information

Article

"Who Should I Trust with My Data?" Ethical and Legal Challenges for Innovation in New Decentralized Data Management Technologies

Haleh Asgarinia [1,*,†], Andres Chomczyk Penedo [2,†], Beatriz Esteves [3,†] and Dave Lewis [4]

1 Behavioural, Management, and Social Science (BMS) Faculty, Department of Philosophy, Universiteit Twente, 7522 DB Enschede, The Netherlands
2 Law, Science, Technology and Society (LSTS), Vrije Universiteit Brussel, 1090 Brussels, Belgium; andres.chomczyk.penedo@vub.be
3 Ontology Engineering Group (OEG), Universidad Politécnica de Madrid, 28006 Madrid, Spain; beatriz.gesteves@upm.es
4 ADAPT Centre, Trinity College Dublin, D02 PN40 Dublin, Ireland; delewis@tcd.ie
* Correspondence: h.asgarinia@utwente.nl
† These authors contributed equally to this work.

Abstract: News about personal data breaches or data abusive practices, such as Cambridge Analytica, has questioned the trustworthiness of certain actors in the control of personal data. Innovations in the field of personal information management systems to address this issue have regained traction in recent years, also coinciding with the emergence of new decentralized technologies. However, only with ethically and legally responsible developments will the mistakes of the past be avoided. This contribution explores how current data management schemes are insufficient to adequately safeguard data subjects, and in particular, it focuses on making these data flows transparent to provide an adequate level of accountability. To showcase this, and with the goal of enhancing transparency to foster trust, this paper investigates solutions for standardizing machine-readable policies to express personal data processing activities and their application to decentralized personal data stores as an example of ethical, legal, and technical responsible innovation in this field.

Keywords: data governance; digital age; transparency; personal data management; identity management

check for updates

Citation: Asgarinia, H.; Chomczyk Penedo, A.; Esteves, B.; Lewis, D. "Who Should I Trust with My Data?" Ethical and Legal Challenges for Innovation in New Decentralized Data Management Technologies. *Information* **2023**, *14*, 351. https://doi.org/10.3390/info14070351

Academic Editor: Haridimos Kondylakis

Received: 30 April 2023
Revised: 4 June 2023
Accepted: 19 June 2023
Published: 21 June 2023

1. Introduction

Data-driven innovations are expected to deliver further economic and societal development [1]. Through the analysis, sharing, and (re-)use of data, business models and governments' processes have been transformed to benefit from those practices [2]. The emergence of a data-driven society is being fostered by policy actions from different governments on a worldwide scale. The European Union (EU) is no exception to this, as the European Commission has put on its agenda the development of "A Europe fit for the Digital Age". The European Commission's strategy and related policy documents can be located at the following link: https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age_en (accessed on 26 May 2023). Regardless of whether it is a Big Tech company based in the United States (US), a large data broker in the EU, or a Chinese government-controlled entity, current data practices have been questioned by different societal sectors, from individuals to nongovernmental organizations (NGOs) or from academics to governments. Trust in many digital services has been compromised [3], which has left individuals asking themselves "who should I trust with my data".

In response to this trust crisis, technology has been looked upon to provide answers. Applied to the field of (personal) data, self-sovereign identity models [4] — as improvements over existing Personal Information Management Systems (PIMS) — have been put under the spotlight due to their potential, but they are also taken with "a grain of salt", as

they are not free from shortcomings [5]. Through them, users would be in direct control of their information and decide when, how, and who can access such information. Certain policy strategies, particularly in the EU, seem to appreciate these new technologies, and despite their infancy, are inclined to include them in the roadmap for the development of new data governance schemes, such as *data trusts* or *data spaces*. As a matter of fact, it is possible to argue that the EU is making a technological bet to secure more democratic and participatory data practices through technology [2]. In this scenario, confidence in these data-intensive practices is promoted by seeking more technologically robust systems that do not depend on a firm's reputation so as to balance the power imbalance between data subjects and data controllers. Particularly, PIMS are supposed to tackle the obscurity found in many complex data flows by promoting "transparency and control measures".

The literature around the notion of trust, while rich, is complex given the different understandings of this concept. In this respect, De Filippi et al. [6] made a distinction between trust and confidence: "trust depends on personal vulnerability and risk-taking, whereas confidence depends on internalized expectations deriving from knowledge or past experiences". Applying this approach to our research object, given the lack of trust over what data controllers will do with personal data, data subjects could be interested in technologies that would allow them to comprehend how their information is involved in actual data flows.

As such, we expect to explore the ethical and legal challenges in building confidence in these technological solutions but also in trusting the operators of these systems to provide a balanced ecosystem for data-sharing practices. In this respect, legal rules can show us which elements a lawmaker considers relevant to the promotion of trust and the building of confidence in these technological solutions. While there are many different regulatory strategies, disclosure-based approaches dominate data-related regulations, particularly in the EU, with the intention to rely on consent [7] as an enabler of the data economy. These types of legal rules are intended to mitigate imbalances, i.e., vulnerabilities, between two or more parties by exposing the potential risks and subsequent harms; however, they are also intended to deliver key information to the decision-making individual, enabling them to make a confident choice [8]. Moreover, from an ethical perspective, various norms can be identified that should be complied with by a person with whom information is shared in order to be trustworthy. These norms include sincerity, competency, and the permissibility of the task that the trustor relies upon the trustee to perform [9].

Given the multitude of factors that can have an influence on both trust and confidence, we limit our analysis to how transparency is pursued as a necessary precondition for the operation of a given technology (PIMS, in this case), the applicable regulatory framework (personal data-related rules, limited to the EU context for this paper), and an ethical discussion around data control, as suggested by Bodó's framework for mediated technological trust [10]. The particular focus on transparency is based on three grounds: (i) from a regulatory perspective, transparency is a cornerstone principle of personal data protection regimes, and it is usually included alongside lawfulness and fairness, as in the General Data Protection Regulation (GDPR) Article 5.1(a); (ii) transparency includes both its *ex-ante* as well as its *ex-post* elements, the latter including the issue of explainability [11]; and (iii) it is possible to crack the "black box" that many Artificial Intelligence (AI) systems present and to identify potential biases towards vulnerable populations by revealing how data flows and nurtures data-driven innovations through transparency [12].

Consequently, this contribution explores the current and upcoming European data protection rules with a focus on transparency. Moreover, the ethical impact of centralized and decentralized technologies on the empowerment of users with respect to the types of control they can exert over their personal data is examined. By focusing on transparency, we explore how technology can enable responsible innovation in the field of personal data management systems through standardizing machine-readable policies for the expression of personal data-handling activities. Through this joint legal, ethical, and technical approach, we expect to identify existing gaps in the literature that can be addressed in

future works. The contributions of this work can be summarized by the following research question: *"What are the indicators that provide trust and confidence over parties and technologies involved in decentralized personal information management systems from an ethical, (EU-oriented) legal, and technological perspective to secure responsible innovation?"*

In this respect, we organize the paper as follows: Section 2 identifies the motivation that guides this contribution—the emergence of decentralized PIMS; Section 3 provides an overview of existing and proposed regulatory data protection regimes; Section 4 takes this legal and technological discussion to the philosophical arena by putting forward the research question of "who should citizens trust with their personal data"; Section 5 discusses the use of one core aspect of data protection—transparency—to foster ethical and legal innovation in the field of personal data management, including a comparison of existing machine-readable solutions for the representation of data protection requirements; Section 6 presents open research directions and ideas for future work; and Section 7 summarizes the conclusions and presents possible future areas of discussion regarding the use of decentralized data governance schemes.

