Taylor & Francis
Taylor & Francis Group

International Journal of Human–Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: www.tandfonline.com/journals/hihc20

Understanding User Perceptions of Personal Data Stores: A Prototype-Driven Multi-Scenario Study

Tim Theys, Peter Mechant, Lieven De Marez & Jelle Saldien

To cite this article: Tim Theys, Peter Mechant, Lieven De Marez & Jelle Saldien (19 Aug 2025): Understanding User Perceptions of Personal Data Stores: A Prototype-Driven Multi-Scenario Study, International Journal of Human–Computer Interaction, DOI: 10.1080/10447318.2025.2545467

To link to this article: https://doi.org/10.1080/10447318.2025.2545467

Published online: 19 Aug 2025.

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2025.2545467

Understanding User Perceptions of Personal Data Stores: A Prototype-Driven Multi-Scenario Study

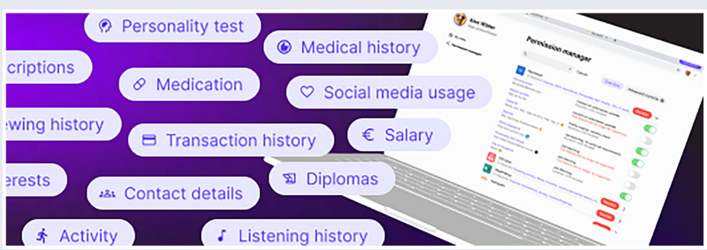
Tim Theys[a], Peter Mechant[b], Lieven De Marez[b] and Jelle Saldien[a,c]

[a]Department of Industrial Systems Engineering and Product Design, Ghent University, Ghent, Belgium; [b]Department of Communication Sciences, Ghent University, Ghent, Belgium; [c]Department of Product Development, University of Antwerp, Antwerp, Belgium

ABSTRACT

While Personal Data Stores (PDSs) aim to support personal data management and enable data portability across applications, user interest remains limited. This study explores the factors influencing user adoption and willingness to share data across four PDS use case scenarios—HR, Media, Finance, and Health—through an iterative research-through-design approach. The study combined a large-scale survey ($n = 2{,}335$), prototype development, expert reviews ($n = 10$), and user testing ($n = 17$). Findings reveal that although security concerns persist, participants responded positively to the PDS concept, particularly for use cases that reduce administrative burdens. Notably, users who engaged with the prototype reported higher intentions to use and greater willingness to share data than those evaluating written scenarios. These findings highlight the importance of tangible user experiences in shaping perceptions of PDSs. The study concludes with actionable recommendations and an illustrative prototype to inform effective user-centred PDS implementation.

GRAPHICAL ABSTRACT



KEYWORDS
Personal data stores; human data interaction; solid; prototyping; technology adoption

1. Introduction

With the growing number of digital platforms requesting, collecting and processing personal data, users struggle to maintain oversight of who has access to their information, let alone exert meaningful control over it (Solove, 2020). As data is scattered across a large number of service providers, it exposes users to significant privacy and security risks (Fallatah et al., 2023; Mansour et al., 2016). Additionally, it restricts users' ability to transfer their data across platforms as data portability remains largely impractical (Drechsler, 2018). This further reinforces the dominance of large corporations that control vast data resources, making it harder for new service providers to compete. As a result, competition is largely driven by data collection rather than providing the best services (Verborgh, 2023).

Personal Data Stores (PDSs) have emerged as a solution to enhance transparency and control over personal data usage (Larsen et al., 2015). By enabling users to aggregate data from various sources—such as media consumption data, demographics, social media data, and health information—into a

CONTACT Tim Theys tim.theys@ugent.be Department of Industrial Systems Engineering and Product Design, Ghent University, Imec-Mict-UGent, Miriam Makebaplein 1, 9000 Ghent, Belgium

user-governed repository, PDSs separate data storage from service providers (Fallatah et al., 2023). Through user-controlled access management, PDSs facilitate data portability, allowing individuals to seamlessly transfer their data across platforms. As a result, PDSs pave the way for new services that leverage cross-domain data to generate new insights and interactions (Bernes, 2022).

However, despite its promising potential, PDS adoption remains limited. This is partly due to ongoing technical and legal challenges faced by companies implementing PDSs (Janssen et al., 2021), but also because end-users' interest is lacking (Fallatah et al., 2023; Van Kleek & O'Hara, 2014). Factors such as the absence of tangible user experiences (Van Kleek & O'Hara, 2014), trust and security concerns (Steedman et al., 2020), and fears of complex usability and time consumption (Chaudhry et al., 2015b) have been identified as key obstacles to user adoption.

Yet, studies that actively involve end-users in evaluating PDSs and provide them with a tangible experience remain scarce (Hartman et al., 2020). Existing user research typically focuses on a single application domain and has shown potential user interest in PDSs for applications such as enhanced media recommendations (Sailaja et al., 2021b; Theys et al., 2024b) and more efficient job application processes (Theys et al., 2024a), despite ongoing user concerns about security and conceptual complexity.

Building on this prior research, this study takes a broader approach by examining both barriers and drivers of PDS adoption across multiple use cases. By focusing on four different application domains, users can experience the broader potential of data portability. Guided by industry interests and academic relevance, four distinct use cases were explored (a more detailed description of the use cases can be found in Table 2):

- HR: An online career platform uses data available in the PDS to match candidates with relevant jobs more effectively by accessing relevant career data with user permission.
- Media: Media platforms such as streaming services can request access to relevant data in the PDS to enable enhanced, more personalized recommendations.
- Finance: A financial dashboarding platform leverages PDS data to provide users with a complete overview of their finances across banks and investment platforms, helping them track income, expenses, and assets to achieve their financial goals.
- Health: A health platform enables users to share data from their PDS with doctors, supporting more informed and personalized health monitoring.

The potential of these use cases were explored by employing an iterative research-through-design approach (Zimmerman et al., 2007) to gain deeper insights into users' attitudes and willingness to share data. Drawing on the Theory of Planned Behaviour (TPB) (Ajzen, 1991), the study specifically examined which factors influence users' intention to adopt them. First, a large-scale survey ($n = 2{,}335$) was conducted to gain initial insights into users' intentions to adopt PDSs based on written use case scenarios, addressing the following research questions:

- **RQ1:** How does the initial adoption potential of the four use cases compare based on a written scenario?
- **RQ2:** What is users' initial willingness to share data across the four use cases based on a written scenario?

The written scenarios were then transformed into a tangible prototype, designed in alignment with the core principles of the Human Data Interaction (HDI) framework—legibility, agency, and negotiability (Mortier et al., 2014). To refine the initial prototype its usability and assess feasibility it underwent a series of expert reviews ($n = 10$). A second iteration was subsequently evaluated through user testing ($n = 17$) using a think-aloud protocol to gain qualitative insights into users' attitudes across the different use case scenarios, addressing:

- **RQ3:** What factors influence user intention to adopt PDSs across the different use case scenarios?

Participants who interacted with the prototype also assessed their behavioural intentions and willingness to share data. This allowed us to compare their responses to those from participants who only evaluated written scenarios. Given that a lack of tangible experience has been identified as a barrier to adoption (Fallatah et al., 2023), this comparison addressed:

- **RQ4:** Does interacting with a prototype influence users' intention to adopt PDSs across the different use case scenarios compared to those who only evaluated a scenario description?
- **RQ5:** Does interacting with a prototype influence users' willingness to share data across the different PDS use case scenarios compared to those who only evaluated a scenario description?

Finally, insights from the user testing informed the design of the third and final prototype iteration. By synthesizing insights from the different research steps, this study seeks to provide a clearer understanding of users' perceptions of PDSs and their potential applications. In doing so, it aims to inform businesses considering PDS implementation while highlighting opportunities to enhance data transparency, user control, and cross-domain data portability.

2. Background

2.1. Personal data stores

A Personal Data Store (PDS), also known as a Personal Online Data Store (Pods) (Dedecker et al., 2022) or data vault (Verbrugge et al., 2021) is a technology that enables individuals to collect, store, update, correct, analyze, and share their personal data. A key feature of a PDS is its granular consent management, allowing users to grant or revoke third-party access at any time (Fallatah et al., 2023; Larsen et al., 2015). Figure 1 illustrates how personal data stores can hold any type of data and decouple data storage from the service providers, thereby supporting data portability.

The PDS model holds several anticipated advantages (Fallatah et al., 2023), primarily empowering individuals by giving them access control over their data while promoting data portability for seamless and user-consented transfers (Bernes, 2022). This decentralized approach not only enhances privacy but also enables new data-driven applications by integrating information across silos, granting access to a diverse range of personal data, such as media consumption habits, health records, shopping history, and financial information. Additionally, some PDS platforms introduce monetization models (Bourgeus et al., 2024) enabling users to earn rewards for sharing their data. PDS systems also support self-quantification and personal informatics, allowing users to analyze and reflect on their data for improved self-awareness and decision-making (Fallatah et al., 2023). Examples of PDS applications include CitizenMe[1], Ockto[2] and my:D[3]—all certified as MyData operators (Langford et al., 2022) for promoting user-centric data practices (Bourgeus et al., 2024).

Despite the anticipated advantages, PDSs have yet to achieve widespread adoption due to several legal, technical, and social challenges. Legal uncertainties, such as defining data controllers and processors (Janssen et al., 2021), remain unresolved, while technical barriers like cross-platform interoperability

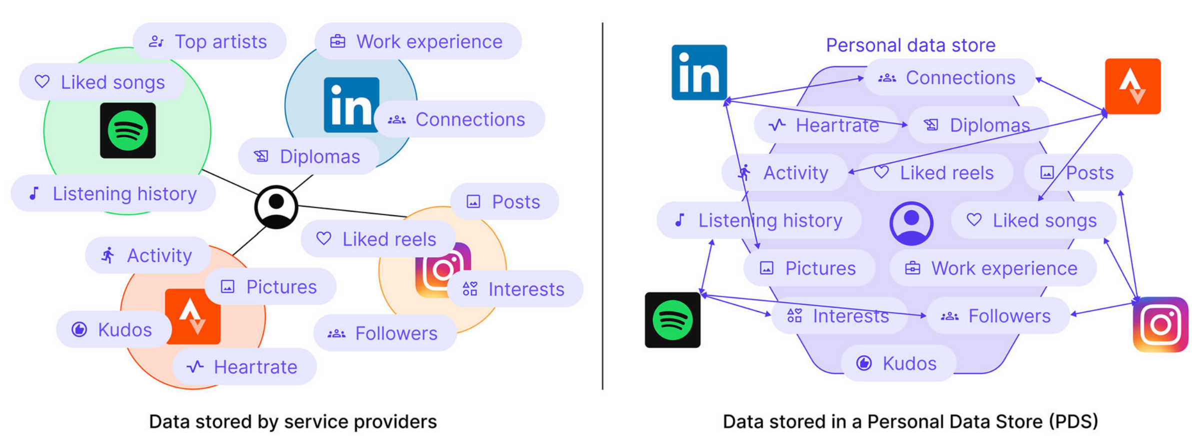


Figure 1. Comparison of the traditional web, where service providers store user data (left), and the decentralized web, where data is stored in personal data stores (right).

Table 1. Social challenges facing PDS adoption.

Challenge	References
Low user interest in PDS.	(Fallatah et al., 2023; Van Kleek & O'Hara, 2014)
Lack of tangible experiences to attract users.	(Fallatah et al., 2023; Van Kleek & O'Hara, 2014)
Trust concerns regarding PDS providers.	(Fallatah et al., 2023; Chaudhry et al., 2015b; Steedman et al., 2020; Theys et al., 2024a)
Fear of complex usability and having too limited technical skills.	(Fallatah et al., 2023; Chaudhry et al., 2015b; Theys et al., 2024a; Theys et al., 2024a)
Perceived time burden in managing data access.	(Steedman et al., 2020; Theys et al., 2024a)
Concerns about PDS security and hacking.	(Steedman et al., 2020; Theys et al., 2024a)
Expected financial cost of having a PDS.	(Chaudhry et al., 2015b)

Table 2. Use case descriptions and related data types as used in the survey (translated from Dutch).

HR
An online career platform tries to better match companies with candidates for certain positions. To do this, they can use data stores. With your permission, such a platform can read your diploma and previous work experience. The platform itself can also store data in your data locker, such as the results of a personality test, what work environment you are looking for, and what type of jobs, with your permission. The added value? With all this information, the platform can look for matches with jobs and companies that can select more specific profiles. If the platform finds a match between you and a job opening, the company can get your contact information to contact you if the position interests you.
Data types: Personality test, contact details, diplomas and, work experience
Media
When you watch, read or listen to media, media platforms collect usage data from you. Media platforms normally keep this data to themselves. Thanks to a personal data store, your media use can be taken with you and shared between different media platforms, for example, from VRT to Netflix and from Spotify to YouTube. You control your personal data store: you grant permission to the media platforms to read and share data from these different sources. If you want, you can also use your personal data store to log into the media platforms. The added value? By bringing types of data together, media platforms can better assess your interests and provide better recommendations tailored to your needs for what to watch, read or listen to.
Data types: Social media connections, social media usage (e.g., following pages and saved posts), location, library history (e.g., books borrowed), listening history (e.g., music and podcasts listened to) and, viewing history (e.g., series/film watched online)
Finance
A personal data store can be connected to financial services. Thus, data about what you own at which bank and other sources of income (stocks, funds, …) can be brought together in your data safe. Your expenses, taxes, subscriptions and so on can also be stored. On a website you can connect your data store and choose which data may be used. The added value? Based on this, the website can immediately give you an overview of all your income, expenses and total assets. On this website you can then find recommendations, for example, on how much you should best save or which expense items might be better eliminated to achieve your financial goals.
Data types: Financial instruments (e.g., stocks, funds and crypto), transaction history (e.g., purchases and expenses), salary and, subscriptions
Health
A personal data store can store medical (and other) data about you. This can be all kinds of data. Think previous lab results, your medical history, weight/bmi or data from your smartphone & smartwatch (heart rate, sleep, …). Once brought together, more personalized health advice can be provided. You can choose what data you store and share in it. The added value? Suppose you are in treatment with a (family) doctor, he or she can use this personal data store and your data that is available here. In this way, more targeted health advice can be given.
Data types: Smart device data (e.g., heart rate, weight, sleep measurements), lifestyle habits (e.g., workouts, smoking and alcohol consumption), medications and, medical history

complicate implementation (Fallatah et al., 2023). In this study we focus however on the various social factors hindering user adoption. Table 1 summarizes the social challenges identified in previous research.

To overcome these challenges and drive adoption, researchers suggests that the focus of PDS initiatives should shift from being a purely privacy-enhancing tool to a solution that delivers clear, practical benefits to individuals (Van Kleek & O'Hara, 2014). PDS platforms must simplify and enhance users' daily lives, offering seamless data management, personal insights, and effortless sharing options while minimizing complexity and time investment (Fallatah et al., 2023).

Solid (Social Linked Data) is a notable web specification that enables Personal Data Stores (PDSs) by facilitating a decentralized web (Sambra et al., 2016). As Solid is gaining traction across various domains, including media (e.g., jouw.id[4] for personalized news recommendations), real estate (e.g., Linkcr[5] for streamlined data sharing with notaries), and government services (e.g., a Flemish initiative for securely sharing diplomas with recruitment agencies), its architecture served as the foundation for the use cases explored in this study.

2.2. The use cases

Our study explored four potential use case scenarios in which PDSs could potentially enhance the user experience. By examining distinct domains—HR, Media, Finance, and Health—we aimed to demonstrate the potential of cross-domain data portability. Table 2 provides a detailed overview of each use case as presented to participants, including the specific data types involved. The selection and design of these use cases were guided and inspired by academic relevance, previous research, active interest from industry stakeholders, and current applications.

Prior research has explored the use of PDSs in HR to streamline the application process by matching candidates with jobs based on a rich variety of data and by simplifying the sharing of official documents such as diplomas (De Meester et al., 2025; de Mildt, 2024). First results showed notable user interest despite initial users hesitancy due to their unfamiliarity with PDSs (Theys et al., 2024a). Initiatives like the Flemish government's system for seamless diploma sharing with recruitment agencies (Vlaanderen, 2024) and Karamel (De Vrieze, 2023), a digital platform that anonymously matches candidates with jobs using PDS data, inspired the design of this use case scenario.

PDSs have also attracted the interest from the media sector, with major companies exploring their potential. The BBC, for instance, developed *Together + Data Pod* (BBC, 2023), a trial application allowing users to store streaming data in a PDS. They also investigated the *Cross Media Profiler*, which leveraged external platform data (e.g., Instagram, Spotify) to recommend BBC content (Sailaja et al., 2021a). Despite positive user evaluations, these initiatives never progressed to full deployment. The Flemish public broadcaster VRT also explored a PDS enhanced streaming service, where the added value was met with scepticism and the willingness to share data seemed to be more dependent on the perceived relevance rather than sensitivity of the data. More recently, the Dutch *jouw.id*[6] introduced PDS-enabled demo for personalized news recommendations (Stichting Nederlandse Datakluis, 2021). This interest is also reflected in Flanders, where the public broadcaster VRT is exploring how PDSs can enhance media consumption experiences (VRT, 2023). An initial study of their PDS-enhanced streaming service found that the added value was met with scepticism, and that users' willingness to share data appeared to depend more on the perceived relevance of the data than on its sensitivity (Theys et al., 2024b).

Likewise, the financial sector has shown growing interest in PDS applications, particularly for consolidating financial information. One key use case is cross-bank data aggregation, which enables users to create a unified dashboard of their personal finances across banks, investment platforms, and other related services. Examples of such tools include the Easy Banking App[7] and Nerdwallet[8]. Additionally, the Belgian bank KBC allows users to store personal documents—such as medical records and driver's licenses—in a secure digital vault, giving them control over access and duration of sharing (KBC, 2025).

Finally, the health sector also presents promising opportunities for PDS integration that are being explored. For instance, PDSs could facilitate the integration of individually generated health data (e.g., from fitness trackers) into personal health monitoring by doctors (Carmichael et al., 2024; Ghayvat et al., 2022). Additionally, they could streamline data donation for health and social care research (Carmichael et al., 2024; Sun et al., 2023). A notable example is *digi.me*[9], a platform that aggregates medical records for Dutch citizens, simplifying access and sharing to improve healthcare management.

