Legal, Technical, and Social Limitations of Data Portability through Decentralized Applications

Yongle Chao[1,2,*], Meihe Xu[2], Aurelia Tamò-Larrieux[3], Kimberly Garcia[4] and Konrad Kollnig[2]

[1]KoGuan School of Law, Shanghai Jiao Tong University, Shanghai, China

[2]Law and Tech Lab, Maastricht University, Maastricht, Netherland

[3]Faculty of Law, Crimial Jutice and Public Administration, Private Law Center, University of Lausanne, Lausanne, Switzerland

[4]Institute of Computer Science, University of St. Gallen, Switzerland

Abstract

A centralized model with a few large entities holding large amounts of personal data is not conducive to data sharing and data reuse, which leads to lock-in effects and data monopoly. Data portability and decentralized applications serve as a tool for dismantling data silos by granting users control over their data. However, both tools encounter legal, technical, and social limitations. Therefore, we propose a framework that divides and integrates regulatory, technical, and social measures into inner and outer circle measures for decentralization. We hope that such a conceptualization helps to push forward a new paradigm for enhanced data portability.

Keywords

GDPR, Data Act, Access, Control, Solid, Decentralization

1. Introduction

The centralization of personal data management by large technology companies, often referred to as big tech, has become a defining feature of the digital ecosystem. The aggregation of personal data by a few companies, such as Meta and Alphabet, has sparked substantial fears about data abuse, privacy violations, and the solidification of monopolistic behaviors [1]. These practices undermine individual control over personal information, exacerbate market concentration, promote user lock-in effects, and stifle competition, challenging the very principles of a free and open digital economy. Recognizing these issues, the European Commission (EC) proposed to break up the centralized data management paradigm by putting in place regulations that foster data access and portability (e.g., Data Act, Digital Markets Act, Digital Governance Act, General Data Protection Regulation (GDPR)). These regulatory developments are important to enable decentralized approaches to data management, which in turn aim to tackle the very challenges

that centralized environments create (e.g., undermining individual control, or lock-in effects).[1]

In the EU, different regulations such as the GDPR and Data Act aim to facilitate the flow and reuse of data, breaking down information silos and thus enabling data decentralization [2], in particular, the right to data portability (RtDP). However, even though the GDPR came into force in 2018, the right has not been fully realized in practice [3]. The lack of practical implementation demonstrates the tension between data portability as a normative ideal and its practical application (Section 2). Not only the legal domain but also the technical community has been working towards means to enhance decentralization, developing several decentralized applications, which stand as a promising tool for the enforcement of data interoperability and data flow [4]. However, this technique has yet to emerge as the mainstream paradigm (Section 3). In order to overcome the challenges of both legal and technical measures to foster decentralized data management, we propose a framework (Section 4) that divides and integrates regulatory, technical, and social measures into inner and outer circle measures.

2. Right to Data Portability and Practical Hurdles

Data portability evolved from a proposed right that can be traced back to a consultation in 2010, as noted in the legislative background documents.[2] This was aimed at decentralizing data control from centralized platforms to individuals, enabling the data subject, instead of a few platforms, to decide to whom access is to be granted [5]. It ultimately led to Article 18 of the 2012 Proposal (after revisions now Art. 20 GDPR). Under this article, data subjects are granted the right to obtain and transfer their personal data between controllers, thereby enhancing access to their own data. The requirements of the RtDP are as follows: (1) subject: data transfer request initiated by the data subject to the data controller; (2) data format: structured , commonly, machine-readable; (3) data scope: personal data concerning data subject; (4) transfer time: real-time. Recital 68 GDPR mentions interoperability as an additional non-mandatory requirement adding to the description of the format in Article 20. Interoperable standards enable transmission of data from one system to another with limited loss of data. To fulfill the requirements of Art. 20 GDPR, various platforms have updated their privacy policies and developed tools to enable easier data transfer (e.g., Facebook, X, and Google). We summarize two distinct approaches to implementing RtDP, evaluate their efficacy in bolstering data subjects' control and pinpoint practical obstacles to these approaches.