2. Motivation—The Emergence of Decentralized PIMS

Managing and governing data flows is not a new issue and, as such, rules to tackle this task, particularly legal ones, have existed for quite some time: from the Fair Information Practice Principles principles (FIPPs) produced back in the early 1970s [13] to Convention 108 (Convention 108 and related documents can be located in the following link: https://www.coe.int/en/web/data-protection/convention108-and-protocol, (accessed on 26 May 2023), to the latest regulatory frameworks for developing countries such as Brazil or India. However, a common trend among these provisions was and is the existence of an accountable and responsible entity, the *data controller*, which determines the means and purposes for processing personal data from somebody else, the *data subject*, who is granted certain rights to ensure the respect for their information. However, as discussed above, this model has shown some shortcomings, particularly due to issues such as information overload that negates any real possibility of consenting to data processing or the use of an unsuitable legal basis for certain activities [8]. In response, new data governance schemes are being proposed [14] that could constitute an improvement towards more egalitarian control over personal data [15], where data subjects are aided in managing their personal data via *data cooperatives*, *data trusts*, or *data commons*, as is explored further on.

However, these solutions still rely on the fact that most personal data are centralized in large databases of data controllers or there is a significant transfer of control over a trusted party. Generally speaking, trust in intermediaries has suffered considerably in the last two decades: from the financial crisis in the late 2000s, political instability due to the surface of scandals on both sides of the political spectrum, or more related to our research, the exploitation and manipulation conducted by these "trusted" data entities and the lack of concrete enforcement from regulators. Given this background, certain technological tools have emerged to terminate these "distrustful" intermediaries and allow for pure peer-to-peer relations where trust in third parties can be circumvented.

In this context, the emergence of decentralized solutions for the Web has gained many supporters since the development of Bitcoin (https://bitcoin.org/, accessed on 26 May 2023) and a myriad of other applications. While blockchain and distributed ledger technologies have received substantial development in the financial services industry, there are other connected developments, such as the Semantic Web and its stack of technologies [16], that seek to decentralize access to the data on the Web. While these technologies are not without their critics, some of them are quite valid, as in the case of the ICOs (Initial Coin Offerings) scams in 2017 or the appropriation by certain companies of the metaverse, since they challenged the status quo and opened the discussion about alternative models in an era where trust was falling.

Among these other applications, we can find certain developments in the field of personal information management systems, particularly under the banner of "self-sovereign

identity" solutions. These decentralized systems can allow users to select who can access their data and, therefore, actually shift the power balance. In practice, decentralized systems detach the data from the service itself, giving the users full control over the data, as they can be kept in individual personal data stores. Secondly, these systems address issues such as data portability and interoperability, making it easier for users to exchange data between applications, again further reinforcing a new power position for them. Therefore, these user-managed data systems can be considered the next step towards an actual negotiation of privacy terms between data subjects and controllers. This represents a considerable change from the current situation where individuals are presented with the service's privacy notice and must usually accept it to access it.

While their promoters back these developments under the premise that data subjects could become controllers of their own personal data, this view, to a certain extent, is incompatible with existing regulations [4]. The whole premise of data protection rules is to identify a responsible party, different from the data subject, and impose a series of duties to ensure that the data processing does not compromise fundamental rights in the process. However, these technologies can play a role in facilitating the exercise of data subjects' rights and provide a much more detailed overview of how personal data are processed in contrast to the existing landscape [5].

In this context, as was previously pointed out, the use of decentralized PIMS may be the answer for users to regain control over who has access to their data and under what conditions. Currently, centralized services dominate the control of data flows on the Web, regardless of whether we are referring to a US-based Big Tech company, such as Facebook or Twitter, a CCP-influenced (CCP—Chinese Communist Party) service provider, such as TikTok or AliPay, or even EU-based entities, such as IAB Europe; while some jurisdictions are working on improving this, such as Europe with its data spaces initiative [2] , the current landscape is dominated by very few "gatekeepers" with significant influence over data. These entities collect large amounts of personal data in their data silos, making it difficult for the users to access their data, let alone reuse it for other Web services according to the data subject's wishes. In addition, the reuse of such data for other purposes, determined by these controllers, remains unchecked by data protection authorities.

With the data stored in decentralized PIMS, users regain control over what information the services have access to, and they can actually reuse the same data across different applications and services. In this context, several personal data stores' models have been emerging in the market, such as the Solid project (Solid Technical Reports: https://solidproject.org/TR/, accessed on 29 May 2023), the Hub of All Things (https://www.hubofallthings.com/, accessed on 29 May 2023), and Meeco (https://www.meeco.me/, accessed on 29 May 2023). While all of these different models were developed with the goal of giving the users control over how their personal data are used, the Solid ecosystem has been gaining greater adoption as it is open-source and based on the Linked Data Collection of *inter-related* datasets on the Web. More information on Linked Data is available at https://www.w3.org/DesignIssues/LinkedData.html (accessed on 29 May 2023). It promotes interoperability by relying on already-existing Web standards for identity management. Identification in Solid is based on the WebID-TLS (WebID Authentication over TLS, https://www.w3.org/2005/Incubator/webid/spec/tls/, accessed on 29 May 2023) or OIDC (OpenID Connect, https://openid.net/connect/, accessed on 29 May 2023) protocols and Linked Data resource usage. The core Solid specification relies on the RDF (Resource Description Framework, hhttps://www.w3.org/TR/rdf11-concepts/, accessed on 29 May 2023) and LDP (Linked Data Platform,http://www.w3.org/TR/ldp/, accessed on 29 May 2023) standards, while providing granular access control to Web resources and collections of resources. To expand on this, we use Solid as a case study of this new technological development.

Solid, a project led by the inventor of the Web, Tim Berners-Lee, is a specification for decentralized personal online data stores ("Pods") based on interoperable Web standards and protocols. Solid allows its users to take ownership of their data by storing and

managing access to them, while Solid applications have access to this information through dynamic access control rules that the users themselves choose. Moreover, Solid applications and Pods are interoperable, since the data generated by any Solid app can be stored in any Pod, independently of the Pod provider. In addition, Solid applications are also interoperable, since the data generated by one application can be reused by another. A list of approved Pod providers is maintained by the Solid Community for the users to choose according to their terms and conditions, and there is also the possibility to self-host their own Pod. Information on Pod providers and self-hosting is available at https://solidproject.org/users/get-a-pod (accessed on 29 May 2023).

Moreover, Solid implements authentication and authorization protocols as two processes to improve users' trust in the privacy and security of their data. When it comes to the authentication protocol (https://solidproject.org/TR/oidc, accessed on 29 May 2023), Solid uses the WebID (https://www.w3.org/2005/Incubator/webid/spec/identity/, accessed on 29 May 2023); https://solid.github.io/authorization-panel/authorization-ucr/#definitions, accessed on 29 May 2023) specification. A WebID is used to authenticate the users when logging into Solid applications and for users to manage the data on their Pods. Additionally, the access rights on a Pod are attached to specific WebIDs, regardless of the user to which they relate.

On the other hand, the authorization protocol (https://solid.github.io/authorization-panel/authorization-ucr/, accessed on 29 May 2023) in Solid is based on the Web Access Control (WAC) (https://solidproject.org/TR/wac, accessed on 29 May 2023) specification. Using WAC, each resource in a Pod can have a set of authorization statements stored in the so-called Access Control List resources (ACLs). The ACL ontology is available at http://www.w3.org/ns/auth/acl#. These statements include the authorized agents and the modes of access that they have for the resources. Additionally, these authorizations can be explicitly set for an individual resource, inherited from the parent folder, or even set at the Pod level, which can be easily set by the users, for instance, using a drag-and-drop solution, so that they do not have to understand what is happening behind the scenes with the ACL code.

A decentralized system such as Solid brings in a few advantages. First, by default, access to the resources is not allowed unless the user gives active consent, which is aligned with the EU's GDPR principle of privacy by default and by design in Article 25. Another strong aspect is related to the fact that permissions can be set to local files or at a broader level, for instance, over folders or even over the whole Pod, and it is much easier to update and revoke access than on the usual centralized applications. On the other hand, some disadvantages still need to be overcome, such as creating a mechanism to specify prohibitions, e.g., if a user wants to state that they do not want their data to be used to develop commercial products. Additionally, in line with the previous point, there is still work to be conducted to be able to write authorizations over specific types of data or specific purposes [17].

