Linking Data Sovereignty and Data Economy: Arising Areas of Tension

Conference Paper · January 2022

CITATIONS
0

READS
126

11 authors, including:

Florian Lauf
Fraunhofer Institute for Software and Systems Engineering ISST
3 PUBLICATIONS **0** CITATIONS
SEE PROFILE

Simon Scheider
Fraunhofer Institute for Software and Systems Engineering ISST
1 PUBLICATION **0** CITATIONS
SEE PROFILE

Jan Bartsch
Karlsruhe Institute of Technology
4 PUBLICATIONS **1** CITATION
SEE PROFILE

Linking Data Sovereignty and Data Economy: Arising Areas of Tension

Florian Lauf[1], Simon Scheider[1,2], Jan Bartsch[3], Philipp Herrmann[4], Marija Radic[4], Marcel Rebbert[5], André T. Nemat[5], Christoph Schlueter-Langdon[6,7], Ralf Konrad[6], Ali Sunyaev[3,8] and Sven Meister[9,1]

[1] Fraunhofer Institute for Software and Systems Engineering ISST, Healthcare, Dortmund, Germany
{florian.lauf,simon.scheider,sven.meister}@isst.fraunhofer.de
[2] TU Dortmund, Chair of Industrial Information Management, Dortmund, Germany
simon.scheider@tu-dortmund.de
[3] Karlsruhe Institute of Technology (KIT), Department of Economics and Management, Karlsruhe, Germany
{jan.bartsch,sunyaev}@kit.edu
[4] Fraunhofer Center for International Management and Knowledge Economy IMW, Corporate Development in International Competition Division, Leipzig, Germany
{philipp.herrmann,marija.radic}@imw.fraunhofer.de
[5] Witten/Herdecke University, Institute for Digital Transformation in Healthcare GmbH, Witten, Germany
{marcel.rebbert,andre.nemat}@transforming-healthcare.com
[6] T-Systems International GmbH, Digital Solutions - Data Intelligence Hub, Frankfurt, Germany
christop.schlueter-langdon@telekom.de, ralf.konrad@t-systems.com
[7] Claremont Graduate University, Drucker School of Management, Claremont, USA
chris.langdon@cgu.edu
[8] KASTEL Security Research Labs, Karlsruhe, Germany
sunyaev@kit.edu
[9] Witten/Herdecke University, Faculty of Health/School of Medicine, Witten, Germany
sven.meister@uni-wh.de

Abstract. In the emerging information economy, data evolves as an essential asset and personal data in particular is used for data-driven business models. However, companies frequently leverage personal data without considering individuals' data sovereignty. Therefore, we strive to strengthen individuals' position in data ecosystems by combining concepts of data sovereignty and data economy. Our research design comprises an approach to design thinking iteratively generating, validating, and refining such concepts. As a result, we identified ten areas of tension that arise when linking data sovereignty and data economy. Subsequently, we propose initial solutions to resolve these tensions and thus contribute to knowledge about the development of fair data ecosystems benefiting both individuals' sovereignty and companies' access to data.

Keywords: Data Sovereignty, Data Economy, Data Ecosystem, Personal Data, Areas of Tension

17th International Conference on Wirtschaftsinformatik,
February 2022, Nürnberg, Germany

1 Introduction

As the recent development of the global economy shows, the importance of data, particularly personal data, is constantly rising and thus data increasingly evolves into an asset [1]. Despite this significant increase of relevance, data is still left out of scope frequently when considering assets from an industrial perspective. However, data has been growing in volume and data as an asset is slowly seeking the attention it deserves [2–4]. In our digital age, we witness the emergence of information economies and societies depicting that digitalization effects both companies and individuals. The Federal Statistical Office of Germany regularly evaluates the maturity of digitalization and integration of technology into everyday life of German citizens [5]. Their surveys revealed that 92% of German citizens use the internet every day or almost every day. Furthermore, 55% of the internet users were active in social networks [5]. Social networks are a common example of data economies as they provide a platform for individuals to share their personal data for the purpose of connecting with others. There are several dominant platform companies already exploiting the potential of their data economies effectively, e.g., Airbnb, Amazon, Facebook, Google, or Uber. The particular business model of these platform providers is to accumulate huge amounts of personal data from their users, to entangle this data, and subsequently generating data-driven business models (e.g., personalized advertising) [2]. However, hyperscalers usually provide insufficient possibilities for the individuals of whom the data is from to manage their personal data sovereignly. Thus, current platforms lack to combine their data economies with aspects of data sovereignty, such as tools and systems for digital rights management or personal information management enabling individuals to self-determine both access and usage of their personal data by third parties [6, 7].

This paper provides insights into areas of tension arising when interweaving the paradigms data economy and data sovereignty. Our work is based on a position paper [8] but goes beyond that by explaining the research methodology, comprehensively introducing the areas of tension, and presenting initial solutions. Our ultimate objective is to support companies in developing data-driven business models by appropriately considering individuals' needs and entitlements of related to their personal data. Thus, our research contributes to current discussions of personal data as a post-industrial opportunity and considerations to (re-)build fair data ecosystems with individuals actively involved. Consequently, we address the following research questions (RQ):

RQ1: *Which areas of tension arise between data sovereignty and data economy given the premise that the needs of both companies and individuals are considered?*

RQ2: *What could be promising solution approaches to solve the identified tensions?*

The paper is structured as follows. In Section 2, we provide a brief theoretical overview in terms of the domains data sovereignty and data economy encompassing the related fields of data ethics and data rights. In Section 3, we outline our research methodology consisting of an adapted design thinking process. The resulting areas of tension are presented in Section 4. Subsequently, we discuss our results and propose initial promising solution approaches to handle the identified challenges. Lastly, we describe our main contributions, appreciate our limitations, and conclude with recommendations for future research. Our research contributes to the development of

fair data ecosystems by promoting individuals' data sovereignty and sensitizing both individuals and companies for the value of (personal) data as well as its responsible handling. We provide essential, although non-conclusive, recommendations for action in terms of politics, society, and technology to ensure data sovereignty of individuals in data ecosystems for the long-term benefit of both individuals and companies.

2 Theoretical Foundation

To foster an understanding for the opportunities and challenges potentially arising when linking data sovereignty and data economy, we clarify these essential terms. Since our research indicated that further domains influence their relation, we also included data ethics and data rights in our theoretical part. We share the view that a holistic interaction of different disciplines enables informational self-determination for individuals [9].