- **Automated data portability:** The first pattern for RtDP execution involves direct data transfers between platforms, exemplified by the use case of Facebook to Koofr. Facebook introduced an option that "transfer a copy of your information", which allows users to transfer notes and posts to Google Docs, Blogger, and WordPress.com,and port photos

---

and videos to Backblaze, Dropbox, Google Photos, and Koofr. However, the transfer destination for receiving information is limited and the data format is not machine-readable. This limitation on data transfer did not alleviate the lock-in effect for users, nor did it address the issue of data monopolies held by large platforms.

- **Manual data portability:** The second pattern involves users manually downloading and updating their data from one platform to another. Empirical evidence indicates that responses to such requests often face disproportionate delays, are marred by deliberate obfuscation and stalling tactics, are incomplete, or in some cases, there is a complete lack of response [6]. This method of data transfer presents the following drawbacks: (1) increased burden on users to manage their data. This data, whether stored in the cloud or on hard drives, can become overwhelming for users when large volumes of personal information are accumulated. (2) it does not support the real-time transmission required by Art. 20 GDPR, falling short of true data portability.

The Data Act (in force since 11 January 2024) lays down certain essential requirements for data access beyond personal data of connected devices and services (e.g., Art. 5 relevant metadata necessary to interpret and use those data) as well as for interoperability (e.g., Art. 30, Art. 33, 34, 35). The Data Act and confers implementation power to the EC to ensure uniform implementation of common interoperability specifications, which helps to address the absence of interoperability that impedes data portability. The main purpose of these provisions is to facilitate the flow and reuse of data, breaking down information silos and thus enabling data decentralization. However, even though the GDPR came into force in 2018, the right has not been fully realized in practice [3] and it remains to be seen, how the provisions of the Data Act will be implemented in practice.

3. Decentralized Technical Support and Social Limitations

Decentralized applications can empower users with enhanced control of their data. We can see that these technologies and legal norms are aligned to empower control of data to the users and data sharing. Firstly, Solid enables decoupling user's data from the platform[7], which could overcome delays by data controllers in responding to data transfer requests and tampering with data. Platforms now store substantial personal data and often withhold this information to strengthen their economic dominance. Moreover, Solid addresses the interoperability issue among different pods and lays a solid foundation for enhanced data interoperability across applications. Finally, data portability through Solid fulfills the legal requirement of real-time data updates. Once access is granted by the user, the platform immediately gains access to the data stored in Solid. Additionally, we reveal the social limitations of decentralized applications in achieving data portability between different platforms.

Firstly, decentralized applications do not address data formats' interoperability (compatibility) across applications. Solid Pod addresses the challenge of data accessibility across different pods and enables users to authorize various applications to access the data stored in their pod. However, it still leaves open the issue of whether the applications are interoperable (compatible) with the data, to which they have access permission [8]. Secondly, Solid applications are still in development. Thus, they lack maintenance and ease of usage for general users. The

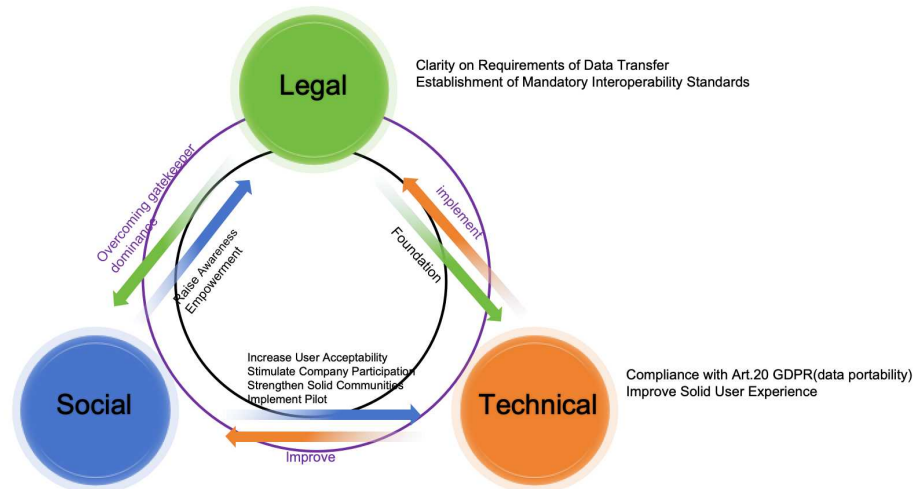**Building a more responsible data economy, and helping legal compliance**

