Human Systems Engineering and Design (IHSED 2024), Vol. 158, 2024, 100–108
https://doi.org/10.54941/ahfe1005533

Solid-Enabled Personal Online Data Stores: Uncovering UX Implications for Online Data Management

Tim Theys[1], Tom Haegemans[2], Peter Mechant[3], August Bourgeus[4], and Jelle Saldien[1]

[1]imec-mict-UGent, Department of Industrial Systems Engineering and Product Design, Ghent University, Ghent, Belgium
[2]Research Center for Management Informatics, Catholic University of Leuven, Leuven, Belgium
[3]imec-mict-UGent, Department of Communication Sciences, Ghent University, Ghent, Belgium
[4]imec-SMIT-VUB, Vrije Universiteit Brussel, Brussels, Belgium

ABSTRACT

Enhanced transparency, control, and usability in online data management systems are in greater demand due to the challenges users face in tracking who accesses and stores their online personal data. In response, the W3C web standard Solid (Social Linked Data) provides users with data vaults to address these issues. This transformation in how individuals interact with their personal data online through data vaults also presents new challenges in UX design. This paper presents an exploratory usability testing study (n = 10), during which participants created and interacted with a data vault within a job application context, aimed at identifying how data vaults impact the user experience. Our study uncovered three key factors introducing UX challenges associated with Solid-enabled data vaults (1) users' mental models shaping expectations, (2) fears, doubts and the mitigating role of brand trust, and (3) a challenge of explainability rather than usability. These empirically derived factors were analysed through established theoretical frameworks on mental models, technology acceptance, and human-data interaction, enabling the formulation of practical implications and future research topics to address these challenges.

Keywords: Solid (social linked data), Data vaults, Personal online data stores, Human-data interaction, UX

INTRODUCTION

Due to the multitude of digital platforms that collect and utilize users' personal data, it has become nearly impossible for individuals to manage their data separately on each platform (Solove, 2013). Data thus ends up siloed in distinct platforms, leading to inefficiencies and potential privacy concerns as users lose transparency and control over how their information is stored and accessed (Mansour *et al.*, 2016).

As a pioneering web standard, Solid (Social Linked Data) aims to address these concerns by providing users with personal online data stores, known as

pods or data vaults (Berners-Lee, Hendler and Lassila, 2001). In their Solid-enabled data vault, users can store any type of data, ranging from official records such as diplomas to their digital music library. The data vault is linked to a WebID, a unique identifier that enables authentication. By logging in with their WebID, users can access and reuse the data stored in their data vault across various platforms (Mechant *et al.*, 2021). In this way, Solid is positioned as a means to enhance data portability, transparency, and user control (Dedecker *et al.*, 2022). However, despite its potential and interest from institutions such as the BBC (BBC, 2022) and the Flemish government (Michiels, 2023), adoption is still in its infancy.

As Solid-enabled data vaults transform how individuals interact with their personal data online, we expect a significant impact on the user experience of platforms that adopt the use of data vaults. This paper presents an exploratory study aimed at identifying the factors by which Solid-enabled data vaults impact the user experience. Therefore we conducted 10 usability tests, during which the participants created and used a data vault to share an official copy of their diploma with a job search platform. This method allowed us to observe how users interact with the Solid-enabled data vault in a practical scenario, which resulted in the identification of three factors in which Solid-enabled data vaults impacted the user experience. By aligning our results with relevant theoretical frameworks and extracting practical implications, we aim to equip organizations considering data vault implementation with the tools to proactively identify and address potential UX challenges. Furthermore, our results lay the base for future in-depth research into the identified effects of Solid-enabled data vaults on the user experience.

METHODOLOGY

To better understand the challenges users encounter while using Solid-enabled applications, we conducted a series of ten usability tests during which the participants interacted with a personal data vault in the context of a job application. The participants were therefore presented with a job opening on Randstad's[1] job search site (**Figure 1**, screen 1) and were tasked with applying for the job. To complete the application they needed to share an official digital copy of their diploma with Randstad. To collect and share their diploma digitally, participants were redirected to the Solid-enabled data vault feature of 'My Citizen Profile'. This platform, operated by the Flemish government, centralizes the collection and display of official data and government administration for citizens. After logging in (**Figure 1**, screen 2), participants received a step-by-step explanation of the data-sharing process (**Figure 1**, screen 3). Initially, they were asked to grant permission to create their personal data vault (**Figure 1**, screen 4). Subsequently, they needed to authorize Randstad to request documents from this vault (**Figure 1**, screen 5). Finally, participants selected the diplomas they wished to include in their data vault and shared them with Randstad (**Figure 1**, screen 6). After making their

[1] https://www.randstad.be/en/

selections, they were redirected to a confirmation screen on the Randstad platform.

