GI-Edition

Lecture Notes in Informatics

Heiko Roßnagel, Christian H. Schunck,
Filipe Sousa (Hrsg.)

Open Identity Summit 2024

20.–21. June 2024
Porto, Portugal

Proceedings

350

Heiko Roßnagel, Christian H. Schunck, Filipe Sousa (Hrsg.): Open Identity Summit 2024

GESELLSCHAFT
FÜR INFORMATIK

Heiko Roßnagel, Christian H. Schunck,
Filipe Sousa (Hrsg.)

Open Identity Summit 2024

20. - 21.06.2024
Porto, Portugal

Gesellschaft für Informatik e.V. (GI)

Lecture Notes in Informatics (LNI) - Proceedings
Series of the Gesellschaft für Informatik (GI)

Volume P-350

ISBN 978-3-88579-744-9
ISSN 1617-5468

Volume Editors
Heiko Roßnagel | Christian Schunck
Fraunhofer IAO, Fraunhofer Institute for Industrial Engineering IAO
Nobelstr. 12, D-70569 Stuttgart, Germany
heiko.rossnagel|christian.schunck@iao.fraunhofer.de
Filipe Sousa
Fraunhofer Portugal AICOS
Rua Alfredo Allen 455/461, 4200-135 Porto, Portugal
filipe.sousa@aicos.fraunhofer.pt

printed by Köllen Druck+Verlag GmbH, Bonn

Preface

Welcome to the “Open Identity Summit 2024”, which has been jointly organized by the Special Interest Groups BIOSIG within the German Computer Science Society (Gesellschaft für Informatik), Fraunhofer Institute for Industrial Engineering IAO and the Fraunhofer Center for Assistive Information and Communication Solutions – AICOS in Portugal.

The international program committee performed a strong review process according to the LNI guidelines with at least three reviews per paper and accepted 46% of the 28 submitted papers as full scientific papers.

Furthermore, the program committee has created a program including selected contributions of strong interest (further conference contributions) for the outlined scope of this conference.

We would like to thank all authors for their contributions and the numerous reviewers for their work in the program committee.

Porto, 20th of May 2024

Heiko Roßnagel
Fraunhofer IAO

Christian H. Schunck
Fraunhofer IAO

Filipe Sousa
Fraunhofer AICOS

Conference Chairs

Heiko Roßnagel, Fraunhofer Institute for Industrial Engineering IAO
Christian H. Schunck, Fraunhofer Institute for Industrial Engineering IAO
Filipe Sousa, Fraunhofer Portugal AICOS

Programme Committee

Jürgen Anke, Germany
Franco Arcieri, Italy
Tamas Bisztray, Norway
Arslan Broemme, Germany
Christoph Busch, Germany
Victor-Philipp Busch, Germany
Jos Dumortier, Belgium
Nicolas Fähnrich, Germany
Daniel Fett, Germany
Waldir Moreira, Portugal
Lothar Fritsch, Norway
Igor Furgel, Germany
Jochen Günther, Germany
Marit Hansen, Germany
Michael Heinl, Germany
Olaf Herden, Germany
Gerrit Hornung, Germany
Detlef Houdeau, Germany
Richard Huber, Germany
Detlef Hühnlein, Germany
Tina Hühnlein, Germany
Ulrike Korte, Germany
Michael Kubach, Germany
Andreas Kühne, Germany
Sebastian Kurowski, Germany
Herbert Leitold, Austria
Luigi Lo Iacono, Germany
Tarvi Martens, Estonia
Nadia Menz, Germany
Sebastian Mödersheim, Denmark
Alexander Nouak, Germany
René Peinl, Germany
Henrich Pöhls, Germany
Daniela Pöhn, Germany
Alexander Roßnagel, Germany
Heiko Roßnagel, Germany
Mirko Ross, Germany
Christian H. Schunck, Germany
Rachelle Sellung, Germany
Detlef Stern, Germany
Jon Shamah, United Kingdom
Filipe Souza, Portugal
Hermann Strack, Germany
Maurizio Talamo, Italy
Karsten Treiber, Germany
Samuel Wairimu, Sweden
Tobias Wich, Germany
Thomas Wieland, Germany
Alex Wiesmaier, Germany
Jan Zibuschka, Germany
Frank Zimmermann, Switzerland

Hosts and Partners

BIOSIG – Biometrics and Electronic Signatures (www.biosig.org)
The special interest group “Biometrics and Electronic Signatures” (BIOSIG) within GI e.V. is dedicated to the fundamentals, methods, techniques, processes and implementations used to guarantee the authenticity and integrity of entities.

Inhaltsverzeichnis

Regular Research Papers

Further Conference Contributions

Regular Research Papers

Towards Robust Trust Frameworks for Data Exchange: A Multidisciplinary Inquiry

Aytaj Badirova [1], Bithin Alangot[2], Theo Dimitrakos[3], and Ramin Yahyapour[4]

Abstract: Data exchange is essential in the fast-changing field of data-driven innovations. In exploring the importance of data and data exchange, this paper highlights the necessity of building trust. For data sharing to be successful, trust is essential for assuring reliability, security, and ethical behaviour. We review the current state of the art where research and real-world applications converge in both academia and industry. Notably, trust framework-based projects are starting to take shape, promoting safe and open data markets that are included in this work. Nonetheless, challenges still exist, such as complex legal, technological, and business issues. Some of the main challenges that are faced while establishing a trust framework have also been briefly mentioned in this paper.