Figure 1: A Synergy Framework for Decentralization

functionalities of the application are hard to understand. For example, the Solid community server is not intuitive. The "Stuff" page features many folders with ambiguous purposes, which duplicate content from the "Storage" page, resulting in confusion. Thirdly, the Solid protocol has yet to achieve mainstream recognition and remains largely unknown to both individuals and platforms. Although some milestones have been reached in making the Flanders government adopt the solid protocol, as well as use cases in topical domains, such as education and career (e.g., Athumi, Karamel). Their practical application and promotion are still lagging conspicuously[9]. Major platforms, such as Amazon, X, and Meta, which handle vast amounts of personal data, have not yet embraced the Solid ecosystem.

4. A Framework: Combining Regulatory, Technical, and Social Measures

To overcome the limitations elaborated in Section 2 and 3 we propose a framework (Figure 1) that divides and integrates regulatory, technical, and social measures into inner and outer circle measures. Theoretically, once the conditions within the inner circle are met, the outer circle benefits from positive feedback. Both circles involve three factors: legal, technical, and social. (1) legal: is the foundation for decentralized application and simultaneously represents the goals that need to be achieved. (2) technical: a tool for implementing Art. 20 and improving social welfare. (3) social: guide and motivate stakeholders to perform better in participation and development, e.g. through education or assessment. The three aforementioned factors are interdependent, with each one propelling another forward while simultaneously allowing itself

to be influenced by the party it supports.

4.1. Inner Circle Measures

4.1.1. Legal and Technical

At the regulatory level, there is a need for further clarification regarding how the data is ported. The current practice of many service providers (e.g., Meta, X, Google) fulfills the data transfer by providing copies of the data and transferring them automated or manually. However, data porting can also be accomplished by authorizing other platforms to directly access data, reducing steps for downloading and uploading. To move towards this new paradigm [10], we need social measures (Section 4.1.2, 4.1.3) as well as standards for interoperability (which is the basis for a functioning right to data portability) [11] . From an EU legal tech point of view, it has a long history of using interoperability as a policy tool to overcome network effects and high switching costs in concentrated markets. For example, the Access Directive (2002/19/EC) and the Framework Directive (2002/21EC) were established and encouraged some services to use communications interoperable standards (in the field of communications). In March 2017 the EC adopted the European Interoperability Framework which gives specific guidance on how to set up interoperable digital public services. Several governments and legislatures have produced draft competition legislation mandating interoperability for large online platforms.

4.1.2. Technical and Social

The Solid protocol has not yet become a mainstream data management technology. On the one hand, many large companies are reluctant to participate in Solid for reasons of financial interest. But on the other aspects is that Solid is not as user-friendly. We believe that Solid has the potential to serve as an alternative to the prevailing model of centralized data storage. However, achieving this requires the acceptance of Solid, or more broadly, decentralized data storage, by users. To measure user acceptance, we propose a mixed-method approach that combines a survey and focus groups. The survey will employ Unified Theory of Acceptance and Use of Technology (UTAUT), which has been used to study the adoption of various technology [12]. UTAUT contains key factors, including performance expectancy, effort expectancy, social influence, facilitating conditions, and moderating factors of gender, age, voluntariness, and experience [13]. In addition to UTAUT, we will use the System Usability Scale (SUS) to evaluate the perceived usability of the technology [14]. This comprehensive survey approach aims to gain a detailed understanding of EU users' acceptance of Solid, if it were made available to them, and to identify potential areas for improvement to enhance its acceptability and usability. Following the survey, focus groups will be conducted to tackle specific aspects regarding users' acceptance of Solid, including addressing any concerns users may have.