All usability tests were administered online following a think-aloud protocol with active interventions (Krug, 2009; Olmsted-Hawala *et al.*, 2010). This involved the interviewer asking additional questions during the task, particularly when the user encountered difficulties, hesitated, or expressed confusion or frustration. Following the completion of the task, an interview was conducted to capture the user's comprehensive feedback and experiences.

In total, we conducted 10 usability tests, which are estimated to identify over 90% of usability issues. This estimate is based on the average probability of encountering a usability problem being $p = 0.31$ (Nielsen and Landauer, 1993; Sauro and Lewis, 2016).

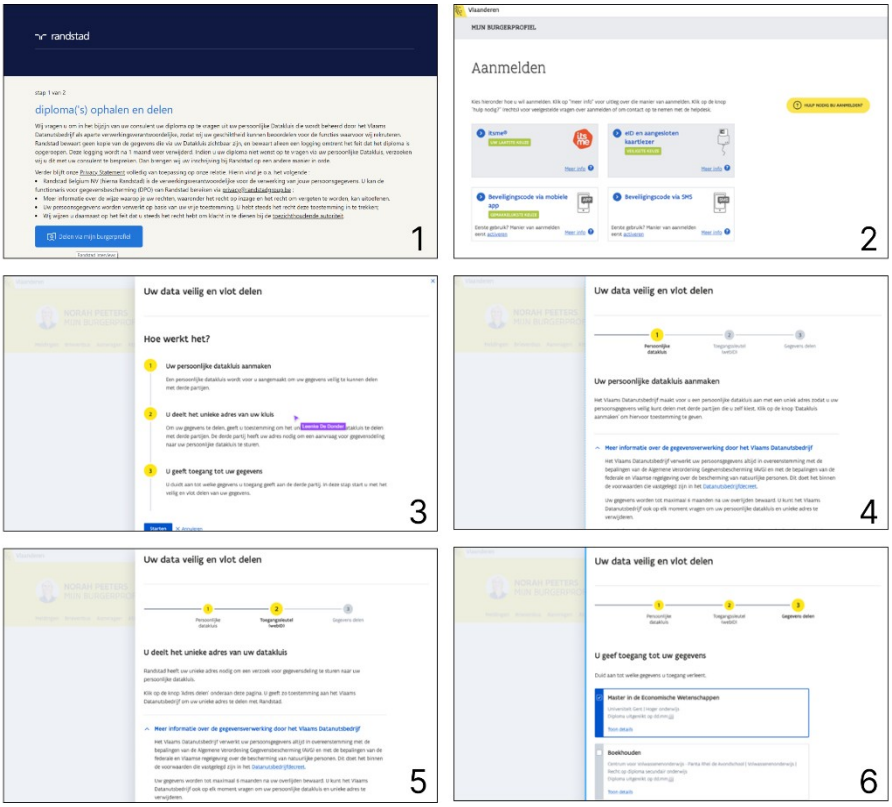

Figure 1: Steps participants followed to complete the task (access high resolution image: https://bit.ly/3wF5LN1).

The average age of the participants was 31.8 years and 50% of the participants had a higher education degree. Due to the predominantly female clientele of the Randstad branch that we partnered up with for the recruitment, we ended up with all female participants. Given the explorative nature of our research, we believe these results are nevertheless valuable

as they provide initial insights into user interactions and experiences with Solid-enabled data vault applications.

The initial transcription of the interviews was performed using the OpenAI Whisper model (OpenAI, 2022; Radford *et al.*, 2023), and verified by comparing them with the original recordings. We then analysed the transcripts using inductive thematic analysis, selected for its exploratory nature as this aligns with the explorative goal of our research.

RESULTS

The analysis of the usability tests revealed three main factors in which Solid-enabled data vaults impacted the user experience. The three factors are (1) users' mental models shaping expectations, (2) fears, doubts and the mitigating role of brand trust, and (3) a challenge of explainability rather than usability.

