Towards a Decentralized Data Privacy Protocol for Self-sovereignty in the Digital World

Rodrigo Falcão[1,*], Arghavan Hosseinzadeh[1]

[1]*Fraunhofer IESE, Fraunhofer-Platz 1, 67663 Kaiserslautern, Germany*

Abstract

A typical user interacts with many digital services nowadays, providing these services with their data. As of now, the management of privacy preferences is service-centric: Users must manage their privacy preferences according to the rules of each service provider, meaning that every provider offers its unique mechanisms for users to control their privacy settings. However, managing privacy preferences holistically (i.e., across multiple digital services) is just impractical. In this vision paper, we propose a paradigm shift towards an enriched user-centric approach for cross-service privacy preferences management: the realization of a decentralized data privacy protocol.

Keywords

personal data, privacy, protocol

1. Motivation

Not long ago, users navigated through Internet-based services (e.g., websites) leaving behind an unnoticed trail of personal data, browsing habits, and online interactions. Tracking technologies such as cookies, the development of data analysis techniques, and the rise of tech giants such as Google and Facebook (whose online presence became pervasive) have given rise to a massive concentration of data, which has supported targeted advertising for diverse purposes, sometimes raising ethical concerns [1]. These concerns led to the emergence of several data protection initiatives around the globe, resulting in regulations such as the GDPR [2] and the Data Governance Act [3], which aim to facilitate data exchange while increasing trust and ensuring data protection. One of the most tangible consequences a user experiences occurs when they visit a website using cookie technology: On the first access, they are asked to accept or reject cookies.

While users have been given some sort of freedom to choose which types of data are used by different services and for what purposes, and have also been entitled to change their preferences and revoke permissions for data usage, it is still hard for them maintain sovereignty over their choices. Gradually, their data and their choices regarding data usage are spread across the dozens, hundreds, or even thousands of digital services they interact with. In the long run, it is easy to lose track of where data is stored and used, by whom, and for what purpose. Taking

Joint Proceedings of RCIS 2024 Workshops and Research Projects Track, May 14-17, 2024, Guimarães, Portugal

*Corresponding author.

rodrigo.falcao@iese.fraunhofer.de (R. Falcão); arghavan.hosseinzadeh@iese.fraunhofer.de (A. Hosseinzadeh)

http://rodrigofalcao.info/ (R. Falcão)

0000-0003-1222-0046 (R. Falcão); 0009-0001-8699-9972 (A. Hosseinzadeh)

back control of the usage of their data becomes a near-impossible endeavor. Even if each digital service were shipped with a comprehensive and easy-to-use privacy preferences management tool, users would still have to remember all the services that use their data, visit them, and, one after another, review their privacy preferences if they want to. In other words, from the end user's perspective, privacy preferences management solutions are distributed across virtually all digital services they interact with. For this very reason, they cannot be effective. To make matters worse, there is a lack of standardization in the field.

In this vision paper, we propose a paradigm shift in how we approach the issue of privacy preferences management. The current paradigm focuses on the interaction between a user and a digital service. It encompasses usable privacy management tools, which are a necessary but insufficient solution. We envision a future where a fully user-centric approach takes a holistic view of all digital services, allowing users to manage their preferences in one place, without any particular service provider monopolizing this space. This is where the decentralized data privacy protocol comes into play. It can be implemented by any party and offers benefits to both end users and service providers. It also opens up opportunities for the development of new privacy-enhancing technologies on top of it. The remainder of this paper is structured as follows: In Section 2, we review recent related work; Section 3 outlines the protocol concept; Section 4 explores the benefits and opportunities; and Section 5 concludes the paper, outlining our next steps.

2. Related work

Several years prior, the World Wide Web Consortium (W3C) developed the Platform for Privacy Preferences (P3P) to enable web browsers to automatically read, interpret, and compare website privacy policies against user preferences or settings [4]. Despite its initial innovation, few websites have adopted P3P, and growing privacy needs have surpassed the protocol's capabilities.