The applications of PDSs are, however, not limited to these domains, as their potential is also being explored in areas such as mobility (D'Hauwers et al., 2025) and the Internet of Things (Pinto et al., 2024). The choice of the four selected use cases is motivated by their maturity and the potential to demonstrate data portability across application domains. The use case were connected in the prototype as they all relied on the data stored in one PDS as illustrated in Figure 2.

2.3. Related work

In the previous sections, we referenced several prior user studies on PDSs. Table 3 provides a comparative overview of these contributions. The prior work includes a usability study within the HR domain (Theys et al., 2024a), three media-focused studies using focus groups (Sailaja et al., 2021b; Steedman et al., 2020; Theys et al., 2024b), and a UK-based survey comparing general data management models, including PDSs (Hartman et al., 2020).

While these studies offer valuable insights—highlighting both the potential and concerns surrounding PDSs—they are limited to a single application domain and rely on either qualitative or quantitative methods. Our study aims to extend this literature in two ways. First, we adopt a mixed-methods

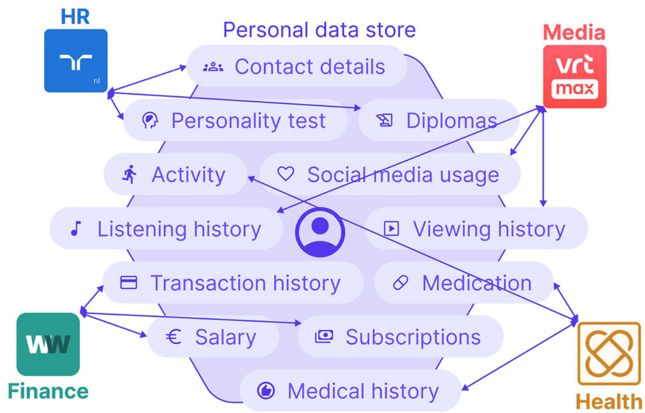


Figure 2. The four use cases all rely on the data stored in the PDS.

Table 3. Overview of prior user studies on PDSs.

Study	Domain	Methodology (sample size)	Main conclusion
Solid-enabled personal online data stores (Theys et al., 2024a)	HR: Diploma sharing with recruitment agencies	Usability testing ($n = 10$) using a functional demo	Users see added value for administrative simplification but struggle to understand the PDS concept.
Human data interaction in data-driven media experiences (Sailaja et al., 2021a)	Media: Cross media profiler	Focus groups ($n = 23$) using a prototype	Participants valued the transparency and control enabled by the PDS, though struggled to grasp the data flows.
Exploring users' perspectives on a solid-enabled personal data store enhanced streaming service (Theys et al., 2024b)	Media: Streaming service	Focus groups ($n = 19$) using a prototype	Participants balanced curiosity with concerns, though transparency and control features eased many concerns.
Complex ecologies of trust in data practices and data-driven systems (Steedman et al., 2020)	Media: BBC personalization	Focus groups ($n = 68$) based on a description of PDSs	PDSs were appreciated for offering control but raised concerns over effort and security.
Public perceptions of good data management: findings from a UK-based survey (Hartman et al., 2020)	General	Survey ($n = 2{,}169$) based on a textual description of PDSs	PDSs emerged as the most preferred data management model among eight options.

approach, combining large-scale quantitative data from a survey evaluating written scenario descriptions with qualitative insights from expert reviews and user testing involving an interactive prototype. Second, we explore four distinct application domains, thereby enabling cross-domain comparison and reflecting the broader promise of data portability of PDSs.

2.4. Theory of planned behaviour

To better understand the factors influencing PDS adoption, both drivers and barriers, we reviewed several relevant theoretical frameworks. Technology adoption can be explained through various models, including the Theory of Reasoned Action (Fishbein & Ajzen, 1975), the Theory of Planned Behaviour (Ajzen, 1991), the Technology Acceptance Model (Davis, 1989), and the Unified Theory of Acceptance and Use of Technology (Venkatesh et al., 2003). Given its broad applicability (Mathieson, 1991), strong predictive power (Godin & Kok, 1996), and focus on user attitudes, we selected the Theory of Planned Behaviour (TPB) as the guiding framework.

According to the TPB, behaviour is driven by individuals' intentions, which are shaped by three key factors (Ajzen, 1991), as displayed in Figure 3: (1) attitudes—the perceived desirability, usefulness, and overall favourability of performing the behaviour; (2) subjective norm—the influence of perceived social pressure and societal expectations; and (3) perceived behavioural control—an individual's perception of their ability to perform the behaviour, considering self-efficacy and ease of use.

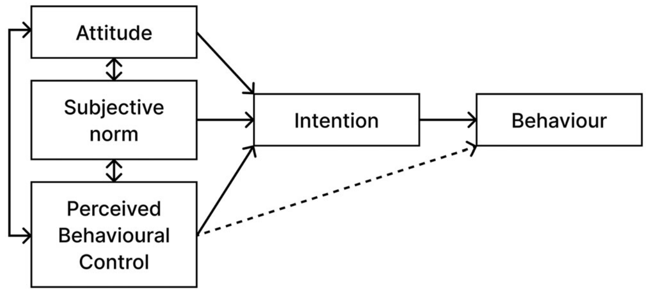


Figure 3. The theory of planned Behaviour (Ajzen, 1991).

In this study, we assess users' behavioural intentions toward PDS adoption using quantitative measures. The qualitative component complements this by exploring the underlying factor that shape those intentions, offering deeper insight into the factors influencing both acceptance and resistance.

2.5. Human data interaction

The multidisciplinary field of Human-Data Interaction (HDI) emerged from Human-Computer Interaction (HCI) in response to the increasingly pervasive processing of personal data (Mortier et al., 2014). HDI focuses on how people interact with data and emphasizes the importance of individuals' understanding of the ways in which their behaviour, the data it generates, and the algorithms that process it shape their daily life (Santos et al., 2018; Victorelli et al., 2020). By placing humans at the centre of data flows, HDI seeks to develop instruments that enable individuals to engage explicitly with data-driven systems rather than remaining passive subjects of automated processes.

Notable examples of such efforts include the Dataware model (McAuley et al., 2011), which provides users with a personal container to oversee and manage access to their data, similar to Solid-enabled personal data stores, and its successor, the Databox (Chaudhry et al., 2015a), a physical device designed to support the collection and controlled sharing of personal data.

At its core, HDI is guided by three key principles: legibility, agency, and negotiability (Mortier et al., 2014). These principles aim to address critical gaps in current data practices, where users often lack transparency, control, and the ability to negotiate how their data is used.

Legibility refers to making systems, data flows, and algorithmic outputs understandable to end users (Mortier et al., 2014). Given that data interactions are often opaque and available information is frequently technical and complex, legibility requires moving beyond basic transparency toward truly comprehensible data processes for those affected by them.

This involves not only ensuring that users are aware of the data collected about them but also helping them understand the implications of this data collection (Mortier et al., 2014). Achieving this goal requires more than simply providing data visualizations; it involves structuring information in a way that is accessible and meaningful without overwhelming users. However, this remains a significant challenge, particularly due to the sheer scale of data collection (Li et al., 2013), a complexity that is also likely to arise within personal data stores. Despite these challenges, legibility should be prioritized, as it serves as a prerequisite for user agency (Sailaja et al., 2021b; Segijn et al., 2021).

Agency involves empowering data subjects to take meaningful action within data systems (Mortier et al., 2014). This goes beyond the possibility to merely provide informed consent; it also includes the ability to revoke previously shared data and correct false inferences drawn from personal information (Whitley, 2009). By enabling users to actively engage with and refine collected information, such mechanisms could even enhance the accuracy and effectiveness of applications that rely on this data (Mortier et al., 2014). While not all users—or even a majority—are expected to regularly engage with advanced control features, the presence of these features remains essential, ensuring that individuals can exercise agency when needed, particularly in response to potential harm or triggered concerns (Barnes, 2006).

Negotiability refers to the ability of users to challenge, modify, or negotiate the sharing of their personal data as contexts evolve (Mortier et al., 2014). These changes may be external, such as shifts in legal jurisdictions or new data policies, or internal, where users' attitudes toward data sharing evolve over time based on novel experiences (Patil et al., 2014). A central challenge in personal data ecosystems is that power is often disproportionately concentrated in the hands of data aggregators and brokers, who control how personal data is stored, traded, and utilized (D'Hauwers & Bourgeus, 2024). To re-balance this power dynamic, the importance of contextual integrity (Nissenbaum, 2004) should be highlighted by ensuring that personal data is processed in ways that align with its original intent and user expectations. Achieving true negotiability requires not only user-facing tools, but also legal and regulatory reforms that reflect the complexity and dynamism of modern data flows (Solove, 2013; Westby, 2011).

In this study, the HDI framework was applied as guidance during the iterative prototyping process, with features designed and evaluated according to its core principles of legibility, agency, and negotiability.

3. Methodology

3.1. Research approach

To explore users' attitudes and their willingness to share data across the different use cases, we followed an iterative research-through-design process (Zimmerman et al., 2007), as visualized in Figure 4.

As a first step to gain initial insights into users' intentions and willingness to share data, a large-scale survey was launched ($n = 2{,}335$) in which potential end-users evaluated the four use cases (HR, Media, Finance and Health) based on written scenarios (Table 2). Following the survey results, these written scenarios were translated into a tangible prototype, which functioned as a research tool throughout the iterative design process. While not a fully functional system, the prototype served as a realistic simulation of user interaction. The primary objective of this prototype was to elicit user reactions regarding the different personal data store use cases and their willingness to share different data types, rather than to serve as a finalized model for PDS interface design.

The initial prototype was evaluated through expert reviews ($n = 10$), involving professionals with expertise in UX design, legal frameworks, privacy, and technical development, evaluating its design and feasibility. Their feedback informed a second iteration of the prototype, refining both functionality and usability. To gain deeper qualitative insights, this revised prototype was reviewed in user testing ($n = 17$), where participants navigated through the prototype following a think-aloud protocol. This stage allowed for a direct comparison between qualitative results and the quantitative findings from the survey.

The combined insights from these research steps led to the development of a third and final version of the prototype, along with the formulation of recommendations for effective personal data store implementation.

3.2. Ethical considerations

As this study involved human participants, ethical approval for the user testing was sought and granted by the faculty's ethical committee, ensuring compliance with the university's General Ethics Protocol for research involving human subjects. In accordance with these guidelines, all participants provided informed consent prior to their involvement in the study. The consent form outlined the purpose, procedures, and potential risks of the study, as well as participants' rights, including their ability to withdraw at any time without consequences. To further protect participants' privacy, the project's data management plan outlined specific measures, including the removal of personally identifiable information from the data.

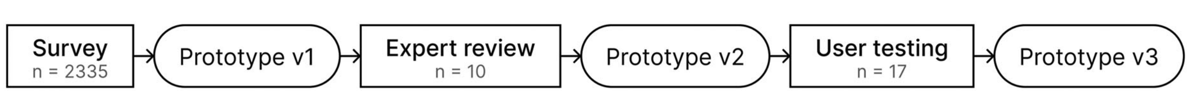


Figure 4. Overview of the different research steps.

4. Survey

4.1. Methodology survey

4.1.1. Participants

The survey, which assessed participants' intentions to use the proposed use case scenarios and their willingness to share data, was conducted with a sample of 2,335 Flemish citizens, following data cleaning procedures. Participants were recruited through a panel management agency, and data collection occurred between July and August 2023. Eligibility required respondents to be at least 18 years old and proficient in Dutch. Respondents with incomplete data or those who failed attention check questions (e.g., "Select 'rather agree' here") were excluded from the final analysis. The evaluation of the use cases was embedded within a larger survey, which had an estimated completion time of 30 min. Participants received compensation in the form of redeemable points through the panel management agency.

The final sample included 1,076 males (46.1%), 1,254 females (53.7%), and 5 individuals (0.2%) who identified as non-binary or preferred not to disclose their gender. Participants' mean age was 51 years ($SD = 16.83$). A detailed breakdown of demographic characteristics is presented in Table 4.

4.1.2. Measures

Participants' behavioural intention was assessed using a three-item scale adapted from prior research (Dwivedi et al., 2017), as displayed in Table 5. Each item was rated on a 7-point Likert scale ranging from "completely disagree" to "completely agree." Internal consistency was assessed using Cronbach's alpha and demonstrated excellent reliability for all use cases: HR ($\alpha = 0.97$), Media ($\alpha = 0.96$), Finance ($\alpha = 0.98$), and Health ($\alpha = 0.98$).

To assess participants' willingness to share data, they were presented with a list of relevant data types for each use case (Table 2) and asked to select the data types they would feel comfortable sharing within the given use case context.

4.1.3. Data analysis

Due to time constraints as this evaluation was part of a larger survey, participants were randomly assigned to one or two use cases. The final subsample sizes were: HR ($n = 835$), Media ($n = 826$), Finance ($n = 830$), and Health ($n = 832$). For each assigned scenario, participants assessed their behavioural intention and willingness to share data based on the use case descriptions provided in Table 2.

Table 4. Demographics survey sample.

	n	Proportion (%)
n (Total)	2335	
Gender		
Male	1076	46.1
Female	1254	53.7
Other/Prefer not to say	5	0.2
Age (average age = 51 years, SD = 16.83)		
18–24	179	7.7
25–34	297	12.7
35–44	385	16.5
45–54	406	17.4
55–64	429	18.4
65+	639	27.4
Education level		
No diploma	22	0.9
Primary education	26	1.1
Lower secondary education (1st and 2nd grade of secondary school)	177	7.6
Upper secondary education (3rd grade of secondary school)	976	41.8
Bachelor's degree	738	31.6
Master's degree/postgraduate	396	17.0
Professional situation		
Student	138	5.9
Working full-time	989	42.4
Working part-time	261	11.2
(Temporarily) not working (e.g., sabbatical, job seeking, …)	84	3.6
(Early) retirement	734	31.4
Not active in the labour market (e.g., homemaker)	129	5.5

Table 5. Behavioural intention scale items (Dwivedi et al., 2017).

BI_1	I plan to use this application in the future.
BI_2	I would use this application in my daily life.
BI_3	I predict I will use this application.

Table 6. Pairwise comparisons behavioural intention ratings between use cases ($^{**}p < 0.01$, $^{***}p < 0.001$).

Comparison	Mean difference	SE	z-Ratio	p-Value
Finance – health	−1.042	0.076	−13.716	<0.001***
Finance – HR	0.030	0.076	0.388	0.980
Finance – media	−0.241	0.076	−3.163	0.009**
Health – HR	1.071	0.076	14.137	<0.001***
Health – media	0.801	0.076	10.534	<0.001***
HR – media	−0.270	0.076	−3.556	0.002**

To compare the behavioural intention ratings between the use cases and to investigate the influence of gender, age, and the use case, a linear mixed-effects model (LMM) was used (Meteyard & Davies, 2020). Given that participants were randomly assigned to the use cases, the data followed a repeated-measures structure, necessitating a model that accounts for within-subject dependencies. To address this, a random intercept for participants was included to control for baseline differences in behavioural intention ratings. The model was specified as follows:

$$Behavioural\ intention \sim gender + age + use\ case + (1|participant\ id)$$

Gender, age and use case were included as fixed effects, and participant ID was included as a random effect to account for individual differences. The LMM was estimated using the lmerTest package in R (Kuznetsova et al., 2017). *Post-hoc* pairwise comparisons between use cases were performed using estimated marginal means (EMMs) with Tukey adjustments for multiple comparisons, as implemented in the emmeans package (Searle et al., 1980).

To assess whether the proportions of participants willing to share different data types varied significantly within each use case, a series of Pearson's Chi-squared tests was performed separately for each use case using the stats package in R (R Core Team, 2021). For use cases with significant results, post-hoc pairwise comparisons of proportions for the individual data types were performed using the Bonferroni correction to adjust for multiple comparisons.

All statistical analyses were conducted using R version 4.4.1 (R Core Team, 2021). Data cleaning and preparation were performed using dplyr (Wickham et al., 2014), and results were visualized with ggplot2 (Wickham, 2016) and ggpubr (Kassambara, 2016).

4.2. Results survey

4.2.1. Behavioural intention

The linear mixed-effects model revealed a significant effect of use case on behavioural intention ratings, indicating that participants' intentions varied depending on the specific use case context. Compared to the Finance scenario (reference), intention ratings were significantly higher for the Health use case ($\beta = 1.04$, SE $= 0.076$, $t = 13.72$, $p < 0.001$) and the Media use case ($\beta = 0.24$, SE $= 0.076$, $t = 3.16$, $p = 0.002$), while no significant difference was observed for the HR use case ($\beta = -0.03$, SE $= 0.076$, $t = -0.39$, $p = 0.700$).

Post-hoc Tukey-adjusted pairwise comparisons between all use cases, as displayed in Table 6, further confirmed that the Health use case had significantly higher behavioural intention ratings than all other use cases ($p < 0.001$). Smaller but statistically significant differences were also observed between media and finance ($p = 0.008$) and between media and HR ($p = 0.002$), with media receiving slightly higher ratings. No significant differences were observed between Finance and HR use cases ($p = 0.98$). Figure 5 visualizes the estimated means including their 95% confidence intervals.