3. Legal Challenges Regarding Data Protection and Decentralized PIMS

As discussed above, new data governance schemes have been proposed to counterbalance existing shortcomings in current rules. While the term "data governance" has several meanings depending on the context and the approach of the involved stakeholders, our focus on this concept for this portion of the article is from a regulatory perspective [18–20]. To further specify our analysis, we look into existing data governance schemes mandated by European regulations as well as proposed future rules.

We limit our analysis to this geopolitical selection for two reasons: (i) the EU provides more of the most robust and developed legal regimes dealing with data processing that position individuals and their protection at the core [21]; and (ii) this has led to the "exportation" of this model under the Brussels effect to many continents, such as Asia, Latin American, or Africa [22]. Considering that we seek to address the intimate nature of individuals and their information, our focus shall be placed exclusively on rules deal-

ing with personal data rather than nonpersonal data. We acknowledge that the line that distinguishes between the two kinds of data is blurry at the current moment, as opinions from the legal literature and certain guidance from authoritative bodies recognize that the technical ability to achieve such separation might not be as effective as originally conceived. Not only that, we also recognize that there is a call for the unification of such distinct legal regimes given that the processing of nonpersonal data could produce similar effects for data subjects as the processing of personal data. As part of our interdisciplinary approach to the research question, this section adopts a legal research methodology by conducting a descriptive analysis of the existing and upcoming legal framework [23]. As for the first portion of the analysis, the method to be used serves the purpose of organizing the existing legal framework alongside its interpretation by authoritative bodies, both administrative and judicial, as well as being enriched by the legal literature. Beyond the scope of this section, and when discussing how to effectively secure legally compliant innovation in this field, an exploratory theoretical analysis [24] on the role of transparency is performed, alongside joint efforts with the two other disciplines at stake. Particularly in this latter stage, we decided to opt for a "law + tech" approach [25] to analyze the particular relation between law and computer science to secure the objective of protecting the right to data protection. Under this lens, the exploratory theoretical analysis seeks to promote the collaboration between law and technology to harness the most desirable solutions to promote the effective protection of the fundamental rights at stake.

3.1. Existing Data Protection Regulations

The principal existing regulatory framework in Europe for data governance is the GDPR [26]. The core of the GDPR is structured around the "data controller–data subject" relation mentioned previously and, from it, all relevant rights and obligations are assigned and structured [27]. In this respect, the GDPR is anchored around the idea of allowing data processing when there is a reason for it, a *legal basis*; the data subject has been informed about it, *transparency*; and a series of provisions are abided by, *principles, duties, and responsibilities*, when handling such information.

While the dichotomy is clearly helpful from a legal point of view, when it comes to the distribution of rights and obligations between the involved parties, certain unforeseen consequences have emerged in recent years regarding how these roles have been interpreted by courts. The GDPR does not forbid the existence of more than one data controller, but rather, acknowledges that, in particular cases, two or more entities might process personal data together for a shared purpose. This reflects the reality of current complex data flows between multiple parties. As such, the GDPR introduced a proper and fully-fledged regime for the notion of joint controllership [28]. While the notion is clearly helpful to provide a clear legal framework for those entities processing personal data jointly, the caselaw from the Court of Justice of the European Union (CJEU) has taken this concept and considerably expanded its area of application in both pre-GDPR and post-GDPR eras by considering the existence of joint control without a previous and formal understanding of the parties [29–31]. From a data governance perspective, this means that certain parties might end up having far more duties and responsibilities than originally intended.

3.2. Upcoming European Regulatory Data Protection Acts

While data controllers join forces to further process and extract the value out of collected personal data on different platforms, data subjects look for new mechanisms to counter the situation and reclaim control over their data, given the citizens' general distrust feeling, for example, in the case of financial services [32]. In this respect, there is a call for new data governance schemes that can improve the current "data controller–data subject" relation. Recent research has mapped the landscape and identified different scenarios in this rapidly changing landscape [14]. In this sense, it is possible to point out the following schemes: (i) data-sharing pools; (ii) data cooperatives; (iii) public data trusts; and (iv) personal data sovereignty.

A data-sharing pool can be defined as a partnership between different data subjects to unite their data and share the responsibility of overseeing their information while distributing the benefits from the use of their data. Data cooperatives are similar to data-sharing pools, although the administration and decision-making activities regarding the data are conducted by a centralized authority, similar to other cooperatives. When this centralized role is assumed or delegated to a public authority, depending on the concrete situation, it is possible to consider the existence of a public data trust. Finally, if these tasks are assumed to be performed directly by the data subject on an individual basis, we can consider the existence of a personal data sovereignty scheme.

While data-sharing pools, data cooperatives, and public data trusts follow the "data controller–data subject" logic, they approach personal data by viewing it as a commodity or something that can be circulated, traded, or donated, depending on the scheme, which is an approach that might not be compatible with the existence of a fundamental right to personal data protection and privacy. On the other hand, personal data sovereignty recognizes this fundamental right characteristic by taking an individual approach to data governance and granting the individual full control over their data [33].

Certain novel legal rules, for example, the Data Governance Act (DGA), show that it is possible to preserve fundamental rights by adopting safeguards, particularly those related to transparency. According to the DGA, these new data governance structures, particularly those regulated as 'data intermediation service providers' as well as 'data altruism organizations', are intended to help data subjects make choices by assisting them with understanding privacy notices and other meaningful details about relevant personal data handling activities. By doing so, it is acknowledged that transparency, and its adequate interpretation, is key to securing the protection of fundamental rights. These would, indeed, help users to navigate the privacy details of different digital online services. Some of the initial criticism from Europe's two main data protection watchdogs—the European Data Protection Board (EDPB) and the European Data Protection Supervisor (EDPS) [34]—has been addressed during its legislative process, and we can now only wait to see how things unfold in the field regarding this new set of rules.

While not dealing directly with personal data governance schemes, the recent proposal to amend the eIDAS Regulation (eIDAS 2) will tackle the issue of personal data sovereignty [35]. In this regard, the failure of eIDAS to provide a secure digital identity for online activities—another data governance failure—has triggered the need for an update [36]. Responding to this, the idea of a European Digital Identity Wallet emerged as *"a product and service that allows the user to store identity data, credentials, and attributes linked to his/her identity to provide them to rely parties on request and to use them for authentication, online and offline"*. As such, the European Digital Identity Wallet would enable further direct control by users of the personal data associated with their identities.

When it comes to PIMS, it is possible to argue that they can be deployed within a multiregulatory context and, consequently, would have to follow numerous rules depending on the links between the data subject, the data controller, and the technology provider. In this context, these services are intended to be highly modular and adaptable to different online services, giving rise to the need for technical tools to achieve these objectives.

4. The Ethical Challenges of Controlling Data and Reclaiming Control over Them

As case law from supervisory authorities demonstrates, the complexity of the data processing activities has been proved to be complicated for data controllers to explain in simple terms and when using limited attention resources from data subjects [27]. The lack of useful information to discern what is happening with the data poses a risk to building trust between the involved stakeholders. Consequently, individuals are seeking to regain control over their data and limit how these entities use it [14]. As such, potential new data governance schemes can improve the knowledge and understanding of data subjects regarding the use of their personal information by introducing new entities/roles that can mediate the relationship between them and the data controllers [14,37]. This legal and

technological debate is again raising the question of "who should I trust with my data?" for all data subjects.

The significance of answering this question can be found in the concept of "control", in the sense that users need someone to trust in order to regain (democratic) control over their data in the digital age. Emerging data governance models provide legal mechanisms to help data subjects to acquire their voice. In the data cooperative model, for example, cooperatives play the role of trustees and manage data on behalf of the data subjects, and, in turn, data subjects retain and preserve democratic control over their data. In this kind of data governance scheme, what matters is that a relationship of trust is established between cooperatives who manage data and data subjects. In some cases, trustees should consult with data subjects. They provide agreements and contracts for data subjects to inform them. Data subjects, on the other hand, can express their preferences and decide how to share their data and for what purposes [14].