The most recent advancements in research with respect to preserving user privacy focus on privacy policy languages and user-friendly settings for privacy preferences. Gharib [5] argues that most data subjects blindly accept the notices, not because they do not value their privacy, but because most privacy policies and terms of services are long, complex, and hard to understand, so the author introduces a model for informed consent. The model involves a Matching Component that compares the privacy preferences of the user that is included within their Personal Privacy Profile (PPPo) and the privacy policies that are published by the service providers and automates the process of giving consent. A dynamic contextual notice is provided to the user when the user preferences and the policies do not match. Accordingly, the user can make an informed decision concerning the consent request. Gharib also proposes an ontology that can be used for realizing preferences and policies. However, where to store the PPPos seems to be out of the scope of this work.

Dehling et al. [6] introduce Privacy Cockpits, which are central dashboards for users to navigate and manage their personal data. This solution aims to ease the enforcement of regulations such as GDPR, although it focuses on protecting the data used across various services within specific digital ecosystems.

In 2020, the European Data Protection Supervisor introduced the Personal Information

Management Systems (PIMS) concept [7]. The PIMS concept offers a new approach in which individuals are the "holders" of their own personal information. PIMS aims to empower users to take charge of their digital identity and the use of their personal information across various services and platforms. In recent years, several initiatives and projects have claimed PIMS features. Among these, the Solid protocol [8] stands out. It proposes a set of conventions and tools for building a decentralized platform for social Web applications. In order to address the challenge of obtaining consent for processing personal data, Florea and Esteves [9] introduced a policy layer into the Solid ecosystem. They integrated the usage of the ODRL Policy Language [10], the ODRL profile for Access Control (OAC) [11], and the Data Privacy Vocabulary (DPV) [12] to allow Solid users to express their privacy preferences. By integrating such a policy layer into the Solid ecosystem, the matching process regarding users' preferences and requests for data can be automated. However, the focus of the Solid protocol remains on data management mechanisms that allow users to store their data such as contacts and photos in Personal Online Datastores (PODs) and control access of applications to this data. We therefore still see an ongoing need for a protocol specifically designed for privacy preferences management.

3. A paradigm shift

We propose a paradigm shift by adjusting the context of the problem. Users need adequate means to exercise their data sovereignty. Tools can implement diverse models to provide such means; however, regardless of whether we consider the implementation of the "notice and consent model", the "informed consent model", or "data protection cockpits", to name but a few strategies, solutions address the data sovereignty challenge in the context of the interaction between the user and *a given digital service (or ecosystem)*. Consider now the exercise of data sovereignty in the context of not only one digital service, but *all of them*. If we have this goal in mind, current strategies – though valuable and necessary – fall short. It is not enough that *each digital service* provides its own means for users to manage their preferences; instead, *each individual* should be able to manage their data preferences *across all digital services that use their data*. A new centralized digital service could fill this gap; however, to be effective, it would require the providers of digital services to adhere to it, and users would still have to rely on a centralized service offered by yet another service provider.

Due to these reasons, we argue that a more adequate solution should be positioned at a higher level of abstraction and be developed as an open protocol that anyone can implement. We envision a decentralized protocol for users and services to manage data privacy preferences. A protocol can be defined as a set of syntactic and semantic rules that standardize communication between two or more entities. Using a defined protocol, users could tell the digital services *where* they want to store and manage their privacy preferences. We call this place the user's *Personal Privacy Preferences Place (P4)*. P4 instances could be hosted by a user's trusted third-party or even be self-hosted. For example, Figure 1 illustrates a scenario where two users, Alice and Bob, use different sets of digital services. Alice manages her privacy preferences through a P4 instance hosted and operated by a trusted company, while Bob manages his using a self-hosted P4 instance. The digital services can interoperate with both P4 instances because the digital services and the P4 instances implement their roles in the protocol.