Additionally, the estimated fixed effects showed a significant negative association between age and behavioural intention ($\beta = -0.016$, SE $= 0.002$, $t = -9.204$, $p < 0.001$), indicating each additional year of age corresponds to a 0.016-point decrease in behavioural intention ratings. Figure 6 illustrates this

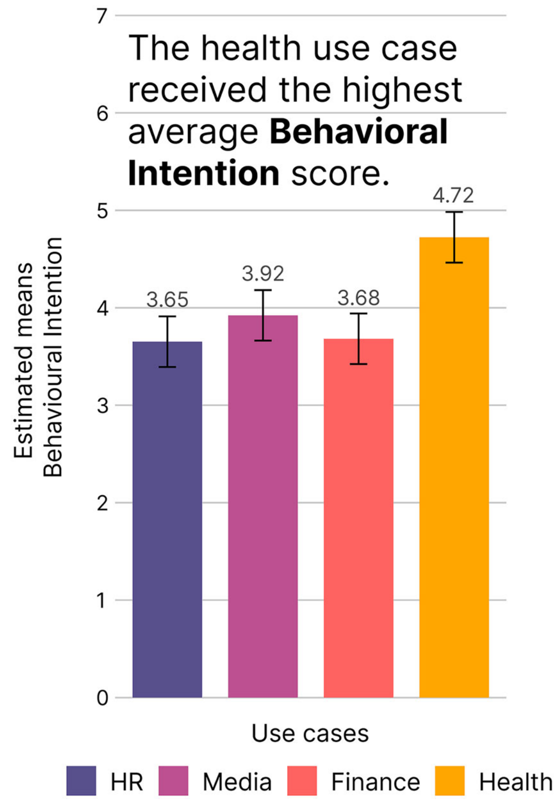


Figure 5. Estimated means for behavioural intention ratings across use cases including the 95% confidence intervals. For accessibility, detailed values are provided in Appendix A.

age-related trend across various use cases, with the most pronounced decreases observed in the HR, Media, and Finance use cases. In contrast, the Health use case appears less sensitive to age, maintaining relatively high behavioural intention scores across all age groups. Gender also had a small but significant effect ($\beta = -0.162$, SE $= 0.057$, $t = -2.855$, $p = 0.004$), with female participants reporting slightly higher behavioural intention ratings than male participants.

4.2.2. Willingness to share data

Pearson's chi-squared tests indicated significant differences in willingness to share data across all use cases: HR ($\chi 2 = 470.84, df = 3, p < 0.001$), Media ($\chi^2 = 391.2, df = 5, p < 0.001$), Finance ($\chi^2 = 32.29, df = 3, p < 0.001$) and Health ($\chi^2 = 441.11, df = 3, p < 0.001$). Figure 7 presents the proportions of participants indicating their willingness to share the individual data types.

For the HR use case, pairwise comparisons (Table 7) indicated that work experience (72%) and diplomas (71%) were significantly more likely to be shared compared to contact details (52%). Notably, only 26% of participants would be willing to share their personality test results.

Regarding the Media use case (see pairwise comparison in Table 8), participants indicated the highest willingness to share viewing history (38%) and listening history (37%), while the willingness to share library history data (22%) was significantly smaller. Participants were especially hesitant to share their location (14%), social media usage (11%), and social media connections (10%).

The results of the Finance use case (see pairwise comparison in Table 9) displayed a general reluctance to share financial data. Less than one third of the participants indicated to be willing to share subscription data (30%) and salary information (27%). Transaction history (25%) was even slightly less accepted, while financial instruments (18%) scored the lowest.

Remarkably, the willingness to share data in the context of the health use case (see pairwise comparison in Table 10) was substantially higher than in the Finance and Media use cases. Participants expressed a high willingness to share medical history (85%) and medication data (82%) and about half of the participants were willing to share their lifestyle habits (52%) and smart device data (48%).

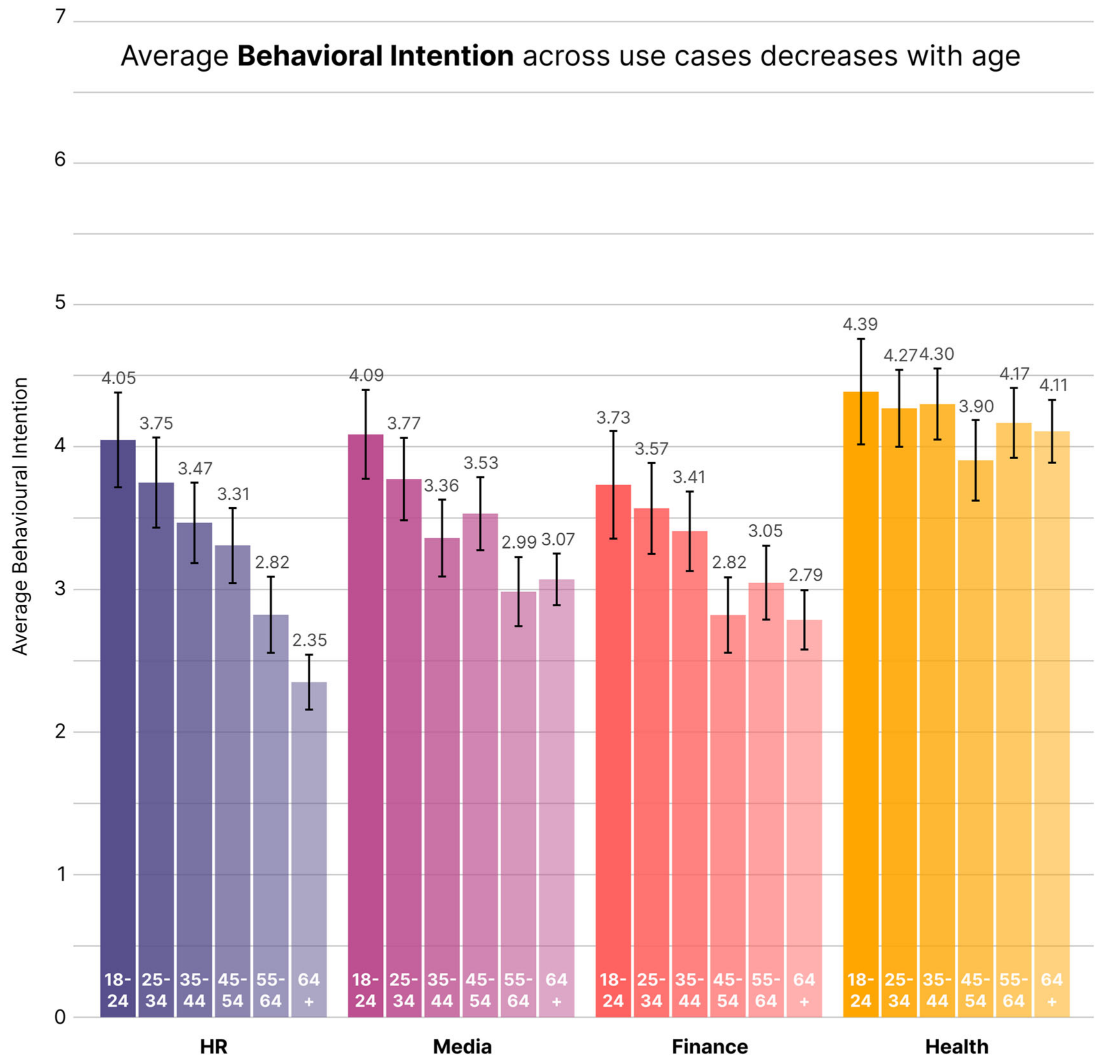


Figure 6. Average behavioural intention ratings across use cases and age groups including 95% confidence intervals. For accessibility, detailed values are provided in Appendix B.

5. Prototype

5.1. Design of the prototype

Following the survey and based on the descriptions of the different use cases in Table 2, we developed a prototype that served as a research tool for the subsequent user testing phase. By making the use cases more tangible, the prototype aimed to elicit user reactions and provide deeper insights into participant attitudes toward each use case.

The prototype was designed as a Figma mock-up, simulating real-world applications for each use case. Participants could navigate through a predefined path, mimicking a realistic user experience. No real-time data processing or backend logic was implemented. Instead, the prototype served as a research probe—its primary function was to elicit user reactions, expectations, and concerns in response to the concept and interface of PDSs. All screenshots of the prototype included in this paper already incorporate expert feedback, as discussed in Section 0. The prototype as presented to the participants can be accessed through the link in the footnotes[10].

The prototype was designed from the perspective of a user who already had a PDS containing a diverse range of data types. To make the participants familiar with the concept of a PDS, the user journey began with an exploration of the data dashboard (Figure 8), which displayed the different data types available in the personal data store. These data types were categorized into the following themes: personal, work and study, mobility, financial, body and health, media, calendar, and socials

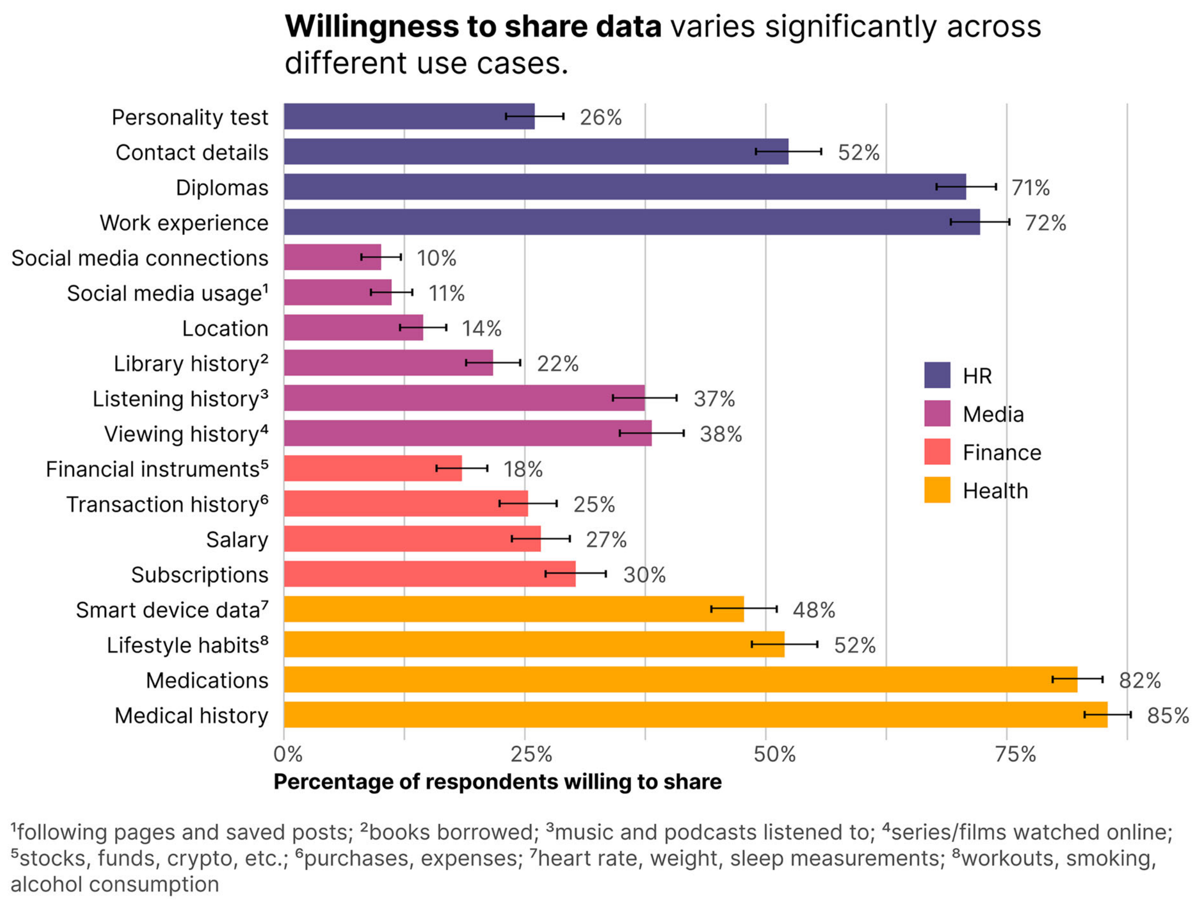


Figure 7. Proportion of participants indicating they are willing to share a certain data type within the given use cases including the 95% confidence intervals. For accessibility, detailed values are provided in Appendix C.

Table 7. HR use case pairwise chi-square comparisons of willingness to share data across different data types.

	Diploma	Work experience	Personality test
Work experience	1.0	—	
Personality test	<0.001***	<0.001***	—
Contact details	<0.001***	<0.001***	<0.001***

Adjusted *p*-values (Bonferroni correction) are presented alongside significance levels (***$p < 0.001$).

Table 8. Media use case pairwise chi-square comparisons of willingness to share data across different data types.

	Viewing history	Listening history	Social media usage	Social media connections	Library history
Listening history	1.0	—			
Social media usage	<0.001***	<0.001***	—		
Social media connections	<0.001***	<0.001***	1.0	—	
Library history	<0.001***	<0.001***	<0.001***	<0.001***	—
Location	<0.001***	<0.001***	0.830	0.129	0.002**

Adjusted *p*-values (Bonferroni correction) are presented alongside significance levels (**$p < 0.01$; ***$p < 0.001$).

Table 9. Finance use case pairwise chi-square comparisons of willingness to share data across different data types.

	Salary	Transaction history	Financial instruments
Transaction history	1.0	—	
Financial instruments	<0.001***	0.005**	—
Subscriptions	0.688	0.170	<0.001***

Adjusted *p*-values (Bonferroni correction) are presented alongside significance levels (**$p < 0.01$ ***$p < 0.001$).

Table 10. Health use case pairwise chi-square comparisons of willingness to share data across different data types.

	Medical history	Smart device data	Lifestyle habits
Smart device data	<0.001***	—	
Lifestyle habits	<0.001***	0.570	—
Medications	0.570	<0.001***	<0.001***

Adjusted *p*-values (Bonferroni correction) are presented alongside significance levels (***$p < 0.001$).

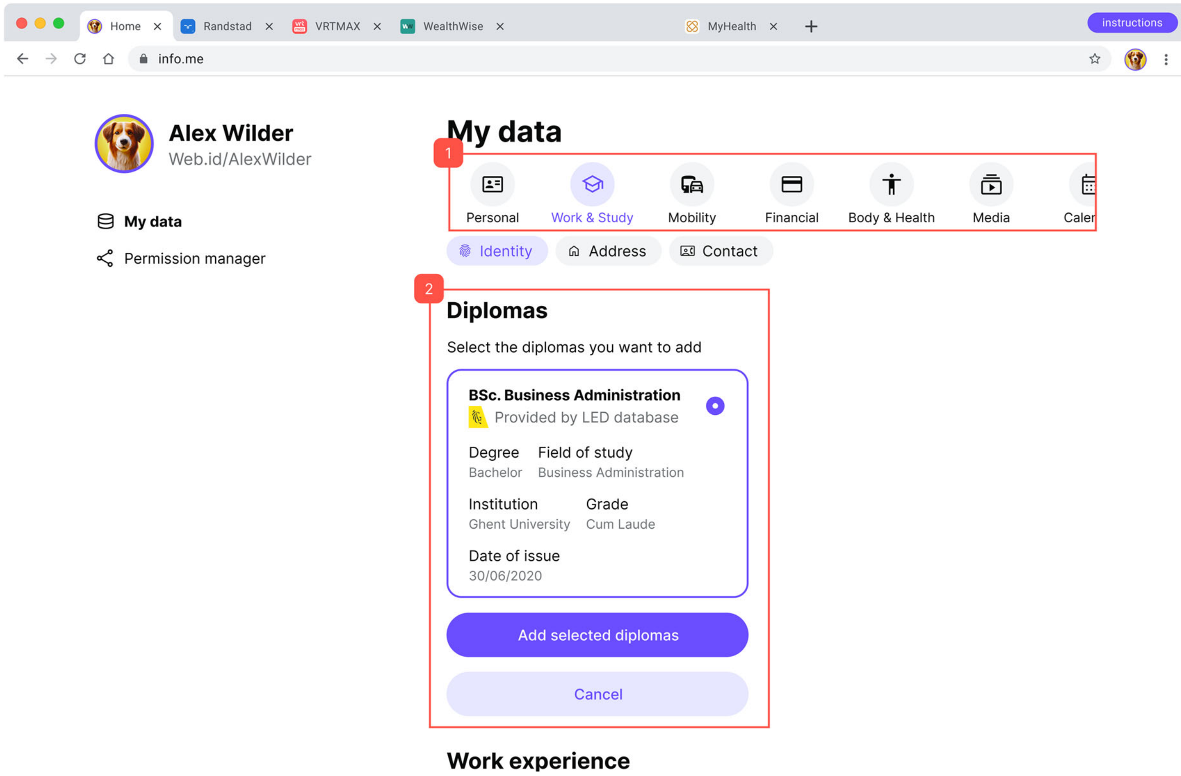


Figure 8. Screenshot of the personal data store data dashboard. (1) Navigation of the different data categories. (2) Example of how a diploma can be added to the personal data store.

(Figure 8, No. 1). As participants navigated through the different data categories, they were prompted to perform two key actions: adding a diploma (Figure 8, No. 2) to simulate the process of uploading formal credentials and connecting their Spotify data to illustrate how external data sources could be integrated into the PDS. These interactions were designed to give participants a more concrete understanding of data management within the platform.

After exploring and familiarizing themselves with the PDS, participants proceeded to the different use case scenarios, beginning with HR. This scenario started on the homepage of Randstad, a well-known recruitment agency, where they were invited to try a new feature that enabled them to share data available in their PDS for job matching. Clicking the button triggered a popup requesting access to specific data from their personal data store, including diploma, work experience, personality test results, and city of residence (Figure 9). Upon granting access, participants were shown a list of relevant job opportunities, each displaying a match percentage based on their data. After selecting a vacancy, users encountered an application form that could either be filled out manually or completed automatically by retrieving the necessary information—contact details, diplomas, and digital ID—from their PDS, thus finalizing the application process.

Secondly, participants were directed to the homepage of VRTMAX, the streaming platform of the Flemish public broadcaster, to explore the Media use case. A popup invited users to share various data types to generate personalized recommendations, including streaming history, music preferences, library borrowing history, followed pages and accounts on social media, contacts labelled as friends, and city of residence. Upon sharing their data, a row of personalized recommendations appeared at the top of the page (Figure 10, No. 1), accompanied by labels identifying the specific data sources used to generate the suggestions (Figure 10, No. 2). Users could modify the recommendation algorithm by adding or excluding specific data types. Additionally, when hovering over a series or movie, a small popup provided further details on why it was suggested (Figure 10, No. 3). Below the main recommendations, the platform also featured a section titled “series your friends like,” showcasing content based on the viewing preferences of their social connections (Figure 10, No. 4).