Data cooperatives or other trustees play a significant role in enabling data subjects to retain control over their data and regain their moral position in the digital age. In particular, personal data sovereignty provides a meaningful return to more democratic and egalitarian governance as individuals reclaim control over their personal data, or at least, in theory, it should have this effect [14,38]. As such, personal autonomy and classical liberal values can be respected once again by fostering trust-based relationships. Moreover, our current democratic experiences can provide guidance to avoid falling within the same cracks as we have in the last two centuries, where a significant portion of the population, particularly in the Global South, have had their rights neglected due to faulty and poor governance safeguards. In this respect, it is possible to highlight the democratic failures in Latin America during the last 50 years due to *coups d'etat* , economic crisis, or environmental disasters and the lack of strong governance mechanisms to cope with these changes and situations. For example, the last Argentinian military dictatorship substantially affected the identity of thousands of individuals who were kidnapped as children and relocated with new families, erasing their true identities in the process. As a response, collective organizations emerged to redress this, given the helpless situation that these people were forced into and their lack of power to push back and reclaim their true identities [39].

Despite the importance of trust in respecting the autonomy and agency of data subjects [8], the existing methods for fostering trust remain controversial, and there are unresolved societal issues in digital services and new digital intermediaries [40]. Given that the issues to be resolved are how to approach trust in practice, how to build trust relationships between data subjects and data cooperatives, and what the necessary conditions for fostering trust are, we decided to address these issues using practical methods rather than theoretical ones. To do so, we organized a public Think-In event where people came together to discuss the implications of shifting the control over the terms under which personal data is shared to personal data spaces operated by trusted data intermediaries. One important advantage of the "citizens' Think-In" approach, which we selected for public discussion, is that, unlike a traditional panel discussion or public lecture, it encourages direct participation from those in attendance. Through small group discussions, a Think-In provides an opportunity for people from diverse walks of life to deliberate and discuss topical societal issues arising from Science, Technology, Engineering, and Mathematics (STEM) innovation. Information about the PROTECT Think-Ins and respective results is available at https://protect.oeg.fi.upm.es/thinkin/ (accessed on 18 June 2023).

While the full results from such a process go beyond the scope of our intended contribution, it is possible to highlight that the general public was sensible about the ethical issues around who to trust and the role of transparency in such a process. Citizens raised the issue of preventing the GDPR turning into the "tick-box" compliance exercise that has produced the current form of privacy notices, which would be the case when deploying a template privacy notice for several different data processing activities. Moreover, disclosure and oversight over the use of personal data in a practical and useful manner is a key topic

that demands attention. In this sense, meaningful transparency to build trust between stakeholders is, consequently, a relevant issue.

The resulting insights gleaned from the Citizens' Think-In discussion provide us with an important basis to think about how to embed transparency in data processing agreement terms for personal data stores in both machine-readable and human-readable forms, allowing data subjects to understand and control the articulation of agreement terms. In this manner, the navigation of data controllers and data subjects in the complex data-sharing environment that the platform economy presents would be more firmly under the control of the subject.

5. Using Transparency to Foster Ethical and Legal Innovation in Personal Data Management

As highlighted in the previous sections, transparency can play a key role in mitigating the ethical and legal challenges around decentralized PIMS, given their relevance. Transparency is the first step for any governance scheme grounded in a human rights approach [41]. Given the role of the right to personal data protection, it is only reasonable to conduct our analysis starting from this element. We begin our analysis by focusing on power and explaining the role of transparency in enhancing trust which, in turn, helps to redress the power asymmetry between data subjects and Big Tech. Emerging data governance models can regulate power relations in the digital age. These models provide legal mechanisms to return power stemming from aggregated data to individuals through bottom-up mechanisms. Trustees or data cooperatives, for example, play a significant role in balancing power between data subjects and big companies that process data and aim to gain value from that processing. They enable data subjects to regain control over their data, regaining their power and moral position in the digital age. The precondition for achieving these benefits is that a relationship of trust should be established between data subjects and data cooperatives or other trustees. In this respect, transparency is a fundamental pillar of any (good) governance scheme, and meaningful and thoughtful implementations can serve to overcome shortcomings caused by power imbalances. Therefore, it is appropriate to ask how to achieve good transparency in these new decentralized PIMS schemes, such as the Solid initiative.

In the field of personal data, transparency, as noted previously, is one of the key principles of data protection regulations worldwide, including the GDPR, and it usually involves certain obligations imposed on data controllers and associated rights for data subjects. Privacy notices are typically used as the instrument to comply with most of the obligations associated with this principle. Through transparency, data subjects can understand how their information is going to be used as well as react to any situation that they do not agree with, if possible.

The legal literature on the matter agrees that the current standard privacy notice consists of a single document that is usually located in an inconvenient location and relies heavily on complex and highly legal explanations [42]. In this regard, privacy notices are seen as an obstacle that must be complied with, rather than as a tool to allow data subjects to monitor how their data are used [43].

Moreover, the GDPR places a great deal of importance on transparency for the data subjects. While certain scholars have pointed out that privacy notices have been improved in terms of the provided information, this has come at a cost in terms of the length, complexity, and user accessibility [44]. In this regard, European authorities, national supervisory authorities, and legal scholars have started to act to change current practices related to drafting privacy notices [45,46]. The risk-based approach demands a careful and tailored design of the information provision at the right stages and in the right medium for data subjects. While a template from a supervisory authority can constitute a starting point for simple data processing activities, the reality, from both economic and regulatory perspectives, is that data processing activities can only be expected to become more complex. Not only that but decisions from supervisory authorities, in particular, high-profile cases

such as WhatsApp [47], demand this approach from data controllers. While the previously mentioned decision regarding WhatsApp constitutes a highly relevant case, several other supervisory authorities have also tackled this matter, particularly the Spanish supervisory authority with heavy fines given to banks [48,49].

In addition, the terms and conditions of these services are sometimes presented to users in a manner that is not transparent regarding their data handling practices and in ways that do not allow users to specify and enforce their privacy preferences. Furthermore, with the emergence of privacy and data protection laws that specify how personal data can be collected, used, and shared, these services can no longer rely on these types of "yes/no" consent management systems to comply with the legislation. However, PIMS provide a fertile ground for integrated "law + tech" approaches, where technology plays a role in facilitating compliance with existing regulations, while at the same actually empowering data subjects, rather than just pretending to do so with spectacular claims but little delivery on the ground. As mentioned before, the fact that new data governance regulations, such as the DGA or eIDAS 2, are moving forward in this direction should be taken as a clear statement from policymakers and regulators regarding the actual adoption of these technologies. The question now is how exactly can these PIMS help data subjects to better understand the involved data flows and their conditions.

Most of the suggested techniques are a consequence of a multidisciplinary approach to the issue of drafting effective privacy policies [50]. Moreover, the involvement of computer science expertise in the field of privacy allows for further interaction and integration between what is happening to personal data and what is being told to data subjects. In any case, the main purpose of this collaborative process is to enhance the effectiveness of privacy notices in the wild.

In this context, semantic machine-readable policy languages have been developed since the early 1990s. The main purpose of these languages revolves around the representation of the handling and sharing practices of an individual or organization in relation to a given data resource. Policy languages are, therefore, a natural candidate for representing privacy notices required by the GDPR's transparency obligations. The pioneer solution—P3P (Platform for Privacy Preferences) [51]—was created in 1998, reached the W3C Recommendation status in 2002, but was turned obsolete in 2018 due to a lack of adoption and enforcement mechanisms. It is an XML-based (XML (Extensible Markup Language) is a markup language for storing, transmitting, and reconstructing arbitrary data (https://www.w3.org/TR/xml/, accessed on 18 June 2023)) specification to allow websites to express their data-collection practices with a limited set of recipients, data categories, retention values, and purposes. Its extension, APPEL (A P3P Preferences Exchange Language) [52], was also developed for users to express their privacy preferences in a similar fashion. In the following years, a set of different solutions emerged:

(i) ODRL (Open Digital Rights Language) [53]—An RDF-based solution that provides a model and vocabulary (Vocabulary available at https://www.w3.org/TR/odrl-vocab/, accessed on 18 June 2023) with deontic concepts, i.e., permissions, prohibitions, and duties, to express actionable policies related to digital assets.
(ii) AIR (Accountability in RDF) [54]—An abstract, rule-based N3Logic [55] policy language that supports rule nesting by assigning a unique URI to each rule which is intended to be domain-agnostic and thus does not provide any taxonomies of terms.
(iii) PPO (Privacy Preference Ontology) [56]—A lightweight, domain-agnostic, RDF-based language that can be used to express permissive and restrictive privacy preferences for RDF documents.
(iv) SPECIAL (Scalable Policy-awarE linked data arChitecture for prIvacy, trAnsparency and compLiance) [57]—SPECIAL (https://specialprivacy.ercim.eu/, accessed on 18 June 2023), an H2020-funded project, developed a usage policy language with the core goal of expressing user consent. It allows for the expression of policies with restrictions on the data, purpose, processing activities, storage, and recipients and

provides a set of taxonomies for each. It has now been largely superseded by the Data Privacy Vocabulary (DPV).

