforms, and data sovereignty. In this interview, we took the perspective of data providers as the core problem owners of data sovereignty. After piloting with two interviews, we found that while participants understood data marketplaces and sovereignty, they needed a more precise explanation of meta-platforms. Thus, we improved our explanation in the final protocol.

The final protocol consisted of three parts. First, we asked about the interviewee's background and knowledge of data marketplaces. Second, we presented the concept of meta-platforms for data marketplaces, specifically with a use case where data providers are not associated with any marketplaces and then join a meta-platform to share their data. In doing so, providers can reach consumers from many participating marketplaces. Participants were allowed to ask for clarifications after the brief presentation. We then asked about potential value propositions related to a meta-platform for data marketplaces. These questions ensured that interviewees had an authentic perspective while answering questions about data sovereignty and helped to ensure that the understanding of meta-platforms was in line with the assumptions in our study. We also inquired about the potential drawbacks of meta-platforms. Often, the interviewees had, at this point, already discussed some aspects related to data sovereignty. Next, we explicitly questioned the participants about data sovereignty concerns in meta-platforms. We did this to encourage unrestrained perspectives on

Table 2 The illustration of the coding schema

No	Excerpt	Applied coding schema
1	"So, to me, data sovereignty is being in control of your data as much as it is over your meta-data. And that you just have non-repudiation, traceability, that sort of thing. So, what I am actually saying with that is traceability, in order to be able to control it [(meta)-data] at all, you first have to know where it is. You have to have insight into that to be able to enforce anything at all."	■ First-order: data flow tracking ■ Second-order: data provenance ■ Facet: data control ■ Higher-level facet: provision
2	"There is a lot of discussion around tagging data contents [to enhance data sovereignty] so you know who the ultimate owner is. And I think that is the answer. So, at the end of the day, you can review where the content comes from. You can say this comes from Data Provider A."	■ First-order: data origin information ■ Second-order: data provenance ■ Facet: data control ■ Higher-level facet: provision
3	"I think that [providing technical enforcements for data control] is the main part where the industry has been struggling in the ideal world: You can share data, You have some control over what is done with the data, You can revoke the rights to use the data at any time."	■ First-order: data access revocation ■ Second-order: data removal ■ Facet: data control ■ Higher-level facet: provision
4	"Data sovereignty is basically the ability of the data provider, for instance, to withdraw the datasets from a particular marketplace at the time of their choosing. That is one example of exercising data sovereignty. My dataset is being traded somewhere. If I no longer want to do it, I can remove the dataset."	■ First-order: dataset retraction ■ Second-order: data removal ■ Facet: data control ■ Higher-level facet: provision

data sovereignty. This approach ensured that participants shared their insights without feeling directed by predefined expected results. We often asked follow-up questions to get detailed elaboration. All interviews, with the participant's consent, were recorded, transcribed anonymously, and sent to interviewees for validation. Five transcriptions underwent minor revisions post-review.

Data analysis

We analyzed each transcript using Atlas.TI 22.4. We chose this software for its superior data visualization capabilities and efficient quotation system, allowing richer data interaction and more intuitive coding than alternatives like NVivo and MAXQDA.[2] The coding process employed *a structured categorization matrix* (Elo & Kyngäs, 2008), incorporating pre-defined categories based on a theoretical understanding of the substantive aspect of Social Contract Theory (SCT). This matrix comprises four levels: higher-level facets, facets, second-order codes, and first-order codes. The codebook of the structured categorization matrix can be found in Appendix 2.

We conducted three rounds of coding. In the first round, we inductively coded a relevant block of statements into first-order code by interpreting what the participants said when discussing data sovereignty. We then revised or merged the first-order codes into second-order codes in the second coding round. These second-order codes represented broader, abstract code groupings that emerged from the data. In the third round, we further grouped the second-order code into relevant data sovereignty facets. Furthermore, we interpreted how these identified facets correlate with the higher-level facets of SCT: the three Ps (protection, provision, and participation). To increase internal validity, the first and the second authors performed an inter-coder reliability assessment to check the consistency of the codes, which the third author then reviewed.

To illustrate the coding process, consider the relation between excerpts and the applied coding schema in Table 2. For the first example shown in the table, we assigned the first excerpt example to the first-order code of *data flow tracking* and the second one to *data origin information*. Furthermore, the third example was coded as *data access revocation*, while the fourth was coded as *dataset retraction.* In the second coding round, *data flow tracking* and *data origin information* were categorized into the second-order code of *data provenance* because both were fundamentally discussing the origins and pathways of data, emphasizing the importance of knowing where the data comes from and how it moves that represent the idea of data provenance. The first-order codes of *data access revocation* and *dataset retraction* were classified under the second-order code of *data removal* because both emphasize the deliberate actions taken to limit or end access to specific data sets. The third coding round further categorized the second-order code of *data provenance* and *data removal* into the sovereignty facet of *data control* because they represent an overarching theme approach to ensuring that data remains under the intended authority and purpose throughout its lifecycle. Afterward, we mapped *data control* as part of the provision's higher-level facet. The theoretical underpinning for this mapping is the expectation that data providers desire data control capabilities as a part of their vertical arrangement with meta-platform operators. In return, operators are expected to provision

[2] https://atlasti.com/research-hub/atlas-ti-alternative-to-other-programs, accessed on November 26, 2023.

such control mechanisms. Data control was not categorized under protection since protection primarily focuses on recognizing rights, a concept distinct from active provisioning. Likewise, data control was not associated with participation as its primary objective is not directly fostering engagement from data providers. The online appendix offers additional examples of how we connected excerpts and codes.

Findings

This section presents the research findings. "Section The substantive aspect of data sovereignty: protection, provision, and participation" delves into the substantive aspect of data sovereignty to address RQ1 and RQ2, highlighting the three Ps: protection, provision, and participation. Facets and their interrelationships are emphasized using *italics* and **bold**, respectively. "Section The spatial aspect of data sovereignty: Contextual conditions" examines the spatial aspect to address RQ3, exploring how contextual conditions impact the difficulty in realizing data sovereignty facets. "Section Data sovereignty conceptual framework: Interactions between data sovereignty facets and contextual conditions" presents a conceptual framework that captures the interactions between data sovereignty facets. Participants are referenced using the identifier (I-X) from Appendix 1.

The substantive aspect of data sovereignty: protection, provision, and participation

Protection of data ownership and privacy

Data ownership and *privacy* emerge as critical facets of data sovereignty, which can be interpreted as part of protection. Considering data ownership, one interviewee (I-21) explicitly stated its correlation with data sovereignty: "I think data sovereignty means control over data ownership." Data ownership is associated with possession (I-28), meaning data providers can retain intellectual property rights for their data products (I-01). One participant stated, "… data ownership should always remain with the provider, and that should be clear through whatever kind of licensing they do" (I-25). However, claiming ownership is not straightforward, with some participants questioning whether data ownership can be as transferrable as physical products (I-06, I-10). With this complexity, defining terms of uses becomes pivotal (I-01), as it defines how data products are used, specifies monetary incentives (I-05), and decides data storage locations (I-22). One participant illustrated this in detail (I-24):

> *"[Data ownership means]: I can define my policies and be sure that no one accesses my data without my consent, I can define how long the access is granted, I can define who is getting access. I have a data contract to define how to use this data for which purposes. So, as long as I define all the conditions, no one other, and not the platform, I am fine. What also is very relevant is to declare how this [a data product] is charged."*

When data products encompass personal, sensitive information, *privacy* is unlocked as a paramount facet of data sovereignty. This facet **redefines** the boundaries of data ownership, mandating data providers to consider the rights and interests of data subjects (I-07). This means data providers must obtain explicit consent from data subjects for approval (I-27) and ensure they get tangible benefits in sharing their data (I-13). Therefore, claiming ownership becomes even more complicated due to the tensions with data subjects (I-13). Participant I-03 highlighted this tension: "I always doubt that we have any data at all because we are maintaining the data of our customers. We have data about their activities. You can ask for consent, but that is always a gray area. So, if you collect the data, it always comes from somebody else."

Drawing on the empirical evidence above, emphasis on defining usage policies and ensuring consent in data ownership and privacy, respectively, demonstrate essential needs for safeguarding. Therefore, it is logical to classify them under the overarching protection facet. This interpretation is aligned with participant I-01's view that linked ownership and protection, explicitly stating the need to "… have strong protection mechanisms of the ownership of the data." Meanwhile, participant I-20 connected privacy and protection: "In practice, when looking at the personal data and digital markets, you should be careful … [Sharing] personal data from one platform to another should comply with GDPR rules or data regulations, with special attention must be paid to privacy and data protection in this case." With this interpretation, protection encapsulates the baseline rights inherently held by data providers (i.e., data ownership) and subjects (i.e., privacy), which are recognized as societal norms. These rights are pre-existing conditions *before* data sharing transactions occur, setting a baseline for subsequent data sharing processes. Data ownership and privacy do not fall under provision, as provision primarily concentrates on service delivery. Similarly, these facets are distinct from participation, measures taken to improve active engagement of data providers.