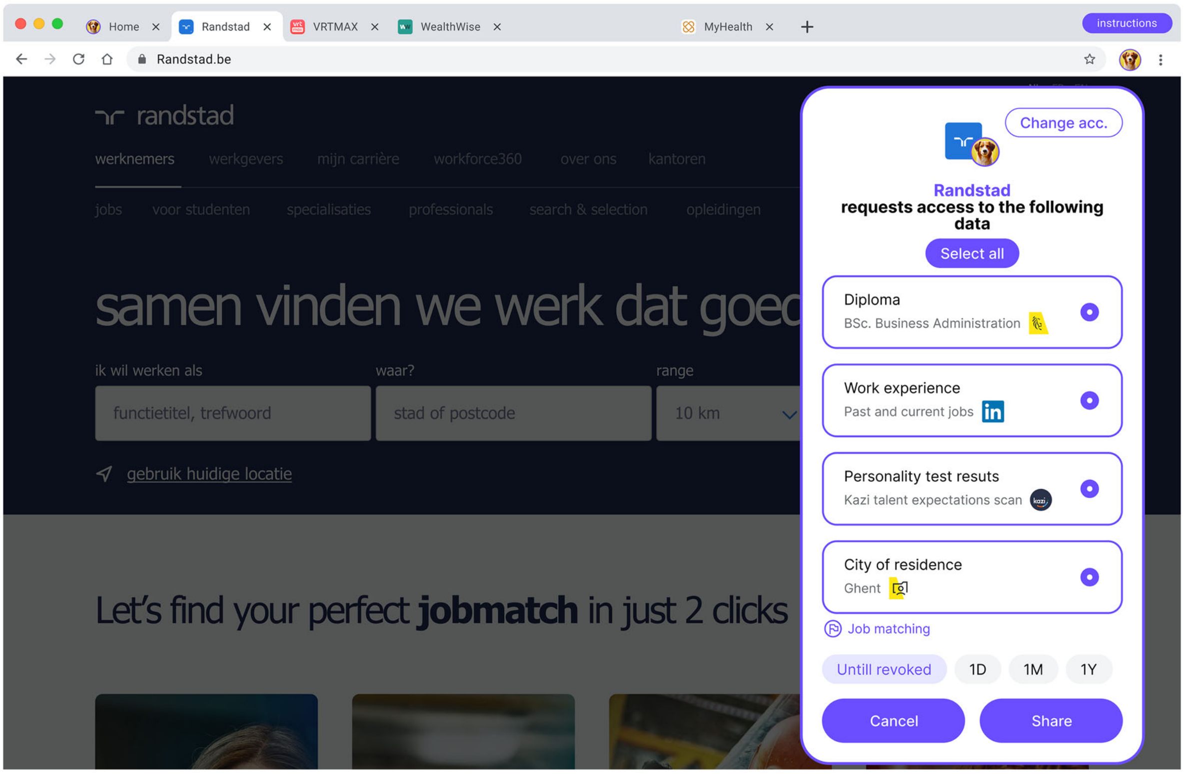


Figure 9. Screenshot HR use case.

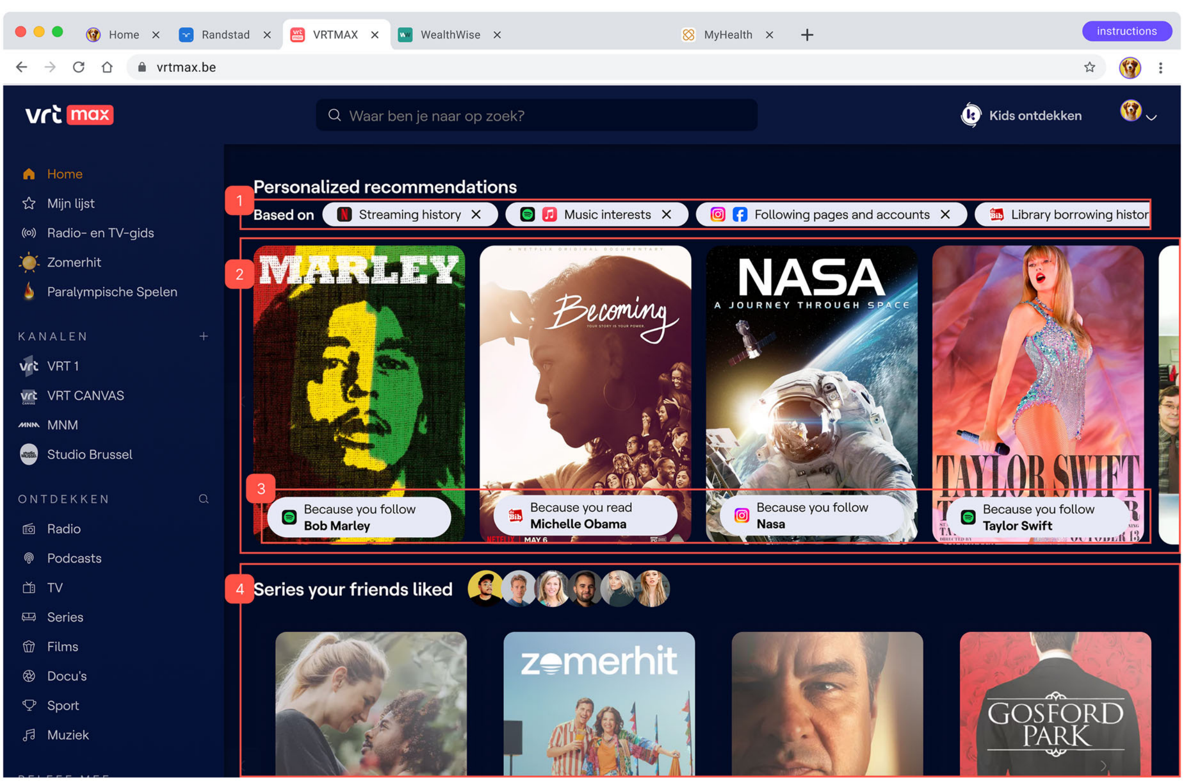


Figure 10. Screenshot media use case.

Next, participants explored the Finance use case. Upon logging in, a popup appeared requesting access to specific financial data, including transaction history, stocks, income, and expenses. Once access was granted, a financial dashboard was generated (Figure 11), displaying an overall balance

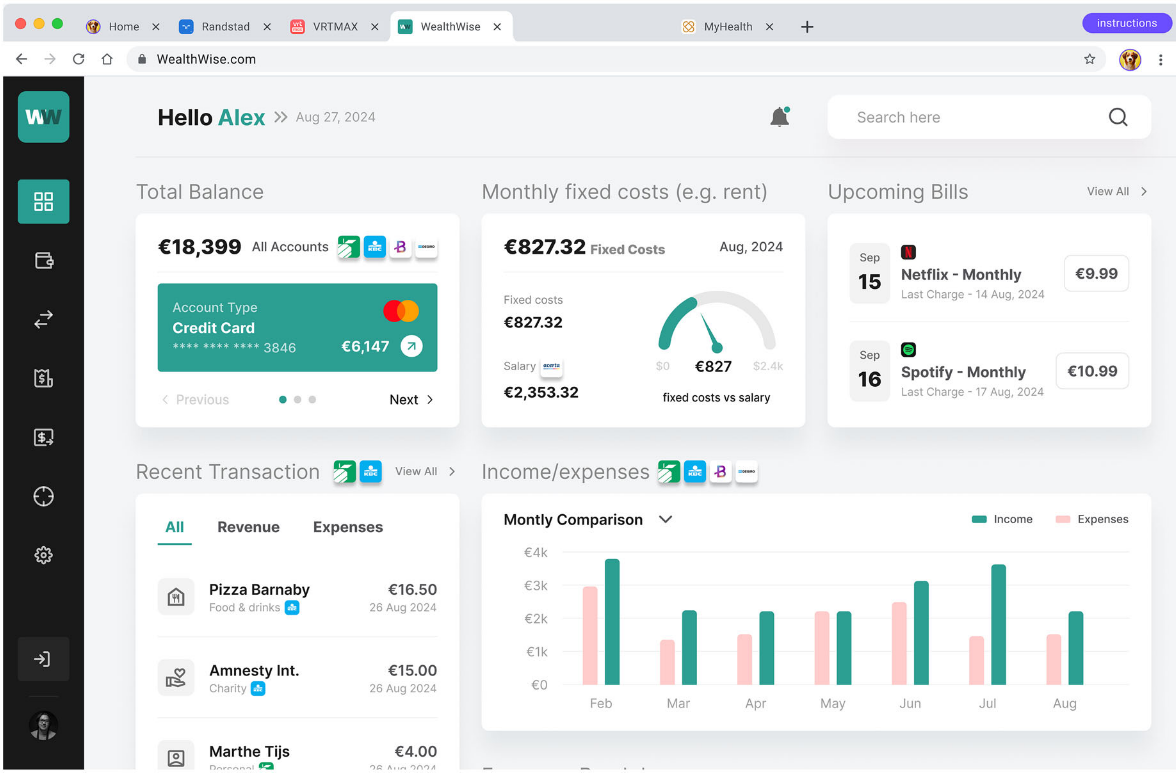


Figure 11. Screenshot finance use case.

across bank accounts, along with a summary of monthly fixed costs and upcoming bills. The dashboard also featured categorized recent transactions and a bar chart comparing monthly income and expenses. Each section was accompanied by logos representing the data sources, ensuring transparency about where the financial information originated.

The final use case participants explored was the Health use case. Upon account creation, they were asked to share their digital ID in order to be matched with their general practitioner. Once participants confirmed the match, a request from their doctor appeared, asking for access to specific health data, including physical activity, vitals, body measurements, medication, and medical history, to monitor their health status. Granting access generated a dashboard displaying these parameters along with interpretations of the data (Figure 12). A first popup illustrated how a doctor could recommend a check-up if certain parameters were outside the normal range (Figure 12, No. 1). A second popup, appearing after a few seconds, notified users of their upcoming tetanus vaccine renewal, demonstrating how the system could proactively support preventive care (Figure 12, No. 2).

Finally, after completing all four use case scenarios, participants were redirected to the PDS platform to reflect on their data-sharing decisions. In the permission manager (Figure 13), they could review which data they had shared, with the option to retract specific data types using a toggle function. Additionally, an advanced permissions tab allowed users to restrict data sharing for broader purposes, such as personalized advertising.

5.2. Incorporation of the HDI principles into the prototype

The Human Data Interaction (HDI) framework (Mortier et al., 2014) guided the prototyping process, ensuring that features were designed, added, and evaluated in alignment with its core principles: legibility, agency, and negotiability.

5.2.1. Legibility

To enhance legibility and improve users' understanding of data processes, we deliberately avoided technical jargon such as personal data store or data vault, as these terms are unfamiliar to most users. Prior

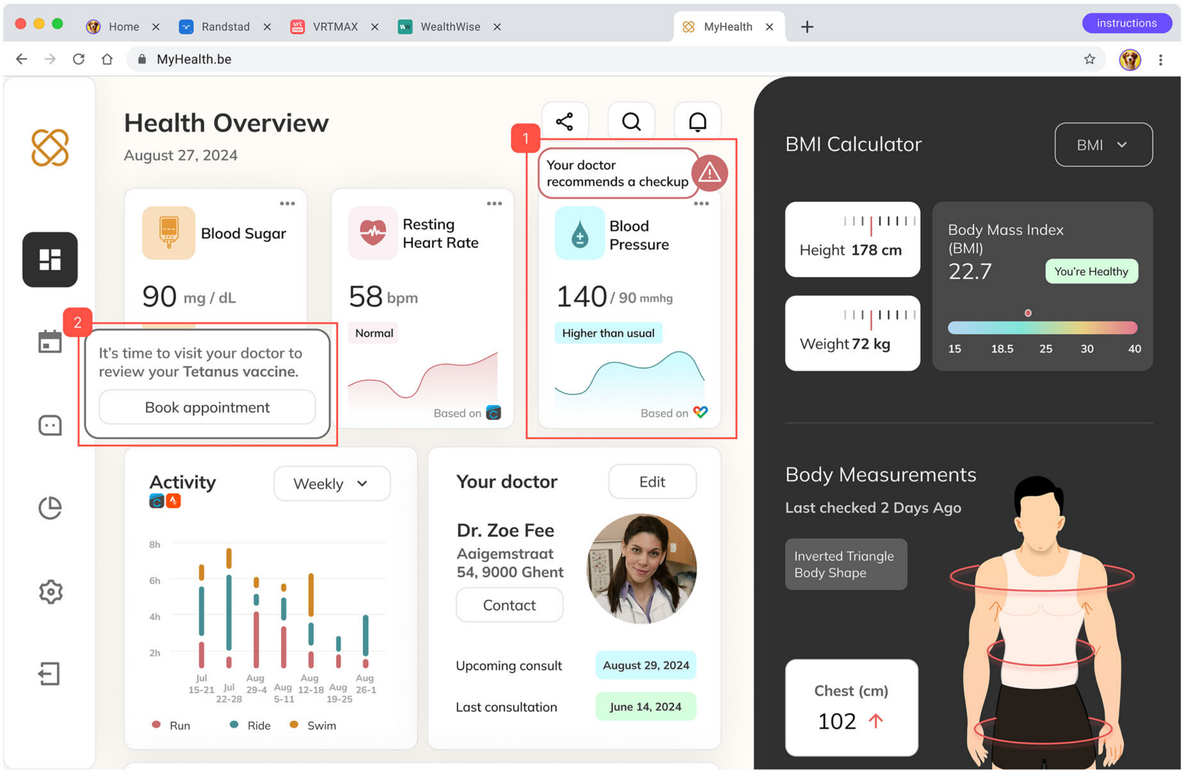


Figure 12. Screenshot health use case.

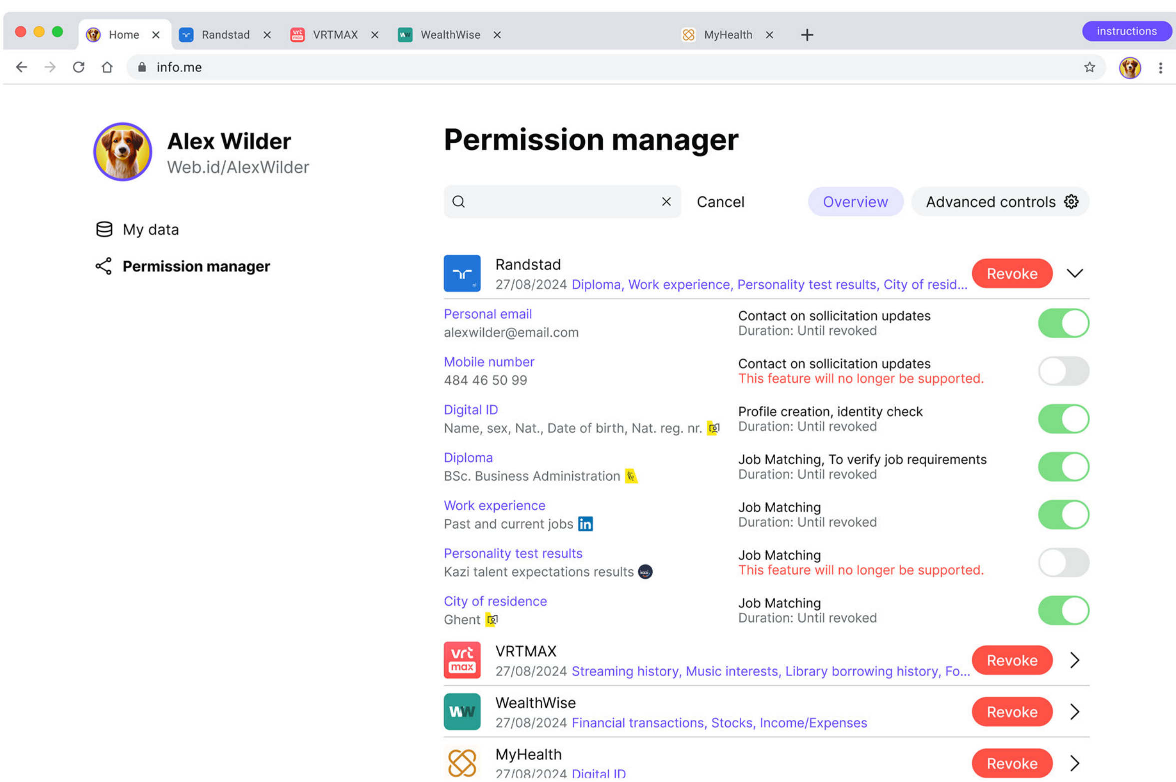


Figure 13. Screenshot permission manager of the personal data store.

research has shown that the concept of a personal data store is difficult for end users to grasp, as they often struggle to form a concrete mental model of a personal data store (Theys et al., 2024a). Rather than emphasizing technical infrastructure, the prototype highlighted what users could *do* with their data. This emphasis on practical affordances over technical explanation aimed to support the development of effective mental models, consistent with the recommendations of Potesnak (1989).

Additionally, given the rich variety of data contained within the personal data store, we sought to prevent cognitive overload by structuring data into hierarchical categories. Data was grouped into main categories and subcategories (Figure 8, No. 1), allowing users to navigate and understand their stored information more easily.

Furthermore we designed the data request popups in alignment with best practices for privacy notices (Feng et al., 2021; Murmann & Karegar, 2021; Schaub et al., 2015). Each popup clearly specified the data requester, the specific data types requested, the purpose of data sharing, and its duration, ensuring optimal transparency. To prevent manipulation, we intentionally avoided dark patterns that might steer users toward specific choices (Habib et al., 2022). A consistent template for the data request popups was used across all use cases, enabling users to quickly recognize and interpret the requests regardless of context.

Similar design interventions were implemented across the use case scenarios to enhance legibility. In the Media use case, users could hover over recommended series to see which data informed the recommendation (Figure 10, No. 3), aligning with best practices in algorithmic transparency (Rader et al., 2018). Similarly, in the HR use case, job recommendations explicitly displayed the matched competences, ensuring that users could easily interpret why certain jobs were suggested (Gutiérrez et al., 2019). To further enhance transparency, the Finance and Health use cases included logos of the original data sources next to each field in the data dashboards, allowing users to clearly identify where their information originated.