(v) BPR4GDPR (Business Process Re-Engineering and Functional Toolkit for GDPR Compliance) [37]—A H2020-funded project (https://www.bpr4gdpr.eu/, accessed on 18 June 2023) that developed an OWL-based (Web Ontology Language, https://www.w3.org/TR/owl2-overview/, accessed on 18 June 2023) policy language and an information model, focused on specifying entities' roles related to organizations processes' life cycles. It provides taxonomies for purposes, actors, roles operations, and organizations.

(vi) DPF (Declarative Policy Framework) [58]—a policy language that is built upon domain-specific OWL ontologies to support the definition of time-limited permissive and prohibitive privacy preferences for specific data categories and data requesters.

Nevertheless, the most relevant solutions for machine-readable policy languages related to privacy, briefly described above, are obsolete or have been without new developments in recent years, as their adoption and application are not on the agenda of the main players in the Big Tech sector. However, since the enforcement of the GDPR, a new wave of vocabularies has emerged with the purpose of providing data-protection-related taxonomies that can be used, for instance, to populate privacy notices. Details about each identified solution are provided below:

(i) DPO (Data Protection Ontology) [59]—Pre-GDPR OWL-based ontology based on the data protection principles of the 1995 Data Protection Directive [60], which models data subject's rights, data processing categories, and entities.

(ii) GDPRov (GDPR Provenance ontology) [61]—RDF-based ontology focused on modeling the provenance of consent and respective data collection, usage, and storage activities for GDPR compliance.

(iii) GDPRtEXT (GDPR text EXTensions) [62]—RDF-based vocabulary that aims to model GDPR concepts and connect them with their respective GDPR chapter, article, and/or point.

(iv) GConsent (GDPR Consent ontology) [63]—RDF-based ontology focused on modeling the consent life cycle, as presented in the GDPR, including terms to represent the status of consent.

(v) PrOnto (Privacy Ontology for legal reasoning) [64]—Closed-access legal ontology that models privacy agents, data types, processing operations, and deontic concepts to support compliance with the GDPR.

(vi) DPV (Data Privacy Vocabulary) [65]—The DPV provides a set of taxonomies to model entities, data, purposes, processing, and their context, technical and organizational measures, legal bases, location and jurisdiction, risks, rules, and rights in RDF and OWL serializations.

In this context, we used the following criteria to analyze the previously identified policy language and vocabulary solutions:

Q1. Does it provide a framework to specify machine-readable privacy policies?
Q2. Does it continue to be maintained, or are new improvements being developed?
Q3. Are the resources available on an open and accessible platform?
Q4. Can it be used to model GDPR concepts and principles?
Q5. Does it provide a vocabulary of terms to populate the policies?
Q6. Does it implement any mechanisms to assist with compliance?

Then, using Table 1 as a reference, it is possible to compare them in terms of their responsiveness to the criteria defined above. The solutions are sorted by the number of supported criteria, in descending order, and then alphabetically to improve the readability.

Table 1. Categorization of privacy and data protection policy languages and vocabularies.

Solution	Q1	Q2	Q3	Q4	Q5	Q6
DPV [65]	Yes	Yes	Yes	Partially	Yes	No
ODRL [53]	Yes	Yes	Yes	Partially	Yes	No
SPECIAL [57]	Yes	No	Yes	Partially	Yes	Yes
BPR4GDPR [37]	Yes	Yes	No	Partially	Yes	No
GConsent [63]	No	Yes	Yes	Partially	Yes	No
GDPRov [61]	No	Yes	Yes	Partially	Yes	No
GDPRtEXT [62]	No	Yes	Yes	Partially	Yes	No
P3P [51]	Yes	No	Yes	Partially	Yes	No
AIR [54]	Yes	No	Yes	No	No	Yes
DPF [58]	Yes	Yes	No	Partially	No	Yes
DPO [59]	No	No	Yes	Partially	Yes	No
APPEL [52]	Yes	No	Yes	No	No	No
PrOnto [64]	No	No	No	Partially	Yes	No
PPO [56]	Yes	No	No	No	No	No

Therefore, as can be concluded from Table 1, of all the examined solutions, only ODRL, DPF, the BPR4GDPR vocabularies, GConsent, GDPRov, GDPRtEXT, and DPV continue to be actively maintained and developed. Apart from BPR4GDPR, DPF, PrOnto, and PPO, most of the presented solutions are open-source.

Furthermore, most policy languages include references to a few GDPR concepts and can partially represent the transparency needs brought on by GDPR Articles 13 and 14, although most of them were developed before this legislation came into force. P3P, ODRL, BPR4GDPR, and SPECIAL developed taxonomies of terms that can be used to partially populate privacy notice terms as well as DPO, GDPRov, GDPRtEXT, PrOnto, GConsent, and DPV. A complete analysis of the vocabularies and policy languages discussed above, as well as their adequacy to cover the representation needs brought by the GDPR, was recently published by [66].

DPV must be highlighted, as it is an extension of the SPECIAL vocabularies that incorporates the most complete list of taxonomies, including vocabularies for personal data categories, purposes, processing categories, technical and organizational measures, legal entities, and legal basis, as described before, and specific extensions for GDPR (DPV-GDPR) (https://www.w3id.org/dpv/dpv-gdpr, accessed on 2 June 2023), which models the GDPR's legal basis for the processing and transfer of personal data and GDPR's data subject rights, DPV-PD (https://www.w3id.org/dpv/dpv-pd, accessed on 2 June 2023), an extension with personal data categories, and the DPV-LEGAL (https://www.w3id.org/dpv/dpv-legal, accessed on 2 June 2023), DPV-TECH (https://www.w3id.org/dpv/dpv-tech, accessed on 2 June 2023), and RISK (https://www.w3id.org/dpv/risk, accessed on 2 June 2023) extensions for jurisdiction-relevant, technology, and risk assessment and management concepts, respectively. Notably, ODRL is a W3C Recommendation for digital rights management that is already being used with DPV to create a profile for access control in Solid [17].

6. Future Research Directions

Through the analysis performed in Section 5, it can be concluded that there is still a gap in the representation of concepts related to privacy notices that needs to be addressed to have a language that can completely model all the terms described in GDPR Articles 13 and 14, as well as to allow data subjects to manage who has access to their data and assist data controllers in the process of compliance with their GDPR obligations. Even though DPV's taxonomies already provide a good basis to represent most terms that are necessary to deal with the GDPR's information requirements, GDPR points 13.2(e), 13.2(f), and 14.2(g), related to statutory and contractual requirements and the existence of automated decision-making, are still not covered and need to be further explored. Concepts to justify the data subjects' right-related requests also need to be included, as well as concepts to represent

the data controller's ground to not comply with such requests. The modeling of personal data breaches and respective compliance documentation is also missing. Moreover, the requirements and information flows brought on by the DGA, eIDAS 2, and other data-related regulations and proposals for the regulation of the EU must also be a target of future research. In this context, the authors recently published an ODRL profile that proposes a model to define transparent access control policies for individual and group-shared PIMS, while tackling DGA requirements [67].