Provision of control, security, and compliance mechanisms

The discussion with participants shifted to provisioning control, security, and compliance mechanisms. With the foundational rights of ownership in place, data providers require mechanisms for *data control* to protect these rights. Thus, data control emerges as another critical facet of data sovereignty. Participant I-18 illustrated the connection between control and sovereignty: "One can imagine data sovereignty

is basically the ability of the data owner, for instance, to withdraw the data set from a particular marketplace at the time of their choosing. That is one example of exercising data sovereignty."

The ownership facet **defines** the data control facet because control over data is exercised based on the agreed-upon terms of use, which are formalized through contracts signed by consumers (I-18). These contracts are technically enforced to enable control over data (I-16). I-24 illustrated this view in the case of technical enforcement using a *connector*:

> *"In the connector world, they [researchers and practitioners] often talk about fully enforced policies. In your data source, you have a connector. You have another connector in your data sink. And you have your offer, you agree on the contract, and then you have all the terms, conditions, and policies. After that, the data gets transferred from the data source to the sink. Technically, we could build this."*

Having the agreement technically enforced, data control thus enables data provenance to track down data usage history for data monitoring (I-10). In all, data control helps **retain** ownership of data providers.

When personal data is involved, providing data subjects with data control mechanisms to **safeguard** their privacy becomes essential. Generally, they demand transparency (I-06, I-10) to know how their data is used (I-09). In extreme cases, data subjects should monitor the data themselves, such as from the dashboard functionalities. "So, more control about your data in a personal space. I see that personal space as a dashboard where I can check who is doing what with my data" (I-06). In addition, data subjects must have a choice to revoke their consent (I-26). However, getting consent for data sharing, if not impossible, is extremely hard, "So a lot of data is shared but under strict conditions. Furthermore, data sharing for commercial purposes is also being done, but it is a lot more difficult. Because, I think, we will get to it later, the consent of the (end-) consumer" (I-08). I-09 also amplified this view by contrasting the difficulties between sharing aggregated and personal data:

> *"If we share personal data, this concept [business data sharing] has some problems. I have to inform clients or customers, the data subjects: for which purpose will I process their data, and to whom will I send their data? If I will put data subjects' data on a meta-platform, and this meta-platform provides the data to an unlimited number of customers, data subjects do not have a clue to whom the data will be shared. So, this concept is okay for aggregated data, but for personal data, I think there is lots of work to ensure that this is GDPR compliance."*

Beyond data control, security stands out as a data sovereignty facet to **safeguard** ownership of providers; as one interviewee highlighted, "Looking at data sovereignty, security is also important because you do not want everything to be put out in the open" (I-26). Security prevents unauthorized parties from accessing the data (I-12), making sure transactions cannot be denied (non-repudiation) (I-05), and ensuring "… availability of data, and then you can use it in a certain application. That is where the added value [of a meta-platform] lies" (I-21). Security entails providing cutting-edge security protections for data sharing (I-05), such as watermarking (I-16), certification (I-19), and smart contracts (I-23), to name a few.

Moreover, discussions with participants highlighted the significance of privacy-enhancing technologies, such as anonymization (I-11) and encryption (I-16). Participant I-05 illustrated the relation between security and privacy, "I would say, for privacy, the meta-platform does not have that privilege of seeing what the data is and possibly putting different data marketplace in competition or selling data that belongs to one marketplace to another marketplace. The meta-platform has to be really secure … So, security, in terms of data transportation and the data sharing from one micro-PC to another until it reaches the buyer, is really important." Such emphasis on privacy technologies underscores the dual role of security: safeguarding data ownership and privacy of data subjects, as and when applicable.

Compliance also emerges as a vital facet of data sovereignty, branching into two areas: external and internal compliance. Firstly, data providers must respect external compliance, encompassing legal and regulatory mandates. For example, I-27 illustrated compliance in the context of the Data Governance Act:

> *"The core principle is that there is always control by the entitled party. So, the entitled party controls what happens to his data and where it is published. The entitled party has data classifications [shareable vs. non-shareable data]. But for [privacy] protection, there must always be explicit consent [from data subjects that] in line with the Data Governance Act."*

Given legal intricacies, several participants voiced the need for guidance. I-07 expressed, "It will be great if data stewardship is established. So, when somebody is unsure about this data point, they can immediately contact data stewardships for further explanation."

Internal compliance involves data providers aligning with the technical aspects enforced by meta-platform operators. I-11 emphasized the technical nature of this compliance, stating, "I believe compliance also has a technical facet, given that your data should be standardized or normalized." To simplify this compliance process, operators should introduce easy-to-follow mechanisms,

such as clear certifications (I-19) or the adoption of a widely recognized reference architecture, like those from the International Data Sharing Association (I-24) or Gaia-X (I-23).

Like control and security, compliance mechanisms serve as tools to **safeguard** the ownership and privacy rights of data providers and subjects, respectively. I-09 stressed the significance of such mechanisms:

> *"We have technical and organizational protection. In my words, you must have some technical and legal skills ... People in security will be looking for some security aspects from a technical point of view. From a contractual point and a legal point, we will try to find some stipulations that are not in line with the GDPR or the national law. For example, you can have a contract between two data controllers. One of the obligations is to inform your data subjects about data processing. I can put in a contract that another data controller is obliged to inform my clients about my data processing. This is possible in contracts. So, you can protect your clients."*

In summary, we infer that provision in data sovereignty encapsulates data control, security, and compliance mechanisms. Data control via technical enforcement provides mechanisms for ongoing monitoring post-transaction to check whether data is shared according to terms and conditions. Security mechanisms, such as encryption and watermarking, aim to continuously safeguard against unauthorized access, even after sharing data. Compliance extends beyond legal frameworks, necessitating consistent alignment with meta-platform technical specifications and the continuing fulfillment of contractual obligations between data providers and consumers. The findings indicate two types of provisions: *control-based*, which facilitates horizontal interactions among societal actors like data subjects, providers, and consumers, ensuring adherence with established terms for data sharing; and *defense-based*, which involves security and compliance, countering breaches from actors outside the standard data sharing processes (e.g., cybercriminals, unauthorized third-party organizations) and ensuring territorial regulations to avoid negative consequences, respectively. In all, provision highlights intentional actions by meta-platforms to safeguard ownership and privacy *during* and *post-data transactions*. Hence, provision differs from protection, primarily concerned with recognizing and establishing data providers and subject rights, rather than actively implementing technical and compliance measures that characterize provision. Furthermore, they are distinct from participation, which aims for the active engagement and involvement of data providers in meta-platforms.

Ensuring participation through clear responsibility division

The connection between data sovereignty, participation, and *responsibility* became a recurrent theme in interviews. Participation represents the myriad opportunities available to diverse societal groups, especially data providers, to articulate concerns, contribute feedback, steer decisions, and participate in meaningful interactions. From the perspective of data providers, participation extends beyond mere outcomes like willingness to share data. For example, participatory engagements mean that providers (and other actors) use meta-platforms in standardized, mutually agreed upon, and approved ways (I-29). Another example is active oversight of other societal groups (I-27).

Yet, concerns emerged about meta-platforms potentially turning monopolistic and non-democratic. I-29 articulated, "… what makes me doubtful is such a meta-platform will always be coupled to commercial aspects and capitalistic systems which are inherently non-democratic." Echoing this, concerns about dominant platforms emerged, with the dilemma of high participation costs on platforms like hotel booking services highlighted by I-26: "… participating comes at a steep cost, sometimes as much as around 20% in charges." Further, concerns about platforms potentially exploiting participants were evident, as I-21 pointed out: "I can well imagine that a (meta-)platform will be created that organizes it very well from a technological perspective, but then starts to exploit participants."

To ensure constructive and meaningful participation, delineating responsibility among sovereign entities and societal groups is crucial. I-23 asked, "…who is going to write the software, and who is just going to install it?" This view resonates with I-28's statement:

> *"Who should provide the infrastructure [for sovereignty]? It could be a meta platform, but it could also be a marketplace. But the governance, from my perspective, has to be some cooperative model—an association or a foundation or any other form. If you want to maintain trust, because that is ultimately what this is about, because you will only participate in it if you know that this is reliable, then it must also be reflected in the way in which you organize it together."*

Nevertheless, responsibility division is not straightforward, especially in the context of meta-platforms where the governance structure between a meta-platform and data marketplace participants remains unclear. One interviewee (I-12) said:

> *"So, for example, if you are a meta-platform and a data marketplace gets data products from you and then sells it to data consumers, and then that data marketplace has security issues or goes down, or the data is*

corrupted, and then the question is, who is responsible for that? Is it the data marketplace itself? Is it the meta-platform?"

To conclude, our findings highlight the criticality of defining clear responsibilities between meta-platform and data marketplace operators. This clarity is essential to foster active participation among all data sharing actors. Such well-defined responsibilities are pivotal for **steering** the provision of data sovereignty measures, which are crucial for safeguarding the fundamental rights of data providers and subjects. Thus, the responsibility facet of data sovereignty belongs to participation rather than provision, which concentrates on delivering mechanisms supporting data sharing rights, or protection, which focuses on recognizing and establishing these rights.