5.2.2. Agency

Beyond interventions to enhance legibility, several data control mechanisms were integrated to strengthen user agency. The most notable of these was the permission manager (Figure 13), embedded within the PDS platform, enabling users to manage data access centrally at any time. The permission manager provided granular control over individual data types rather than enforcing an all-or-nothing approach, which is considered a dark pattern (Habib et al., 2022). This principle of granular control was also reflected in the data request popups, ensuring that users could selectively grant access to specific data fields.

Additionally, given that previous research has identified the risk of consent fatigue (Choi et al., 2018), which can lead to click-to-dismiss behaviour (Habib et al., 2022) and uninformed consent (Utz et al., 2019), we introduced (semi-)automated consent features in the advanced controls tab of the permission manager. These features, such as the ability to automatically deny all data requests for personalized advertising purposes, aimed to minimize users' data management burden while maintaining control over their information. This approach aligns with recommendations from previous research on automated consent management (Bourgeus & Vandercruysse, 2024; Colnago et al., 2020; Florea & Esteves, 2023).

Several functionalities were also integrated into the use cases to further enhance users' data control. For example, in the Media use case, a user-controllable recommender system was implemented (Wang et al., 2022), allowing users to easily exclude specific data from the recommendation algorithm (see Figure 10, No. 1).

5.2.3. Negotiability

The PDS model inherently incorporates negotiability by decoupling data storage from platforms. This separation allows users to switch platforms more easily without losing their data, thereby reducing platform dependency and strengthening their negotiation position with respect to the proper use of their data.

From a business perspective, data portability across applications fosters a more competitive and balanced market, as different platforms can request access to the same pool of user data rather than relying on exclusive data collection. This shift refocuses competition from data ownership to service

Table 11. Overview of the experts involved in the expert review of the prototype.

Field of expertise	Professional experience with PDS	Academic/industry	Objective of the review
Product design	No	Academic	Optimization of the UI/UX of the prototype.
UX design	Yes	Industry	
Data visualization	No	Academic	
Privacy	Yes	Academic	Assessment of the implemented privacy controls.
Privacy	Yes	Academic	
Privacy	Yes	Academic	
Technical development	Yes	Industry	Assessment of the technical feasibility of the prototype.
Technical development	Yes	Academic	
Governance	Yes	Academic	Assessment of the prototype's compliance with existing governance frameworks
Legal	Yes	Academic	Assessment of the legal compliance of the data requests with European law.

quality, encouraging businesses to provide the best possible user experience based on available data (Verborgh, 2023), ultimately benefiting end users.

Additionally, users retain continuous control over their data-sharing preferences through the permission manager, which operates independently of the platforms. This feature enables users to revisit and modify their data permissions at any time, further enhancing their ability to negotiate and regulate how their data is used.

6. Expert reviews

6.1. Methodology expert reviews

To refine the prototype before user testing, we conducted a series of expert reviews (Tory & Möller, 2005) involving ten professionals from diverse relevant domains, ranging from product and UX designers to privacy experts. To be considered an expert, individuals were required to have a minimum of three years of professional experience in their respective fields. Table 11 provides an overview of the experts' areas of expertise and the specific objectives of the review process.

The expert evaluation followed a practical walkthrough approach (Light et al., 2018), in which experts navigated the prototype as an end user would. Throughout this process, they provided qualitative feedback on various aspects of the prototype, identifying potential usability issues, suggesting design improvements, and indicating areas requiring further refinement. The evaluation followed a think-aloud protocol with active interventions, as the experts verbally expressed their thoughts while interacting with the prototype (Krug, 2009; Olmsted-Hawala et al., 2010). Feedback was collected through written observations to document their comments and reactions. The collected comments were analyzed using affinity diagramming (Lucero, 2015) to identify recurring themes and insights, which informed the second iteration of the prototype.

6.2. Results expert reviews

The expert review phase resulted in 23 concrete design modifications to the prototype. These changes included adding icons to the Health use case's data dashboard to indicate the original data sources, introducing a time-limiting option for data sharing within the data request popups, and adjusting the colours of certain data visualizations to enhance interpretability.

Beyond specific interface elements, the experts also assessed the overall feasibility of the prototype. Technical development experts agreed that while the prototype surpasses the current capabilities of commercial PDS applications, such as jouw.ID[11]—a Solid-enabled personal data store platform that allows users to store news interests for personalized recommendations—there are no fundamental technical barriers preventing the development of such a solution.

However, significant challenges emerged at the legal and governance levels. Experts emphasized that many businesses remain hesitant to move away from centralized data control, as data portability and interoperability could challenge existing profit models. Additionally, legal uncertainties persist, particularly regarding how current data protection regulations apply to the PDS model, raising concerns about compliance, accountability, and enforcement.

7. User evaluation of the prototype

7.1. Methodology user evaluation of the prototype

7.1.1. Participants

To gain qualitative insights into the willingness to share and the perceptions of potential end-users towards the different use cases, we conducted 17 user tests. These tests specifically targeted participants aged 18–44 years, as this group demonstrated the highest adoption potential in the survey results (Section 0). Participants were recruited through a combination of convenience sampling and a panel management agency. Each participant received €15 compensation for their time, as the study session lasted approximately one hour.

In addition to the qualitative evaluation of the prototype, participants provided quantitative assessments of their behavioural intention and willingness to share data. This enabled an exploratory comparison between participants who interacted with the prototype and those in the survey who assessed the use cases based solely on textual scenario descriptions. To ensure a valid comparison, we excluded survey participants aged 45 and older from the survey results, as only individuals aged 18–44 participated in the prototype evaluation. The average age of both samples was 33 years ($SD = 8.76$ for the prototype group, $SD = 7.77$ for the scenario-based survey group). The proportions for gender, age, and professional situation were closely aligned between the two samples. Regarding education, the prototype group had a slightly higher average level of education.

Although efforts were made to align demographic characteristics, the significantly smaller sample size in the prototype group means these results should be interpreted with caution. Rather than being confirmatory, these comparisons are exploratory and aim to highlight the potential effects of interacting with a tangible prototype on user perceptions.

A detailed comparison of the two samples is presented in Table 12.

7.1.2. Procedure

The prototype evaluation followed a structured three-part procedure to assess participants' perceptions towards the personal data store use case scenarios.

Table 12. Demographics of the participants that interacted with the prototype compared to the demographics of the participants under 45 who evaluated the prototypes based on a description in the survey.

	Prototype		Scenario (age < 45)	
	N	Proportion (%)	*N*	Proportion (%)
n (Total)	17		861	
Gender				
Male	7	41.2	322	37.4
Female	10	58.8	538	62.5
Other/Prefer not to say	0	0.0	1	0.1
Age				
18–24	3	17.6	179	20.8
25–34	7	41.2	297	34.5
35–44	7	41.2	385	44.7
Education				
No diploma	0	0.0	6	0.7
Primary education	0	0.0	5	0.6
Lower secondary education (1st and 2nd grade of secondary school)	0	0.0	30	3.5
Upper secondary education (3rd grade of secondary school)	4	23.5	349	40.5
Bachelor's degree	4	23.5	277	32.2
Master's degree/postgraduate	9	52.9	194	22.5
Professional situation				
Student	2	11.8	137	15.9
Working full-time	13	76.5	552	64.1
Working part-time	1	5.9	101	11.7
(Temporarily) not working (e.g., sabbatical, job seeking, …)	1	5.9	32	3.7
(Early) retirement	0	0.0	1	0.1
Not active in the labour market (e.g., homemaker)	0	0.0	38	4.4

Table 13. Shapiro-Wilk Normality check for behavioural intention ratings (*$p < 0.05$; ***$p < 0.001$).

Use case	Group	W	df	p	Normal distribution
HR	Scenario	0.944	314	<0.001***	No
	Application	0.920	16	0.150	Yes
Media	Scenario	0.957	289	<0.001***	No
	Application	0.902	16	0.074	Yes
Finance	Scenario	0.936	311	<0.001***	No
	Application	0.885	16	0.038*	No
Health	Scenario	0.932	326	<0.001***	No
	Application	0.740	16	<0.001***	No

First, participants navigated through the four use cases, following the predefined flow described in Section 0. A think-aloud protocol with active interventions was employed (Krug, 2009; Olmsted-Hawala et al., 2010), allowing the interviewer to ask follow-up questions whenever participants hesitated, encountered difficulties, or expressed confusion or frustration. At the end of each use case, participants were asked to reflect on why they would or would not consider using the application in the future.

After completing the full prototype flow, participants engaged in a semi-structured reflective interview designed to capture their overall feedback. This interview provided deeper insights into their general experience with the prototype beyond the specific usability issues encountered during the navigation phase.

In the final stage, participants completed a short survey assessing their behavioural intention and willingness to share data for each use case. The survey employed the same measures as described in Section 0, ensuring comparability with the survey sample.

All user tests were conducted online using Microsoft Teams.

7.1.3. Data analysis

7.1.3.1. Quantitative data analysis. To compare behavioural intention ratings between participants who interacted with the prototype and those who evaluated the written scenario, we first assessed the normality of the behavioural intention score distributions using the Shapiro-Wilk test for each group (prototype and scenario) within each use case. As shown in Table 13, only two groups met the normality assumption. Since most groups violated the normality assumption, a non-parametric Mann-Whitney U test was selected for all comparisons between the prototype and scenario groups.

To assess whether the proportion of participants willing to share data differed significantly between the prototype and scenario groups, a series of Pearson's chi-squared tests were conducted. To control for multiple comparisons, Bonferroni corrections were applied to the p-values.

All statistical analyses were conducted using R version 4.4.1 (R Core Team, 2021). The Shapiro-Wilk normality tests, Mann–Whitney U tests, and chi-squared tests were performed using the stats package (R Core Team, 2021).

7.1.3.2. Qualitative data analysis. All interview recordings were transcribed using the transcription software Scribewave[12]. The interviews were conducted and transcribed in Dutch. To preserve the original meaning as accurately as possible, only the quotes included in this paper were translated into English. To ensure participant anonymity, participants names were replaced by numerical identifiers.

The transcripts were analyzed using a deductive coding approach, as described by Mezmir (2020). The analysis examined participants' attitudes, subjective norms, and perceived behavioral control across the different use cases, based on the Theory of Planned Behavior (Ajzen, 1991), as well as their willingness to share data. The complete codebook is available in Appendix F.

7.2. Results user evaluation of the prototype

7.2.1. Behavioural intention

The Mann–Whitney U test revealed statistically significant differences in behavioural intention scores between the scenario and prototype groups across all four use cases: HR ($U = 4775.5$, $z = 5.44$, $p < 0.001$), Media ($U = 3573.5$, $z = 3.12$, $p = 0.002$), Finance ($U = 3770$, $z = 2.93$, $p = 0.003$) and Health

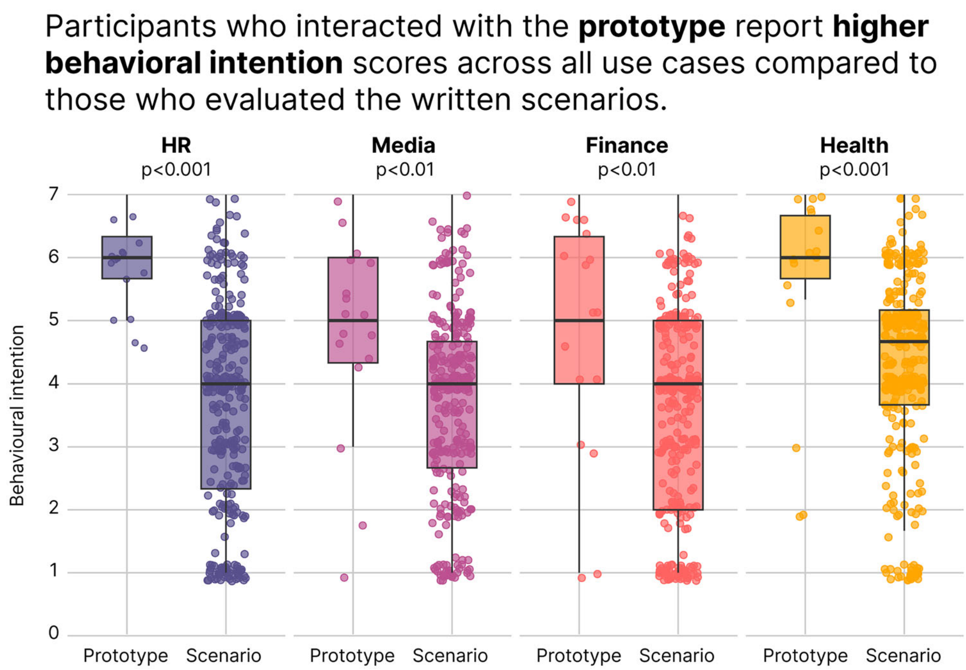


Figure 14. Boxplots presenting the behavioural intention scores compared between participants who interacted with the prototype and those who evaluated written scenarios in the survey. For accessibility, detailed values are provided in Appendix D.

($U = 4315$, $z = 3.84$, $p < 0.001$). As visualized in Figure 14, participants who interacted with the prototype consistently reported higher behavioural intention scores compared to those who only evaluated the textual scenarios in the survey. This effect was most pronounced for the HR and Health use cases.

7.2.2. Willingness to share

As visualized in Figure 15, the willingness to share data was higher in the prototype condition compared to the scenario condition across all data types. The Pearson chi-squared results displayed in Table 14 indicate that these differences are statistically significant for 10 of the 16 data types.

For HR, the willingness to share personality test results nearly doubled, rising from 36 to 77% ($\chi^2(1) = 9.83$, $p = 0.002$), while contact details increased from 55 to 94% ($\chi^2(1) = 8.71$, $p = 0.003$) and work experience from 76 to 100% ($\chi^2(1) = 3.96$, $p = 0.047$). For Media, the largest increases were observed for library history (26–65%, $\chi^2(1) = 10.4$, $p = 0.001$) and viewing history (53–94%, $\chi^2(1) = 9.33$, $p = 0.002$). In Finance, willingness to share salary information more than doubled from 27 to 67% ($\chi^2(1) = 5.52$, $p = 0.019$), transaction history increased from 25 to 66% ($\chi^2(1) = 10.3$, $p = 0.001$), and subscription data from 30 to 68% ($\chi^2(1) = 6.45$, $p = 0.011$). The most substantial increase was observed for financial instruments, which more than tripled from 18 to 64% ($\chi^2(1) = 17.9$, $p < 0.001$). Finally, for Health, willingness to share smart device data rose significantly from 59 to 88% ($\chi^2(1) = 4.71$, $p = 0.030$), while other health-related data types remained consistently high.

Overall, interacting with the prototype substantially increased participants' willingness to share data, particularly for sensitive financial and professional information.

However, due to big difference in sample sizes between the prototype and scenario group, these results should be interpreted with caution. Rather than confirmatory, these comparisons are exploratory in nature that suggest a possible trend, but require replication in larger studies.

7.2.3. Qualitative feedback

7.2.3.1. HR. Participants demonstrated a remarkably high willingness to share data in order to be matched with relevant job opportunities. The only data type that elicited some hesitation was personality test results, as some participants doubted the relevancy for job matching. However, participants acknowledged that if personality assessments were a standard part of the job application process, it

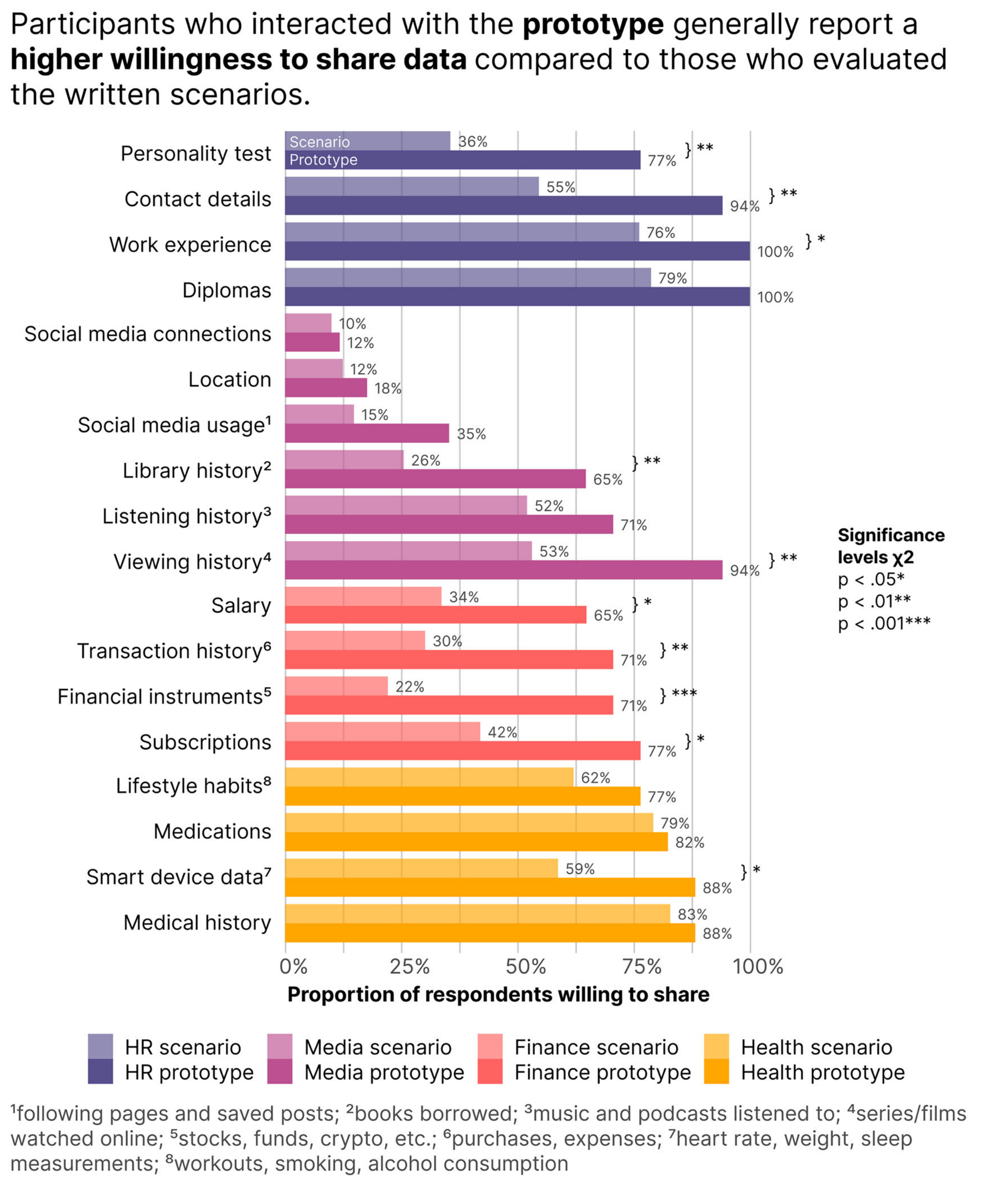


Figure 15. Comparison between the scenario and prototype group of the proportion of participants indicating they are willing to share a certain data type within the given use cases. For accessibility, detailed values are provided in Appendix E.

would be beneficial to reuse previous test results: "It could be useful that you don't have to redo those a hundred times on every application." (Participant 2).