We interpret the term **Data Sovereignty** as a branch of digital sovereignty of individuals and companies and thus also of informational self-determination, which explicitly focuses on data [10–12]. In the context of data usage, issues in the area of hardware, software, and infrastructures are commonly addressed [13], albeit the term data sovereignty relates to the data itself [11]. Even though the notion of data sovereignty is yet not uniformly defined in literature [14], we tend towards a definition resembling privacy as control of communicating information within the context of current data-sharing platforms [15]. To this end, we state that data sovereignty involves making independent, controlled, and self-determined decisions about what happens to one's own data [16]. On the one hand, such considerations lead to individuals being able to view, store, track, and delete their personal data. On the other hand, companies are encouraged to incentivize personal data sharing of individuals in a self-determined way, because they want to exchange, share, and use individuals' personal data. This situation points out the connection to the data economy. Nevertheless, since current solutions do not generally guarantee individuals' data sovereignty, strictly blocking data sharing appears to be the safest way to maintain data sovereignty [17]. However, fair value creation from personal data involves individuals enabled to participate of the economic recovery potential gained from their data. Furthermore, data sovereignty comprises the knowledge of who can access individuals' data and where this data is transferred [18]. Thereby, an important aspect in practice is certainly the condition that such determinations are also enforced by the system used [19], e.g., by a policy, referable to as policy enforcement [19–21]. This implies to guarantee the implementation of control mechanisms required by the system as a prerequisite to permit data sovereignty for all actors. Conclusively, we define data sovereignty as the ability to decide in a self-determined way, at any time, and by means of preferences, which entity can use one's own (personal) data for selected purposes.

Data is already considered as an economic asset representing the basis to develop entirely new digital business models [2, 22, 23]. Platforms applying such digital business models, like Amazon, Facebook, or Uber, benefit from a large amount of data [24]. The ability to generate, collect, analyze, process, and link data creates a **Data Economy** which is definable as a market trading with data [3]. The tremendous amount

of generated data is certain to increase exponentially in the future, accelerating the emergence of data ecosystems [25]. Naturally, personal data plays an important role in this process since this kind of information is increasingly applied to develop personalized products and services tailored to the individual. Noteworthy, an area of growing importance of the data economy is pricing personal data since a company is increasingly encouraged to reward individuals for sharing personal information due to their rising awareness for its value [26]. Thereby, individuals can be incentivized (non) monetarily. Both variants strengthen the position of the individual in data ecosystems.

Ethical issues also play an important role in linking of data sovereignty and data economy. **Data Ethics** is commonly considered as a subset of ethics examining and evaluating moral problems in data access and use [27, 28]. Data ethics aims to identify possible solutions to these moral problems and, consequently, to define a responsible handling of data [27, 28]. Companies are often criticized for misusing personal data and sharing it without sufficient consent given by individuals, causing mistrust among the latter. An example is the scandal of Facebook and Cambridge Analytica [29] which drew the attention of media and public towards data-driven companies aiming to use personal data for profit maximization. Discussions on data collection and usage practices appeared on blogs, social media platforms, and political discussions [30]. As a result, this scandal showed the critically of data economy and its mismatch between profit maximization and individuals' data sovereignty. To prevent data abuse and the subsequent creation of mistrust among individuals, the concept of corporate digital responsibility (CDR) arose following the example of corporate social responsibility [31]. CDR describes principles for the responsible handling of data by companies. Although legislation already defines provisions for handling personal data, for instance, the GDPR in Europe [32], but data ethics and CDR frequently exceed data protection. In addition, common values such as autonomy, transparency, responsibility, or explicability are considered important ethical pillars in today's (information) society [33]. Even though the criteria for the responsible use of digital technologies are already established, the hurdle exists in transferring these values to digital products or services. Likewise, individuals need to gain insights into a responsible handling of data to make ethical decisions related to their own data. In this context, the digital literacy describes not only the ability to use digital media but also aspects such as how to handle data responsibly or building awareness for data protection and safety issues [34–36]. Further, an important part of digital literacy is also the ability to understand one's role in data ecosystems and to assess threats and opportunities arising from being involved in these systems [37]. We consider this ability a crucial premise for data sovereignty.

Data Rights deal with legal issues related to data. Currently, there is no legal basis for ownership [16] or exclusive right to data in many countries. In German jurisdiction, according to § 90 of the German Civil Code, the owner of a non-physical object cannot be determined. However, a right to data can be granted selectively. A popular legal case is a database owner who 'owns' a self-created database according to §§ 87a ff. UrhG [38]. Furthermore, data can be the subject of contracts under the law of obligations [39]. This enables the creation of data markets where participants can offer, sell, and share their data, as well as obtain data themselves. If data is traded, the purchaser neither becomes the owner, due to a lack of data ownership regulations nor receives an absolute

legal position [40], and thus merely gains data access. When personal data is considered, the legal situation is determined by distinct laws for data protection [32]. For instance, in Europe, the GDPR provides binding instructions and restrictions for the legitimate handling of personal data. Vital restrictions are that, firstly, consent of a citizen must be obtained before processing its data and, secondly, the corresponding processing procedure must be dedicated to a specific purpose. Hence, the GDPR lays the foundation for informational self-determination and thus for the individuals' data sovereignty. However, the GDPR does not define an ownership right as well.

3 Research Methodology

To answer the RQs, we opted for a **Design Thinking** approach. Design thinking can be described as a strategy to solve complex questions with the aid of multidisciplinary researchers [41–43]. In our study, an important advantage of design thinking is empathizing with the role of the target group [44], i.e., we focus on the role of individuals and companies. Furthermore, various process models of design thinking exist, facilitating the formation of our research design. A common model by HPI School of Design Thinking comprises an iterative process consisting of six phases with the opportunity to return to previous phases: understand, observe, define the point of view, ideate, prototype, and test [45]. We altered this original design by transforming the stated phases as follows to adjust the approach to our research purposes: *Awareness Building*, *Knowledge Building*, *Point of View*, *Ideate*, *Concept Development*, and *Validation*. These adaptations were essential due to our approach to a conceptual research methodology instead of a rather technical procedure commonly used in design thinking. Specifically, we applied literature analyses in the first two phases while changing phases five and six from a prototyping focus to concept developments and plenum discussions. The adapted iterative approach is shown in Figure 1 with an embedded loop from the last to the first phase to support an agile research process.