Keywords: trust framework, dataspace, data exchange

1 Introduction

Data is the cornerstone of innovations and a fundamental asset for organizations. A tremendous amount of data is being gathered and maintained. From business to health, each sector continuously gathers and maintains data, primarily internally and in accordance with local norms and regulations. On the other hand, interconnectivity is necessary to advance development in collaborative processes and innovations. The significance of data lies in the potential of bringing novelties. Data kept in silos cause different obstacles in various sectors creates a barrier for success in data exchange. The most frequent problems are those with trust, compliance (e.g., technical, legal, or policy level), and laws. The foundational prerequisite for the data exchange process is establishing an appropriate trust architecture that parties can adhere to. Encouraging a trustworthy framework will draw in a higher number of partners, increasing the volume and variety of data exchanged and broadening the scope. Trust architecture plays therefore a crucial role in data management and is a basement for innovations. To enable trust in a large data exchange where there are numerous participants from diverse backgrounds (e.g., different policies, access management schemas, or legal variations in

[1] German Research Center, Huawei Technologies Düsseldorf GmbH, Munich, Germany, email: aytaj.badirova@huawei.com,
[2] German Research Center, Huawei Technologies Düsseldorf GmbH, Munich, Germany, email: bithin.alangot@huawei.com
[3] School of Computing, University of Kent, Canterbury, UK, email: theo.dimitrakos@huawei.com
[4] Institute of Informatik, University of Goettingen, Goettingen, Germany, email: ramin.yahyapour@gwdg.de

doi:10.18420/OID2024_01

a case of cross-border collaborations) is a very challenging process, and to have an architecture that works for all is nearly unattainable. However, for an adequate number of participants with more similar interests, domain-specific trust can still be enabled. In order to address the aforementioned difficulties in establishing a trustworthy data exchange environment, Fraunhofer Institute developed a novel strategy known as Dataspace. The participants of a dataspace or dataspaces can publish and share their data with other participants in a secure way while being compliant with various regulations. Dataspace provides interconnectivity while encouraging data sovereignty – participants decide to give access to others in the ecosystem. The process of establishing such a system starts with trust. However, achieving a shared sense of trust without depending on a single, central decision-making authority is an incredibly challenging task. Though very difficult, a number of studies and projects aim to create a decentralised, unbiased, and equitable trust structure that can benefit all parties equally. One of the recent initiatives called the International Dataspace Association - IDSA [IDS24] is formed to support and govern the adoption of such a trustworthy environment. The role of IDSA is to coordinate and support the processes at a business level, however, there is still a need to define more granular approaches to define a secure architecture and guidelines. Gaia-X [Gai24] is another important initiative to foster data sovereignty, interoperability, and secure data exchange by providing reference architectures and trusted frameworks. Both initiatives have been covered in this report together with current implementations and projects where the main focus is establishing a trust framework that helps participants follow the EU's data strategy. Industry and academia propose diverse approaches for building domain-specific trust. This work aims to review the major models from different fields, both conceptual and practical with their pros and cons, and discuss the significant challenges that still need to be resolved.

This is the first thorough study of trust in a data exchange domain. The study's classification of various data exchange strategies into different categories according to their features, characteristics, and needs, such as industrial/international, cyber-threat intelligence, personal, and transportation dataspace models, is another significance distinction. Furthermore, by outlining potential benefits and flows, this paper questions the current solutions and offers valuable guidance for future research and ecosystem development.

2 Scholarly Foundations and Practical Implications

Although the process of exchanging data is not new, the idea of "dataspace" seeks to expand it into the direction of sovereign data. This data management approach focuses on uniting data/service providers under one roof and offering them a reliable, seamless infrastructure for data exchange while encouraging data sovereignty. In order to achieve its primary objective, there are several proposed approaches for data exchange architecture and to establish a trust framework to make the exchange process reliable in dataspace system. Scholarly foundations and the most prominent architectures and

frameworks have been explored together with leading initiatives in this section.

2.1 Trust in Data Exchange

To establish a reliable data sharing environment, proper data exchange mechanisms are crucial. Data exchange models provide a framework for defining the roles and responsibilities of the parties involved in data exchange, as well as the technical mechanisms for securely transferring and storing data. There is a slight gap between the state of the art in academics and industry, which is also influenced by strict regulations, given the rapid advancement of technology. At this point, it is important to note that the trust framework in dataspace is a very broad concept, and covering every component in a single study is quite difficult. Thus, the majority of the works that are suggested concentrate on a particular aspect of building trust. This section's first half discusses the state of research regardless of a particular domain, and the latter portion focuses on the state of the industry.

State of the Art in Research: Fraunhofer IAO suggested the Trust mAnagement Infrastructure (TRAIN) [Jey22] as a trust schema to build trust in decentralised identity data management. The suggested methodology provides a scalable mechanism to confirm the credential's issuer and determine its reliability. The DNS list serves as the foundation for both the model and the verification process. The proposed has approach many great benefits, but it also has several drawbacks, like being limited to a static list and missing a trust level variation. Establishing trust in a decentralized environment for identity data highly depends on the used technology on personal devices – so called digital wallets. There are relatively small amount of work that focuses on security in wallets such as the apporach called DOOR [La23]. The proposed model facilitates the establishment of both identity integration and wallet validity while being aligned with the current regulations (e.g., eIDAS). It applies Attribute-based Direct Anonymous Attestation (DAA-A) cryptographic protocol to ensure anonymity, unforgeability, and unlinkability.