The spatial aspect of data sovereignty: Contextual conditions

Drawing from "Section Social contract theory," the spatial aspect of SCT identifies conditions that shape the substantive aspect (i.e., the three Ps). In data sovereignty, contextual conditions serve a similar function because they influence the difficulty of realizing data sovereignty facets. To explore this further, the following sections examine three key conditions: data type, business data sharing setting, and organizational size.

Data type

Beyond personal data, diverse *data type formats* complicate *control*. Some consumers prefer single dataset purchases (I-03), while others seek continuous streams (I-05) with potential time constraints (I-03). Data products, especially when transformed into machine learning models (I-18), make control even more challenging. Given these varying needs for different data types, providing suitable control mechanisms to safeguard all data types presents a formidable challenge.

Data origin, particularly the *industry* it originated from, plays a pivotal role in the complexities of *compliance* and *ownership*. Unique characteristics across industrial sectors necessitate tailored regulations. This necessitates (meta-) platform operators to guide data providers in adapting ownership definitions according to pertinent policies (I-31). Furthermore, regulatory and law requirements exhibit considerable discrepancies across various industries. For example, I-07 argued about "… over-regulation of the banking industry. So, there are a lot of regulations on the table" compared to the telecommunication industry. I-10 from a finance industry mentioned, "If you just looked through our IT portfolio … there you see the part of legal is increasing every year. Now, we have European regulation; we have European bank law; we have our national regulators." This highlights the expansion of regulations in certain sectors, spotlighting the role of industry-specific data in shaping compliance and data ownership paradigms.

Certain characteristics of *industry-specific data* influence the complexities associated with *data ownership* and *control*. Consider the data sharing practices prevalent in the capital markets industry as an illustrative example. This sector's maturity in data practices and regulations has led to a sophisticated understanding and effective management of data sharing. I-25 stated, "The capital markets as a data provider area are fairly mature … And what is also interesting is that the financial and capital markets industry is highly regulated. So, they are mature in compliance practices." Therefore, the nature of industry-specific data shapes the industry practices, regulations, and awareness around data sharing. This, in turn, fosters an increased level of data literacy, thereby supporting data providers in defining ownership and meta-platforms in provisioning control measures.

Business data sharing setting

Data sovereignty complexities intensify in *the meta-platform setting*, mainly due to ambiguous governance between meta-platforms and marketplace participants. "Section Ensuring participation through clear responsibility division" highlights the challenge of pinpointing data control *responsibility*, especially during security breaches or system failures. Furthermore, meta-platform operators are responsible for selecting trustworthy data marketplace participants, which adds another layer of complexity given the diverse operation rules and security standards across platforms. Data providers are suspicious if specific marketplaces are disreputable intermediaries (I-04). One interviewee (I-01) said: "If a channel [a data marketplace], for instance, is ruled by mafias, you will try to avoid it." Evaluating such marketplaces is problematic because each has unique operation rules (I-01). For example, while some marketplaces have decent security, others do not (I-12). These challenges highlight the importance of defining clear responsibilities to enforce data sovereignty measures effectively.

The meta-platform context also increases *control* complexity. While achieving complete control is conceivable, technical hindrances persist (I-02, I-03). For example, data sharing via a meta-platform raises concerns related to data provenance. Meta-platforms allow providers to share their business data with multiple data marketplaces. Hence, data lineage from providers to consumers becomes more complex and blurry. An interviewee (I-12) asked: "Who is responsible for providing the lineage from supplier to buyer if you have two stops, which are two separate entities? … We have two parts in the chain." Therefore, there is a possibility of having blind spots in the data lineage, making data tracing difficult

(I-7). In addition, data providers may need to withdraw data for specific reasons. Nevertheless, retrieving shared data is difficult: data providers must identify which data marketplace shares their data (and to which data consumers) (I-09).

The meta-platform setting raises difficulties in providing *compliance mechanisms*. A meta-platform commonly aims to be interoperable across data marketplaces in different countries or industries. Nevertheless, different work rules depend on specific areas (I-01), and translating diverse legal instruments between countries is difficult (I-02, I-08). For example, in extreme cases where a meta-platform is interoperable with data marketplaces outside the European Union, some regulations like GDPR may not be applicable (I-01). Hence, meta-platforms may not help data providers understand what they can (and cannot) do with the data (I-13).

Organizational size

Organizational size is pivotal in data providers' capacity to define *ownership*. Larger entities are often more prepared to share data. As I-26 put it, "But I do not think Small-Medium Enterprises (SMEs) will share raw data on such a marketplace." Typically, bigger organizations have enhanced capabilities for ownership definition (I-21). In contrast, smaller providers face challenges due to inadequate data skills and awareness. I-10 voiced a concern, asking: "What happens to the ownership of the data?" For SMEs lacking data skills, a solution is to outsource processes and draw insights from external parties. I-21 suggested, "Larger organizations have those [data sharing] capabilities. The smaller ones can rely on external parties, for instance, for data storage."

Organizational size is essential when addressing *non-compliance* data sharing cases, especially regarding legal consequences. Given their substantial market presence, larger organizations often experience the implications more intensely. I-13 noted, "It is the bigger player in the market that is always going to bear the brunt of it." I-20, who shared a personal experience of a security breach at a small enterprise, further agreed with this view. Although this incident occurred in a small-scale context, it caused considerable distress: "We were careless, and it happened. So, these security breaches were very painful for us even though we are a small business. For a big business, I think it is even more painful."

Nevertheless, large organizations that handle extensive datasets are often more ready to conduct data sharing due to their rigorous liability measures. They are more vigilant about potential infringements, thereby minimizing risks. I-27 illustrated this point, noting, "Because with that, you also have the liability taken seriously. The chances of violation are smaller than with many small players." Hence, collaborating with larger companies often signifies a more secure data sharing than partnering with multiple smaller firms with potentially inadequate data practices.

Data sovereignty conceptual framework: Interactions between data sovereignty facets and contextual conditions

This section summarizes the previous findings about data sovereignty facets, their interactions, and how contextual conditions determine the difficulty in realizing sovereignty facets (Fig. 2). Figure 2 shows the relationships between higher-level facets (i.e., participation, provision, and protection) and lower level facets. Regarding the higher-level facets, participation *steers* how meta-platforms and data marketplace operators divide responsibilities for developing provision mechanisms, which are either control or defense-based. Control-based provision facilitates horizontal interactions among actors such as data subjects and providers to ensure sovereignty. Meanwhile, defense-based provision consists of security and compliance mechanisms, *safeguarding* against breaches (e.g., by cybercriminals) and ensuring adherence to territorial regulations, respectively. Meta-platforms develop these mechanisms to support the foundational rights that need protection: data ownership of providers and privacy of data subjects.

At the facet level, every data sharing should start with providers describing data ownership, thus *defining* the data control facet. Control over data is exercised based on terms of use formalized through contracts, aiding ownership *retention*. When personal data is shared, the privacy facet *redefines* ownership definitions, requiring providers to factor in data subjects' rights. Moreover, it becomes essential to provision data subjects, too, with control mechanisms to *safeguard* their privacy.

The conceptual framework clarifies the relationship between sovereignty facets and specific actors. For instance, the *responsibility division* under participation belongs to the operator perspective. Meanwhile, the *data control* facet mediates between the perspectives of data providers and data subjects. The overlap in Fig. 2 highlights this, showing that data control under control-based provision belongs to both data providers and consumers. This positioning demonstrates that data control serves as a *bridge*, potentially resolving the tensions of ownership claims between data subjects and data providers. Security and compliance facets fall under defense-based provision, indicating that data providers utilize these mechanisms, whereas data subjects do not directly use them. Figure 2 also shows that each (higher-level) facet interrelates with multiple others rather than simply exhibiting one-on-one interrelationships like mutual interdependence. At the same time, it is also not the case that all facets are connected to everything else. Instead, the facets interrelate variously.

Table 3 summarizes how contextual conditions influence the complexity of realizing data sovereignty facets.