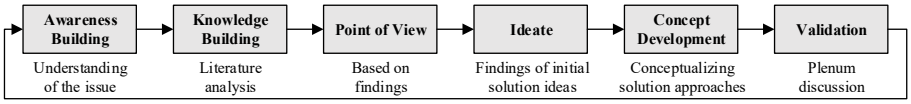

Figure 1. Methodology based on a design thinking process with six phases [45]

Our research group consisted of 14 members with different research foci, encompassing the authors and three additional participants. We classified those foci as follows: data sovereignty (6 members), data economy (4), data ethics (3), and data rights (1). Each researcher passed our methodology at least once but typically multiple times to find appropriate concepts contributing our RQs. By starting at *Awareness Building*, the researchers developed their own understanding for possible tensions related to our domains described in Section 2. Subsequently, in the second phase, the researchers applied individual methods of *Knowledge Building*, typically a structured or unstructured literature analysis within their research fields, to acquire comprehensive information about potential tensions. This phase deviated among researchers, due to

differences in terms of existing expertise. After that, each researcher formed an own *Point of View*, based on the accumulated findings from previous phases. Consequently, in the phase of *Ideate*, the researchers relied on their rather subjective insights gained to generate various ideas. These first ideas were transformed into concrete areas of tension and appropriate solution approaches in the following *Concept Development* phase. Each researcher handled the five phases explained so far independently. However, the final phase of *Validation* was carried out in focus groups with all members involved. In this context, we adapted the parameters to guide focus groups based on recommendations of Merton and Kendall [46]. This particularly includes the necessity to address the existence of specific experiences of participants in the focus group about the topic under investigation and their systematic exploration [46]. Noteworthy, the applied method is suited if the researcher aims to obtain a multitude of eventually divergent perspectives about a selected topic in order to capture the issue at hand as holistically as possible [47, 48]. This was especially useful in our *Validation* phase. We gained the following advantages of focus groups for our research. Firstly, we accumulated insights into different understandings of researchers related to a given topic in terms of specific areas of tension and solution approaches. Secondly, researchers could alter their initial understandings due to plenum discussions [49]. The discussion with all members involved in the *Validation* phase resulted in new findings within the interdisciplinary research group in each iteration. After an iteration, members returned to the first phase of *Awareness Building* considering insights from the focus group. In the plenum discussions, we used recommended collaborative tools such as digital whiteboards or presentations to support our research process [44]. After the plenum discussion in the third iteration, our research process terminated as there was a high perceived congruence among all researchers in terms of identified concepts for areas of tensions and solution approaches.

4 Results

Through multiple iterations of our research methodology, we conceptualized ten areas of tension arising in the intersection of data economy and data sovereignty. Our final results are briefly summarized in Table 1, representing our answer to RQ1. In the following, we explicate them in more detail.

If a subject cannot be identified within a cluster, the subject has **Anonymity** in this particular cluster [50]. Due to legal restrictions for handling personal data [32], many companies anonymize personal data to be less restricted in processing and, consequently, enhance their data availability. By unlinking the connection between individuals and their data, it appears to be complicated to remunerate the individual after several processing steps. This aspect is conflictive with our notion of data sovereignty, since it requires the individuals being able to orchestrate their data stored in data ecosystems or any other environment provided by the data processor.

In communication technology, a carrier-wave is defined by a certain frequency modulated with a signal to transmit encoded data. This physical principle can be observed in data sharing processes. That is, since every individual behaves similar to a

carrier-wave transmitting various information intentionally or unintentionally, frequently sharing more data than actually intended. We refer to this phenomenon as **Carrier-Wave Principle** [51]. Accordingly, companies can extract more information of personal data than individuals have intended and are aware of, resulting in the disclosure of information which they might not want to reveal (e.g., information from genome data [52]). Additionally, new technology evolves that is increasingly able to leach more information from the same data than past technologies. As a consequence, it becomes more difficult to predict which kind of information can be gained from a given dataset. Hence, a data sharing decision can be compromised in the future.

Table 1. Areas of Tension

Area of Tension	Description
Anonymity	Companies tend to anonymize data to avoid legal restrictions. This can be conflictive with our notion of data sovereignty since the data source is unknown and cannot be remunerated for sharing its data.
Carrier-Wave Principle	Sharing data indirectly provides more information than initially intended due to future technological progress [51]. Thus, companies extract more information than citizens are aware of.
Data Processing	Most companies offer individuals little to no insights into or influences in their data processing procedures [53, 54].
Intangibility	Since data is not a physical asset, it requires separate economic and ownership consideration [55].
Lock-In Effects	Companies want to lock users in their ecosystems [54, 55], e.g., due to the number of users denotes an important aspect for the success. Hence, companies are incentivized to limit citizens' ability to switch platforms or services sovereignly.
Manipulation	To increase profit margins, companies might manipulate individuals' sovereign behavior in order to align it with company objectives [54, 58], also referable to as nudging [59].
Mistrust	Emerging knowledge about manipulation and data scandals result in mistrust among individuals. Hence, individuals could hesitate to share their data in data ecosystems [60].
Privacy Paradox	It describes the phenomenon of individuals stating a claim to data privacy or protection, albeit sharing carefree personal data without concerns [61]. Companies benefit from the privacy paradox because they receive more data than intended.
Responsibility	The use of data implies responsibility for the handling of this data [27]. Consequently, data sovereignty requires knowledge about one's own data and tactics for responsible handling. However, most companies do not take over this responsibility, although they use citizens' data.
Unraveling Effects	Data sovereignty of one individual can have implications for another individual's data sovereignty or influence the ability of sovereignly deciding upon personal data sharing in data ecosystems [16].

In the context of **Data Processing**, in data collection as the initial phase, data is made accessible and is generated by companies [62]. Analytical procedures, especially artificial intelligence as a service [63] or deep learning algorithms [64], require a pool of high qualitative data. Subsequently, in the phase of information creation, information is extracted from data relying on the created data collection [62]. Lastly, in the phase of value creation, the extracted information is shared, combined, and used by data consumers to develop novel business models built upon product and service innovations which basis is given by the extracted information [62]. However, most platform ecosystems offer individuals hardly any insights into their data processing operations [53, 54]. Consequently, both sovereignty in and transparency of data usage are highly limited from the individual's perspective.