Different sectors have diverse approaches when it comes to data sharing. The purpose of the article [CFH22] is to assess and examine common automotive HTAs (hardware trust anchors) in terms of their suitability for application in contemporary and future vehicle architectures. The authors use the automotive domain analysis and associated studies to establish thorough evaluation criteria in order to do this. Key technologies evaluated are Trusted Platform Module, Hardware Security Module, and Secure Hardware Extension. By bridging the gap between theoretical discussions and real-world requirements in the automotive arena, the study offers a thorough review of HTAs. Another study focuses on trust establishment for data exchange in the Internet of Vehicles (IoV)[AlM24]. It uses Naive Bayes machine learning to propose a classification-based trust model (CTM) that is specifically designed for IoV. By classifying cars as trusted or untrusted, this concept improves secure communication throughout the Internet of Vehicles. In the era of connected automobiles, the research advances safer data transmission by improving the

effectiveness of trusted and untrusted vehicle recognition.

In diverse ecosystems, the quantity of participants in data exchange may differ. Whether centralised or decentralised, the exchange procedure becomes increasingly difficult as the number of participants rises due to the variations in access control methods. It complicates the process of building trust. Thus, it is necessary to build a bridge between two different systems, especially between decentralised and legacy systems. The goal of the Fed2SSI [Ku23] suggested architecture is to build this bridge by converting legacy credentials into the appropriate verifiable presentation (VP), which is comprehensible to all parties involved in the data exchange ecosystem. The suggested method boosts participant trust while enhancing interoperability. The study [SSA21] focuses on a novel approach to data exchange that is inspired by Channel Island legislation. The main focus is on data governance and trust. The study aims to establish trust in personal data exchange by including citizens as stakeholders. The key guidelines have been proposed to implement them in practice for organizations. The guidelines define the boundaries, goals, stakeholders, limitations, and responsibilities. On the other hand, it has limitations, such as not covering diversities in trust levels and different data types.

Exploring Industrial Initiatives: IDSA and Gaia-X provide frameworks for trust establishment, negotiation, and data exchange. IDS-RAM [IDS24] is the initial data exchange framework that is provided by IDSA that is aligned with Gaia-X trust framework. The main component of the architecture is a Connector. In the secure dataspace formed by the Connector connection, the data provider and the data consumer transfer data including data metadata, data + usage policies, and data processed in the DataApp. Parties in an IDS can exchange data over secure channels using the IDS Communications Protocol (IDSCP).

The IDS-RAM data exchange model is primarily designed for industrial data and is not well-suited for personal data, which is more sensitive and requires a more privacy-focused approach. One such conceptual model is the W3C SOLID Data Pod [Sa16], a decentralized data storage system that gives users full control over their data. SOLID Data Pods offer advantages for personal data exchange, including decentralization, user control, and fine-grained access control. SOLID aims to provide private spheres – PODs for individuals to keep the control over their data. Instead of a centralized systems, the data can be stored in decentralized way. Users have full control of their data, they decide who can access what, to what degree.

Distributed Data Store Mesh is another proposed approach for securing personal data exchange process. It puts control over personal data governance, disclosure and usage in the hands of the data owner. These ledgers are well suitable for providing the immutable foundation for Decentralized Identifiers, but should not be used to store personal identity data. This would be at odds with the goals of privacy by design and with existing and emerging data regulations, such as the GDPR. Instead, we need a different solution for secure storage of personal data and information. Identity Hubs are that solution. Identity

Hubs are decentralized, off-chain, personal data stores that put control over personal data in the hands of users. They allow users to store their sensitive data-identity information, official documents, app data, etc.-in a way that prevents anyone from using their data without their explicit permission. Users can use their Identity Hubs to securely share their data with other people, apps, and businesses, providing access to the minimum amount of data necessary, while retaining a record of its use.

A significant portion of the data-focused industries are covered by industrial and personal dataspaces, but not nearly enough to address problems in other specialized industries, like transportation. To improve the transportation systems for passengers and transports the projects C3ISP [C324] and E-CORRIDOR [Ed24] established another data exchange framework that serves Collaborative Cyber Threat intelligence (C3IS). They employ a data exchange pattern that is based on the transfer of a bundle consisting of protected data and the associated agreement about sharing and usage of the protected data. This is referred to as Data Bundle or (Data Protected Object). Data exchange relies on Data Sharing Agreements (DSAs) to manage access and usage of data. Each DSA incorporates information about data provider, data consumer, the validity period of the DSA, a list of parties that can use this DSA and map it to their data and a set of policies.

The described models emphasise a secure method of data exchange. However, maintaining data control both during transmission and after exchange is a difficult challenge. Therefore, a new approach was proposed by MIT OPAL [Ce23] that focuses on moving algorithms, not data. Data needs to be stored in an encrypted form and computations should be done on the data side on encrypted data. OPAL ("Open Algorithms") is a non-profit social technology innovation founded in 2017 by the MIT Media Lab. The core idea is that data is not copied or moved. Algorithms are deployed on computing nodes of data providers after privacy compliance review – bring algorithm to data, not vice versa. The access control scheme in this concept is called consent for execution instead of consent for access.

The aforementioned architectures, which include the identity hub in distributed data mesh, the pod in W3C Solid, and the connector in IDS architecture, center on the flow of data exchange. These methods are insufficient on their own to provide a safe and reliable environment. It emphasizes the necessity of the trust frameworks where the primary dataspace trust frameworks have been covered in the next section.