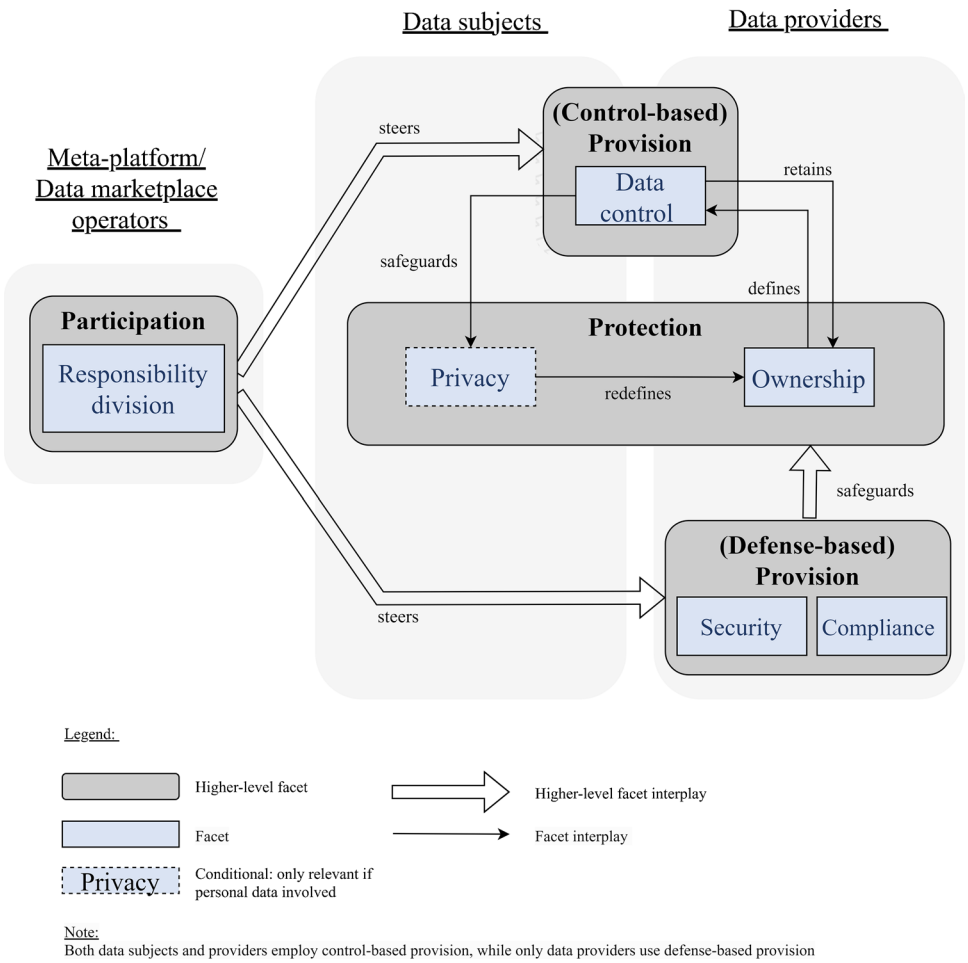

Fig. 2 A multi-faceted conceptual framework of data sovereignty

Table 3 Contextual conditions affecting data sovereignty facets

Contextual condition	Influence on data sovereignty facet
Data type (personal, sensitive data)	• Sharing personal, sensitive data unlocks the *privacy* facet, triggering difficulties in defining *data ownership* and provisioning *control* and *compliance* mechanisms
Data type (format variations)	• Data format variations raise technical challenges for provisioning *control* mechanisms
Data type (industry-specific data)	• The diversity in industry-specific laws and regulations mandates (meta-)platform operators to provision *compliance* mechanisms for data providers in tailoring data *ownership* definitions following applicable policies • In some industrial types, such as capital markets, data providers and consumers generally know about business data sharing practices, increasing the overall data skills and awareness to define *ownership* and exercise *control*
Business data sharing setting	• An unclear governance structure between meta-platform and data marketplace operators amplifies the complexity of realizing *responsibility division* • Meta-platform architecture raises technical challenges in realizing *control* mechanisms • The meta-platform setting raises difficulties in adhering to *compliance*, primarily due to aiming for cross-industry and cross-border data sharing
Organizational size	• SMEs lack data skills and awareness to define *data ownership* • Larger enterprises are more liable to the *consequences of non-compliance*

Discussion

In "Section Interpretation of the substantive aspect of SCT in data sovereignty," we discuss the interpretation of the substantive aspect of Social Contract Theory (SCT) in data sovereignty. "Section Reflecting on contextual conditions affecting data sovereignty" discusses the contextual conditions. "Section Framework applicability to other business data sharing settings: Hierarchal and network mode" discusses the framework's applicability in other business data sharing settings, and "Section Theoretical implications" outlines the theoretical implications.

Interpretation of the substantive aspect of SCT in data sovereignty

Our framework in Fig. 2 shows that provision and protection facets have many mutual influences, suggesting that the other is indirectly affected when one facet is addressed. These interplays suggest that addressing data sovereignty in provision affects protection and vice versa. In contrast, participation influences provision and protection facets but not vice versa. Hence, exploring scenarios where protection and provision co-vary and integrating them with participation may be valuable. By connecting these scenarios to the various philosophical views in SCT from "Section Social contract theory," we may explain the prevalence of control-centric conceptualizations in the data sovereignty discourse. We discuss four potential scenarios as follows.

In the first scenario, data sovereignty is absent in a data sharing setting. This absence is marked by a lack of recognition of data ownership rights for providers and privacy for data subjects to be protected. Moreover, provisioning is also non-existent due to the absence of clearly defined responsibilities to do that. Data sharing without data sovereignty risks violating the fundamental rights of data providers and subjects, indicating a lack of social contracts. For example, consider data sharing settings lacking control and security mechanisms, combined with unclear responsibility for the decision-making process (Fassnacht et al., 2023). Such shortcomings could be the reasons why data sharing settings, such as data marketplaces, struggle to gain traction (Spiekermann, 2019).

The second scenario showcases sufficient provision mechanisms for control, security, and compliance, safeguarding the rights of data providers and subjects. However, this scenario has poorly defined and communicated responsibility divisions, resulting in minimal engagement from the relevant actors. This scenario aligns with Hobbes's (1651) philosophical view, where pronounced institutional controls exist, but citizen engagement is relatively low.

The third scenario represents a configuration with medium levels of protection–provision and participation. Provision mechanisms are sufficient, and relevant actors in a data sharing setting display moderate engagement due to a clear understanding of their roles. This scenario reflects Locke's (1689) viewpoint on social contracts: rights are sufficiently safeguarded, roles are explicitly articulated, and citizen involvement is evident, although not maximized.

In the fourth scenario, participation receives pronounced emphasis, in contrast to the more modestly developed provision mechanisms. Here, data sharing actors are actively involved, with a distinct understanding of their roles. However, provision mechanisms might not be extensive. This aligns with Rousseau's (1762) stance, valuing citizen participation even when provision mechanisms are less rigid.

Exploring the four scenarios of social contracts uncovers the implicit philosophical assumptions that apparently underlie existing literature on data sovereignty. As noted in "Section Introduction," data sovereignty literature takes a one-sided focus on control, which falls under provision facets while overlooking protection and participation facets. This focus prompts a preference for technical and authoritative measures over inclusive, participatory approaches, but is only justifiable if one assumes that actors within data sharing are inclined toward self-interest and disorder (i.e., in a Hobbesian view). Acknowledging Rousseau's view, for instance, would imply that participation facets should be considered. As such, the scenario guides data sovereignty researchers to what facets to consider, depending on their assumptions of the nature of data sharing actors.

Reflecting on contextual conditions affecting data sovereignty

While acknowledging the possibility of data sovereignty at both organizational and individual levels (Otto, 2019), the current literature focuses on the individual level (e.g., Lauf et al., 2022; Schinle et al., 2021). This emphasis may stem from the historical focus on *indigenous data sovereignty*, which centers on the self-determination of native communities regarding data rights and privacy (Kukutai & Taylor, 2016). Our research brings the organizational perspective to the forefront. Operators must provision control, security, and compliance mechanisms to protect data providers' ownership. Providers must navigate ownership tensions with data subjects in scenarios involving personal data. Hence, achieving sovereignty necessitates clear

communication about benefits for data subjects (Spiekermann et al., 2015).

Meta-platforms as emergent business data sharing settings amplify data sovereignty concerns, particularly in the control and responsibility facets. Meta-platforms enable data asset movement across multiple marketplaces. Consequently, data provenance becomes more difficult because we now have two separate entities, amplifying the probability of having blind spots in the data lineage. Thus, data providers generally do not know what happens when data is transferred. Blockchain-based smart contracts are one of the (future) promising means for data provenance (Moyano et al., 2021), but these may be incompatible with other marketplaces. To address this issue, interoperable solutions like *side-chain* or *interchain* technologies are essential for meta-platforms that federate existing data marketplaces. However, these initiatives are still in their early stages and can be challenging to implement in practice (Singh et al., 2020).

A meta-platform aims to federate existing data marketplaces. Nevertheless, knowledge about the interactions and proper institutional arrangements between these entities is generally unexplored. For example, meta-platforms may have sufficient power to dictate enabling technologies and infrastructures if they are horizontally integrated by the same parent company (such as Alipay, WeChat, or Tencent) (Coe & Yang, 2022; Zhang & Williamson, 2021). Nonetheless, meta-platforms for business data sharing often emerge from consortium efforts (Floetgen et al., 2021). Hence, having a consensus on governance mechanisms and joint efforts to develop technological standards is challenging (Gelhaar & Otto, 2020). Each data marketplace participant may have internal (technically enforced) governance mechanisms to ensure data sovereignty, but such mechanisms may be incompatible with others. Consequently, responsibilities to provide provision mechanisms remain unclear.