Another challenge is **Intangibility** of data. Since data is not a physical asset, it is denoted as intangible and thus requires separate economic and ownership consideration [55]. For instance, data can be copied, used multiple times by different actors without any depreciation in value, and is neither affected by wear nor aging [55]. Moreover, value of data depends on its timeliness, which means that outdated information is usually worthless. This implies that data ecosystems must define boundaries within the shared data can only be used as contracted. Furthermore, there must be a link between individuals and their data in order to be sovereign and remunerated. To this end, clearly defined ownership is required from a legal perspective, but there is yet no property right on data [65]. This ambiguity in terms of data rights impedes data platforms from growing even faster than they already do.

Lock-In Effects refer to the unwillingness or the inability of users, respectively, to switch services, platforms, or products sovereignly due to high switching costs [56, 57]. These costs are not just comprised by money, but also encompass time expenditures or general platform issues whereas the latter can be subdivided into three main aspects: the number of platform users not reaching a critical amount (i.e., less personal network effects), the risk of losing combined information of shared data over time, and a low degree of interoperability with respect to alternative solutions [17, 66]. Thereby, the number of users denotes an important aspect for platform success, since the value of information is mainly determined by the ability to link different pieces of data. To this end, companies want to bind users to their platforms or services. Using various services from the same company, the generated values are rising and, as a result, lock-in effects are amplified [67]. Therefore, data sovereignty is compromised when high switching costs impede users from moving to other platforms or service providers.

In order to maximize the revenue generated from monetizing data, companies have an incentive to collect as much data as possible from citizens [67]. This predisposition may tempt companies to the **Manipulation** of individuals' behavior in order to align it with company objectives [54, 58, 68], referable to as nudging [59]. Hence, manipulation contradicts informational self-determination and data sovereignty. As long as companies have a strong incentive to manipulate individuals' decision-making for the sake of collection more data, data sovereignty is impaired. A common example for user manipulation is to place a privacy notice on a website leading to the conviction of users that the website's data protection standard is reliable [69].

The success of data ecosystem depends on the amount of data gathered [22, 23, 67]. Consequently, companies have an interest in citizens continuously contributing to their ecosystem. However, emerging knowledge about manipulation techniques results in **Mistrust** among individuals as shown, for instance, by the Facebook and Cambridge Analytica scandal [70]. This misconduct in handling personal data attained high popularity and pointed out the need for enhanced protection of personal data. Such scandals increase individuals' distrustfulness and thus strengthen reluctance concerning sharing personal data [60]. Ultimately, mistrust hinders the success of data ecosystems depending on high amounts of data being (sovereignly) shared.

A situation in which an individual states a claim to data privacy and data protection, respectively, albeit sharing personal data at the same time without concerns is referred to as the **Privacy Paradox** [61]. There are several reasons why this phenomenon exists [71]. Firstly, it is manifested in the human behavior by weighing both the benefits and the risks of data disclosure, commonly referred to as privacy calculus [72]. Secondly, cognitive biases influence the assumed risk-benefit calculation of users [73, 74]. For example, users apply mental models favoring benefits [75]. These mental models are heuristics and, for instance, include shortcuts to ease decision-making in situations perceived by the individual as being too extensive for rational risk-benefit calculation. Thirdly, users might not have enough information to behave in a privacy-preserving way to protect their personal data [76]. The privacy paradox describes a tension, because it indicates individuals do not act carefully with their data and therefore, exercising data sovereignty may overwhelm them. Additionally, companies might want to perpetuate this phenomenon, so they still receive data from citizens who would not share data if more information about the data usage were available. Hence, the privacy paradox favors data economy over data sovereignty.

The orchestration of data by an entity involves the entity's **Responsibility** for the handling of this data. But data scandals show that companies are not fulfilling this responsibility [29]. However, the GDPR determines how companies have to handle personal data and thus assigns them the responsibility for the safety and security of personal data used [32]. In addition, individuals must be aware of how to handle own data responsibly if they want to be sovereign in data ecosystems.

Finally, we define **Unraveling Effects** as an individuals' data sovereignty compromising the data sovereignty of another individual [16]. For instance, sharing health information with an insurance company in exchange for lower monthly fees leads to a more detrimental price model for other consumers. In an extreme case, there are individuals in a data ecosystem who disclose their entire personal data to optimize their economic profit while other individuals do not reveal their personal data at all. Subsequently, data platform providers possessing personal data of the first group can analyze this data in a way to infer characteristics about a second group by means of generalizing results from a representative sample of data [77]. Hence, a platform provider can bypass data sovereignty of individuals by inferring their characteristics from using data of others. Ultimately, an individual may feel compelled to share data because not sharing data might directly results in detrimental treatment [77].

5 Discussion and Implications

To strengthen the individual's position in data ecosystems for the benefit of both individuals and companies, solutions are required resolving or at least emasculating the identified areas of tension. Therefore, we propose initial solution approaches and answer RQ2. An overview of our solution approaches is shown in Table 2.

Table 2. Solution Approaches

Area of Tension	Solution Approaches
Anonymity	pseudonymization, raw data sharing
Carrier-Wave Principle	contracts, expiration date, policies
Data Processing	certification, data governance, data provenance, policies, remuneration, transparency
Intangibility	contracts, data ownership, data provenance, policies
Lock-In Effects	data portability
Manipulation	certification, code of conduct, policies, transparency
Mistrust	certification, policies, transparency
Privacy Paradox	digital literacy, UX design
Responsibility	certification, digital literacy, GDPR, transparency
Unraveling Effects	digital literacy, transparency

When considering the challenges of **Anonymity**, there is a less restrictive mechanism called pseudonymization [78, 79]. In pseudonymization, an intermediary instance is applied to assign data to the original data source while, at the same time, enabling data processing for the data consumer. This intermediary instance might be a data trustee or the platform owner. With the concept of pseudonymization, individuals can be identified and thus remunerated for their shared data. Furthermore, empowering individuals to sharing their raw data sovereignly is also an opportunity.

The **Carrier-Wave Principle** can be solved by enabling the deletion of data by the individual or adding an expiration date for data, which might be anchored in usage policies included in the aforementioned contracts. Expiration dates enable the automatic deletion of shared data after a certain period. Therefore, a time constraint on data access and usage prevents uncertainty concerning future technologies.