2.2 Trust Frameworks

Sharing and accessing digital data is a way of building robust systems, and analytics, dealing with obstacles in business, and opening new horizons for organizations. Security is one of the key features of digital data sharing. Regardless of academia or industry, the goal is to share/access required data in a trustworthy way where the degree of security must be agreed upon. This agreement should be based on a set of rules – not only company-based, but also high level such as national and international. This systematic

collection of rules is called a Trust Framework. By enacting the necessary legislation and acts, the trust framework creates a reliable data sharing environment for all participants. A trust framework architecture has been initiated by many participants in the dataspace domain. This section covers the primary approaches. The Gaia-X Association developed its Gaia-X Framework, which enables the transition from disjoint data and infrastructure ecosystems, to composable, interoperable, and portable cross-sector data sets and services. Gaia-X Framework builds on top of the X-Model in order to enable trust and interoperability within and across dataspaces and federations. The Gaia-X Trust Framework is the set of rules that define the minimum baseline to be part of the Gaia-X Ecosystem. Those rules provide common governance and the basic level of interoperability across individual ecosystems while letting the users be in full control of their choices.

While Gaia-X offers a trust framework for specifically dataspace, Trust Over IP [Da19] trust model concentrates on internet-wide trust via verifiable credentials, trust registries, and decentralized identity. ToIP trust framework is supported by Linux Foundation. The goal is to provide a “trust layer” that can be achieved internet-wide. The model merges technology and governance in four layers of stack where the technology stack includes Public Utilities, Peer-to-Peer Communication, Trust Task Protocols, Application Ecosystem and the governance stack contains Utility Governance Framework, Agent/Wallet Governance Framework, Trust Tasks Framework, Ecosystem Governance Framework. The greater and more comprehensive scope of ToIP increases complexity and slows down the adaption process.

DSBA Trust Framework [DS24] is another trust model approach that provides a Trust Anchor for a secure ecosystem. DSBA Trust Anchor adopts Gaia-X Trust Framework and sets additional rules on top of it which allow organizations to use their digital identities for interactions. This makes it easier for organizations to trust each other and share data securely. DSBA trust framework focuses more on business compliance while previously mentioned approaches are technical compliance oriented. DSBA TF addresses he issues mentioned in Figure 1.

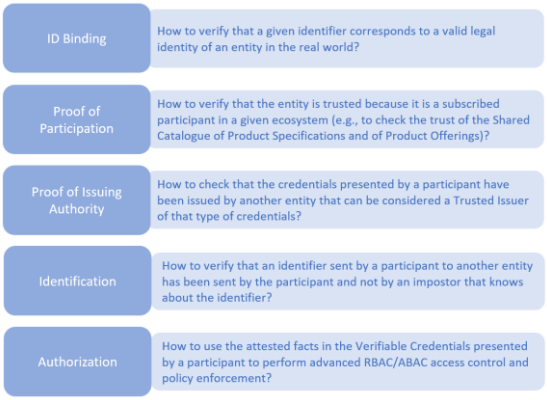


Figure 1 DSBA TF focused issues

Compared to other data types, the trust framework for identity data demands a more fine-grained and privacy-preserving trust architecture, especially in the EU. eIDAS [eID24] aims to provide a secure trust framework architecture for primarily identity data exchange in the EU among organizations, institutions, medical, and banking sectors etc. The goal is to make the digital identity available EU-wide. However, the concept is not limited to the EU, but can be adapted to other regions as well. The project supports the sovereign identity concept where individuals are in charge of their personal data, they should be able to share only the required data via a selective disclosure method and should be compatible with GDPR rules. An enhanced version of eIDAS, known as eIDAS2, was proposed in an effort to increase trustworthiness. Increased user experience, cross-border adaptability, and e-ID schema convergence are the goals of the new version. The trust frameworks that have been discussed encompass a variety of industries and data kinds. Overall, they all work toward the same objective of facilitating trustworthy and safe data transfers across the EU and beyond while abiding by national, international, and local laws and regulations. These architectures have been applied in the implementation of the initial dataspace projects that are discussed in detail in the following section. The first dataspace initiatives, which are covered in detail in the section that follows, were implemented using these architectures.

2.3 Initiatives and Trends

The design and acceptance of secure data exchange efforts across multiple industries and data sectors have been advocated for by a variety of public and commercial entities. Several notable initiatives by the main actors are covered in this section. Industrial/international dataspaces are designed for scalable enterprise data exchange and collaboration among businesses and organizations. They must handle a wide variety of structured and unstructured data, often at high volume and velocity. Industrial/international dataspaces provide robust data ingestion, processing,

management, and usage control mechanisms of the data based on commonly agreed policies. These are built to facilitate data exchange and collaboration at an industrial or organizational level. In these environments, data from various sources, including IoT devices, operational data, business data, etc., is pooled and made accessible to participating entities under specific conditions. The governance model in industrial dataspaces ensures a secure, controlled and equitable sharing of data, respecting data privacy laws, and facilitating collaboration and value creation. The typical data exchange pattern for industrial/international dataspace is that of an IDS connector architecture and the IDSA Dataspace Protocol. The emerging reference architecture for such dataspaces is based on a convergence of the IDS-RAM and the Gaia-X Architecture on a baseline defined by the DBSA convergence activity. Most of the current lighthouse projects follow this model. Catena-X [Cat24], German Mobility Dataspace (MDS) [MD24] and (more recently) Manufacturing-X [MX24] are the most widely recognized initiatives in this domain.