Additionally, participants appreciated data minimization efforts, such as commute time calculations being based on their city of residence rather than requiring them to disclose their exact address. This preference for limiting data disclosure was also evident in their concerns regarding the final step of the process, where detailed identity information—such as a national identification number, sex, and date of birth—was requested. Some participants found this level of detail excessive at an early stage of the application process, as reflected in the following statement: "You have not been accepted yet, and they are actually already asking for very far-reaching data. I don't like that very much" (Participant 8).

Table 14. Results of the chi-squared tests comparing the willingness to share the different data types between the scenario and prototype group (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$).

Use case	Data type	X^2	*df*	*p*-Value
HR	Personality test	9.83	1	0.002**
	Contact details	8.71	1	0.003**
	Work experience	3.96	1	0.047*
	Diploma	3.31	1	0.069
Media	Social media connections	0.00	1	1.000
	location	0.07	1	0.799
	Social media usage	3.61	1	0.058
	Library history	10.40	1	0.001**
	Listening history	1.53	1	0.216
	Viewing history	9.33	1	0.002**
Finance	Salary	5.52	1	0.019*
	Transaction history	10.30	1	0.001**
	Financial instruments	17.90	1	<0.001***
	subscriptions	6.45	1	0.011*
Health	Lifestyle habits	0.88	1	0.347
	Medications	0.00	1	0.996
	Smart device data	4.71	1	0.030*
	Medical history	0.06	1	0.808

Additionally, some participants expressed concerns about potential discrimination resulting from sharing information about sex and age.

Despite some data sharing concerns, overall, participants evaluated the HR use case very positively, which was reflected in the high behavioural intention scores. The main benefits cited were administrative simplification, ease of use, and the efficiency of the process. As one participant summarized: "I just think it goes really fast, and it's just super simple like that. Check, check, check. […] It's really super simple. It doesn't get any simpler than this in my opinion." (Participant 8).

Nevertheless, various participants also expressed a desire for greater control over job recommendations, particularly the ability to refine recommendations based on personal preferences. Some suggested the inclusion of customizable filters or keyword-based preferences to tailor job suggestions more effectively. One participant explained: "I would add some personal preferences, maybe through keywords, like 'I like to work with data'. […] Because I also think that your degree doesn't say everything either." (Participant 3).

7.2.3.2. Media. Initially, the data request on the streaming platform prompted numerous questions regarding the added value of sharing specific data types. While participants generally recognized the potential benefits of sharing data from other streaming platforms to receive personalized recommendations, they were more sceptical about the relevance of other requested data types. Many struggled to envision how these additional data points, such as their listening history or city of residence, would meaningfully enhance their recommendations.

The data type that elicited the greatest reluctance to share was social media connections, as participants were uncomfortable with the idea of others seeing their viewing history. Concerns were expressed about potential judgment based on their viewing choices, as illustrated by one participant: "Sometimes you just want to watch a no-brainer, and then I wouldn't share because I think, wow, people would think I only watch silly series." (Participant 15).

Despite these initial doubts about the relevance of certain data types, participants were often positively surprised by the outcomes of the personalized recommendations. Many indicated that, upon seeing the results, they would be more willing to share data than they had originally anticipated, as the purpose of data sharing became clearer. Another key factor influencing willingness to share was the perceived trustworthiness of the platform, particularly in relation to its affiliation with the government. One participant contrasted their willingness to share data with public versus private entities: "VRT is still governmental, but VTM should not try that with me. They get nothing from me. That's a private company." (Participant 16).

Overall, reactions to the media use case were divided. Some participants appreciated the convenience of personalized recommendations, as reflected in the following statement: "I would use it because it is indeed convenient. Yes. I call that the laziness syndrome. Instead of actively looking." (Participant 8). Others, however, perceived limited added value, expressing concerns that such recommendation systems could contribute to the formation of filter bubbles, potentially restricting content diversity rather than enhancing it.

7.2.3.3. Finance. The financial use case was initially met with considerable caution, as participants were hesitant to share financial information with an unfamiliar platform. This scepticism was encapsulated by one participant: "If I don't know a website, I don't share anything either." (Participant 14). Compared to the HR and Media use cases, participants expressed notably higher concerns about data security. For example, one participant emphasized the importance of security in their decision to adopt the platform: "It must be really secure, because otherwise I would never use it. […] If this doesn't look secure, then I would rather scroll through the various banking apps than open this." (Participant 8). Security and trustworthiness appeared to be crucial factors influencing adoption, with participants expecting features such as a direct connection to a bank and two-factor authentication.

Despite these concerns, most participants acknowledged the potential benefits of having a consolidated financial overview. The dashboard was generally perceived as clear and well-organized, with participants appreciating its ability to present financial information in an accessible manner. One participant remarked: "It's very low-key, and you look at it, you get it immediately, without thinking about it." (Participant 4). This positive evaluation, once participants became familiar with the platform, was also reflected in their willingness to share data. Given the sensitivity of financial information, the reported willingness to share was relatively high, ranging from 65% for salary information to 77% for subscription details. This shift in perception was exemplified by one participant: "In retrospect, I would share everything actually, because it compiles your full financial report." (Participant 3).

Participants who were less enthusiastic generally cited limited personal relevance as their primary reason for disinterest. Those who only used a single bank account or had no investments perceived little added value compared to their existing banking applications. However, some participants anticipated that their perspective might change in the future as their financial situation evolved: "I wouldn't use that too quickly. Maybe also because I'm still a student and don't have a lot of expenses. Maybe that will be different later." (Participant 6).

7.2.3.4. Health. The request to share health information with a doctor elicited mixed reactions among participants. While some believed that full transparency with a healthcare provider was essential—"With your doctor you should be able to share everything." (Participant 12)—others expressed hesitation, particularly regarding the sharing of physical activity data. One participant noted concerns about continuous monitoring: "I don't know if I would also want my doctor to constantly monitor me. […] I might find it a bit too private. If I do not have any issues, then he should not know when I'm exercising." (Participant 10). Despite these reservations, overall willingness to share health-related data remained notably high, ranging from 77% for lifestyle habits to 88% for smart device data and medical history.

Similar to the HR use case, participants valued the potential for administrative simplification, particularly in light of previous inefficiencies in healthcare administration: "I had surgery recently, and I had to do things on a thousand platforms and so on. So if those things were all clustered, that would be super practical." (Participant 6). Additionally, proactive health notifications—such as reminders for vaccination renewals or alerts for abnormal blood pressure—were generally well received. As in the financial use case, participants appreciated the clear and well-organized data visualizations, which facilitated an easily interpretable health overview.

However, not all participants were convinced of the platform's benefits. Some argued that similar alternatives already existed, while others raised concerns that continuous health monitoring might induce stress. A particular point of concern was the potential misuse of aggregated health information by insurance companies. One participant highlighted this issue in the context of family medical history: "My mother had cancer all her life, and so did my father. And yes, all that information gets bundled, and sometimes in insurance files, that gets played against you." (Participant 8).

7.2.3.5. Personal data store. The personal data store, which provided participants with an overview of their stored data and the ability to manage permissions, was generally well received. Participants appreciated its structured design, particularly the use of data categorization and icons, which contributed to "one very clear overview" (Participant 1). Additionally, the idea of having one central platform for all data types was perceived as highly efficient, reducing the need to manage data across multiple applications: “Interesting to have that in one location instead of having 456 different apps for it.” (Participant 17).

However, the comprehensive nature of the data store—particularly its inclusion of sensitive information such as financial and health data—also raised significant security concerns. Many participants expressed fears of hacking: “If they have the one password from that platform, they can get to all your data, but really everything. To your medical records, to your banking details, to your music preferences. Where you live, what newspaper you read. Yeah, you do not want that.” (Participant 6). As a result, participants carefully weighed the practical benefits of having all their data in one location against these potential security risks when considering whether they would use the platform.

A key condition mentioned by the majority of participants was that the platform should be provided by—or at least connected to—a governmental institution to ensure credibility and security. However, one participant questioned the appropriateness of government involvement in less formal data types: “Should the government get involved in your choice of music? I don’t think so.” (Participant 11). The inclusion of leisure-related data, such as media preferences, was frequently debated, with many participants finding more value in bundling official governmental, financial, and health data rather than combining it with entertainment-related information.

Finally, most participants responded positively to the permission manager, appreciating the intuitive controls it provided for managing data access. However, the advanced settings were perceived as overwhelming and overly complex by most users, with only a small subset of participants valuing these additional customization options, suggesting they may cater to a niche user group.

7.2.4. SWOT analysis by domain

To support organizations considering PDS implementation, we translated our findings into a domain specific SWOT analysis (Gürel & Tat, 2017), as displayed in Table 15. This analysis highlights key Strengths, Weaknesses, Opportunities, and Threats observed across the four evaluated domains. It provides a practical synthesis of the user perceptions. By summarizing the empirical insights into actionable points, this overview can help stakeholders identify potential adoption enablers and barriers.

8. Discussion and recommendations for effective PDS implementation

Despite previous research indicating low user interest in PDSs (Fallatah et al., 2023; Van Kleek & O’Hara, 2014), our study found a general enthusiasm for the concept, especially among those who interacted with the PDS prototype. Users appreciated the idea of a well-organized platform that enables cross-platform data portability, addressing common personal data challenges.

However, this enthusiasm was conditional. Consistent with prior findings (Steedman et al., 2020; Theys et al., 2024a), security concerns remained the primary expressed barrier to adoption, with users

Table 15. Domain specific SWOT analysis for PDS implementation.

Domain	Strengths	Weaknesses	Opportunities	Threats
HR	High willingness to share; administrative simplification; ease of use; data minimization appreciated	Concerns over sharing sensitive data early in the process (e.g., ID number, age); desire for more control over job recommendations	Integration with existing job platforms to enhance trust; filtering tools for job preferences	Potential discrimination based on sensitive data (e.g., age or ethnicity)
Media	Enhanced personalization; high trust in public broadcaster; convenience of recommendations	Initial scepticism about data relevance; discomfort with social features; mixed perceptions of added value	User controlled recommendation systems; curated content	Risk of filter bubbles; lower trust in private media providers; low perceived added value for certain data types
Finance	Consolidated financial overview; dashboard clarity; growing willingness to share after hands-on interaction	High initial distrust; strong security concerns; limited perceived value for users with simple financial profiles	Link with banks can increase trust; AI integration to analyze expense patterns	High perceived risk of data breaches; dependency on institutional trust
Health	High willingness to share; administrative efficiency; proactive notifications; clear data visualizations	Privacy concerns about continuous monitoring; stress from over-tracking	Preventive care support; integration with national healthcare systems	Potential insurance misuse of aggregated data; stress from medical awareness

fearing that a single breach could expose all their personal data. Dividing data into thematic PDSs (e.g., health, media, government) could help reduce these risks and lessen the context collapse (Marwick & Boyd, 2011). This is because data that was once scattered across various platforms—without users perceiving any connection between them—is now brought together in a single PDS. However, this strategy may limit the benefits of cross-domain integration.

These reflections highlight a deeper tension: while users desire convenience and unification, they also fear the risks of over-centralization. What is gained in simplicity may be perceived as lost in security. This tension between simplicity and perceived vulnerability complicates PDS design by creating a delicate balance—finding the sweet spot between offering enough data diversity to enhance PDS relevance and avoiding an overabundance that could trigger users' concerns. In these cases, designing for gradual trust-building—by starting with low-stakes data sharing before progressing to sensitive domains—may help overcome initial hesitation. Defaulting to data minimization—by requesting only the information necessary for a specific use case and explicitly justifying and warning users about the impacts of each request (Carpenter et al., 2018)—may further help alleviate these concerns.

Furthermore, trust in the PDS provider emerged as a crucial factor to mitigate security concerns, aligning with previous studies (Fallatah et al., 2023; Chaudhry et al., 2015b; Steedman et al., 2020). In our study, conducted in Flanders, government-backed PDSs enjoyed high levels of trust, though this finding may not be generalizable, as trust in government varies significantly across countries (OECD, 2024). Previous research has also identified trust in government as a major determinant of personal data disclosure (Beldad et al., 2012). Beyond the provider itself, partnerships with established and trusted institutions, such as banks or public broadcasters, could further enhance user confidence in PDS adoption. Yet even here, tensions surfaced. Some participants appreciated government affiliation for its perceived legitimacy, while others questioned whether the government should be involved in more informal data domains like entertainment. This boundary of institutional trust reflects how data sensitivity is tied to users' expectations of institutional roles.

The comparison of different use cases in the initial survey indicated a slightly higher intention to use the Health use case compared to HR, Media, and Finance. The latter three appeared to be more affected by the general decline in PDS adoption intention with increasing age. The qualitative feedback from participants who interacted with the prototype reinforced the broad interest in the Health use case, while also revealing strong enthusiasm for the HR scenario. Reactions to the Media and Finance use cases were more polarized, with some participants finding them highly valuable while others saw little added benefit or personal relevance. Overall, participants expressed a preference for use cases that facilitate administrative simplification and save time over those primarily focused on leisure.

Here too, participants negotiated a trade-off between automation and control. They appreciated features such as auto-filled forms and proactive health reminders, though some voiced concerns about systems that seemed overly anticipatory or relied on opaque recommendations that required a high level of trust. Personalization was appreciated, but only when participants felt they could see and influence the underlying data logic.

This greater interest in use cases that reduce administrative overhead over leisure applications was also reflected in users' willingness to share data across different scenarios. Among participants who interacted with the prototype, willingness to share data in the HR, Financial, and Health use cases remained relatively high, with even the lowest-scoring data point receiving 65% approval. In contrast, the Media use case faced significantly more reluctance, particularly for sharing social media connections (12%), location (18%), and social media usage (35%). Notably, the use case scenario itself appeared to have a far greater influence on users' willingness to share data than the perceived sensitivity of the data. For instance, willingness to share medical history data (88%) within the Health scenario was nearly 2.5 times higher than the willingness to share social media usage data (35%) within the Media use case. Considering the parameters of the contextual integrity theory (Nissenbaum, 2004), which shape contextual information norms, our findings suggest that the recipient of the data (e.g., a doctor versus a streaming platform) has a significantly greater impact on users' willingness to disclose information than the nature of the data itself.

Furthermore, the lack of tangible experiences with PDSs has been identified as a key barrier to user adoption (Fallatah et al., 2023; Van Kleek & O'Hara, 2014). Previous studies have largely examined

PDSs conceptually, often presenting them primarily as privacy-enhancing tools without concrete applications, making it difficult for users to develop a clear mental model of their functionality (Theys et al., 2024a). Our study took a different approach by focusing on four concrete use case scenarios that demonstrated the practical applications of PDSs.

By comparing intention-to-use scores between participants who evaluated the written scenarios in the survey and those who interacted with the prototype, our results suggests that tangible prototype experiences can increase users' intention to adopt PDSs across all use cases. A similar trend was observed in willingness to share data, which was also consistently higher among participants who engaged with the prototype. While these results are exploratory due to the substantial difference in sample sizes between the prototype and survey groups they illustrate the importance of providing users with concrete, hands-on experiences when introducing complex concepts like PDSs, which many find difficult to conceptualize based solely on descriptions. We therefore advocate for a "show, don't tell" approach, emphasizing how PDSs can practically benefit users by enabling cross-application data portability in a privacy-friendly manner, rather than framing them merely as privacy-enhancing tools (Fallatah et al., 2023; Van Kleek & O'Hara, 2014).

Finally, while previous research has identified complex usability, the need for technical skills (Fallatah et al., 2023; Chaudhry et al., 2015b), and the perceived time burden of managing permissions (Steedman et al., 2020) as potential hurdles to PDS adoption, our user testing yielded contrasting results. Participants generally appreciated how the uniformity in data requests across applications, having a central location to manage permissions, and eliminating the need for multiple logins could simplify their online experience. Rather than usability, the greater challenge appeared to lie in explainability (Theys et al., 2024a)—specifically, effectively communicating the various possibilities of a PDS and clarifying how data flows within this ecosystem. This finding underscores the importance of legibility in data processes, as emphasized by the Human-Data Interaction (HDI) framework (Mortier et al., 2014). We anticipate, however, that understanding of the still largely unfamiliar concept of PDSs may take time and will evolve as new use cases emerge and demonstrate their potential.

In conclusion, while previous research identified several societal barriers to PDS adoption, our findings suggest that, despite some persistent concerns and tensions, potential end users generally hold a positive attitude toward PDSs, especially when presented with a tangible experience. The consolidated insights from our different research steps informed the development of a third iteration of the prototype[13], which may serve as a reference for frequently recurring interaction patterns, such as data requests, data dashboarding, and permission management.

9. Limitations and future research

While this study provides valuable insights into user attitudes towards PDSs and their adoption potential across different domains, several limitations should be acknowledged. We reflect on these in relation to established validity threats in empirical HCI research, including internal, external, ecological, construct, and statistical conclusion validity (Kieffer & Université catholique de Louvain, 2017).