To guarantee data sovereignty for the individual in **Data Processing** and to remunerate the data sharing, it is necessary to consider tracing back the individual as initial data source. However, the increasing number of refinement steps within a data-driven value chain complicates proper data provenance tracking [80]. Additionally, the continuous sharing and refinement of data may influence the ownership role, because the claim of the individual on the final product after multiple refinements is questionable. We ascertained that a promising solution must consider mechanisms to trace back and remunerate the individual that contributed to the final product for as long as possible. Furthermore, data governance aspects must define processing steps, where remuneration claims are transferred. By means of such mechanisms, the individual as data source receives remuneration for its data but may lose claims after specific

refinement steps. Hence, it is vital to ensure transparency in data processing and enforcement of individuals' preferences and policies. A trustworthy authority can certify data processing steps to foster transparency and trust.

Intangibility addresses a missing property right on data [65]. This lack of legal clarity results in (digital) contracts as the main opportunity to systematically share data. In such contracts, both the data source and the data consumer negotiate data usage and access policies applied to the underlying data, but without the data recipient becoming the data owner from a legal perspective. Both parties have to be complaint to the conditions they agreed upon, so that enable the data consumer to use the purchased data. We identified literature suggesting a simple kind of immaterial property right (e.g., data ownership) as an enabler for individuals' data sovereignty in data ecosystems [81] based on clearly defined regularities allowing for consistently tracking the data source. Nevertheless, we suggest contracts as mandatory starting points of data sharing.

Lock-In Effects hinder individuals from changing platforms. A solution to this area of tension requires data sharing ecosystems emphasizing transparency, availability, openness, and, consequently, mobility of data across its boundaries (i.e., data portability [32, 82]). Transferring information by connecting data to another platform is complicated, but easier in terms of an interoperable and standardized approach to data portability. Thus, data portability represents a solution to reduce switching costs and to foster data sovereignty in data economies.

To avoid **Manipulation** by companies, data ecosystems need to provide a high level of transparency in data processing to comprehend promises of data sovereignty. A system must ensure that the conditions and obligations attached to data by individuals are obeyed and data processing entities adhere to designed codes of conducts. This could implemented by means of certifications and usage policies.

Enabling individuals to view, store, track, alter, and delete their data being stored by on platforms permanently and consistently, strengthens data sovereignty and, consequently, reduces **Mistrust**. Such transparency is an opportunity for companies to build trust with individuals since they trust ecosystems that clearly communicate how data sovereignty is ensured [27, 83]. Another possibility is to use certified control mechanisms (e.g., policies) or infrastructures provided by trustworthy authorities to build trust among individuals. Finally, we state that companies must build trust of individuals to create a successful data ecosystem, while effectively counteracting the corresponding reluctance of sharing data by means of appropriate methods.

To avoid the **Privacy Paradox**, we suggest fostering individuals' digital literacy to increase their awareness about the intrinsic value of their personal data. Since the mere facilitation of digital literacy is surely not enough to strengthen data sovereignty of individuals decisively, we propose novel and user-friendly applications. Examples are clearly arranged user interfaces (i.e., UX design) and fine granular consent mechanisms to manage data sharing, e.g., by means of privacy icons or similar user aids.

Regarding the identified area **Responsibility**, certifications by trusted third parties can be a mechanism to demonstrate efforts on responsibilities of companies. However, knowledge about a responsible handling of one's own data is required to provide individuals with tools that empower them to be sovereign over their data. Considering a holistic approach to data sovereignty of individuals, skills of digital literacy are

essential as well, as individuals are currently hardly aware of the value concerning their personal data [84]. Such an awareness is important for making responsible decisions about data. To this end, individuals must be properly informed about implications accompanied with sharing (personal) data.

Since the identification of the types of data affected by **Unraveling Effects** is challenging, a consideration is required which data can be shared safely by an individual and when consent of other individuals might be required prior to sharing [85]. Basically, providers of data ecosystems should point out to the individuals the implications of disclosing their own data. After that, it is in the responsibility of both companies and individuals to handle data in an ethically and morally manner. However, avoiding unraveling effects entirely seems impossible, since both inside and outside of data ecosystems decisions made by individuals affect others.

6 Conclusion and Outlook

We identified challenges arising when bringing together data sovereignty and data economy, resulting in ten distinct areas of tension. These areas represent hurdles for the emergence of fair data ecosystems. Though, we state that data sovereignty and data economy are linkable if a platform succeed in implementing appropriate measures. To support in the conceptualization of such measures, we propose initial solution approaches that represent first recommendations for action in terms of politics, economy, society, and technology in order to resolve the identified issues. We contribute to data ecosystem research as we identified challenges for individuals' data sovereignty and provide a set of conceivable directions to search for possible solutions.

Since we focused on the domains introduced in Section 2 to identify areas of tension, our research is limited in terms of considered literature. This naturally results in both a topical bias and a certain degree of incompleteness regarding our results. Furthermore, we faced subjectivity issues in concept elicitation from literature. However, we counteracted this problem by conducting plenum discussions and validating concepts in focus groups. In addition, we mainly relied on desk research, which inevitably means that our results built on what was publicly available. Though, we state that the limitations do not diminish the validity and meaningfulness of our results to a considerable extent albeit they emphasize that we only provided a snapshot on the broad topics of data sovereignty and data economy.

Future research should focus on examining the areas of tension in-depth to elaborate more concrete solution approaches to develop data ecosystems perceived as fair from the viewpoint of individuals and companies. Ultimately, the objective must be to entangle data sovereignty and data economy in practice and to contribute to a trustworthy, liberal, and fair information society.

7 Acknowledgement

This research was carried out in the DaWID project funded by the German Federal Ministry of Education and Research (BMBF; funding reference number: 16SV8381).