Mobility Dataspace is another striking dataspace idea that has drawn interest in the EU. Particularly, the German government is interested in developing this industry. More than 200 German mobile stakeholders from science, business, and public administration participated in its vision. Its goal is to further develop the Mobility Dataspace as a business organization and ensure its technological advancement. Traffic optimization, multimodal transportation, and new mobility services are among the main objectives of the Mobility Dataspace while establishing a sustainable, efficient, and user-friendly mobility system.

Intelligent Transportation System (ITS) Dataspaces and mobility dataspaces are designed to offer an open ecosystem that enables the trustworthy exchange of data between different traffic participants, providers, and operators in order to optimize traffic flows, increase safety and protect the environment. The ITS dataspace addresses use-cases that are related to data generated by vehicles, privately owned mobile devices, as well as may be collected and processed by public transport providers, navigation service providers, fleet operators (OEMs) and mobile communication providers (MNOs) which is of sensitive in nature and need to be handled uniquely (with separate policies and technical requirements).

Automotive and mobility dataspace schemas follow industrial data exchange patterns. However, this approach is insufficient for person data. Therefore, a new dataspace architecture model was established called Personal Dataspace. Personal Dataspaces are designed around individual users, enabling them to maintain control and portability over their own data, deciding who has access and for what purpose. The architecture of personal dataspaces usually prioritizes ease of use, transparency, and privacy. MyData Global [MG24] is the main forum where new projects focusing on personal dataspaces are concentrated. The typical characteristic of the dataspace is targeting specific data sectors such as industry, mobility, health, green hub etc. It brings out one of the main differences of personal dataspaces – to play a bridge role among dataspaces. aNewGovernance is one of those examples that focuses on this feature [AG24]. Its goal

is to move from platform-centric data model to human-centric one – individuals should be able to use their data on diverse dataspaces without additional adjustment. Since aNewGovernance targets personal data it has to follow GDPR and other personal data regulations strictly. Personal dataspace covers data more than personal identity information, such as skills and education. Dataspace for Skills (DS4Skills) [DS24] and Prometheus (from DASES - Dataspace Education and Skills) [Px24] both aim to collect, store, and securely share educational and skill data with businesses with the permission of data owners. Individuals will be able to quickly enter the job market in this way, while organizations will have better opportunities to locate skilled staff.

As of now, the presented dataspaces serve similar purposes such as increasing connectivity, establishing new business models, supporting business growth, encouraging sovereignty while preserving security. However, dataspace can also be established to increase security in digital ecosystem – preventing cyber-crimes. In this case, security is the goal not the feature. Sensitive Dataspaces (SDS) are designed with that goal in mind. Sensitive Dataspace's consists of building blocks that can help with secure storage, management, exchange and analysis of data of critical nature, such as cyber-incidents information and more detailed Cyber Threat Intelligence (CTI). As with any dataspace, SDS relies on a data governance paradigm, which comprises a set of rules and policies determining the rights to access process, use and share data in a trustful way. Since data used and shared within this particular dataspace is considered to be sensitive and contains confidential information, data providers must have complete control over who can have access to their data, for which purpose, and under which conditions it can be used. The C3ISP and CyCLONe projects [C324] project are of examples of a sensitive dataspace.

3 Challenges

A reliable data exchange ecosystem has given businesses new avenues for growth. Thus far, the present implementations and initiatives, together with the concept of dataspace and its primary characteristics and components, have been discussed. Nevertheless, like other advancements, it has its share of obstacles and challenges. This new strategy, given its infancy, presents a number of challenges and risks from many angles that need to be handled. To provide an overview, this section addresses the main challenges in the dataspace domain, ranging from technological to legal, business, and sectoral (Figure 1). Sovereign data management gives companies more control and security, but it also presents new difficulties. It takes a significant adjustment to create new business models that succeed in a decentralised data environment. Complexity is increased by creating and sustaining a dynamic network of cooperative partners. It might be difficult to negotiate data usage with different parties under different legislation. To fully realize the benefits of data sovereignty, businesses must carefully negotiate these obstacles, weighing opportunities against the reality of this shifting paradigm.

3.1 Domain Specific Challenges

Health: Health data is one of the most crucial data types that requires high security and privacy. Therefore EU aims to bring sovereignty to this data type where the individuals will be in full control of the data. The data sovereignty approach seems promising to prevent security incidents. Hence, it completely aligns with the first purpose of health data. However, when it comes to the secondary use of health data - which is research, the new approach creates barriers such as getting consent from users, translating the language of the data when changing the country, security of the shared data in the case of doctor visits etc. A more comprehensive evaluation of the security concerns related to health data has been covered in the study [Ma23]. These challenges call for an appropriate foundation of trust wherein the requirements of both sides of the health data can be met. The answer has not yet been developed.

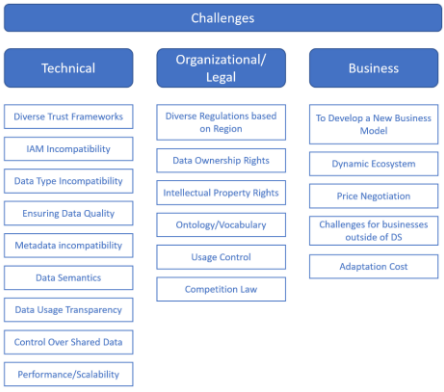


Figure 2 Challenges in Data Exchange

Personal data: Person dataspace is one of the most challenging, considering its criticality. In the EU, eIDAS 2.0 is one of the main trust frameworks for organisations, regardless of whether they are governmental or non-governmental, to deal with data in a secure way. It aims to give individuals control over their data, and in this way, it is possible to avoid massive personal data breaches or honeypots. However, the suggested paradigm is incomplete and contains security flaws. The aforementioned points are supported by the German IDWallet incidents, in which sensitive personal data was released to unaffiliated parties without the required authentication [SDI22].