Considering another contextual condition, data providers at the Small Medium–sized Enterprise (SME) level often face challenges in defining ownership and exercising control. This is partly due to their limited data skills and resource constraints in implementing the technical infrastructure (Scaria et al., 2018). SMEs also have limited bargaining power when negotiating data sharing agreements (Touboulic et al., 2014). Moreover, SMEs find it hard to keep up with novel requirements because business data sharing has not yet become their core business (Fassnacht et al., 2023). In contrast, large companies have more resources and bargaining power to negotiate favorable data sharing agreements. Large companies are also more likely to invest in data analytics and employ specialists to support their data management efforts (Martens et al., 2020). Specifically, large companies specializing in the data economy already have data sharing capabilities and are, therefore, capable of defining ownership that deals with regulations (Schmidt et al., 2021). To tackle this issue, (meta-)platform operators can explore alternative value offerings for data providers in the SME category, for example, by providing consultation services or training. Such offerings can foster robust customer relationships, allowing data providers to define data ownership and exercise control over their data seamlessly (Virkar et al., 2019).

Framework applicability to other business data sharing settings: Hierarchal and network mode

The conceptual framework of data sovereignty can be generalized to two other data sharing settings: hierarchy (e.g., supply chains) and network (e.g., data ecosystems). As a contextual factor, unique characteristics of each setting play a crucial role in realizing data sovereignty facets.

For instance, considering the responsibility division, focal partners in supply chains generally must provide provision mechanisms, dictating the infrastructure for data control, which low-power partners must comply with (Ke & Wei, 2007). In contrast, the responsibility for provisioning data ecosystem infrastructure becomes the joint responsibility of its members (Otto & Jarke, 2019). Because data ecosystems are still in their infancy, the proven value of developing such ecosystems is unclear (Otto & Jarke, 2019). To speed up the emergence of a data ecosystem, data providers and consumers are often involved in the development; thus, allocating extra money to engage with these processes is required (Martin et al., 2021). This example illustrates how data sharing settings influence the difficulty of responsibility division.

Another example of the influence of the business data sharing setting is illustrated in defining data ownership in hierarchical-based data sharing. Low-power partners often depend on focal partners who may unilaterally change data sharing agreements (Hewage, 2018; Kembro et al., 2017). Consequently, despite the risk of disclosing too much information and potential exploitation, low-power partners might feel compelled to comply with data sharing requests to maintain their relationships with focal partners (Ke & Wei, 2007). Likewise, power dynamics also happen in the network mode, where keystone members coordinate data ecosystem developments. These keystone members can influence data ownership, such as monetary incentives (Otto & Jarke, 2019). Consequently, non-keystone actors may have limited negotiating power (Scaria et al., 2018), further complicating their ability to define ownership within the ecosystem.

The above examples show that the conceptual framework of data sovereignty is relevant and can be extended to other data sharing settings, although there may be slight differences in the influence specification of each facet.

Theoretical implications

Our study makes a primary contribution to the existing Information Systems literature, particularly in research streams related to data sharing and data sovereignty, by *proposing an alternative conceptualization of data sovereignty as a multi-faceted construct*. Contrary to current literature that predominantly links data sovereignty to data control (see a review by Hellmeier & von Scherenberg, 2023), we propose a framework that takes a multi-faceted perspective on data sovereignty drawing from Social Contract Theory (SCT) (see Fig. 2). With this alternative conceptualization, researchers and practitioners can identify complementary approaches to achieving data sovereignty beyond mere control measures. The framework also helps understand how addressing one facet impacts other facets. Data control measures are in high demand (e.g., usage and access control), but their implementation is generally still in its infancy (Tao et al., 2022). Thus, a more expansive viewpoint on data sovereignty allows us to analyze the potential trade-off and complex interplay between facets. For instance, (meta-)platform operators can explore potential solutions by weighing the trade-off between data type, ownership, and control. In cases where data is used for training machine learning models via a federated approach (Li et al., 2020), control over the data may not be required if the data is anonymized and used only for training purposes. When involving personal data, Multi-Party Computation (MPC) enables data sharing actors to collaboratively compute a function over their inputs without revealing private data (Koch et al., 2021). Hence, MPC may make data control less relevant because data providers cannot lose ownership. These examples illustrate how, paradoxically, data sovereignty may become easier to achieve when addressing non-control-related facets.

The conceptual framework offers researchers and practitioners enhanced precision when developing (and claiming) data sovereignty solutions. For instance, while Schmidt et al. (2022) primarily emphasize privacy technologies and Pedreira et al. (2021) delve into security mechanisms, their claims about exploring sovereignty solutions are not misguided. This is because, in fact, privacy and security are one of data sovereignty facets. Our framework helps make more precise claims on which sovereignty facets are targeted when designing solutions.

Our interpretation of the substantive aspect of SCT in data sovereignty uncovers the implicit philosophical assumptions inherent in the current data sovereignty literature. We showed that the overly focused on control in the literature implicitly resonates with Hobbes's viewpoint, suggesting that societal groups in data sharing lean toward self-interest and disorder by their inherent nature. This explains the literature's tendency toward technical, authoritative solutions over inclusive participation in dealing with data sovereignty. This perspective introduces a compelling counterpoint: What if, drawing from Locke (1689) and Rousseau (1762), societal groups engaged in data sharing are assumed to be fundamentally altruistic? Such an assumption might suggest that the path to achieving data sovereignty need not be heavily anchored in stringent protective measures. Instead, there could be a shift toward strengthening responsible participatory mechanisms. This alternative viewpoint might recalibrate the trajectory for building data sovereignty solutions, addressing participation-related facets rather than control-based provision. As such, our framework contributes by uncovering hidden assumptions behind work on data sovereignty and facilitating a reflective discourse on whether this literature is going in a direction that aligns with Locke's and Rousseau's views on human nature.

We frame our secondary contribution as follows. *First, we provide empirical evidence of the descriptive knowledge of data sovereignty.* Hummel et al. (2021) identify 16 *notions* correlated with sovereignty, with the leading themes encompassing control, security, privacy, and ownership. This aligns with our empirical findings, emphasizing the centrality of these facets in data sovereignty discourses in business data sharing. Interestingly, although responsibility ranks low regarding its co-occurrence with data sovereignty in their study, suggesting lesser importance, our research highlights its critical role in data sharing. Moreover, the facet of compliance emerges as another significant consideration in our study, which is yet to receive its deserved prominence in the existing literature. Thus, our research not only confirms the importance of these notions as data sovereignty facets but also elevates the importance of responsibility and compliance in contemporary data sovereignty discussions.

Second, we propose causal mechanisms between data sovereignty (higher-level) facets: We go beyond Hummel et al. (2021) work that correlates data sovereignty with certain notions. As shown in Fig. 2, we propose a causal mechanism to explain how data sovereignty (higher-level) facets correlate to each other's. These causal mechanisms provide clearer hypotheses for empirical investigations, directing future studies toward understanding the forces that drive changes in data sovereignty facets rather than just observing that changes occur. In experimental research, awareness of facets that co-vary with treatment and the selected facet can help researchers avoid drawing misleading conclusions based on correlations that may not reflect causality.

Third, we highlight contextual conditions for specifying boundary conditions: Prior to this study, data sovereignty literature did not explore the contextual conditions that influence the difficulty of realizing its facets. Consequently, boundary conditions for precisely theorizing data sovereignty are unexplored. Therefore, we contribute to the literature by identifying three contextual conditions that serve as boundary conditions: data type, business data sharing setting, and organizational size. This is particularly

important for future studies that aim to select facets of data sovereignty relevant to their study context, ensuring that they focus on the aspects that are both contextually significant and challenging to achieve; otherwise, data sovereignty may be incorrectly deemed unimportant.

Conclusion

Our research concludes by summarizing the findings in "Section Summary of the findings," discussing limitations and future research in "Section Limitations and future research," and elaborating on societal relevance in "Section Societal relevance."

Summary of the findings

In the data economy literature, data sovereignty is often narrowly interpreted as data providers' ability to control their shared data. This narrow focus on control might (1) lead to solutions that address only fragments of the overarching issues, (2) result in inconsistent research outcomes and potential biases, and (3) imply a conceptual basis is missing for higher levels of theory development. To address these shortcomings, we propose a multi-faceted conceptual framework that explains the relationships between data sovereignty facets (see Fig. 2).

Building on Social Contract Theory (SCT), we identify three higher-level facets of data sovereignty. First, the protection higher-level facets encompass the baseline rights inherently held by data providers and subjects for their data ownership and privacy, respectively. These rights stand as a pre-existing condition before any data sharing transactions occur. Second, the provision higher-level facets encapsulates data control, security, and compliance mechanisms. These mechanisms are provided by meta-platforms to safeguard ownership and privacy during and after data sharing transactions. Third, the participation higher-level facet requires clear responsibility division between sovereign entities (e.g., meta-platforms and data marketplace participants) to ensure active engagements of societal groups (e.g., data providers).