References

1. Dong, X., Guo, B., Duan, X., Shen, Y., Zhang, H., Shen, Y.: DSPM: A Platform for Personal Data Share and Privacy Protect Based on Metadata. In: 2016 13th International Conference on Embedded Software and Systems (ICESS), pp. 182–185. Conference Publishing Services, IEEE Computer Society, Los Alamitos, CA (2016). doi: 10.1109/ICESS.2016.10
2. World Economic Forum: Personal Data: The Emergence of a New Asset Class. Geneva, CH (2011)
3. S. Oliveira, M.I., Barros Lima, Glória de Fátima, Farias Lóscio, B.: Investigations into Data Ecosystems: a systematic mapping study. Knowledge and Information Systems, vol. 61, 589–630 (2019). doi: 10.1007/s10115-018-1323-6
4. Opher, A., Chou, A., Onda, A. and Sounderrajan, K.: The Rise of the Data Economy: Driving Value through Internet of Things Data Monetization. A Perspective for Chief Digital Officers and Chief Technology Officers (2016), https://www.ibm.com/downloads/cas/4JROLDQ7, (Accessed: 29.04.2021)
5. Federal Statistical Office of Germany (Destatis): Private Haushalte in der Informationsgesellschaft - Nutzung von Informations- und Kommunikationstechnologien. Fachserie 15 Reihe 4 - 2020 (2021), https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/IT-Nutzung/Publikationen/Downloads-IT-Nutzung/private-haushalte-ikt-2150400197004.pdf, (Accessed: 05.11.2021)
6. Liu, Q., Safavi-Naini, R., Sheppard, N.P.: Digital Rights Management for Content Distribution. In: Proceedings of the Australasian Information Security Workshop Conference on ACSW Frontiers 2003 - Volume 21. ACSW Frontiers '03, pp. 49–58. Australian Computer Society, Inc, AUS (2003)
7. Jones, W.: Personal Information Management. Ann. Rev. Info. Sci. Tech., vol. 41, 453–504 (2007). doi: 10.1002/aris.2007.1440410117
8. Lauf, F., Scheider, S., Meister, S., Radic, M., Herrmann, P., Schulze, M., Nemat, A.T., Becker, S.J., Rebbert, M., Abate, C., et al.: Data Sovereignty and Data Economy—Two Repulsive Forces? Fraunhofer Institute for Software and Systems Engineering ISST, Dortmund (2021). doi: 10.24406/isst-n-634865
9. Meister, S., Otto, B.: Digital Life Journey. Framework for a self-determined life of citizens in an increasingly digitized world (basic research paper). Dortmund (2019). doi: 10.24406/ISST-N-559377
10. Adonis, A.A.: Critical Engagement on Digital Sovereignty in International Relations: Actor Transformation and Global Hierarchy. Global: Jurnal Politik Internasional, vol. 21, 262–282 (2019). doi: 10.7454/global.v21i2.412
11. Couture, S., Toupin, S.: What does the notion of "sovereignty" mean when referring to the digital? New Media & Society, vol. 21, 2305–2322 (2019). doi: 10.1177/1461444819865984
12. Boris Otto: Digitale Souveränität: Beitrag des Industrial Data Space. Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V., Munich (2016). doi: 10.13140/RG.2.2.35125.68321
13. Aydin, A., Bensghir, T.K.: Digital Data Sovereignty: Towards a Conceptual Framework. In: 2019 1st International Informatics and Software Engineering Conference (UBMYK), pp. 1–6 (2019). doi: 10.1109/UBMYK48245.2019.8965469

14. Hummel, P., Braun, M., Tretter, M., Dabrock, P.: Data sovereignty: A review. Big Data & Society, vol. 8, 205395172098201 (2021). doi: 10.1177/2053951720982012
15. Westin, A.F.: Privacy And Freedom. Washington and Lee Law Review, vol. 25 (1968)
16. Hummel, P., Braun, M., Augsberg, S., Dabrock, P.: Sovereignty and Data Sharing. ITU Journal: ICT Discoveries, vol. 1 (2018)
17. Filippi, P. de, McCarthy, S.: Cloud Computing: Centralization and Data Sovereignty. European Journal of Law and Technology, vol. 3 (2012)
18. Posch, R.: Digital sovereignty and IT-security for a prosperous society. In: Werthner, H., van Harmelen, F. (eds.) Informatics in the Future, pp. 77–86. Springer, Cham, CH (2017). doi: 10.1007/978-3-319-55735-9_7
19. Bartsch, J., Dehling, T., Lauf, F., Meister, S., Sunyaev, A.: Let the Computer Say NO! The Neglected Potential of Policy Definition Languages for Data Sovereignty. In: Friedewald, M., Kreutzer, M., Hansen, M. (eds.) Selbstbestimmung, Privatheit und Datenschutz. Gestaltungsoptionen für einen europäischen Weg. Springer Fachmedien Wiesbaden (2022)
20. Han, W., Lei, C.: A survey on policy languages in network and security management. Computer Networks, vol. 56, 477–489 (2012). doi: 10.1016/j.comnet.2011.09.014
21. OASIS: eXtensible Access Control Markup Language (XACML) Version 3.0, http://docs.oasis-open.org/xacml/3.0/xacml-3.0-core-spec-os-en.html, (Accessed: 03.08.21)
22. Nachira, F., Nicolai, A., Dini, P.: Digital business ecosystems. Publications Office, Luxembourg (2007)
23. Otto, B., Lis, D., Jürjens, J., Cirullies, J., Opriel, S., Howar, F., Meister, S., Spiekermann, M., Pettenpohl, H., Möller, F.: Data Ecosystems. Conceptual Foundations, Constituents and Recommendations for Action. Fraunhofer Institute for Software and Systems Engineering ISST, Dortmund (2019). doi: 10.13140/RG.2.2.35125.68321
24. Piepenbrink, J.: Datenökonomie. Aus Politik und Zeitgeschichte, vol. 69 (2019)
25. Azkan, C., Goecke, H., Spiekermann, M.: Forschungsbereiche der Datenökonomie. Wirtschaftsdienst, vol. 100, 124–127 (2020). doi: 10.1007/s10273-020-2582-x
26. Li, C., Li, D.Y., Miklau, G., Suciu, D.: A Theory of Pricing Private Data. ACM Transactions on Database Systems, vol. 39, 1–28 (2014). doi: 10.1145/2691190.2691191
27. Floridi, L., Taddeo, M.: What is data ethics? Philosophical transactions. Series A, Mathematical, physical, and engineering sciences, vol. 374 (2016). doi: 10.1098/rsta.2016.0360
28. German Data Ethics Commission: Opinion of the Data Ethics Commission (2019), https://datenethikkommission.de/wp-content/uploads/DEK_Gutachten_engl_bf_200121.pdf, (Accessed: 29.04.2021)
29. Hu, M.: Cambridge Analytica's black box. Big Data & Society, vol. 7 (2020). doi: 10.1177/2053951720938091
30. González, F., Yu, Y., Figueroa, A., López, C., Aragon, C.: Global Reactions to the Cambridge Analytica Scandal: A Cross-Language Social Media Study. In: Liu, L., White, R. (eds.) Companion Proceedings of The 2019 World Wide Web Conference, pp. 799–806. ACM, New York, NY, USA (2019). doi: 10.1145/3308560.3316456
31. Lobschat, L., Mueller, B., Eggers, F., Brandimarte, L., Diefenbach, S., Kroschke, M., Wirtz, J.: Corporate digital responsibility. Journal of Business Research, vol. 122, 875–888 (2021). doi: 10.1016/j.jbusres.2019.10.006