Furthermore, as we mentioned before, transparency plays a key role in a (good) data governance scheme. The existing regulatory data governance scheme demands that data controllers disclose all relevant required elements regarding the data processing activity to the data subjects. In this sense, the regulation places a lot of weight on these stakeholders to ensure that other parties are adequately informed regarding what is happening with their personal data in any given situation. Under the accountability principle, data controllers (and also data joint controllers) are left on their own to figure this out. While it is true that large platform gatekeepers would be able to take all relevant measures to comply with these data governance requirements, smaller firms and individuals that are caught in a joint controllership with them are not in the same situation. Regardless, these situations do not radically change the existing status quo regarding securing adequate data governance. While data processing activities might become more complex, from a technical and/or organizational perspective, the involved stakeholders have a consolidated legal tradition that can guide them in this process. Nevertheless, the development of open-source tools to support the controllers in the reporting of compliance documentation, which can be automated through the usage of semantic vocabularies and Linked Data, should be further investigated.

The real questioning comes from these new governance schemes proposed by the Data Governance Act. In this respect, we can still see the same governance logic of a data controller disclosing details to a data subject. All three, data-sharing pools, data cooperatives, and public data trusts, engage in the same process. As certain scholars argue, the real challenge is making sense of the terms used in the DGA and the GDPR to achieve sensible systematic application of the rules [68]. However, personal data sovereignty entails a truly new governance scheme where the individual recovers a relevant voice and has a considerable say in how their personal data are managed. Through these legislative proposals, European regulators are envisaging a future where data subjects are assisted by technology in their data-related choices. As future work, the development of a digital personal assistant, together with a personal data dashboard, might be an important tool to help users to make informed choices, for instance, regarding which health data they want to provide to a data altruism organization or which intermediary they want to use to govern the access to their location data.

This shift in how personal data are managed also entails a change in stakeholders' power. Individuals are now the ones who choose which personal data are associated with them, and, by design, any decision regarding them should be decided, either manually or automatically, by the person. While the exact details of how personal data sovereignty fits with existing regulatory data governance schemes, such as the GDPR, are still an open and unanswered question [4], it is clear that data flows and, consequently, decisions about data are going to change. Taking this into account, as future work, we also plan to research how these new laws fit together and how they can be applied to such decentralized systems.

New governance schemes mean that we need to engage in new relations for which the existing categories and terminologies might be sufficient. As such, common terminology is needed to allow communication between the involved parties, whether it is person-to-person, person-to-machine, or machine-to-machine. In this sense, a common ontology could serve as a bridge for identifying elements that serve to foster trust through transparency about data protection practices in a context where data-sharing activities are increasing within the platform and data subjects seek to reclaim their data and control their digital identity via personal sovereignty. Consequently, and given the novelty of these mechanisms,

it is impossible to identify meaningful work that looks at how individuals feel regarding these alternative data governance schemes, in particular, whether or not these can be trusted with a sensible task, such as advising on personal data choices and preferences, making it also an excellent candidate for further research.

7. Conclusions

Data subjects are exposed to a considerable number of data processing activities about themselves and, in some cases, those related to them. In this sense, they can be considered stakeholders in many different situations. However, it is possible to question whether they can actually engage in a significant manner in these processes. As such, there are incentives (mainly regulatory) placed on the other stakeholders to help data subjects to exercise their control to govern their personal data, mainly through data rights and information.

Technological developments have emerged in the form of personal data stores that may help data subjects have more choice and say in the management and use of their data. In this regard, our focus was placed on transparency measures that enable data subjects to understand data activities and explanations, allowing them to make the informed decisions required to give consent for data processing. Furthermore, the new legislative proposals in the EU, either adopted as the DGA or, in other works, as eIDAS 2, contain a provision to classify and encourage trust in data-sharing intermediaries, such as the providers of personal data spaces' software and servers. With this stage in mind, we provided a foundational discussion that will assist in the further development of common models for the drafting of data protection notices, using technological resources such as ODRL, DPV, and other ontological models, to address the requirements of data subjects in fully understanding and controlling the articulation of these agreement terms (human-readable forms) and for capturing data processing agreement terms in machine-readable forms. In this context, the adoption of legally-aligned machine-readable policy languages and vocabularies in a decentralized environment, such as the one provided by the Solid ecosystem, might be the answer to the research question posed in this work, as they provide a secure and responsible innovation environment that respects the data protection laws and ethical values that guide the EU and represent a trustworthy and transparent solution to give individuals more direct and stricter control over how their data are used by different controllers. In this sense, some emerging solutions in this field have been recently proposed to bridge these domains discussed in this work [67].

Ultimately, the discussion on new data governance schemes belongs to a broader debate regarding power structures in the digital age and the underlying political perspectives on society. While our existing data governance regulations were inspired by traditional liberal perspectives on human autonomy, and individuals are presumed to be capable of making informed decisions, reality shows us that certain stakeholders have stronger positions in comparison to them. In this respect, these large gatekeepers and consolidated platforms can guide and nudge human behavior to their benefit. Moreover, as this work shows, a gap in the literature still exists in the representation of knowledge related to transparency and trust in the context of a joint technical, legal, and ethical approach to data protection. Considering this, the development of common data models and vocabularies that can provide the relevant shared criteria and terminology to both data controllers and data subjects should be encouraged so that all parties understand what is happening with the involved personal data at any given time in a highly modular, multistakeholder data-sharing environment where individuals have a greater degree of control over their data.

Author Contributions: Conceptualization, H.A., A.C.P., B.E. and D.L.; Methodology, A.C.P.; Investigation, H.A.; Resources, B.E.; Writing—original draft, H.A., A.C.P. and B.E.; Supervision, D.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement No. 813497 (PROTECT).