Finance: The challenges that financial data interchange must overcome are numerous and need careful planning. Considering that financial data is one of the most powerful data that gives a lot of information about individuals, it requires strong encryption, authentication, and consent management are essential to guaranteeing data privacy and security. Interoperability, consistency, and data validation are common needs shared by various industries. Dealing with the date and time is very important, especially when there is cross-border data interchange where is requires specific adjustment to the

established ecosystem.

4 Conclusion

Data is a valuable asset, thus gathering more of it will increase its value. On the other hand, data in silos cannot be used to its full potential, so data exchange is a requirement. A trusted environment that facilitates seamless transactions is necessary for the data exchange process. In the EU, data protection is given top attention, and numerous laws, rules, and data acts are in place. The drive for greater cooperation is hampered by how difficult it is to follow each guideline precisely for each partnership. In order to ease this process dataspace gives stakeholders high interoperability while enabling data exchange in a safe and trustworthy environment. Gaia-X provides a framework to follow these sets of rules and regulations while starting a Dataspace. Gaia-X regulations are very well aligned with EU data and cloud strategies which means all the regulations are taken into account. As providing regulation frameworks is not enough for building a Dataspace, IDS-RAM proposed a blueprint architecture for data exchange. Gaia-X and IDSA architectures led several organisations to launch dataspace projects in various industries including German Mobility Dataspace, Catena-X, MyData Global and others that have been covered in this paper. This paper provided a thorough analysis of the trust establishment in data exchange, including reference architectures, adaptable areas, involved organisations, and state-of-the-art. However, because the data providers come from a variety of backgrounds, including technical and legal differences, it is highly challenging to develop perfect interoperability together with high security.

Bibliography

[ID24] International Data Space Association. International Data Spaces. URL: https://internationaldataspaces.org/ (acc. 02/01/2024)

[Ga24] Gaia-X. Gaia-X. URL: https://gaia-x.eu/ (acc. 02/01/2024).

[Je22] Jeyakumar, J., et al. "A novel approach to establish trust in verifiable credential issuers in Self-sovereign identity ecosystems using TRAIN." In Open Identity Summit 2022, 2022.

[La23] Larsen, B., et al. "Achieving Higher Level of Assurance in Privacy Preserving Identity Wallets." URL: www.ioanniskrontiris.de/publications/2023937.pdf (acc. 02/01/2024).

[CFH22] Plappert, C., Andreas, F., & Ronald H. "Analysis and Evaluation of Hardware Trust Anchors in the Automotive Domain." In Proceedings of the 17th International Conference on Availability, Reliability and Security, 2022.

[EM24] Alalwany, E., Imad, M. "Security and Trust Management in the Internet of Vehicles (IoV): Challenges and Machine Learning Solutions." Sensors 24.2 (2024): 368.

[Ku23] Kubach, M., et al. "A shared responsibility model to support cross border and cross organizational federation on top of decentralized and self-sovereign identity: Architecture and first PoC." In Open Identity Summit 2023, 2023.

[Sa16] Sambra, A.V., et al. "Solid: a platform for decentralized social applications based on linked data." MIT CSAIL & Qatar Computing Research Institute, Tech. Rep., 2016.

[Co24] Collaborative and Confidential Information Sharing. URL: https://c3isp.eu/ (acc. 02/01/2024).

[Ed24] Edge enabled Privacy and Security Platform for Multi Modal Transport. URL: https://e-corridor.eu/ (acc. 02/01/2024).

[Ch23] Chen, Z., et al. "OPAL: Ontology-Aware Pretrained Language Model for End-to-End Task-Oriented Dialogue." Transactions of the Association for Computational Linguistics 11 (2023): 68-84.

[Da19] Davie, M., et al. "The trust over ip stack." IEEE Communications Standards Magazine 3.4 (2019): 46-51.

[DS24] Data Spaces Business Alliance. Data Spaces Business Alliance. URL: https://data-spaces-business-alliance.eu/ (acc. 02/01/2024).

[Ei24] eIDAS Regulation. URL: https://digital-strategy.ec.europa.eu/en/policies/eidas-regulation (acc. 02/01/2024).

[Ca24] Catena-X. Catena-X. URL: https://catena-x.net/en/ (acc. 02/01/2024).

[GM24] German Mobility Dataspace. URL: https://mobility-dataspace.eu/ (acc. 02/01/2024).

[Ma24] Manufacturing-X. Manufacturing-X. URL: https://www.plattform-i40.de/IP/Navigation/EN/Manufacturing-X/Manufacturing-X.html (acc. 02/01/2024).

[My24] My Data Global. My Data Global. URL: https://mydata.org/ (acc. 02/01/2024).

[En24] aNewGovernance. aNewGovernance. URL: https://www.anewgovernance.org/(acc. 02/01/2024).

[Da24] Data Space for Skills. Data Space for Skills. URL: https://www.skillsdataspace.eu/(acc. 02/01/2024).

[Pr24] Prometheus-x. Prometheus-x. URL: https://prometheus-x.org/?locale=en(acc. 02/01/2024).

[Ma23] Marelli, L., et al. "The European Health Data Space: Too Big To Succeed?." Health Policy (2023): 104861.