32. European Parliament and Council of European Union: Regulation (EU) 2016/679 (2016), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2016:119:TOC, (Accessed: 29.04.2021)
33. Becker, S.J., Nemat, A.T., Rebbert, M.: Der Ruf nach operationalisierbarer Ethik – Unternehmensverantwortung in der digitalen Welt. In: Bertelsmann Stiftung, Wittenberg-Zentrum für Globale Ethik (eds.) Unternehmensverantwortung im digitalen Wandel. Ein Debattenbeitrag zu Corporate Digital Responsibility, pp. 28–34 (2020)
34. Gilster, P.: Digital literacy. John Wiley, New York, Chichester (1997)
35. Iordache, C., Mariën, I., Baelden, D.: Developing Digital Skills and Competences: A Quick-Scan Analysis of 13 Digital Literacy Models. Italian Journal of Sociology of Education, vol. 9, 6–30 (2017). doi: 10.14658/pupj-ijse-2017-1-2
36. Reddy, P., Sharma, B., Chaudhary, K.: Digital Literacy. International Journal of Technoethics, vol. 11, 65–94 (2020). doi: 10.4018/IJT.20200701.oa1
37. Grefen, P.: Digital Literacy and Electronic Business. Encyclopedia, vol. 1, 934–941 (2021). doi: 10.3390/encyclopedia1030071
38. Mitterer, K., Wiedemann M., Zwissler, T.: BB-Gesetzgebungs- und Rechtsprechungsreport zu Industrie 4.0 und Digitalisierung (2017)
39. Palandt, O., Ellenberger, J., Götz, I., Grüneberg, C., Herrler, S.: Bürgerliches Gesetzbuch. Mit Nebengesetzen insbesondere mit Einführungsgesetz (Auszug) einschließlich Rom I-, Rom II und Rom III-Verordnungen sowie EU-Güterrechtsverordnungen, Haager Unterhaltsprotokoll und EU-Erbrechtsverordnung, Allgemeines Gleichbehandlungsgesetz (Auszug), Wohn- und Betreuungsvertragsgesetz, Unterlassungsklagengesetz (PalHome), Produkthaftungsgesetz, Erbbaurechtsgesetz, Wohnungseigentumsgesetz, Versorgungsausgleichsgesetz, Lebenspartnerschaftsgesetz (PalHome), Gewaltschutzgesetz. C.H.Beck, München (2021)
40. Paal, B., Hennemann, M.: Big Data im Recht – Wettbewerbs- und daten(schutz)rechtliche Herausforderungen. NJW, vol. 70, 1697–1701 (2017)
41. Plattner, H., Meinel, C., Leifer, L.J.: Design Thinking. Understand - Improve - Apply. Springer, Heidelberg, London (2011)
42. Wölbling, A., Krämer, K., Buss, C.N., Dribbisch, K., LoBue, P., Taherivand, A.: Design Thinking: An Innovative Concept for Developing User-Centered Software. In: Maedche, A., Botzenhardt, A., Neer, L. (eds.) Software for People. Management for Professionals, pp. 121–136. Springer Berlin Heidelberg, Berlin, Heidelberg (2012). doi: 10.1007/978-3-642-31371-4_7
43. Wylant, B.: Design Thinking and the Experience of Innovation. Design Issues, vol. 24, 3–14 (2008). doi: 10.1162/desi.2008.24.2.3
44. Brown, T.: Design Thinking. Harvard Business Review, vol. 86, 84–92 (2008)
45. HPI School of Design Thinking: The six phases of the Design Thinking process, https://hpi.de/en/school-of-design-thinking/design-thinking/background/design-thinking-process.html, (Accessed: 05.08.21)
46. Merton, R.K., Kendall, P.: The focused interview. Chicago (1946)
47. Morgan, D.L., Krueger, R.A.: When to use focus groups and why Successful focus groups : advancing the state of the art, pp. 3–19. Sage Publications, Newbury Park, Calif (1993)
48. Stewart, D.W., Shamdasani, P.N.: Focus groups. Theory and practice. Sage, Los Angeles (2015)