Data Availability Statement: Information about the PROTECT Think-Ins and respective results is available at https://protect.oeg.fi.upm.es/thinkin/ (accessed on 18 June 2023).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jacobides, M.G.; Sundararajan, A.; Van Alstyne, M. *Platforms and Ecosystems: Enabling the Digital Economy*; World Economic Forum: Cologny, Switzerland, 2019.
2. European Commission. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions—A European Strategy for Data*; COM(2020) 66 Final ed.; European Commission: Brussels, Belgium, 2020.
3. Waldman, A.E. *Industry Unbound: The Inside Story of Privacy, Data, and Corporate Power*, 1st ed.; Cambridge University Press: Cambridge, UK, 2021; ISBN 9781108591386. [CrossRef]
4. Chomczyk Penedo, A. Self-sovereign identity systems and European data protection regulations: An analysis of roles and responsibilities. In *Open Identity Summit 2021*; Gesellschaft für Informatik e.V.: Bonn, Germany, 2021; pp. 95–106.
5. Janssen, H.; Cobbe, J.; Singh, J. Personal information management systems: A user-centric privacy utopia? *Internet Policy Rev.* **2020**, *9*, 1–25. [CrossRef]
6. De Filippi, P.; Mannan, M.; Reijers, W. Blockchain as a confidence machine: The problem of trust & challenges of governance. *Technol. Soc.* **2020**, *62*, 101284. [CrossRef]
7. Chomczyk Penedo, A. Towards a technologically assisted consent in the upcoming new EU data laws? *Priv. Ger.* **2022**, *5*, 180–187. [CrossRef]
8. Ben-Shahar, O.; Schneider, C.E. *More Than You Wanted to Know: The Failure of Mandated Disclosure*; Princeton University Press: Princeton, NJ, USA, 2014. [CrossRef]
9. Hawley, K. *How To Be Trustworthy*; Oxford University Press: Oxford, UK, 2019.
10. Bodó, B. Mediated trust: A theoretical framework to address the trustworthiness of technological trust mediators. *New Media Soc.* **2021**, *23*, 2668–2690. [CrossRef]
11. Felzmann, H.; Villaronga, E.F.; Lutz, C.; Tamò-Larrieux, A. Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns. *Big Data Soc.* **2019**, *6*, 2053951719860542. [CrossRef]
12. Pasquale, F. *The Black Box Society: The Secret Algorithms That Control Money and Information*; Harvard University Press: Cambridge, MA, USA, 2015.
13. Cate, F.H. The Failure of Fair Information Practice Principles. In *Consumer Protection in the Age of the Information Economy*; Routledge: London, UK, 2006.
14. Craglia, M.; Scholten, H.; Micheli, M.; Hradec, J.; Calzada, I.; Luitjens, S.; Ponti, M.; Boter, J. *Digitranscope: The Governance of Digitally Transformed Society*; Publications Office of the European Union: Luxembourg, 2021.
15. Viljoen, S. A Relational Theory of Data Governance. 2020. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3727562 (accessed on 20 September 2022).
16. Berners-Lee, T.; Hendler, J.; Lassila, O. The Semantic Web. *Sci. Am.* **2001**, *284*, 34–43. [CrossRef]
17. Esteves, B.; Pandit, H.J.; Rodríguez-Doncel, V. ODRL Profile for Expressing Consent through Granular Access Control Policies in Solid. In Proceedings of the 2021 IEEE European Symposium on Security and Privacy Workshops (EuroS PW), Vienna, Austria, 6–10 September 2021; pp. 298–306. ISSN: 2768-0657. [CrossRef]
18. Abraham, R.; Schneider, J.; vom Brocke, J. Data governance: A conceptual framework, structured review, and research agenda. *Int. J. Inf. Manag.* **2019**, *49*, 424–438. [CrossRef]
19. Alhassan, I.; Sammon, D.; Daly, M. Data governance activities: A comparison between scientific and practice-oriented literature. *J. Enterp. Inf. Manag.* **2018**, *31*, 300–316. [CrossRef]
20. Mahanti, R. Data Governance and Compliance. In *Data Governance and Compliance: Evolving to Our Current High Stakes Environment*; Mahanti, R., Ed.; Springer: Berlin/Heidelberg, Germany, , 2021; pp. 109–153. [CrossRef]
21. Celeste, E. Digital Sovereignty in the EU: Challenges and Future Perspectives. In *Data Protection Beyond Borders: Transatlantic Perspectives on Extraterritoriality and Sovereignty*; Fabbrini, F.; Quinn, J.; Celeste, E., Eds.; Hart Publishing: London, UK, 2020; pp. 211–228. [CrossRef]
22. Bradford, A. *The Brussels Effect: How the European Union Rules the World*; Oxford University Press: Oxford, UK, 2019. [CrossRef]
23. Smits, J.M. *The Mind and Method of the Legal Academic*; Edward Elgar Publishing: Cheltenham, UK, 2012.
24. Ballin, E.H. *Advanced Introduction to Legal Research Methods*; Edward Elgar Publishing: Cheltenham, UK, 2020.
25. Schrepel, T. *Blockchain + Antitrust: The Decentralization Formula*; Edward Elgar Publishing: Cheltenham, UK, 2021.
26. Shabani, M. The Data Governance Act and the EU's move towards facilitating data sharing. *Mol. Syst. Biol.* **2021**, *17*, e10229. [CrossRef] [PubMed]
27. European Data Protection Board. *Guidelines 07/2020 on the Concepts of Controller and Processor in the GDPR*; 2020. Available online: https://edpb.europa.eu/our-work-tools/documents/public-consultations/2020/guidelines-072020-concepts-controller-and_en (accessed on 18 June 2023).

28. Millard, C.; Kamarinou, D. Article 26. Joint controllers. In *The EU General Data Protection Regulation: A Commentary*; Oxford University Press: Oxford, UK, 2020; pp. 582–588.
29. Court of Justice of the European Union. Tietosuojavaltuutettu v Jehovan Todistajat—Uskonnollinen Yhdyskunta. 2018. ECLI:EU:C:2018:551. Available online: https://curia.europa.eu/juris/liste.jsf?language=en&num=c-25/17&td=ALL (accessed on 18 June 2023).
30. Court of Justice of the European Union. Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein v Wirtschaftsakademie Schleswig-Holstein GmbH, 2018. ECLI:EU:C:2018: 388. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62016CJ0210 (accessed on 18 June 2023).
31. Court of Justice of the European Union. Fashion ID GmbH & Co. KG v Verbraucherzentrale NRW eV. 2019. ECLI:EU:C:2019: 629. Available online: https://curia.europa.eu/juris/liste.jsf?num=C-40/17 (accessed on 18 June 2023).
32. Armantier, O.; Doerr, S.; Frost, J.; Fuster, A.; Shue, K. *Whom Do Consumers Trust with Their Data? US Survey Evidence*; BIS Bulletins 42; Bank for International Settlements: Basel, Switzerland, 2021.
33. Wang, F.; De Filippi, P. Self-Sovereign Identity in a Globalized World: Credentials-Based Identity Systems as a Driver for Economic Inclusion. *Front. Blockchain* **2020**, *2*, 28. [CrossRef]
34. European Data Protection Board. EDPB-EDPS Joint Opinion 03/2021 on the Proposal for a Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act), 2021. Available online: https://edpb.europa.eu/our-work-tools/our-documents/edpbedps-joint-opinion/edpb-edps-joint-opinion-032021-proposal_en (accessed on 18 June 2023).
35. Ortalda, A.; Jasmontaite, L.; Tsakalakis, N. *The European Commission Proposal Amending the eIDAS Regulation: A Personal Data Protection Perspective*; Technical Report; Vrije Universiteit Brussel, Brussels Privacy HUB: Brussel, Belgium, 2021.
36. Domingo, I.A. La propuesta de Reglamento eIDAS 2: La identidad digital autosoberana y la regulación de Blockchain. *Diario La Ley* **2021** .Available online: https://diariolaley.laleynext.es/dll/2021/06/24/la-propuesta-de-reglamento-eidas-2-la-identidad-digital-autosoberana-y-la-regulacion-de-blockchain (accessed on 18 June 2023).
37. Papagiannakopoulou, E.I.; Koukovini, M.N.; Lioudakis, G.; Dellas, N.; Garcia-Alfaro, J.; Kaklamani, D.I.; Venieris, I.S.; Cuppens-Boulahia, N.; Cuppens, F. Leveraging Ontologies upon a Holistic Privacy-Aware Access Control Model. In *Foundations and Practice of Security. FPS 2013*; Danger, J., Debbabi, M., Marion, J., Garcia-Alfaro, J., Zincir Heywood, N., Eds. Springer: Cham, Switzerland, 2014; Lecture Notes in Computer Science; Volume 8352, pp. 209–226.
38. Giannopoulou, A. Digital Identity Infrastructures: A Critical Approach of Self-Sovereign Identity. *Digit. Soc.* **2023**, *2*, 18. [CrossRef] [PubMed]
39. Gesteira, S. Más allá de la apropiación criminal de niños: El surgimiento de organizaciones de personas "adoptadas" que buscan su "identidad biológica" en Argentina. *RUNA Arch. Para Las Cienc. Hombre* **2014**, *35*, 61–76. [CrossRef]
40. Carovano, G.; Finck, M. Regulating Data Intermediaries: The Impact of the Data Governance Act on the EU's Data Economy, 2023. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4422263 (accessed on 18 June 2023).
41. De Hert, P. Globalisation, crime and governance. Transparency, accountability and participation as principles for global criminal law. In *Transitional Justice and Its Public Spheres: Engagement, Legitimacy and Contestation*; Brants, C., Karstedt, S., Eds.; Hart Publishing: London, UK, 2017; pp. 91–123.
42. Terpstra, A.; Schouten, A.P.; Rooij, A.d.; Leenes, R.E. Improving privacy choice through design: How designing for reflection could support privacy self-management. *First Monday* **2019**, *24*. [CrossRef]
43. Mohan, J.; Wasserman, M.; Chidambaram, V. Analyzing GDPR Compliance Through the Lens of Privacy Policy. In *Heterogeneous Data Management, Polystores, and Analytics for Healthcare*; Gadepally, V., Mattson, T., Stonebraker, M., Wang, F., Luo, G., Laing, Y., Dubovitskaya, A., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2019; pp. 82–95. [CrossRef]
44. Linden, T.; Khandelwal, R.; Harkous, H.; Fawaz, K. The Privacy Policy Landscape After the GDPR. *Priv. Enhancing Technol.* **2020**, *1*, 47–64. [CrossRef]
45. European Data Protection Board. Guidelines 05/2020 on Consent under Regulation 2016/679 Version 1.1, 2020. Available online: https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202005_consent_en.pdf (accessed on 18 June 2023).
46. Article 29 Data Protection Working Party. Guidelines on Transparency under Regulation 2016/679, 2018. Available online: https://ec.europa.eu/newsroom/article29/items/622227 (accessed on 18 June 2023).
47. Data Protection Commission. WhatsApp Ireland Limited—IN-18-12-2. 2021. Available online: https://gdprhub.eu/index.php?title=DPC_(Ireland)_-_WhatsApp_Ireland_Limited_-_IN-18-12-2 (accessed on 18 June 2023).
48. Agencia Española de Protección de Datos. Banco Bilbao Vizcaya Argentaria, S.A., 2020. PS/00068/2020. Available online: https://www.dataguidance.com/sites/default/files/ps-00068-2020.pdf (accessed on 18 June 2023).
49. Agencia Española de Protección de Datos. CAIXABANK, S.A., 2021. PS/00477/2019. Available online: https://www.aepd.es/es/buscador?f%5B0%5D=sectorial%3A903&search=&page=0 (accessed on 18 June 2023).
50. Brennan-Marquez, K.; Susser, D. Obstacles to Transparency in Privacy Engineering. In Proceedings of the 2016 IEEE Security and Privacy Workshops (SPW), San Jose, CA, USA, 22–26 May 2016; pp. 49–52. [CrossRef]
51. Cranor, L.; Langheinrich, M.; Marchiori, M.; Presler-Marshall, M.; Reagle, J. The Platform for Privacy Preferences 1.0 (P3P1.0) Specification, 2002. W3C Recommendation 16 April 2002 obsoleted 30 August 2018. Available online: https://www.w3.org/TR/P3P/ (accessed on 2 July 2020).