First, the study was conducted within a specific geographical and cultural context (Flanders, Belgium), which may limit its external validity and thus generalizability of certain findings, such as the high trust in governmental institutions as PDS providers. Since trust is a key factor influencing adoption, future research should explore which entities foster trust in PDSs across diverse cultural and regulatory environments.

Another key limitation is that the prototype, while visually rich and interactive in structure, was not a fully functional application. Participants navigated a scripted experience rather than exploring freely or interacting with live data. This may have influenced their perceptions, for example by reducing perceived friction, latency, or potential for errors, affecting ecological validity. As such, participants' willingness to share data or perceived usefulness may have been shaped more by expectations than actual use. While these results provide valuable early-stage insights, future research should explore whether the expressed attitudes are also reflected in users' actual behaviour with fully functional systems.

Additionally, due to the large survey sample and the significantly smaller prototype user test group, the results from this comparison are exploratory and should be seen as indicative rather than definitive. This disparity in sample sizes may introduce limitations to statistical conclusion validity, potentially affecting the generalizability of observed differences. While the results provide valuable early insights into how tangibility may influence user perceptions, further studies with larger prototype user groups are needed to validate these findings.

Furthermore, the study focused on four distinct use cases—HR, Media, Finance, and Health—to explore the potential of cross-domain data portability. However, other domains, such as e-commerce or mobility, may present unique challenges and opportunities. Future research could expand the scope of PDS applications to explore their impact in additional contexts.

In addition to these methodological limitations, we acknowledge a theoretical limitation related to the use of the Theory of Planned Behaviour (TPB). While TPB offers a useful framework for capturing internal determinants of adoption intention, it does not explicitly account for external constraints that may hinder actual adoption. In the context of PDSs, such constraints include regulatory uncertainty (Fallatah et al., 2023), a lack of viable business models (D'Hauwers et al., 2025), and technical implementation challenges (Fallatah et al., 2023). Legally, for example, data protection frameworks such as the GDPR define distinct roles for data subjects, controllers, and processors; however, PDS users may simultaneously embody elements of all three—or none clearly—which challenges existing legal models and raises questions about accountability and compliance (Fallatah et al., 2023; Van Kleek & O'Hara, 2014). While this study focuses on user-related factors, future research should further explore these external barriers to complement user-centred insights into PDS adoption.

Rather than presenting a definitive conclusion, this research thus serves as a foundation for future research and development of PDS applications that can leverage cross-platform data portability while ensuring data transparency and control.

10. Conclusion

Building on previously identified challenges to widespread PDS adoption, this study employed an iterative research-through-design approach—including an initial survey, prototype development, expert reviews, and user testing—to gain deeper insights into user attitudes and willingness to share data within PDS-enabled services. By grounding the research in four concrete use case scenarios—HR, Media, Finance, and Health—we provided participants with a tangible experience that elicited a diverse range of responses.

Our findings indicate that while concerns around security and trust, particularly regarding the choice of PDS provider, remain significant, users are generally receptive to the concept of PDSs. This is especially true when PDSs provide clear value, such as simplifying administrative processes. Notably, comparing survey responses based on written scenarios with those from participants who engaged with the PDS prototype revealed that users who explored the prototype reported greater intention to adopt and higher willingness to share data, suggesting that tangible experiences may positively influence perceptions of PDSs.

Beyond these insights, the study also produced an illustrative prototype showcasing common interaction patterns such as data requests, dashboarding, and permission management. This example can serve as a practical reference for the ongoing development of PDS applications that balance cross-platform data portability with transparency and user control.

Notes

1. CitizenMe homepage: https://citizenme.com/.
2. Ockto homepage: https://ockto.eu/.
3. My:D homepage: https://myd.world/.
4. Jouw.id homepage: https://jouw.id/.
5. https://www.linckr.com/.
6. Jouw.id homepage: https://jouw.id/.
7. Easy banking App homepage: https://www.bnpparibasfortis.be/en/public/individuals/daily-banking/payments/easy-banking-app.
8. Nerdwallet homepage: https://www.nerdwallet.com/.
9. Digi.me homepage: https://digi.me/.
10. Link to the prototype version used for the user testing: https://www.figma.com/proto/89Jn0Ucdewxxv47U6PEs7I/The-Solid-application?node-id=18381-29692&t=ShLFQRQPFladZ99E-1&scaling=scale-down&content-scaling=fixed&page-id=17975%3A10116&starting-point-node-id=17907%3A4356.
11. Jouw.id homepage: https://jouw.id/.

12. Scribewave homepage: https://scribewave.com/nl.
13. Third iteration of the prototype: https://www.figma.com/proto/89Jn0Ucdewxxv47U6PEs7I/The-Solid-application?node-id=19560-51596&t=bYOmUfjf4Wr1OdfV-1&scaling=scale-down&content-scaling=fixed&page-id=19560%3A26496&starting-point-node-id=19560%3A26497.

Acknowledgements

Sections of this manuscript were proofread using ChatGPT-4o (OpenAI) to improve the language. Additionally, DeepL was used to assist the author with translating participant quotes from Dutch to English.

Author contributions

CRediT: **Tim Theys**: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing; **Peter Mechant**: Project administration, Supervision; **Lieven De Marez**: Supervision; **Jelle Saldien**: Supervision.

Ethical approval

As the study involved human participants, ethical approval for the user testing was obtained from the Ethics Committee of the Faculty of Engineering and Architecture at Ghent University (File No. 241120001).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was supported by SolidLab Vlaanderen (Flemish Government, EWI and RRF project V023/10).

ORCID

Tim Theys http://orcid.org/0000-0003-1392-773X
Peter Mechant http://orcid.org/0000-0002-5283-5806
Lieven De Marez http://orcid.org/0000-0001-7716-4079
Jelle Saldien http://orcid.org/0000-0003-2557-3764

Data availability statement

The data supporting this study can be accessed through the Open Science Framework at https://osf.io/7pmjd/?view_only=b63eae5b2c074f2ebe3aaca3faceea2d.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Barnes, S. B. (2006). A privacy paradox: Social networking in the United States. *First Monday*, *11*(9). https://doi.org/10.5210/fm.v11i9.1394

BBC. (2023). Together + Data Pod. https://www.bbc.co.uk/taster/pilots/together-pod.

Beldad, A., van der Geest, T., de Jong, M., & Steehouder, M. (2012). Shall i tell you where i live and who i am? Factors influencing the behavioral intention to disclose personal data for online government transactions. *International Journal of Human-Computer Interaction*, *28*(3), 163–177. https://doi.org/10.1080/10447318.2011.572331

Bernes, A. (2022). Enhancing transparency of data processing and data subject's rights through technical tools: The PIMS and PDS solution. In R. Senigaglia Roberto, C. Irti, & A. Bermes (Eds.), *Privacy and data protection in software services* (pp. 197–208). Springer Singapore.

Bourgeus, A., Theys, T., Verhulst, N., Vandercruysse, L., & Mechant, P. (2024). Correction to: Towards an affordance-based typology of personal data stores. In Felix Bieker, S. de Conca, N. Gruschka, M. Jensen, & I. Schiering (Eds.), *Privacy and identity management. Sharing in a digital world* (pp. C1). Springer Nature Switzerland.

Bourgeus, A., & Vandercruysse, L. (2024). Reinvigorating consent: Exploring new paradigms for privacy and data sharing. *Proceedings of 22nd European Conference on Computer-Supported Cooperative Work*. European Society for Socially Embedded Technologies. https://doi.org/10.48340/ecscw2024_ep08

Carmichael, L., Hall, W., & Boniface, M. (2024). Personal data store ecosystems in health and social care. *Frontiers in Public Health*, *12*, 1348044. https://doi.org/10.3389/fpubh.2024.1348044

Carpenter, S., Shreeves, M., Brown, P., Zhu, F., & Zeng, M. (2018). Designing warnings to reduce identity disclosure. *International Journal of Human–Computer Interaction*, *34*(11), 1077–1084. https://doi.org/10.1080/10447318.2017.1413792

Chaudhry, A., Crowcroft, J., Howard, H., Madhavapeddy, A., Mortier, R., Haddadi, H., & McAuley, D. (2015a). Personal data: Thinking inside the box. *Aarhus Series on Human Centered Computing*, *1*(1), 4. https://doi.org/10.7146/aahcc.v1i1.21312

Chaudhry, A., Crowcroft, J., Howard, H., Madhavapeddy, A., Mortier, R., Haddadi, H., & McAuley, D. (2015b). Personal data: Thinking inside the box. *Proceedings of The Fifth Decennial Aarhus Conference on Critical Alternatives* (pp. 29–32). Aarhus University Press.

Choi, H., Park, J., & Jung, Y. (2018). The role of privacy fatigue in online privacy behavior. *Computers in Human Behavior*, *81*, 42–51. https://doi.org/10.1016/j.chb.2017.12.001

Colnago, J., Feng, Y., Palanivel, T., Pearman, S., Ung, M., Acquisti, A., Cranor, L. F., & Sadeh, N. (2020). Informing the design of a personalized privacy assistant for the internet of things. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376389

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319. https://doi.org/10.2307/249008

De Meester, B., de Mildt, M., Henon, P., Lagauw, M., Mechant, P., Raymaekers, T., Ryckaert, K., Schupp, C., Theys, T., & Verbrugge, S. (2025). SHARCS – findings and conclusions. https://solidlab.be/wp-content/uploads/2025/02/End-report.pdf

de Mildt, M. (2024). The added value of using Solid data pods in a lifelong learning context. *Belgian Data Spaces Day 2024, abstracts*. https://solidlab.be/wp-content/uploads/2024/06/The-added-value-of-using-Solid-personal-data-pods-in-a-life-long-learning-context.pdf.

De Vrieze, M. (2023, November 23). Belgische hr-tech op de werkvloer. https://datanews.knack.be/carriere/belgische-hr-tech-op-de-werkvloer/.

Dedecker, R., Slabbinck, W., Wright, J., Hochstenbach, P., Colpaert, P., & Verborgh, R. (2022). What's in a Pod?∼– A knowledge graph interpretation for the Solid ecosystem. In M. Saleem & A.-C. Ngonga Ngomo (Eds.), *Proceedings of the 6th Workshop on Storing, Querying and Benchmarking Knowledge Graphs* (vol. 3279, pp. 81–96).

D'Hauwers, R., & Bourgeus, A. (2024). *Empowering data sovereignty: Strategies of data intermediaries in data ecosystems* [Paper presentation]. 37th Bled EConference (pp. 243–256). University of Maribor Press. https://doi.org/10.18690/um.fov.4.2024.

D'Hauwers, R., Vandercruysse, L., & Ballon, P. (2025). Data sovereignty in personal data mobility ecosystems: A business model perspective. *Technology in Society, 83*. Elsevier. https://doi.org/10.1016/j.techsoc.2025.102953

Drechsler, L. (2018). Practical challenges to the right to data portability in the collaborative economy. *Proceedings of the 14th International Conference on Internet, Law & Politics.* Huygens Editorial.

Dwivedi, Y. K., Rana, N. P., Janssen, M., Lal, B., Williams, M. D., & Clement, M. (2017). An empirical validation of a unified model of electronic government adoption (UMEGA). *Government Information Quarterly*, *34*(2), 211–230. https://doi.org/10.1016/j.giq.2017.03.001

Fallatah, K. U., Barhamgi, M., & Perera, C. (2023). Personal data stores (PDS): A review. *Sensors*, *23*(3), 1477. https://doi.org/10.3390/s23031477

Feng, Y., Yao, Y., & Sadeh, N. (2021). A design space for privacy choices: towards meaningful privacy control in the internet of things. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery. *64*, 1–16. https://doi.org/10.1145/3411764.3445148.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behaviour: An introduction to theory and research* (vol. 27). Addison-Wesley.

Florea, M., & Esteves, B. (2023). Is automated consent in solid GDPR-compliant? An approach for obtaining valid consent with the solid protocol. *Information*, *14*(12), 631. https://doi.org/10.3390/info14120631

Ghayvat, H., Sharma, M., Gope, P., & Sharma, P. K. (2022). SHARIF: Solid pod-based secured healthcare information storage and exchange solution in internet of things. *IEEE Transactions on Industrial Informatics*, *18*(8), 5609–5618. https://doi.org/10.1109/TII.2021.3136884

Godin, G., & Kok, G. (1996). The theory of planned behavior: A review of its applications to health-related behaviors. *American Journal of Health Promotion*, *11*(2), 87–98. https://doi.org/10.4278/0890-1171-11.2.87

Gutiérrez, F., Charleer, S., De Croon, R., Htun, N. N., Goetschalckx, G., & Verbert, K. (2019). Explaining and exploring job recommendations: A user-driven approach for interacting with knowledge-based job recommender systems. *Proceedings of the 13th ACM Conference on Recommender Systems* (pp. 60–68). Association for Computing Machinery. https://doi.org/10.1145/3298689.3347001.

Gürel, E., & Tat, M. (2017). SWOT analysis: A theoretical review. *Journal of International Social Research, 10*(51), 994–1006. https://doi.org/10.17719/jisr.2017.1832

Habib, H., Li, M., Young, E., & Cranor, L. (2022). "Okay, whatever": An evaluation of cookie consent interfaces. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery. *621*, 1–27. https://doi.org/10.1145/3491102.3501985.

Hartman, T., Kennedy, H., Steedman, R., & Jones, R. (2020). Public perceptions of good data management: Findings from a UK-based survey. *Big Data & Society*, *7*(1), 205395172093561. https://doi.org/10.1177/2053951720935616

Janssen, H., Cobbe, J., Norval, C., & Singh, J. (2021). Decentralized data processing: Personal data stores and the GDPR. *International Data Privacy Law*, *10*(4), 356–384. https://doi.org/10.1093/idpl/ipaa016

Kassambara, A. (2016). ggpubr: "ggplot2" based publication ready plots. In *CRAN: Contributed Packages*.

KBC. (2025). Your digital safe. Retrieved March 26, 2025, from https://www.kbc.be/retail/en/products/payments/self-banking/on-your-smartphone/mobile/digital-safe.html.

Kieffer, S, Université catholique de Louvain. (2017). ECOVAL: Ecological validity of cues and representative design in user experience evaluations. *AIS Transactions on Human-Computer Interaction*, *9*(2), 149–172. https://doi.org/10.17705/1thci.00093

Krug, S. (2009). *Rocket surgery made easy: The do-it-yourself guide to finding and fixing usability problems*. New Riders.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Langford, J., Poikola, A., Janssen, W., Lähteenoja, V., Rikken, M. (2022). Understanding MyData operators. https://mydata.org/wp-content/uploads/2022/07/Understanding-MyData-Operators-2022-1.pdf.

Larsen, R., Lewis, D., Liu Eisma, F. (2015). European Commission: Report *on personal data stores*. https://www.academia.edu/20193979/European_Commission_Report_on_Personal_Data_Stores.

Li, I., Froehlich, J., Larsen, J. E., Grevet, C., & Ramirez, E. (2013). *Personal informatics in the wild: Hacking habits for health & happiness* [Paper presentation]. CHI '13 Extended Abstracts on Human Factors in Computing Systems (pp. 3179–3182), Paris, France. https://doi.org/10.1145/2468356.2479641.

Light, B., Burgess, J., & Duguay, S. (2018). The walkthrough method: An approach to the study of apps. *New Media & Society*, *20*(3), 881–900. https://doi.org/10.1177/1461444816675438

Lucero, A. (2015). Using affinity diagrams to evaluate interactive prototypes. In J. Abascal, S. Barbosa, M. Fetter, P. Palanque, & M. Winckler (Eds.), *Human-computer interaction – INTERACT 2015* (pp. 231–248). Springer International Publishing.

Mansour, E., Sambra, A. V., Hawke, S., Zereba, M., Capadisli, S., Ghanem, A., Aboulnaga, A., & Berners-Lee, T. (2016). A demonstration of the solid platform for social web applications. *Proceedings of the 25th International Conference Companion on World Wide Web* (pp. 223–226). International World Wide Web Conferences Steering Committee. https://doi.org/10.1145/2872518.2890529.

Marwick, A. E., & Boyd, D. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, *13*(1), 114–133. https://doi.org/10.1177/1461444810365313

Mathieson, K. (1991). Predicting user intentions: Comparing the technology acceptance model with the theory of planned behavior. *Information Systems Research*, *2*(3), 173–191. https://doi.org/10.1287/isre.2.3.173

McAuley, D., Mortier, R., & Goulding, J. (2011). The dataware manifesto. *2011 Third International Conference on Communication Systems and Networks (COMSNETS 2011)* (pp. 1–6). IEEE. https://doi.org/10.1109/COMSNETS.2011.5716491.

Meteyard, L., & Davies, R. A. I. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, *112*, 104092. https://doi.org/10.1016/j.jml.2020.104092

Mezmir, E. A. (2020). Qualitative data analysis: An overview of data reduction, data display and interpretation. *Research on Humanities and Social Sciences*, *10*(21), 15–27. https://doi.org/10.7176/RHSS/10-21-02

Mortier, R., Haddadi, H., Henderson, T., McAuley, D., & Crowcroft, J. (2014). Human-data interaction: The human face of the data-driven society. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2508051

Mortier, R., Haddadi, H., Henderson, T., McAuley, D., Crowcroft, J., & Crabtree, A. (2014). Human-*data interaction*. https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/human-data-interaction.

Murmann, P., & Karegar, F. (2021). From design requirements to effective privacy notifications: Empowering users of online services to make informed decisions. *International Journal of Human–Computer Interaction*, *37*(19), 1823–1848. https://doi.org/10.1080/10447318.2021.1913859

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review*, *79*(1), 119–158. https://digitalcommons.law.uw.edu/wlr/vol79/iss1/10.

OECD. (2024). *OECD survey on drivers of trust in public institutions – 2024 results*. OECD.

Olmsted-Hawala, E. L., Murphy, E. D., Hawala, S., & Ashenfelter, K. T. (2010). Think-aloud protocols: A comparison of three think-aloud protocols for use in testing data-dissemination web sites for usability. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2381–2390). Association for Computing Machinery. https://doi.org/10.1145/1753326.1753685.

Patil, S., Schlegel, R., Kapadia, A., & Lee, A. J. (2014). Reflection or action? How feedback and control affect location sharing decisions. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 101–110). Association for Computing Machinery. https://doi.org/10.1145/2556288.2557121.