Users' Mental Models Shaping Expectations

Since none of the participants had previous experience with a data vault, we asked about their expectations to gain insights into their initial mental models of what a data vault is. Users generally viewed a data vault as a comprehensive data repository. This perception of a data vault as a primary data storage facility was exemplified by one participant's comment: *"I just think that it includes all the data and such, mine, of other people, where that is stored, where everything is: my phone number, name, national registry number, etc."*. Next to data storage, the second most anticipated functionality was the ability to control data access, as described by another participant: *"A piece on the internet that belongs to you and stores data about you and where you can choose yourself with whom you share it."*

Notably, various participants compared the data vault to familiar cloud storage services, as illustrated in the following quote: *"I had something like Dropbox in mind or Google Drive. A page where you can upload folders and files."* This comparison highlights a mental model in which users view data vaults as similar to familiar cloud services, thereby influencing their expectations.

Fears, Doubts and the Mitigating Role of Brand Trust

During the use of the platform, participants expressed significant concerns regarding the data security of the data vault. Due to frequent phishing attempts that some participants had previously experienced, many have grown a sense of scepticism and hesitation towards new online data-sharing processes. This hesitance was further intensified by their fear of accidentally giving access to sensitive personal information. Although a diploma or certificate was not generally seen as very sensitive information, participants feared unknowingly consenting to the sharing of important personal data due to the unclarity in the interface about what exactly was being shared with whom. Although data vaults hold the potential to address common data security issues, confronting the participants with the concept of a data vault

often triggered scepticism and doubt due to the users' unfamiliarity with this concept, leading to a hesitancy to engage with the application for some:

"In itself, I have no problem with sharing my diploma. If my diploma, say, leaks from your database, then I think: who cares? It's just my diploma. It's more about the process behind it, like okay, I'm giving access to someone via a link, what am I sharing then? And I also entered my code from my Itsme[2]*, can they also find out about that."*

However, the impact of the aforementioned fears and doubts on users' behaviour appears to be mitigated by their trust in the brand displayed in the interface. As all of the interviewees were already familiar with Randstad's platform, and generally perceived it as a trustworthy party, it appeared to be a strong mitigator for their expressed fears. One of the observed behaviours was that users generally just skimmed the explanations presented to them, without really questioning what was being asked for. When questioned about this, interviewees generally replied having trust in Randstad for safeguarding a safe process:

"The thing is, if I were to get that screen from Randstad, I would assume it's trustworthy, so you read it quickly and don't think about it too much. If I have to stop and understand what's written, then you've lost me."

During the later stages of interaction with the government's 'My Citizen Profile' interface, an analogous pattern of trust was observed among participants who were already familiar with the platform. This trust seemed to be influenced by the platform's association with the Flemish government, which acted as a trust-enhancing factor for most participants. In contrast, users who were not familiar with 'My Citizen Profile' exhibited more cautious behaviour. They generally paid closer attention to the information displayed and were more likely to raise questions about the security of the data vault.

A Challenge of Explainability Rather Than Usability

After completing their tasks, participants were invited to reflect on their experience sharing their diplomas via the data vault-enabled platform. The feedback predominantly highlighted the process's ease of use and usefulness, with users appreciating the reduced effort needed to share official documents by just clicking a few buttons. However, the ease of the process was not immediately apparent to all users. Initial perceptions of complexity were often attributed to unfamiliar terminology, such as 'data vault'. This initial perceived complexity of the process, followed by a realization of the process's simplicity, is captured in the following participant's statement:

"I think the information that is there could be more targeted, more specific. Just purely not to scare people off. Because as I see it now, I think wow, if I had known that, I probably would have just clicked through and finished it. But the fact that something is built up like, oh, a data vault and then the Data Utility Company[3] *[sighs]."*

[2]Itsme is a digital identification app that enables citizens to log in to government services, banks, insurers, and various private companies.

[3]The data utility company, as the provider of the Solid-enabled data vaults, is mentioned in the interface when users consent to create a data vault.

One strategy employed in the interface to demystify the concept of data vaults for end users involved using metaphors, like referring to a WebID[4] as 'the unique address of your data vault'. However, this metaphor led to confusion, with interpretations varying from physical address to email address or URL. This disparity between the intended metaphor and users' interpretation underscores the need for terminology that is more intuitive and familiar to the users.