[SAA22] Schwalm, S., Daria, A., & Ignacio, A. "eIDAS 2.0: Challenges, perspectives and proposals to avoid contradictions between eIDAS 2.0 and SSI." In Open Identity Summit 2022, 2022.

[SLW21] Stalla-Bourdillon, S., Laura, C., & Alexsis, W. "Fostering trustworthy data sharing: Establishing data foundations in practice." Data & Policy 3 (2021).

Economically Viable Identity Ecosystems: Value Capture and Market Strategies

Michael Kubach[1], Heiko Roßnagel[1]

Abstract: Prevailing digital identity solutions are facing widespread dissatisfaction, prompting political and business stakeholders to advocate for the development of novel identity (ID) ecosystems. This paper diverges from the traditional focus on the usability, security, and privacy shortcomings of current solutions, directing attention instead to the economic dimensions that are critical for the successful adoption of digital identity management (IdM) systems. The analysis extends beyond the incentives for end-user adoption, considering the roles, motivations, and ability of other key stakeholders to capture value through the ecosystem, particularly service providers, who are anticipated to be the primary financial contributors to ID services. This examination leads to the pivotal inquiry of whether a market for digital identities can materialize and what strategies for market entry could be viable, especially in scenarios involving public sector participation.

Keywords: adoption, digital identity, digital ecosystems, eID, eIDAS 2.0, identity ecosystems, SSI

1 Introduction

Technically powerful and privacy-friendly solutions for digital identities (IDs) have existed for some time. Solutions based on public key infrastructures (PKIs) are established and powerful, and attribute-based credentials have been demonstrated as being able to achieve the highest standards of privacy protection [RCS15]. The development of privacy-friendly IdM solutions and attribute-based credentials has even been supported by large IT companies such as Microsoft (Cardspace [Mi09], U-Prove [Mi14] and IBM [CH02]. With the European eIDAS regulation (Regulation No. 910/2014, in force since 2016) and the Trust Services Act (VDG, since 2017), there is an established legal framework for privacy-friendly and secure digital identities. Nevertheless, a lack of widespread, secure, on a European level interoperable, and easy-to-use digital identity solution is widely recognized. This is considered a major hurdle for the digitalization of public and private organizational processes in Europe [EC22], and the development of corresponding solutions is associated with great development potential for entire economies [Wh19] and correspondingly significant growth potential for the identity management market [Ma23].

In fact, digital identities are already used by the majority of the population on a daily basis, for example for social media, online shopping and online banking. However, ID silos and big tech IDs (Facebook login, Google login, Apple ID, Amazon account) currently dominate widespread practical use. Despite, or perhaps because of, their widespread use

[1] Fraunhofer IAO, Nobelstr. 12, 70569 Stuttgart firstname.lastname@iao.fraunhofer.de

doi:10.18420/OID2024_02

by end users and service providers, these platform IDs are heavily criticized. On the one hand, this criticism comes from a privacy and IT security perspective as these identity providers are able to track their users' login behavior and use the data for their own purposes. Moreover, this information could be of interest to potential attackers [SC22, TA19]. On the other hand, in view of their broad success with users and service providers, the market power of these platform identity providers and their central position in the digital value chain, is increasingly seen as a threat to European (digital) sovereignty [Gi23, TCL23]. For use cases that require a high level of assurance (i.e. banking, healthcare), procedures such as photo or video identification based on an analog governmental ID are widely used, but their security is controversial [Ts22], user-friendliness is limited and individual identifications are rather expensive for the service providers.

In view of these challenges, new approaches emerged. Initially based on blockchain and distributed ledger technologies (DLT), they combined these hype technologies with apps to store and manage identity information on smartphones in so-called wallets. Over time being more strongly marketed under the term self-sovereign identities (SSI), these approaches have received plenty of attention in the last 3-5 years. They promise to create a new paradigm for digital identities and solve the aforementioned "identity crisis" [TA19, p.19] by focusing on users and the protection of their data. Accordingly, the European Commission is currently developing the eIDAS 2.0 regulation with EU Digital Identity Wallets [Ei23]. In addition, countless other start-ups, projects and initiatives, as well as large IT companies, are active in this field - overviews can be found at [PGH21, SN21, TCL23]. However, these approaches also face the challenge of having to establish themselves on the market (many are set up as public private partnerships or driven solely by private companies). This means that they must assert themselves against established technologies and build economically viable ecosystems[2]. This has not yet happened.

Hence, instead of adding to the extensive body of literature focusing on potential deficits in the usability, security, and data protection of existing solutions and discussing potential technical architectures, this paper looks at the economic aspects that are just as important for the spread of digital identity management solutions. To this end, chapter two highlights key principles for the adoption of ID solutions and the operation of an ID ecosystem. Chapter three then analyzes value chains for ID solutions. To this end, this study also draws on results of a qualitative survey service providers. Building on these elements, the fourth chapter discusses strategies for building economically viable ID ecosystems.

This article thus expands the perspective on digital identities from a functional-technical one with a strong emphasis on security and privacy to include considerations of the economic success factors. The term digital identities, hereinafter referred to as digital IDs, is deliberately defined broadly. It includes identity data and other verified and unverified proofs and attribute certificates. The article attempts to answer the question of how an

[2] The value proposition of an ID system cannot be achieved by a single company, but only results from the interaction of several organizations and users in an ecosystem and thus corresponds roughly to the characterization of an innovation ecosystem in [Ta20].

economically viable identity ecosystem can emerge and whether this can be based on a market for digital IDs – now or in the foreseeable future. The article builds on established theoretical concepts such as diffusion and adoption theory and empirical results from a qualitative survey. Developers of identity management solutions, as well as (potential) users and political decision-makers, are thus provided with a more holistic, theoretically, and empirically supported basis for their investment and design decisions regarding digital identities as a fundamental digitalization technology.