4.1.3. Legal and Social

Data portability focuses on the transfer of data between services, thereby enabling decentralization as it decouples the data from the platform. The RtDP facilitates the decoupling of data by granting individuals the ability to transfer their data and imposing an obligation on data

controllers. We can engage users and entities more actively by increasing literacy on RtDP, i.e. educating the general public about the data economy and their data-related rights. In fact, a stronger focus on data literacy has been pushed by policymakers in the EU: according to Recital(19) and Article 5(a) of the Data Act, the competent authorities should promote tools and adopt measures to advance data literacy among users and entities. Data literacy refers to the skills, knowledge, and understanding that allow users, consumers, and businesses to gain awareness of the potential value of the data they generate, produce, and share and that they are motivated to offer and provide access to by relevant legal rules. To begin with, competent authorities need to reinforce individual's awareness of data portability. Individuals should understand how data portability can be realized and the benefits of it (Article 5(a)). In addition, competent authorities should measure how decentralization, for realizing data portability, dismantles data silos and enhances data sharing. Finally, competent authorities should provide opportunities for individuals and entities to learn and practice how to use decentralized applications.

4.2. Outer circle measures: how the inner circle measures influence the outer circle

When inner circle measures are effectively implemented, the outer circle is influenced by the inner circle measures to produce positive results. Von Hippel introduced and validated the concept of democratizing innovation, demonstrating that user-centered innovation becomes more innovative than traditional manufacturer innovation when users have access to the same data and resources as manufacturers [15]. The fundamental condition for enabling user innovation is ensuring users have access to resources concentrated within manufacturers, advocating for resource democratizing instead of resource monopoly. This allows users to create better products based on these resources (because users often have a deeper understanding of their needs than manufacturers), and in doing so, provide feedback and supplementation to the larger entities [15]. Data portability and decentralization technologies provide a level playing field for a few entities and data-poor users (both individuals and businesses) that originally held abundant resources. In theory, their approach aligns with the principles of democratizing innovation, similarly fostering the development of user-centric innovations. In addition, the right to repair empowers the user to modify the copies of the software [16] and its service to customize the product to fit their need, which can enable a more circular economy that makes more efficient use of resources [17], are also movements that can further foster if the inner circle measures are fulfilled.

In practice, FoodCoach demonstrates how data portability enhances consumer welfare. Food-Coach can customize a health plan for the user, after users authorize the platform to access their shopping history[18][19]. Users adjust their purchasing habits based on the healthy plan, and this information, in turn, drives the platform to update the plan. It can be seen data portability contributes to enhancing consumer welfare. In summary, with the interaction of the three elements of the inner circle, data portability will become a reality that can enhance social welfare (e.g., fostering innovation and improving consumer welfare).

5. Conclusion

The right to data portability serves as a crucial gateway for enabling data flows and sharing within the EU.[3] Currently, the control of data portability rests predominantly with platforms, which in various ways obstruct the full realization of this right. This paper argues that Solid is an alternative approach to data portability. As a decentralized technology, Solid transfers the control of data portability from platforms to users, offering greater potential for data interoperability. It effectively facilitates data transfer across diverse platforms. However, the adoption of Solid still encounters social limitions. Therefore, we propose a framework that divides and integrates regulatory, technical, and social measures into inner and outer circle measures for decentralization. We hope that such a conceptualization helps to push forward a new paradigm for enhanced data portability.