49. Gibbs, A.: Focus Groups, vol. 19, 1–7 (1997)
50. Pfitzmann, A., Köhntopp, M.: Anonymity, Unobservability, and Pseudonymity — A Proposal for Terminology. In: Federrath, H. (ed.) Designing Privacy Enhancing Technologies. Lecture Notes in Computer Science, pp. 1–9. Springer, Berlin, Heidelberg (2001). doi: 10.1007/3-540-44702-4_1
51. Sinnreich, A., Gilbert, J.: The Carrier Wave Principle. AoIR Selected Papers of Internet Research, vol. 2019 (2019). doi: 10.5210/spir.v2019i0.11035
52. Thiebes, S., Dehling, T., Sunyaev, A.: One Size Does Not Fit All: Information Security and Information Privacy for Genomic Cloud Services. ECIS 2016 Proceedings, vol. (2016)
53. Sunyaev, A., Dehling, T., Taylor, P.L., Mandl, K.D.: Availability and quality of mobile health app privacy policies. Journal of the American Medical Informatics Association, vol. 22, e28-e33 (2015). doi: 10.1136/amiajnl-2013-002605
54. Zuboff, S.: The age of surveillance capitalism. The fight for a human future at the new frontier of power. PublicAffairs, New York (2019)
55. Atkinson, K., McGaughey, R.: Accounting for data: a shortcoming in accounting for intangible assets. Academy of Accounting and Financial Studies Journal, vol. 10, 85-95 (2006)
56. Liebowitz, S.J., Margolis, S.E.: Network Externality: An Uncommon Tragedy. Journal of Economic Perspectives, vol. 8, 133–150 (1994). doi: 10.1257/jep.8.2.133
57. Swire, P.P., Lagos, Y.: Why the Right to Data Portability Likely Reduces Consumer Welfare: Antitrust and Privacy Critique. Maryland Law Review, vol. 72 (2013). doi: 10.2139/ssrn.2159157
58. Acquisti, A., Brandimarte, L., Loewenstein, G.: Secrets and Likes: The Drive for Privacy and the Difficulty of Achieving It in the Digital Age. Journal of Consumer Psychology, vol. 30, 736–758 (2020). doi: 10.1002/jcpy.1191
59. Thaler, R.H., Sunstein, C.R.: Nudge. Improving decisions about health, wealth, and happiness. Penguin Books, New York (2009)
60. Rantanen, M.M.: Towards Ethical Guidelines for Fair Data Economy – Thematic Analysis of Values of Europeans Proceedings of the Third Seminar on Technology Ethics, vol. 2505, pp. 27–38. CEUR Workshop Proceedings (2019)
61. Engels, B., Grunewald, M.: Das Privacy Paradox: Digitalisierung versus Privatsphäre (2017)
62. Lim, C., Kim, K.-H., Kim, M.-J., Heo, J.-Y., Kim, K.-J., Maglio, P.P.: From data to value: A nine-factor framework for data-based value creation in information-intensive services. International Journal of Information Management, vol. 39, 121–135 (2018). doi: 10.1016/j.ijinfomgt.2017.12.007
63. Lins, S., Pandl, K.D., Teigeler, H., Thiebes, S., Bayer, C., Sunyaev, A.: Artificial Intelligence as a Service. Bus Inf Syst Eng, vol. 63, 441–456 (2021). doi: 10.1007/s12599-021-00708-w
64. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press (2016)
65. Stepanov, I.: Introducing a property right over data in the EU: the data producer's right – an evaluation. International Review of Law, Computers & Technology, vol. 34, 65–86 (2020). doi: 10.1080/13600869.2019.1631621
66. Shapiro, C., Varian, H.R.: Information rules: A strategic guide to the network economy. Harvard Business School Press, Boston, Mass. (1998)

67. Moody, D.L., Walsh, P.: Measuring the Value Of Information - An Asset Valuation Approach European Conference on Information Systems, vol. 7, pp. 496–512 (1999)
68. Waldman, A.E.: Cognitive biases, dark patterns, and the 'privacy paradox'. Current opinion in psychology, vol. 31, 105–109 (2020). doi: 10.1016/j.copsyc.2019.08.025
69. Turow, J., Hennessy, M., Draper, N.: Persistent Misperceptions: Americans' Misplaced Confidence in Privacy Policies 2003–2015. Journal of Broadcasting & Electronic Media, vol. 62, 461–478 (2018). doi: 10.1080/08838151.2018.1451867
70. Cadwalladr, C. and Graham-Harrison, E.: Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach (2018), https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election, (Accessed: 29.04.2021)
71. Barth, S., Jong, M.D. de: The privacy paradox – Investigating discrepancies between expressed privacy concerns and actual online behavior – A systematic literature review. Telematics and Informatics, vol. 34, 1038–1058 (2017). doi: 10.1016/j.tele.2017.04.013
72. Dienlin, T., Metzger, M.J.: An Extended Privacy Calculus Model for SNSs: Analyzing Self-Disclosure and Self-Withdrawal in a Representative U.S. Sample. Journal of Computer-Mediated Communication, vol. 21, 368–383 (2016). doi: 10.1111/jcc4.12163
73. Acquisti, A., Grossklags, J.: Privacy and rationality in individual decision making. IEEE Secur. Privacy Mag., vol. 3, 26–33 (2005). doi: 10.1109/msp.2005.22
74. Simon, H.A.: Models of bounded rationality: Empirically grounded economic reason. MIT Press, Cambridge (1997)
75. Pötzsch, S.: Privacy-Awareness Information for Web Forums: Results from an Empirical Study. In: Hvannberg, E.T. (ed.) Proceedings of the 6th Nordic Conference on Human-Computer Interaction. Extending Boundaries, pp. 363–372. ACM, New York, NY (2010). doi: 10.1145/1868914.1868957
76. Meynhardt, T.: Public Value Inside: What is Public Value Creation? International Journal of Public Administration, vol. 32, 192–219 (2009). doi: 10.1080/01900690902732632
77. Peppet, S.R.: Unraveling privacy: The personal prospectus and the threat of a full-disclosure future. Nw. UL Rev., vol. 105, 1153 (2011)
78. Neubauer, T., Heurix, J.: A methodology for the pseudonymization of medical data. International journal of medical informatics, vol. 80, 190–204 (2011). doi: 10.1016/j.ijmedinf.2010.10.016
79. Stella-Bourdillon, S., Knight, A.: Anonymous Data v. Personal Data - A False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data. Wisconsin International Law Journal, vol. 34 (2016)
80. Buneman, P., Khanna, S., Wang-Chiew, T.: Why and Where: A Characterization of Data Provenance. In: van den Bussche, J.E., Vianu, V. (eds.) Database theory. ICDT 2001. Lecture Notes in Computer Science, vol. 1973, pp. 316–330. Springer Berlin Heidelberg, Berlin, Heidelberg (2001). doi: 10.1007/3-540-44503-X_20
81. Riechert, A.: Dateneigentum - ein unauflösbarer Interessenkonflikt? Datenschutz und Datensicherheit, vol. 43, 353–360 (2019). doi: 10.1007/s11623-019-1121-7
82. Hert, P. de, Papakonstantinou, V., Malgieri, G., Beslay, L., Sanchez, I.: The right to data portability in the GDPR: Towards user-centric interoperability of digital services. Computer Law & Security Review, vol. 34, 193–203 (2018). doi: 10.1016/j.clsr.2017.10.003

83. Taddeo, M., Floridi, L.: The case for e-trust. Ethics and Information Technology, vol. 13, 1–3 (2011). doi: 10.1007/s10676-010-9263-1
84. Acquisti, A., John, L.K., Loewenstein, G.: What Is Privacy Worth? The Journal of Legal Studies, vol. 42, 249–274 (2013). doi: 10.1086/671754
85. Humbert, M., Trubert, B., Huguenin, K.: A Survey on Interdependent Privacy. ACM Computing Surveys, vol. 52 (2019). doi: 10.1145/3360498