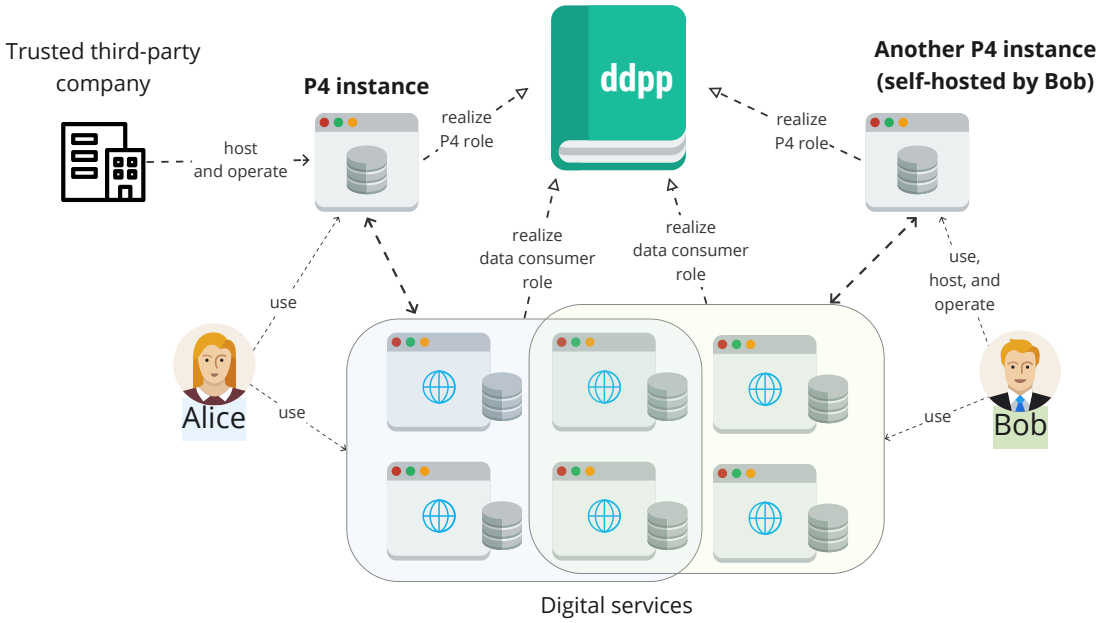

Figure 1: Overview of the decentralized data privacy protocol (ddpp).

Shaping the requirements. The goal of the protocol is to improve the usability and self-sovereignty of privacy preferences from the point of view of end users in the context of numerous digital services. To achieve this goal, the protocol must fulfill certain quality and functional requirements. From a quality perspective, the list includes (but is not limited to) *openness* (the protocol shall have an open specification in order to enable operational independence, meaning that anyone could implement its elements), *adaptability* (The protocol should not prevent providers from setting their privacy preferences freely, i.e., they shouldn't be required to change the way they define privacy preferences), and *confidentiality* (the personal privacy preferences place shall not store nor exchange private user data, but only users' privacy preferences data). Concerning the functional aspects, we highlight three fundamental constructs: the *data structure* (the protocol shall specify a privacy preferences meta-model whose instances express privacy preferences using generic elements), the *behavior* (the protocol shall specify the interaction flows between the participants, namely the user, their P4, and the digital service), and the *interfaces* (syntactic and semantic description of each interface that the participating systems must implement to enable the desired behavior).

On the interaction flows. The protocol must support at least two key flows: *handshake* and *update*. Using the handshake flow, the user informs their digital service about their P4 instance. After the user sets their initial privacy preferences and after the authorization process, the digital service can communicate with the user's P4 instance to exchange privacy preferences data. Using the update flow, every change that the user makes in the privacy preferences on their P4 should be reflected in the affected digital services, and every change made in a digital service should be reflected in the user's P4.

4. Benefits and opportunities

The envisioned approach is an additional step towards enabling data sovereignty as the implementation of the protocol takes the exclusive control of privacy management means from the service providers and gives it to the actual data subjects, i.e., the users. Also, following the same specification language and a standard ontology can facilitate the matching process between privacy policies and user's preferences and further automate it

From the point of view of the service provider, and given that we are living in a regulated society concerning data privacy, non-compliance poses a significant financial risk due to security, safety, and trust issues. Therefore, adherence to an open standard would help service providers transfer at least part of the risk beyond the boundaries of their companies. Furthermore, adopting an open privacy management standard would increase transparency and help build trust in the service providers.