Adding to the confusion was the users' exposure to multiple platforms throughout the process. This variety of platforms is a characteristic aspect of Solid's decentralized architecture, where data storage, identity, and services are being decoupled. In the diploma use case, users encounter four distinct brands. Starting from Randstad's platform, they were then directed to log into the government's 'My Citizen Profile' to share their digital diploma. This process necessitated that users verify their identity using Itsme and subsequently agree to the terms and services of the Data Utility Company, the provider of the data vault, to create their data vault. This multiplicity of platforms led to confusion and caused users to lose track of the process: *"I still don't really understand which profile I'm in. If I now give access to Randstad, how do they get on that? Because... in what way can they see that?"*.

In conclusion, while the process of creating a data vault and sharing data was generally evaluated as useful and easy-to-use, the main challenge lies in clearly explaining the concept of a data vault to end users, avoiding complex terminology and confusion due to the multiplicity of brands presented to the users.

ANALYSIS OF THE RESULTS: LINKING THEORETICAL FRAMEWORKS AND PRACTICAL IMPLICATIONS

To gain deeper insights into our results, we searched for relevant theoretical frameworks that align with the identified factors and can provide theoretical guidance for future research.

The first identified factor—that users' mental models and expectations of data vaults are mainly shaped by familiar cloud storage services—aligns with Jakob's Law of Internet User Experience (Nielsen and Loranger, 2006). This law states that users, who spend most of their time on other websites, have their expectations and mental models shaped by these previous experiences. As a result, they tend to expect new services, such as the data vault, to function similarly to those they are already familiar with. This underscores the importance of designing interaction patterns in Solid-enabled data vault applications that align with those in widely adopted cloud storage services to optimize the ease of use of data vaults. Furthermore, it is considered more essential that users' mental models are practical yet simplified, rather than technically precise (Gentner and Stevens, 2014). Therefore, since the technical concept of a data vault appeared to be difficult for users to grasp, the interface

[4] A WebID is a unique identifier that links to the user's data vault and enables authentication.

should primarily focus on communicating the practical implications of using data vaults rather than conveying the underlying Solid technology.

Secondly, it became evident that users have significant fears and doubts about the security of data vaults. However, the presence of a trusted brand displayed in the interface seemed to mitigate these concerns. Similar to findings in the context of e-commerce (Pavlou, 2003; Suh and Han, 2003) and social network services (Shin, 2010), perceived risk and trust thus appear to influence the acceptance of data vaults. This finding could guide the development of hypotheses for future research focused on the adoption of data vaults, where the precise impact and relationships between these factors can be explored in detail. Practically, companies could benefit by proactively addressing security concerns through the interface by displaying associations with well-known, trusted entities like the government.

Lastly, our findings suggest that the challenges faced by users were less about usability in the traditional sense and more about the clarity and understandability of the processes involved. This observed need to move beyond mere transparent data processes and progress towards truly legible and understandable data processing aligns with the directive outlined in the Human-Data Interaction framework (Mortier *et al.*, 2014). This framework highlights legibility, agency, and negotiability in data processes and could therefore guide the design of interaction patterns between users and their data vaults. In practice, this could involve steps toward standardizing terminology and explanations related to data vaults and the standardization of recurring interaction patterns, such as the process of creating a data vault and granting consent to access data within the vault.

CONCLUSION

Through a series of ten usability tests involving the creation and use of a Solid-enabled data vault, we identified three factors in which Solid-enabled data vaults impacted the user experience: (1) users' mental models shaping expectations, (2) fears, doubts and the mitigating role of brand trust, and (3) a challenge of explainability rather than usability. By connecting these factors to established theoretical frameworks on mental models, technology acceptance, and human-data interaction, and by coupling these insights with practical implications, we aim to provide organizations considering Solid implementation with the tools to proactively identify and tackle potential UX challenges. Furthermore, our exploratory findings pave the way for more in-depth future research on explainability, trust dynamics and technology acceptance, specifically applied to Solid-enabled data vaults, enhancing the potential for widespread adoption and optimization in various contexts.

It is however important to note that the conclusions of our exploratory research are based on a single use case within the HR domain. Additional studies involving different use cases are necessary to confirm these findings and broaden their applicability.

ACKNOWLEDGMENT

This research is supported by SolidLab Vlaanderen (Flemish Government, EWI and RRF project VV023/10); and imec ICON project SHARCS (Agentschap Innoveren en Ondernemen project nr. HBC.2022.0543).

REFERENCES

BBC (2022) 'Together+Data Pod'. Available at: https://www.bbc.co.uk/taster/pilots/together-pod (Accessed: 2 June 2023).