Our conceptual framework shows the interrelation between higher-level facets of sovereignty. The higher-level facet of participation determines how (and by whom) the provision mechanisms are provided, which, in turn, ensures baseline rights protections. Considering the lower-level facet, we find that data ownership defines how data providers can exercise control over their shared data, thereby retaining ownership rights. When personal data is involved, the privacy facet redefines ownership structure, necessitating a balance between providers and subjects and further reinforcing the role of data control in safeguarding privacy.

We find three contextual conditions determining the difficulty in realizing sovereignty facets: data type, business data sharing setting, and organizational size. Data sharing involving personal data type unlocks privacy as part of the protection facet, leading to challenges of ownership between data providers and data subjects, compliance demands, and control enforcement. Different data types, like structured and live-streamed data, present technical challenges for control mechanisms, while industry-specific data intensifies compliance complexity. The meta-platform setting raises ambiguity in determining the responsibility between such a meta-platform and data marketplace operators. Moreover, aligning multiple architectures of data marketplaces raises technical challenges in provisioning control mechanisms. Furthermore, meta-platform settings aiming for cross-border sharing complicate compliance mechanisms. Finally, small to medium-sized enterprises struggle to define ownership and maintain control, often due to resource and expertise constraints.

Limitations and future research

We propose several future research avenues considering our research limitations. First, the unit of analysis of this study focuses on data providers. Future work can explore the data consumers' perspective to provide a more balanced standpoint, possibly leading to new insights. For instance, when data consumers have more purchasing power, data providers may have less influence in creating horizontal agreements (i.e., data sharing agreements). This could require providers to give up some ownership rights. Second, our study focuses on the market-mode data sharing setting. Future studies can empirically confirm the applicability of the proposed framework to the two other settings: the hierarchy and network mode. Third, as this study focuses on the business perspective, exploring data sovereignty from the angle of consumer-to-business data sharing (e.g., personal data monetization) may offer a fresh insight into the facet interrelation. For example, data subjects voluntarily joining personal data marketplaces may be willing to sacrifice a certain level of privacy for direct monetary incentives, potentially reducing the need for stringent security and control mechanisms.

Fourth, as our study is qualitative in nature and generalizes to theoretical facets, future studies could employ quantitative methods. Specifically, the explanatory power of our facets on constructs such as willingness to share data would strengthen the nomological validity of our framework. Alternatively, quantitative configurational analysis methods could be relevant to empirically test whether our protection, provision, and participation configurations occur in reality. Fifth, although the contextual conditions (data type, business

data sharing setting, and organizational size) emerged from our empirical data and signal their importance, we do not claim them to be exhaustive. They serve as a foundational basis, and we see an opportunity for future research to delve deeper, exploring other conditions and possibly conducting prioritization studies. Sixth, our investigation targets data sovereignty post-2018, a pivotal time marked by transformative regulatory shifts within the European Union pertinent to the data economy. It would be insightful for future research to expand the temporal horizons, examining, for instance, periods when significant policy changes or events occurred. This would facilitate a deeper understanding of how the temporal dynamics of SCT influence the three Ps of data sovereignty.

Future research should also investigate how data sovereignty, underpinned by social contracts, can drive innovation to lower the inherent transaction costs of these contracts. Social contracts, like all contracts, incur transaction costs, such as those for (a) searching information, (b) bargaining, and (c) policing and enforcing (Dahlman, 1979). These costs go beyond monetary (Brouthers, 2002), such as the time and effort to police horizontal agreements between data providers and consumers. Innovation in data sovereignty, especially in the provision higher-level facet, may cut down these transactional costs. For example, using smart contracts for data control and compliance mechanisms can reduce the policing costs for data providers. This efficiency is achieved because smart contracts automate the enforcement of data product usage terms, which are encoded in computer-based languages (Petersen, 2022). Finally, future research could explore the connection between data sovereignty and broader political issues, such as state surveillance of data subject activities via sharing global telecommunication systems (e.g., Lashmar, 2017). The potential for state surveillance (e.g., Sweden's Tidö Agreement) might compel data providers to report data subjects who violate state regulations. This scenario indicates that the sovereignty of data providers in a business data sharing context, typically maintained through protection, provision, and participation, might be overtaken by state-imposed mandates.

The multi-faceted framework of data sovereignty developed in this research provides a foundation for further theory development, including explanatory theory (Gregor, 2006), which seeks to explain how and why certain phenomena occur. Further research can develop propositions based on the configuration of data sovereignty facets and test them through empirical research. For example, one could compare different configurations: an emphasis on protection and provision with low participation (akin to Hobbes's perspective), balanced levels of protection, provision, and participation (reflecting Locke's view), and a high emphasis on participation with low protection and provision (mirroring Rousseau's stance). These configurations can then be related to outcomes like trust, perceived risk, and willingness to share business data to strengthen the nomological validity of the data sovereignty construct. The framework can also guide the development of prescriptive theory, which offers principles for designing artifacts (Gregor et al., 2020). Data sovereignty facets can serve as requirements and evaluation indicators to design future data sovereignty solutions, e.g., in design science research (Hevner et al., 2004). Meanwhile, contextual conditions can be used as control variables in design studies. Besides contributing to theory development, our framework also provides a means to reflect on the philosophical assumptions underlying the work on data sovereignty, whereas we specifically identify a lack of Locke's and Rousseau's views.

Societal relevance

The findings of this study offer insights for policymakers, shedding light on how specific regulations influence various data sovereignty facets. For instance, the General Data Protection Regulation (GDPR) directly targets the protection facet, placing a stringent emphasis on safeguarding the privacy rights of data subjects. On the other hand, the Digital Services Act focuses on the provision facet, ensuring the accountability of digital service operators to equip data providers with robust control and compliance mechanisms. Finally, the Data Governance Act (DGA) shapes the participation facet. By instituting a framework for neutral data intermediaries, DGA clarifies the division of data responsibilities (e.g., what platform operators and third-party providers can do) that is vital in complex scenarios in the data economy. In sum, while these regulations collectively address all three higher-level facets of data sovereignty, there may be areas of overlap between these regulations that future policies need to recognize. Hence, future research is advised.

In addition, policymakers should be aware that focusing on a single facet may inadvertently create unintended consequences for other facets, potentially hindering the overall effectiveness of the regulation. Moreover, when regulation aims to be context-independent, the contextual conditions affect whether businesses can realistically deal with the regulations regarding sovereignty. For example, when involving personal data, DGA heavily emphasizes fully informed consent as a control and compliance mechanism to retain privacy. Yet, DGA, in its pursuit of promoting responsible data sharing, may unintentionally impose too stringent expectations (Ruohonen & Mickelsson, 2023). To realize the full potential of the data economy, businesses need the flexibility to adapt and innovate in their data processing activities. However, there are situations where it is challenging to precisely define the purpose for sharing personal data, especially when subsequent processing activities might evolve (Mantelero & Vaciago, 2015).

Consequently, this condition may hinder potential actors from participating in the data economy (Fassnacht et al., 2023). Therefore, policymakers must balance addressing specific facets and maintaining the flexibility to accommodate diverse contexts and their unique challenges. In doing so, policymakers can better facilitate a thriving data economy that fosters innovation, supports diverse business needs, and ensures responsible business data sharing across various contexts.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12525-024-00695-2.

Acknowledgements The research leading to these results has received funding from the European Union's Horizon 2020 Research and Innovation Programme, under Grant Agreement no 871481 – Trusted Secure Data Sharing Space (TRUSTS), from the H2020-ICT-2018-20/H2020-ICT-2019-2 Call. We extend our gratitude to the senior editor, two guest editors, and three anonymous reviewers for their constructive feedback. The first author also thanks Aga and Gilang for their critical insights on the initial version of our data sovereignty conceptualization. Additionally, this paper, an extension of Abbas et al. (2022) presented at the 35th Bled eConference, benefited greatly from feedback received there. This interview data is partly based on the thesis work of van Velzen (2022).

Data Availability The complete raw transcripts of the interviews conducted in this study are confidential and cannot be disclosed in accordance with data privacy regulations. Select excerpts for analytical purposes are accessible in the supplementary materials provided online.

References

Aaltonen, A., Alaimo, C., & Kallinikos, J. (2021). The making of data commodities: Data analytics as an embedded process. *Journal of Management Information Systems, 38*(2), 401–429. https://doi.org/10.1080/07421222.2021.1912928

Abbas, A. E., Agahari, W., van de Ven, M., Zuiderwijk, A., & de Reuver, M. (2021). Business data sharing through data marketplaces: A systematic literature review. *Journal of Theoretical and Applied Electronic Commerce Research, 16*(7), 3321–3339. https://doi.org/10.3390/jtaer16070180

Abbas, A. E., Ofe, H., Zuiderwijk, A., & de Reuver, M. (2022). Preparing future business data sharing via a meta-platform for data marketplaces: Exploring antecedents and consequences of data sovereignty. In *35th Bled eConference - Digital Restructuring and Human (Re-Action)*, Bled, Slovenia.