References

[1] P. De Filippi, S. McCarthy, Cloud computing: Centralization and data sovereignty, European Journal of Law and Technology 3 (2012).

[2] S. Viljoen, A relational theory of data governance, Yale LJ 131 (2021) 573.

[3] P. Dewitte, J. Ausloos, Chronicling GDPR Transparency Rights in Practice: The Good, the Bad and the Challenges Ahead, International Data Privacy Law (2024) ipad026. URL: https://doi.org/10.1093/idpl/ipad026. doi:10.1093/idpl/ipad026.

[4] A. V. Sambra, E. Mansour, S. Hawke, M. Zereba, N. Greco, A. Ghanem, D. Zagidulin, A. Aboulnaga, T. Berners-Lee, Solid: a platform for decentralized social applications based on linked data, MIT CSAIL & Qatar Computing Research Institute, Tech. Rep. (2016).

[5] I. Van Ooijen, H. U. Vrabec, Does the gdpr enhance consumers' control over personal data? an analysis from a behavioural perspective, Journal of consumer policy 42 (2019) 91–107.

[6] A. A. Habu, T. Henderson, Data subject rights as a research methodology: A systematic literature review, Journal of Responsible Technology (2023) 100070.

[7] V. Vizgirda, R. Zhao, N. Goel, Socialgenpod: Privacy-friendly generative ai social web applications with decentralised personal data stores (2024).

[8] A. J. Bokolo, Exploring interoperability of distributed ledger and decentralized technology adoption in virtual enterprises, Information Systems and e-Business Management 20 (2022) 685–718.

[9] E.-M. Arvanitou, D. Gagoutis, A. Ampatzoglou, N. Mittas, I. Deligiannis, A. Chatzigeorgiou, What does matter in the success of a decentralized application? from idea to development, Information and Software Technology (2024) 107414.

[10] H. Janssen, J. Cobbe, C. Norval, J. Singh, Decentralized data processing: personal data stores and the GDPR, International Data Privacy Law 10 (2020) 356–384. doi:10.1093/idpl/ipaa016.

[11] I. Brown, Interoperability as a tool for competition regulation (2020).

---

[3] Data act: member states agree common position on fair access to and use of data. https://www.consilium.europa.eu/en/press/press-releases/2023/03/24/data-act-member-states-agree-common-position-on-fair-access-to-and-use-of-data/

[12] M. D. Williams, N. P. Rana, Y. K. Dwivedi, The unified theory of acceptance and use of technology (utaut): a literature review, Journal of enterprise information management 28 (2015) 443–488.

[13] V. Venkatesh, M. G. Morris, G. B. Davis, F. D. Davis, User acceptance of information technology: Toward a unified view, MIS quarterly (2003) 425–478.

[14] J. R. Lewis, The system usability scale: past, present, and future, International Journal of Human–Computer Interaction 34 (2018) 577–590.

[15] E. Von Hippel, Democratizing innovation, the MIT Press, 2006.

[16] K. Kollnig, S. Datta, T. Şerban von Davier, M. Van Kleek, R. Binns, U. Lyngs, N. Shadbolt, 'we are adults and deserve control of our phones': examining the risks and opportunities of a right to repair for mobile apps, Association for Computing Machinery, 2023, pp. 22–34.

[17] A. Tamò-Larrieux, Z. Zihlmann, K. Garcia, S. Mayer, The right to customization: conceptualizing the right to repair for informational privacy, in: Annual Privacy Forum, Springer, 2021, pp. 3–22.

[18] K. L. Fuchs, Empowering Diet-Related Health Behavior Change Interventions via Digital Receipts and Food Composition Databases, Ph.D. thesis, ETH Zurich, 2020.

[19] B. T. Kaveladze, S. D. Young, S. M. Schueller, Antifragile behavior change through digital health behavior change interventions, JMIR formative research 6 (2022) e32571.