Pinto, G. P., Donta, P. K., Dustdar, S., & Prazeres, C. (2024). A systematic review on privacy-aware IoT personal data stores. *Sensors*, *24*(7), 2197. https://doi.org/10.3390/s24072197

Potesnak, K. (1989). Mental models: Helping users understand software. *IEEE Software*, *6*(5), 85–86. https://doi.org/10.1109/52.35592

R Core Team. (2021). R: A *language and environment for statistical computing*. https://www.R-project.org/.

Rader, E., Cotter, K., & Cho, J. (2018). Explanations as Mechanisms for Supporting Algorithmic Transparency. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3173574.3173677

Sailaja, N., Jones, R., & McAuley, D. (2021a). *Designing for human data interaction in data-driven media experiences* [Paper presentation]. Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan. *271*, 1–7. https://doi.org/10.1145/3411763.3451808.

Sailaja, N., Jones, R., & McAuley, D. (2021b). Human data interaction in data-driven media experiences: An exploration of data sensitive responses to the socio-technical challenges of personal data leverage. *Proceedings of the 2021 ACM International Conference on Interactive Media Experiences* (pp. 108–119), Association for Computing Machinery. https://doi.org/10.1145/3452918.3458797.

Sambra, A. V., Mansour, E., Hawke, S., Zereba, M., Greco, N., Ghanem, A., Zagidulin, D., Aboulnaga, A., & Berners-Lee, T. (2016). Solid: A *platform for decentralized social applications based on linked data*. https://api.semanticscholar.org/CorpusID:49564404

Santos, P., Salgado, L., & Viterbo, J. (2018). Assessing the communicability of human-data interaction mechanisms in transparency enhancing tools. *Proceedings of the 2018 Federated Conference on Computer Science and Information Systems* (pp. 897–906). IEEE.

Schaub, F., Balebako, R., Durity, A. L., & Cranor, L. F. (2015). A design space for effective privacy notices. *Proceedings of the Eleventh USENIX Conference on Usable Privacy and Security* (pp. 1–17). USENIX Association.

Searle, S. R., Speed, F. M., & Milliken, G. A. (1980). Population marginal means in the linear model: An alternative to least squares means. *The American Statistician*, *34*(4), 216–221. https://doi.org/10.1080/00031305.1980.10483031

Segijn, C., Strycharz, J., Riegelman, A., & Hennesy, C. (2021). A literature review of personalization transparency and control: Introducing the transparency-awareness-control framework. *Media and Communication*, *9*(4), 120–133. https://doi.org/10.17645/mac.v9i4.4054

Solove, D. J. (2013). Introduction: Privacy self-management and the consent dilemma. *Harvard Law Review*, *126*(7), 1880–1903. https://harvardlawreview.org/print/vol-126/introduction-privacy-self-management-and-the-consent-dilemma/

Solove, D. J. (2020). The myth of the privacy paradox. *SSRN Electronic Journal*, *98*(1), 1–51. https://doi.org/10.2139/ssrn.3536265

Steedman, R., Kennedy, H., & Jones, R. (2020). Complex ecologies of trust in data practices and data-driven systems. *Information, Communication & Society*, *23*(6), 817–832. https://doi.org/10.1080/1369118X.2020.1748090

Stichting Nederlandse Datakluis. (2021). Data *Weer* in *Eigen* hand. Retrieved March 26, 2025, from https://www.datakluis.nl/.

Sun, C., van Soest, J., & Dumontier, M. (2023). Analyze decentralized personal health data using solid, digital consent, and federated learning. *Semantic Web Applications and Tools for Health Care and Life Sciences*, *3415*, 169–170. https://bbuelens.solidcommunity.net/public/html/SolidSymposium2023_ChangSun.pdf

Theys, T., Bourgeus, A., Mechant, P., De Marez, L., & Saldien, J. (2024a). Expert perspectives on UX design challenges for solid-enabled personal data store applications. In Stephanidis, M. Antona, S. Ntoa, & G. Salvendy (Ed.), *HCI International 2024 posters* (pp. 136–145). Springer Nature Switzerland.

Theys, T., Haegemans, T., Mechant, P., Bourgeus, A., Saldien, J. (2024b). Solid-enabled personal online data stores: Uncovering UX implications for online data management. In T. Ahram, W. Karwowski, D. Etinger, & T. Mijač (Eds.), *Human Systems Engineering and Design (IHSED2024): Future Trends and Applications. AHFE (2024) International Conference. AHFE Open Access* (Vol. 2). AHFE International. http://doi.org/10.54941/ahfe1005533

Tory, M., & Möller, T. (2005). Evaluating visualizations: Do expert reviews work? *IEEE Computer Graphics and Applications*, *25*(5), 8–11. https://doi.org/10.1109/MCG.2005.102

Utz, C., Degeling, M., Fahl, S., Schaub, F., & Holz, T. (2019). (Un)informed consent: Studying GDPR consent notices in the field. *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security* (pp. 973–990). Association for Computing Machinery. https://doi.org/10.1145/3319535.3354212.

Van Kleek, M., & O'Hara, K. (2014). The future of social is personal: The potential of the personal data store. In D. Miorandi, V. Maltese, M. Rovatsos, A. Nijholt, & J. Stewart (Ed.), *Social collective intelligence: Combining the powers of humans and machines to build a smarter society.* (pp. 125–158). Springer International Publishing.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425. https://doi.org/10.2307/30036540

Verborgh, R. (2023). Re-decentralizing the web, for good this time. In *Linking the world's information: Essays on Tim Berners-Lee's invention of the world wide web* (1st ed., pp. 215–230). Association for Computing Machinery. https://doi.org/10.1145/3591366.3591385.

Verbrugge, S., Vannieuwenborg, F., der Wee, M., Colle, D., Taelman, R., & Verborgh, R. (2021). Towards a personal data vault society: An interplay between technological and business perspectives. *2021 60th FITCE Communication Days Congress for ICT Professionals: Industrial Data – Cloud, Low Latency and Privacy (FITCE)* (pp. 1–6). IEEE. https://doi.org/10.1109/FITCE53297.2021.9588540.

Victorelli, E. Z., Dos Reis, J. C., Hornung, H., & Prado, A. B. (2020). Understanding human-data interaction: Literature review and recommendations for design. *International Journal of Human-Computer Studies*, *134*, 13–32. https://doi.org/10.1016/j.ijhcs.2019.09.004

Vlaanderen. (2024). Persoonlijke datakluis in Mijn Burgerprofiel. Retrieved March 26, 2025, from https://www.vlaanderen.be/uw-overheid/mijn-burgerprofiel/privacyverklaring-mijn-burgerprofiel/persoonlijke-datakluis-in-mijn-burgerprofiel.

VRT. (2023). Solid4Media. https://www.vrtinternational.com/innovation/projects/solid4media.

Wang, W., Feng, F., Nie, L., & Chua, T.-S. (2022). User-controllable recommendation against filter bubbles. *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 1251–1261). Association for Computing Machinery. https://doi.org/10.1145/3477495.3532075.

Westby, J. R. (2011). Legal issues associated with data collection & sharing. *Proceedings of the First Workshop on Building Analysis Datasets and Gathering Experience Returns for Security* (pp. 97–102). Association for Computing Machinery. https://doi.org/10.1145/1978672.1978684.

Whitley, E. A. (2009). Informational privacy, consent and the "control" of personal data. *Information Security Technical Report*, *14*(3), 154–159. https://doi.org/10.1016/j.istr.2009.10.001

Wickham, H. (2016). *ggplot2*. Springer International Publishing.

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2014). dplyr: A grammar of data manipulation. In *CRAN: Contributed packages*.

Zimmerman, J., Forlizzi, J., & Evenson, S. (2007). Research through design as a method for interaction design research in HCI. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 493–502). Association for Computing Machinery. https://doi.org/10.1145/1240624.1240704.

About the authors

Tim Theys is a PhD candidate at Ghent University, affiliated with the research group for Media, Innovation and Contemporary Technologies (MICT). His interdisciplinary research focuses on the user experience of Personal Data Stores, aiming to obtain a joint doctoral degree in Industrial Design Engineering Technology and Communication Sciences.

Peter Mechant is a professor and senior researcher at the research group for Media, Innovation and Contemporary Technologies (MICT), Ghent University. His research areas include social media, human-data interaction, human-computer interaction, and digital technologies.

Lieven De Marez is Research Director at the research group for Media, Innovation and Contemporary Technologies (MICT), Ghent University, and Scientific Director Societal Innovation AI & Data at Imec, a leading research centre for digital and nanotechnologies.

Jelle Saldien is a professor at the research group of Product Development, Faculty of Design Sciences, University of Antwerp and at the research group for Media, Innovation and Contemporary Technologies (MICT), Department of Industrial Systems Engineering and Product Design, Ghent University.

Appendix A

Table A1. Estimated means behavioural intention for the four use cases including 95% confidence intervals as displayed in Figure 5.

Use case	Estimated mean	Standard error	95% CI lower	95% CI upper
HR	3.65	0.13	3.39	3.91
Media	3.92	0.13	3.66	4.18
Finance	3.68	0.13	3.42	3.94
Health	4.72	0.13	4.46	4.98

Appendix B

Table B1. Estimated means behavioural intention for the four use cases by age category including 95% confidence intervals as displayed in Figure 6.

Use case	Age category	Mean	SE	95% CI lower	95% CI upper
HR	18–24	4.05	0.17	3.72	4.38
	25–34	3.75	0.16	3.43	4.06
	35–44	3.47	0.14	3.19	3.75
	45–54	3.31	0.13	3.05	3.57
	55–64	2.82	0.14	2.56	3.09
	64+	2.35	0.10	2.16	2.54
Media	18–24	4.09	0.16	3.77	4.40
	25–34	3.77	0.15	3.48	4.06
	35–44	3.36	0.14	3.09	3.63
	45–54	3.53	0.13	3.28	3.79
	55–64	2.99	0.12	2.74	3.23
	64+	3.07	0.09	2.89	3.25
Finance	18–24	3.73	0.19	3.36	4.11
	25–34	3.57	0.16	3.25	3.89
	35–44	3.41	0.14	3.13	3.68
	45–54	2.82	0.13	2.56	3.09
	55–64	3.05	0.13	2.79	3.31
	64+	2.79	0.11	2.58	3.00
Health	18–24	4.39	0.19	4.02	4.76
	25–34	4.27	0.14	4.00	4.54
	35–44	4.30	0.13	4.05	4.55
	45–54	3.90	0.14	3.62	4.19
	55–64	4.17	0.12	3.92	4.41
	64+	4.11	0.11	3.89	4.33

Appendix C

Table C1. Willingness to share various data types including 95% confidence intervals as displayed in Figure 7.

Use case	Data type	Percentage (%)	95% CI lower (%)	95% CI upper (%)
HR	Personality test	26.0	23.0	29.0
	Contact details	52.3	48.9	55.7
	Diploma	70.8	67.7	73.9
	Work experience	72.2	69.2	75.3
Media	Social media connections	10.0	8.0	12.1
	Social media usage	11.1	9.0	13.3
	Location	14.4	12.0	16.8
	Library history	21.7	18.9	24.5
	Listening history	37.4	34.1	40.7
	Viewing history	38.1	34.8	41.4
Finance	Financial instruments	18.4	15.8	21.1
	Transaction history	25.3	22.3	28.3
	Salary	26.6	23.6	29.6
	Subscriptions	30.2	27.1	33.4
Health	Smart device data	47.7	44.3	51.1
	Lifestyle habits	51.9	48.5	55.3
	Medications	82.3	79.7	84.9
	Medical history	85.5	83.1	87.9

Appendix D

Table D1. Behavioural intention scores compared between participants (aged 18–44) who interacted with the prototype and those who evaluated written scenarios in the survey.

Use case	Group	BI median	Q1	Q3	Min	Max	Significance level
HR	Prototype	6.00	5.67	6.33	4.67	7	$p < .001^{***}$
	Scenario	4.00	2.33	5.00	1.00	7	
Media	Prototype	5.00	4.33	6.00	1.00	7	$p < .01^{**}$
	Scenario	4.00	2.67	4.67	1.00	7	
Finance	Prototype	5.00	4.00	6.33	1.00	7	$p < .01^{**}$
	Scenario	4.00	2.00	5.00	1.00	7	
Health	Prototype	6.00	5.67	6.67	2.00	7	$p < .001^{***}$
	Scenario	4.67	3.67	5.17	1.00	7	

The values correspond with the boxplots displayed in Figure 14 ($^{*}p<.05$; $^{**}p<.01$; $^{***}p<.001$)

Appendix E

Table E1. Proportion of respondents willing to share a certain data type a displayed in Figure 15.

Use case	Data type	Group	Proportion willing to share (%)
HR	Personality test	Scenario	35.6
		Prototype	76.5
	Contact details	Scenario	54.6
		Prototype	94.1
	Work experience	Scenario	76.2
		Prototype	100
	Diploma	Scenario	78.7
		Prototype	100
Media	Social media connections	Scenario	10
		Prototype	11.8
	Location	Scenario	12.4
		Prototype	17.6
	Social media usage	Scenario	14.8
		Prototype	35.3
	Library history	Scenario	25.5
		Prototype	64.7
	Listening history	Scenario	52.1
		Prototype	70.6
	Viewing history	Scenario	53.1
		Prototype	94.1
Finance	Salary	Scenario	33.7
		Prototype	64.7
	Transaction history	Scenario	30.1
		Prototype	70.6
	Financial instruments	Scenario	22.1
		Prototype	70.6
	Subscriptions	Scenario	42
		Prototype	76.5
Health	Lifestyle habits	Scenario	62.1
		Prototype	76.5
	Medications	Scenario	79.2
		Prototype	82.4
	Smart device data	Scenario	58.7
		Prototype	88.2
	Medical history	Scenario	82.9
		Prototype	88.2

Appendix F

Table F1. Codebook used for the data analysis of the user evaluation of the prototype.

Theme	Codes	Example
HR		
Willingness to share	Personality test	"Maybe not expected, but could be useful that you don't have to redo those a hundred times on every application." (Participant 2)
	Contact details	/
	Work experience	"Did I do holiday work at Colruyt? Then I would find less relevant to communicate it to Randstad." (Participant 3)
	Diplomas	"Your degree. Yes, that probably just revolves around what kind of job is looking for." (Participant 6)
	General	"When you used to go to an interim agency, just had to write on paper." (Participant 14)
Attitudes	Perceived benefits	"The biggest advantage? That you don't have to enter everything yourself." (Participant 5)
	Perceived drawbacks	"Somewhere I would add some personal preference or something like keywords." (Participant 3)
Subjective norm		"Yes, I'm also fine with sharing everything by the way. Yes, and that's also mainly because Randstad and VRT are also sites I use often and do trust to some extent." (Participant 1)
Perceived behavioural control		"Yes, I think it is just a lot simpler. It is a lot quicker." (Participant 3)
Media		
Willingness to share	Social media connections	"I wouldn't necessarily share that social media either, if that's an option, okay." (Participant 2)
	Location	"A hometown also seems less relevant to me." (Participant 5)
	Social media usage	"Yes, Facebook pages and Instagram pages, yes, I see little added value in that." (Participant 5)

(*continued*)

Table F1. Continued.

Theme	Codes	Example
	Library history	"The books, yes, that's also a little less relevant I think. Unless that's then linked to film adaptations of well-known books or something." (Participant 5)
	Listening history	"I think my music interest, for example, is not always linked to what I would like to see on television." (Participant 2)
	Viewing history	"Streaming history is to so link some programmes to what you've seen before. So that seems logical to me." (Participant 7)
	General	"But I wouldn't share any of that. Because. I'd rather decide what to watch myself." (Participant 6)
Attitudes	Perceived benefits	"That way, you can search more specifically for a programme that suits you. So I think that's something I would use." (Participant 7)
	Perceived drawbacks	"I also think it should still have a bit of free choice and that is restricted a little bit here I think." (Participant 14)
Subjective norm		"Maybe some people are like 'I don't want my friends to know that what I watch'." (Participant 13)
Perceived behavioural control		"I also think it should still have some of the free choice and that is restricted a little bit here I think." (Participant 14)
Finance		
Willingness to share	Salary	"Putting your income somewhere is always a doubt anyway." (Participant 4)
	Transaction history	"The cost? […] Rather not. I find it rather personal." (Participant 9)
	Financial instruments	/
	Subscriptions	/
	General	"In hindsight, I would share everything actually, because that compiles your full financial report anyway." (Participant 3)
Attitudes	Perceived benefits	"Again, I find it all very clear and easy to use, so mostly positives I think about this platform." (Participant 2)
	Perceived drawbacks	"I wouldn't use that super-quickly. Also. Maybe also because I'm still a student and don't have that many expenses." (Participant 6)
Subjective norm		"That actually depends on who this is actually from." (Participant 14)
Perceived behavioural control		"Whereas here, that is very accessible and you look at it, you get it immediately, without thinking." (Participant 4)
Health		
Willingness to share	Lifestyle habits	"My GP does not necessarily need to know my activity." (Participant 4)
	Medications	/
	Smart device data	"Or via Garmin your heart rate. […] I actually think it all makes sense for you to share that." (Participant 15)
	Medical history	"Medication and medical history I would share." (Participant 5)
	General	"With your doctor, you should be able to share everything." (Participant 12)
Attitudes	Perceived benefits	"I find it convenient that it includes your appointments." (Participant 2)
	Perceived drawbacks	"I don't know if I would also want my doctor to follow up with me constantly." (Participant 10)
Subjective norm		"So if my doctor recommended this, without a doubt, I would already be using that and I would have no problem sharing all this then." (Participant 1)
Perceived behavioural control		/
PDS		
Attitudes	Perceived benefits	"Yes, maybe interesting to have that in one location then instead of having 456 different apps for it." (Participant 11)
	Perceived drawbacks	"Plus if they have one password from that platform, they can get to all your data, but really everything." (Participant 6)
Subjective norm		"If other friends or colleagues or family would use it and recommend it to me, I am definitely open to using it." (Participant 11)
Perceived behavioural control		"I would find it easy indeed that you have some control over what is on and what is off." (Participant 2)