Azcoitia, S. A., & Laoutaris, N. (2022). A survey of data marketplaces and their business models. *SIGMOD Record, 51*(3), 18–29. https://doi.org/10.1145/3572751.3572755

Bodin, J. (1576). *Les six livres de la Republique*. Chez Iacques du Puys.

Brouthers, K. D. (2002). Institutional, cultural, and transaction cost influences on entry mode choice and performance. *Journal of International Business Studies, 33*(1), 203–221. https://doi.org/10.1057/palgrave.jibs.8491013

Coe, N. M., & Yang, C. (2022). Mobile gaming production networks, platform business groups, and the market power of China's tencent. *Annals of the American Association of Geographers, 112*(2), 307–330. https://doi.org/10.1080/24694452.2021.1933887

Cook, K. S. (2015). Exchange: Social. In J. D. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)* (pp. 482–488). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.32056-6

Dahlman, C. J. (1979). The problem of externality. *The Journal of Law and Economics, 22*(1), 141–162. https://doi.org/10.1086/466936

De Filippi, P., & McCarthy, S. (2012). Cloud computing: Centralization and data sovereignty. *European Journal of Law and Technology, 3*(2), 1–23. Available at SSRN: https://ssrn.com/abstract=2167372

Dubin, R. (1978). *Theory Building* (Rev). The Free Press.

Duisberg, A. (2022). Legal aspects of IDS: Data sovereignty—what does it imply? In *Designing Data Spaces: The Ecosystem Approach to Competitive Advantage* (pp. 61–90). Cham: Springer. https://doi.org/10.1007/978-3-030-93975-5_5

Edwards, R., & Holland, J. (2013). *What Is qualitative interviewing?* A&C Black.

Ellis, E. (2006). Citizenship and property rights: A new look at social contract theory. *The Journal of Politics, 68*(3), 544–555. https://doi.org/10.1111/j.1468-2508.2006.00444.x

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62*(1), 107–115. https://doi.org/10.1111/j.1365-2648.2007.04569.x

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics, 5*(1), 1–4. https://doi.org/10.11648/j.ajtas.20160501.11

Fassnacht, M., Benz, C., Heinz, D., Leimstoll, J., & Satzger, G. (2023). Barriers to data sharing among private sector organizations. In *Proceedings of the 56th Hawaii International Conference on Systems Sciences*, Maui, Hawaii, the United States.

Floetgen, R. J., Strauss, J., Weking, J., Hein, A., Urmetzer, F., Böhm, M., & Krcmar, H. (2021). Introducing platform ecosystem resilience: Leveraging mobility platforms and their ecosystems for the new normal during COVID-19. *European Journal of Information Systems, 30*(3), 1–18. https://doi.org/10.1080/0960085x.2021.1884009

Friend, C. (2004). *Social contract theory*. Internet Encyclopedia of Philosophy.

Fruhwirth, M., Rachinger, M., & Prlja, E. (2020). Discovering business models of data marketplaces. In *Proceedings of the 53rd Hawaii International Conference on System Sciences*, Maui, Hawaii, the United States.

Furness, M., & Trautner, B. (2020). Reconstituting social contracts in conflict-affected mena countries: Whither Iraq and Libya? *World Development, 135*(1), 1–12. https://doi.org/10.1016/j.worlddev.2020.105085

Gantz, J., & Reinsel, D. (2012). *The Digital Universe in 2020: Big data, bigger digital shadows, and biggest growth in the Far East*. https://datastorageasean.com/sites/default/files/idc-the-digital-universe-in-2020.pdf. Accessed 26 Nov 2023.

Gelhaar, J., & Otto, B. (2020). Challenges in the emergence of data ecosystems. *PACIS 2020 Proceedings*, Dubai, the United Arab Emirates.

Glennon, M., Kolding, M., Sundbland, M., Croce, C. L., Micheletti, G., Raczko, N., Freitas, L., Moise, C., & Osimo, D. (2023). *D2.4 second report on facts and figures (European DATA Market Study 2021–2023).* https://digital-strategy.ec.europa.eu/en/library/results-new-european-data-market-study-2021-2023. Accessed 26 Nov 2023.

Glesne, C. (2016). *Becoming qualitative researchers: An introduction* (Fifth ed.). Pearson Boston.

Gregor, S., Kruse, L. C., & Seidel, S. (2020). The anatomy of a design principle. *Journal of the Association for Information Systems, 21*(6), 1622–1652. https://doi.org/10.17705/1jais.00649

Gregor, S. (2006). The nature of theory in information systems. *MIS Quarterly, 30*(3), 611–642. https://doi.org/10.2307/25148742

Hellmeier, M., & von Scherenberg, F. (2023). A delimitation of data sovereignty from digital and technological sovereignty. In *European Conference on Information Systems* (*ECIS 2023) Research Papers*, Kristiansand, Norway.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly, 28*(1), 75–105. https://doi.org/10.2307/25148625

Hewage, U. (2018). Factors influencing the effective information sharing in Sri Lankan export-led manufacturing supply chains. In *2018 International Conference on Production and Operations Management Society (POMS),* Rio de Janeiro, Brazil. https://doi.org/10.1109/poms.2018.8629440

Hickey, S. (2011). The politics of social protection: What do we get from a 'social contract' approach? *Canadian Journal of Development Studies/revue Canadienne D'études Du Développement, 32*(4), 426–438. https://doi.org/10.1080/02255189.2011.647447

Hinsley, F. H. (1986). *Sovereignty* (Second ed.). Cambridge University Press.

Hobbes, T. (1651). *Leviathan*. London: Penguin Books (reprint 1985).

Homans, G. C. (1958). Social behavior as exchange. *American Journal of Sociology, 63*(6), 597–606. https://doi.org/10.1086/222355

Hummel, P., Braun, M., Tretter, M., & Dabrock, P. (2021). Data sovereignty: A review. *Big Data & Society, 8*(1), 1–17. https://doi.org/10.1177/2053951720982012

Jussen, I., Schweihoff, J., Dahms, V., Möller, F., & Otto, B. (2023). Data sharing fundamentals: characteristics and definition. In *Proceedings of the 56th Hawaii International Conference on System Sciences,* Maui, Hawaii, the United States.

Ke, W., & Wei, K. K. (2007). Factors affecting trading partners' knowledge sharing: using the lens of transaction cost economics and socio-political theories. *Electronic Commerce Research and Applications, 6*(3), 297–308. https://doi.org/10.1016/j.elerap.2006.06.006

Kembro, J., Näslund, D., & Olhager, J. (2017). Information sharing across multiple supply chain tiers: A delphi study on antecedents. *International Journal of Production Economics, 193*, 77–86. https://doi.org/10.1016/j.ijpe.2017.06.032

Koch, K., Krenn, S., Pellegrino, D., & Ramacher, S. (2021). Privacy-preserving analytics for data markets using MPC. In M. Friedewald, S. Schiffner, & S. Krenn (Eds.), *Privacy and Identity Management* (Vol. 619, pp. 226–246). Cham: Springer. https://doi.org/10.1007/978-3-030-72465-8_13

Kukutai, T., & Taylor, J. (2016). *Indigenous data sovereignty: Toward an agenda*. ANU Press.

Lashmar, P. (2017). No More Sources? *Journalism Practice, 11*(6), 665–688. https://doi.org/10.1080/17512786.2016.1179587

Lauf, F., Scheider, S., Bartsch, J., Herrmann, P., Radic, M., Rebbert, M., Nemat, A. T., Schlueter Langdon, C., Konrad, R., & Sunyaev, A. (2022). *Linking data sovereignty and data economy: Arising areas of tension. Wirtschaftsinformatik 2022 Proceedings*, Nuremberg, Germany.

Li, L., Fan, Y., Tse, M., & Lin, K.-Y. (2020). A review of applications in federated learning. *Computers & Industrial Engineering, 149*(1), 1–15. https://doi.org/10.1016/j.cie.2020.106854

Locke, J. (1689). *Two treatises of government.* New Haven: Yale University Press (reprint 2003).

Loewe, M., Zintl, T., & Houdret, A. (2021). The social contract as a tool of analysis: introduction to the special issue on "Framing the Evolution of New Social Contracts in Middle Eastern and North African Countries". *World Development, 145*(1), 1–16. https://doi.org/10.1016/j.worlddev.2020.104982

Lyon, D. (2014). Surveillance, snowden, and big data: Capacities, consequences, critique. *Big Data & Society, 1*(2), 1–13. https://doi.org/10.1177/2053951714541861

Mantelero, A., & Vaciago, G. (2015). Data protection in a big data society. Ideas for a future regulation. *Digital Investigation, 15*(1), 104–109. https://doi.org/10.1016/j.diin.2015.09.006

Manyika, J., Chui, M., Bisson, P., Woetzel, J., Dobbs, R., Bughin, J., & Aharon, D. (2015). *The internet of things: Mapping the value beyond the hype*. https://globaltrends.thedialogue.org/publication/the-internet-of-things-mapping-the-value-beyond-the-hype/. Accessed 26 Nov 2023.