52. Cranor, L.; Langheinrich, M.; Marchiori, M. A P3P Preference Exchange Language 1.0 (APPEL 1.0) Specification, 2002. Available online: https://www.w3.org/TR/2002/WD-P3P-preferences-20020415/ (accessed on 2 July 2020).
53. Iannella, R.; Villata, S. ODRL Information Model 2.2, 2018. Available online: https://www.w3.org/TR/odrl-model/ (accessed on 30 May 2023).
54. Khandelwal, A.; Bao, J.; Kagal, L.; Jacobi, I.; Ding, L.; Hendler, J. Analyzing the AIR Language: A Semantic Web (Production) Rule Language. In *Web Reasoning and Rule Systems*; Hitzler, P., Lukasiewicz, T., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2010; Volume 6333, pp. 58–72.
55. Berners-Lee, T.; Connolly, D.; Kagal, L.; Scharf, Y.; Hendler, J. N3Logic: A logical framework for the World Wide Web. *Theory Pract. Log. Program.* **2008**, *8*, 249–269. [CrossRef]
56. Sacco, O.; Passant, A. A Privacy Preference Ontology (PPO) for Linked Data. In Proceedings of the Linked Data on the Web Workshop at 20th International World Wide Web Conference, Hyderabad India, 28 March–1 April 2011.
57. Kirrane, S.; Fernández, J.D.; Dullaert, W.; Milosevic, U.; Polleres, A.; Bonatti, P.A.; Wenning, R.; Drozd, O.; Raschke, P. A Scalable Consent, Transparency and Compliance Architecture. In *The Semantic Web: ESWC 2018 Satellite Events*; Gangemi, A., Gentile, A.L., Nuzzolese, A.G., Rudolph, S., Maleshkova, M., Paulheim, H., Pan, J.Z., Alam, M., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2018; Volume 11155, pp. 131–136. [CrossRef]
58. Martiny, K.; Elenius, D.; Denker, G. Protecting Privacy with a Declarative Policy Framework. In Proceedings of the 2018 IEEE 12th International Conference on Semantic Computing (ICSC), Laguna Hills, CA, USA, 31 January–2 February 2018; pp. 227–234. [CrossRef]
59. Bartolini, C.; Muthuri, R. Reconciling Data Protection Rights and Obligations: An Ontology of the Forthcoming EU Regulation. In Proceedings of the Workshop on Language and Semantic Technology for Legal Domain, Hissar, Bulgaria, 10 September 2015.
60. European Commission. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data. 1995. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX (accessed on 18 June 2023).
61. Pandit, H.J.; Lewis, D. Modelling Provenance for GDPR Compliance using Linked Open Data Vocabularies. In Proceedings of the Society, Privacy and the Semantic Web—Policy and Technology (PrivOn 2017), Co-Located with ISWC 2017, Vienna, Austria, 22 October 2017; Volume 1951.
62. Pandit, H.J.; Fatema, K.; O'Sullivan, D.; Lewis, D. GDPRtEXT—GDPR as a Linked Data Resource. In *The Semantic Web*; Gangemi, A., Navigli, R., Vidal, M.E., Hitzler, P., Troncy, R., Hollink, L., Tordai, A., Alam, M., Eds.; Lecture Notes in Computer Science, Springer International Publishing: Cham, Switzerland, 2018; Volume 10843, pp. 481–495. [CrossRef]
63. Pandit, H.J.; Debruyne, C.; O'Sullivan, D.; Lewis, D. GConsent—A Consent Ontology Based on the GDPR. In *The Semantic Web*; Hitzler, P., Fernández, M., Janowicz, K., Zaveri, A., Gray, A.J., Lopez, V., Haller, A., Hammar, K., Eds.; Lecture Notes in Computer Science, Springer International Publishing: Cham, Switzerland, 2019; Volume 11503, pp. 270–282. [CrossRef]
64. Palmirani, M.; Martoni, M.; Rossi, A.; Bartolini, C.; Robaldo, L. PrOnto: Privacy Ontology for Legal Reasoning. In *Electronic Government and the Information Systems Perspective (EGOVIS 2018)*; Kő, A., Francesconi, E., Eds.; Lecture Notes in Computer Science, Springer International Publishing: Cham, Switzerland, 2018; Volume 11032, pp. 139–152. [CrossRef]
65. Pandit, H.J.; Polleres, A.; Bos, B.; Brennan, R.; Bruegger, B.; Ekaputra, F.J.; Fernández, J.D.; Hamed, R.G.; Kiesling, E.; Lizar, M.; et al. Creating a Vocabulary for Data Privacy: The First-Year Report of Data Privacy Vocabularies and Controls Community Group (DPVCG). In *On the Move to Meaningful Internet Systems: OTM 2019 Conferences*; Panetto, H., Debruyne, C., Hepp, M., Lewis, D., Ardagna, C.A., Meersman, R., Eds.; Springer International Publishing: Cham, Switzerland, 2019; Volume 11877, pp. 714–730. [CrossRef]
66. Esteves, B.; Rodríguez-Doncel, V. Analysis of Ontologies and Policy Languages to Represent Information Flows in GDPR. *Semant. Web J.* **2022**, 1–35. [CrossRef]
67. Esteves, B.; Asgarinia, H.; Penedo, A.C.; Mutiro, B.; Lewis, D. Fostering trust with transparency in the data economy era: An integrated ethical, legal, and knowledge engineering approach. In Proceedings of the 1st International Workshop on Data Economy, Rome, Italy, 9 December 2022; pp. 57–63.
68. Baloup, J.; Bayamlıoğlu, E.; Benmayor, A.; Ducuing, C.; Dutkiewicz, L.; Lalova, T.; Miadzvetskaya, Y.; Peeters, B. *White Paper on the Data Governance Act, 2021*; Working Paper. CiTiP Working Paper Series. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3872703 (accessed on 18 June 2023).