This idea can be boosted by the adoption of self-sovereign identities (SSI), which have gained increasing traction through initiatives such as the eIDAS regulation [13]. The location of the user's privacy preferences could be tied to their identities. When a user provides their identity to an arbitrary digital service, the service can directly read the user's privacy preferences from the P4 instance.

Customized and optimized P4 instances can give users extended privacy management capabilities in comparison to those offered by their digital services. For example, while a certain digital service may only allow for either consenting or denying access to certain data for a given purpose, P4 instances can provide users with the ability to set dynamic rules or constraints on their preferences (e.g., revoke consent for data X for the purpose Y in all digital services located in Z 30 days after consent was granted). The development, customization, and operation of P4 instances opens up new business opportunities for companies.

5. Conclusion and outlook

So far, it has been hard for users to maintain sovereignty over their privacy preferences because privacy preferences management is not centered on the users but scattered across numerous digital services. In this paper, we sketched our vision of a decentralized data privacy protocol. We acknowledge that the vision does not provide details, which can make a difference in realizing a robust protocol. The challenges to overcome will become more evident as the details are added.

The implementation of this protocol will allow users to manage their privacy preferences effortlessly. It primarily enables communication between web services and P4 instances but is also available for any digital service to implement and utilize. Research plays a key role in the fulfillment of this vision. From our point of view, the next steps include a review of extensible data privacy meta-models, the design of the protocol flow, the design of a reference architecture for P4, and prototyping of a reference implementation.

Acknowledgments

This work has been funded by the German Federal Ministry of Education and Research (BMBF) (grant numbers 16KIS1507 and 16KIS1510).

References

[1] M. K. Daoud, D. M. Al-Qeed, J. A. Al-Gasawneh, A. Ziani, Examining the ethical implications of data privacy and targeted advertising in digital marketing: Consumer perceptions, in: SNAMS-2023, IEEE, 2023, pp. 1–6.

[2] European Parliament and Council of the European Union, Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679, 2016. Accessed on 2024-02-15.

[3] European Union, Data governance act (2022), https://eur-lex.europa.eu/eli/reg/2022/868/oj, 2022. Accessed on 2024-02-12.

[4] World Wide Web Consortium (W3C), Platform for privacy preferences (p3p) project, Accessed 2024. URL: https://www.w3.org/P3P/, accessed on 25 March 2024.

[5] M. Gharib, Toward an architecture to improve privacy and informational self-determination through informed consent, Information & Computer Security 30 (2022) 549–561.

[6] F. Dehling, S. Ludborzs, A. Weßner, R. Falcão, Konzepte für gebrauchstaugliche Datenschutzfunktionen in digitalen Ökosystemen, Datenschutz und Datensicherheit-DuD 48 (2024) 95–102.

[7] European Data Protection Supervisor, Techdispatch 3/2020: Personal information management systems, 2021. URL: https://data.europa.eu/doi/10.2804/096824, accessed on 2024-02-12.

[8] 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).

[9] M. Florea, B. Esteves, Is Automated Consent in Solid GDPR-Compliant? An Approach for Obtaining Valid Consent with the Solid Protocol, Information 14 (2023) 631.

[10] World Wide Web Consortium, ODRL Version 2.0, https://www.w3.org/ns/odrl/2/, 2021. Accessed on 2024-02-15.

[11] B. Esteves, H. J. Pandit, V. Rodríguez-Doncel, ODRL Access Control Profile, https://besteves4.github.io/odrl-access-control-profile/oac.html, 2021. Accessed on 2024-02-15.

[12] World Wide Web Consortium, Data Privacy Vocabulary (DPV), https://w3c.github.io/dpv/dpv/, 2023. Accessed on 2024-02-15.

[13] European Commission, eIDAS Regulation, https://digital-strategy.ec.europa.eu/en/policies/eidas-regulation, 2023. Accessed on 2024-02-22.