Martens, B., De Streel, A., Graef, I., Tombal, T., & Duch-Brown, N. (2020). *Business-to-business data sharing: An economic and legal analysis (JRC Digital Economy Working Paper 2020-05).* https://joint-research-centre.ec.europa.eu/publications/business-business-data-sharing-economic-and-legal-analysis_en. Accessed 26 Nov 2023.

Martin, S., Gautier, P., Turki, S., & Kotsev, A. (2021). *Establishment of sustainable data ecosystems: Recommendations for the evolution of spatial data infrastructures.* Accessed on November 26, 2023. https://publications.jrc.ec.europa.eu/repository/handle/JRC124148

Moyano, J. P., Avital, M., Bühler, M., & Schmedders, K. (2021). Fostering peer-to-peer blockchain-based data markets. In *25th Pacific Asia Conference on Information Systems (PACIS)*, Dubai, the United Arab Emirates.

Munoz-Arcentales, A., López-Pernas, S., Pozo, A., Alonso, Á., Salvachúa, J., & Huecas, G. (2019). An architecture for providing data usage and access control in data sharing ecosystems. *Procedia Computer Science, 160*(1), 590–597. https://doi.org/10.1016/j.procs.2019.11.042

Olufowobi, H., Engel, R., Baracaldo, N., Bathen, L. A. D., Tata, S., & Ludwig, H. (2017). *Data provenance model for internet of things (IoT) systems.* Banff, Canada: Service-Oriented Computing–ICSOC 2016 Workshops: ASOCA, ISyCC, BSCI, and Satellite Events.

Opriel, S., Fraunhofer, I., Skubowius, G. E., Fraunhofer, I., & Lamberjohann, M. (2021). *How usage control fosters willingness to share sensitive data in inter-organizational processes of supply chain.* Bremen, Germany: International Scientific Symposium on Logistics 2021.

Óskarsdóttir, M., Bravo, C., Sarraute, C., Vanthienen, J., & Baesens, B. (2019). The value of big data for credit scoring: Enhancing financial inclusion using mobile phone data and social network analytics. *Applied Soft Computing, 74*(1), 26-39. https://doi.org/10.1016/j.asoc.2018.10.004

Otto, B. (2019). Interview with reinhold achatz on "Data Sovereignty and Data Ecosystems." *Business & Information Systems Engineering, 61*(5), 635–636. https://doi.org/10.1007/s12599-019-00609-z

Otto, B., & Jarke, M. (2019). Designing a multi-sided data platform: Findings from the international data spaces case. *Electronic Markets, 29*(4), 561–580. https://doi.org/10.1007/s12525-019-00362-x

Pedreira, V., Barros, D., & Pinto, P. (2021). A review of attacks, vulnerabilities, and defenses in industry 4.0 with new challenges on data sovereignty ahead. *Sensors, 21*(15), 1–21. https://doi.org/10.3390/s21155189

Petersen, D. (2022). Automating governance: Blockchain delivered governance for business networks. *Industrial Marketing Management, 102*(1), 177–189. https://doi.org/10.1016/j.indmarman.2022.01.017

Richter, H., & Slowinski, P. R. (2019). The data sharing economy: On the emergence of new intermediaries. *IIC - International Review of Intellectual Property and Competition Law, 50*(1), 4–29. https://doi.org/10.1007/s40319-018-00777-7

Rossmannek, O., & Chen, M. (2023). Why people use the sharing economy: A meta-analysis. *Journal of Cleaner Production, 387*(1), 1–16. https://doi.org/10.1016/j.jclepro.2022.135824

Rousseau, J.-J. (1762). *Du Contrat Social: Ou, Principes du Droit Politique*. Chez Marc Michel Rey.

Ruohonen, J., & Mickelsson, S. (2023). Reflections on the data governance act. *Digital Society, 2*(10), 1–9. https://doi.org/10.1007/s44206-023-00041-7

Scaria, E., Berghmans, A., Pont, M., Arnaut, C., & Leconte, S. (2018). *Study on data sharing between companies in Europe.* (A study prepared for the European Commission Directorate-General for Communications Networks, Content and Technology). https://op.europa.eu/en/publication-detail/-/publication/8b8776ff-4834-11e8-be1d-01aa75ed71a1/language-en. Accessed 26 Nov 2023.

Scheider, S., Lauf, F., Möller, F., & Otto, B. (2023). A reference system architecture with data sovereignty for human-centric data ecosystems. *Business & Information Systems Engineering, 65*(1), 577–595. https://doi.org/10.1007/s12599-023-00816-9

Schinle, M., Erler, C., & Stork, W. (2021). Data sovereignty in data donation cycles-requirements and enabling technologies for the data-driven development of health applications. In *Proceedings of the 54th Hawaii International Conference on System Sciences*, Kauai, Hawaii, the United State

Schmidt, K., Ullrich, A., & Eigelshoven, F. (2021). From exploitative structures towards data subject-inclusive personal data markets-a systematic literature review. In *European Conference on Information Systems (ECIS 2021) Research Papers*, Marrakesh, Morocco.

Schmidt, K., Munilla Garrido, G., Mühle, A., & Meinel, C. (2022). Mitigating sovereign data exchange challenges: A mapping to apply privacy-and authenticity-enhancing technologies. In *International Conference on Trust and Privacy in Digital Business*, Vienna, Austria.

Schweihoff, J., Jussen, I., & Möller, F. (2023). Trust me, I'm an intermediary! exploring data intermediation services. In *Proceedings of the 18th International Conference on Wirtschaftsinformatik*, Paderborn, Germany.

Sekaran, U., & Bougie, R. (2016). *Research methods for business: A skill building approach*. John Wiley & Sons.

Sestino, A., Kahlawi, A., & De Mauro, A. (2023). Decoding the data economy: A literature review of its impact on business, society and digital transformation. *European Journal of Innovation Management,* 1-26 https://doi.org/10.1108/EJIM-01-2023-0078

Shah, N., Coathup, V., Teare, H., Forgie, I., Giordano, G. N., Hansen, T. H., Groeneveld, L., Hudson, M., Pearson, E., Ruetten, H., & Kaye, J. (2019). Motivations for data sharing—views of research participants from four European Countries: A DIRECT study. *European Journal of Human Genetics, 27*(5), 721–729. https://doi.org/10.1038/s41431-019-0344-2

Singh, A., Click, K., Parizi, R. M., Zhang, Q., Dehghantanha, A., & Choo, K.-K.R. (2020). Sidechain technologies in blockchain networks: An examination and state-of-the-art review. *Journal of Network and Computer Applications, 149*(1), 1–16. https://doi.org/10.1016/j.jnca.2019.102471

Sobhy, H. (2021). The lived social contract in schools: From protection to the production of hegemony. *World Development, 137*(1), 1–15. https://doi.org/10.1016/j.worlddev.2020.104986

Spiekermann, M. (2019). Data marketplaces: Trends and monetisation of data goods. *Intereconomics, 54*(4), 208–216. https://doi.org/10.1007/s10272-019-0826-z

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

Tao, Y., Yang, S., & Ge, H. (2022). Comparative study on data sovereignty guarantee technology. In *2022 IEEE 13th International Symposium on Parallel Architectures, Algorithms and Programming (PAAP)*, Beijing, China.

Touboulic, A., Chicksand, D., & Walker, H. (2014). Managing imbalanced supply chain relationships for sustainability: A power perspective. *Decision Sciences, 45*(4), 577–619. https://doi.org/10.1111/deci.12087

van den Broek, T., & van Veenstra, A. F. (2015). Modes of governance in inter-organizational data collaborations. In *European Conference on Information Systems (ECIS 2015)*, Münster, Germany.

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*(1), 330–338. https://doi.org/10.1016/j.techfore.2017.09.040

van Velzen, T. (2022). *Business-to-business data sharing via data marketplace meta-platforms: Exploring governance mechanisms to enhance data sovereignty.* MSc thesis. Delft, the Netherlands: Delft University of Technology.

Virkar, S., Viale Pereira, G., & Vignoli, M. (2019). Investigating the social, political, economic and cultural implications of data trading. In I. Lindgren, et al. (Ed.), *Electronic Government. EGOV 2019. Lecture Notes in Computer Science* (pp. 215–229). Cham: Springer. https://doi.org/10.1007/978-3-030-27325-5_17

Williamson, O. E. (1989). Chapter 3 transaction cost economics. In *Handbook of Industrial Organization* (Vol. 1, pp. 135–182). Elsevier. https://doi.org/10.1016/S1573-448X(89)01006-X

Young, S. (2013). Transaction cost economics. In S. O. Idowu, Capaldi, N., Zu, L., Gupta, A. D. (Ed.), *Encyclopedia of Corporate Social Responsibility* (pp. 2547–2552). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-28036-8_221

Zhang, M. Y., & Williamson, P. (2021). The emergence of multiplatform ecosystems: Insights from China's mobile payments system in overcoming bottlenecks to reach the mass market. *Technological Forecasting and Social Change, 173*(1), 1–14. https://doi.org/10.1016/j.techfore.2021.121128