Berners-Lee, T., Hendler, J. and Lassila, O. (2001) 'The Semantic Web: A New Form of Web Content That is Meaningful to Computers Will Unleash a Revolution of New Possibilities', *ScientificAmerican.com* [Preprint].

Dedecker, R. *et al.* (2022) 'What's in a Pod?~– A knowledge graph interpretation for the Solid ecosystem', in M. Saleem and A.-C. Ngonga Ngomo (eds) *Proceedings of the 6th Workshop on Storing, Querying and Benchmarking Knowledge Graphs*. (CEUR Workshop Proceedings), pp. 81–96. Available at: https://solidlabresearch.github.io/WhatsInAPod/.

Gentner, D. and Stevens, A. L. (2014) *Mental Models*. Edited by D. Gentner and A. L. Stevens. Psychology Press. Available at: https://doi.org/10.4324/9781315802725.

Krug, S. (2009) *Rocket surgery made easy*.

Mansour, E. *et al.* (2016) 'A Demonstration of the Solid Platform for Social Web Applications', in *Proceedings of the 25th International Conference Companion on World Wide Web*. Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee (WWW '16 Companion), pp. 223–226. Available at: https://doi.org/10.1145/2872518.2890529.

Mechant, P. *et al.* (2021) 'Saving the web by decentralizing data networks? A socio-technical reflection on the promise of decentralization and personal data stores', *2021 14th CMI International Conference - Critical ICT Infrastructures and Platforms (CMI)*, pp. 1–6. Available at: https://doi.org/10.1109/CMI53512.2021.9663788.

Michiels, C. (2023) 'Vlaanderen lanceert Athumi: bepalen we binnenkort zelf wat er met onze online data gebeurt?', 4 May. Available at: https://www.vrt.be/vrtnws/nl/2023/05/04/vlaanderen-lanceert-athumi-bepalen-we-binnenkort-zelf-wat-er-me/ (Accessed: 2 June 2023).

Mortier, R. *et al.* (2014) 'Human-Data Interaction: The Human Face of the Data-Driven Society'. Available at: https://doi.org/10.48550/arXiv.1412.6159.

Nielsen, J. and Landauer, T. K. (1993) 'A Mathematical Model of the Finding of Usability Problems', in *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery (CHI '93), pp. 206–213. Available at: https://doi.org/10.1145/169059.169166.

Nielsen, J. and Loranger, H. (2006) *Prioritizing Web Usability*. USA: New Riders Publishing.

Olmsted-Hawala, E. L. *et al.* (2010) 'Think-Aloud Protocols: A Comparison of Three Think-Aloud Protocols for Use in Testing Data-Dissemination Web Sites for Usability', in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery (CHI '10), pp. 2381–2390. Available at: https://doi.org/10.1145/1753326.1753685.

OpenAI (2022) *Whisper [Python]*. Available at: https://github.com/openai/whisper (Accessed: 9 January 2024).

Pavlou, P. A. (2003) 'Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model', *International Journal of Electronic Commerce*, 7(3), pp. 101–134. Available at: https://doi.org/10.1080/10864415.2003.11044275.

Radford, A. *et al.* (2023) 'Robust Speech Recognition via Large-Scale Weak Supervision', in A. Krause et al. (eds) *Proceedings of the 40th International Conference on Machine Learning*. PMLR (Proceedings of Machine Learning Research), pp. 28492–28518. Available at: https://proceedings.mlr.press/v202/radford23a.html.

Sauro, J. and Lewis, J. R. (2016) *Quantifying the User Experience, Second Edition: Practical Statistics for User Research*, *ACM SIGSOFT Software Engineering Notes*.

Shin, D.-H. (2010) 'The effects of trust, security and privacy in social networking: A security-based approach to understand the pattern of adoption', *Interacting with Computers*, 22(5), pp. 428–438. Available at: https://doi.org/10.1016/j.intcom.2010.05.001.

Solove, D. J. (2013) 'Introduction: Privacy Self-Management and the Consent Dilemma Symposium: Privacy and Technology', *Harvard Law Review*, 126(7), pp. 1880–1903. Available at: https://heinonline.org/HOL/P?h=hein.journals/hlr126&i=1910.

Suh, B. and Han, I. (2003) 'The Impact of Customer Trust and Perception of Security Control on the Acceptance of Electronic Commerce', *International Journal of Electronic Commerce*, 7(3), pp. 135–161. Available at: https://doi.org/10.1080/10864415.2003.11044270.