2 Basic market structure for digital IDs

To ensure that the ID technologies currently in development realize the potential benefits highlighted in the introduction, and effectively safeguard the digital sovereignty of individuals and the European economy, widespread adoption is essential. For such adoption to occur organically through market forces rather than by mandatory government implementation, it is crucial to understand and address the unique structure and dynamics of the digital ID market.

As demonstrated in previous studies on federated identities [ZR12] and Self-Sovereign Identity (SSI) [KS21], the market entry of ID solutions is particularly challenging because it operates as a multi-sided market characterized by network effects. This type of market caters to at least two distinct customer groups whose needs and actions are deeply interrelated [Ev03]. An ID ecosystem, therefore, comprises a minimum of users and service providers. Service providers are organizations that rely on digital IDs to offer their services, while users provide their digital IDs to access these services. Additionally, the ecosystem includes various other participants and stakeholders such as ID providers, credential issuers, trust service providers, technology providers, standardization bodies, and governmental entities. The composition of these actors can vary based on the specific ID solution and regulatory environment. The value derived by one participant group within the ID ecosystem is contingent upon the engagement of the other group, leading to network effects that create a positive feedback loop: the more service providers that adopt an ID solution, the greater its appeal to users, and conversely, increased user adoption enhances the solution's value for service providers. This interdependence in adoption dynamics leads to a classic *chicken-and-egg conundrum*: an ID ecosystem with scant participation from service providers holds little appeal for users, despite having the foundational capabilities for utilization. Consequently, users are hesitant to make any form of (intangible) investment, such as adapting to new interaction paradigms, downloading specific software, or acquiring new hardware[3].

On the other hand, if an ID solution has few users, the motivation for service providers to bear the investment risks associated with its implementation remains minimal. The low user reach through such a solution diminishes the prospects for recouping the initial

[3] An example is the German national ID card with eID function. It is widespread, but hardly used in practice.

investments required for setup and ongoing operations. Consequently, a sustainable ID ecosystem necessitates an ID solution that is broadly adopted by users and backed by a multitude of service providers. However, due to the network effects, the initial barrier to establishing this ecosystem is substantial. To surmount this barrier, various strategies can be employed to catalyze rapid ecosystem expansion through positive feedback loops.

As previously mentioned, an ID ecosystem encompasses more than just users and service providers. Therefore, it is critical to evaluate the value exchange not solely between these two groups but also considering the broader ecosystem. Additional participants develop and offer ID technology and associated services. In a traditional “centralized” ID ecosystem featuring a single ID provider that supplies the necessary components to both users and service providers, a triangular relationship emerges (user - service provider - ID provider) [ZR12]. Contrastingly, in decentralized ID ecosystems, various entities develop components (such as wallets, agents), provide them, and perform certain operational roles (like trust services). Nevertheless, average users or small service providers with limited IT capabilities may not be able to grasp this decentralized composition with various collaborating entities. Instead, they might be under the impression of interacting with a unified ID system, i.e., a specific SSI wallet from manufacturer X, that stands in for the entire system. Therefore, at this juncture, adopting the simplified perspective of a triangular relationship (user - service provider - ID system) is deemed reasonable.

The relationships between users and the ID system, as well as between service providers and the ID system, are predominantly founded on trust. Users might be reluctant to entrust their personal identity data to an ID system if they lack confidence in its (or the underlying technology's and actors') ability to manage or safeguard their data in a manner aligned with their interests. Similarly, a service provider must trust in the ID system's capacity to consistently provide data of the requisite quality[4], though regulated trust levels, such as those defined by eIDAS, may only be pertinent to certain applications. Consequently, research and development in digital IDs are intensively directed towards enhancing the security and privacy of these solutions. However, economic considerations are equally critical. The various entities involved in the ID system expect to be compensated for their services to offset costs or generate profits. Such compensation might stem from the willingness of certain market participants to pay. In theory, both users and service providers might finance commercial ID services. For instance, service providers currently incur costs for video identification methods. Users, too, might be willing to pay for certain identity features (like credit cards), provided these services offer added value over the free systems that dominate the current market. The feasibility of delivering such value and the potential for commercial viability are further explored in the subsequent chapter.

[4] Users and service providers must also be able to trust that the ID system will work reliably when they need it.

3 Analysis of the value chains of ID ecosystems

Building on the theoretical and conceptual analysis of the digital ID market presented in the previous chapter, the following section will delve into the value creation mechanisms of ID ecosystems. This aims to feed into the development of strategies for economically sustainable ID ecosystems, which will be elaborated upon in the subsequent chapter.

3.1 Conceptual framework of the analysis

The value chain analysis builds on work of [Ta20], who introduced the Ecosystem Pie Model (EPM) — a framework that can be utilized to scrutinize value capture and interrelations within innovation ecosystems. An innovation ecosystem according to [Ta20] is an environment where no single firm possesses all necessary resources to independently develop and commercialize a complex offering from inception to market launch. Consequently, firms must collaborate with others within their ecosystem to collectively construct a shared value proposition for the ecosystem. Given the parallels with ID ecosystems, particularly in the context of the multi-sided market discussed earlier, the EPM is deemed appropriate for analyzing the value chains of ID solutions. It is pertinent to note that [Ta20] do not differentiate between actors as organizations and the various roles these organizations might play within an ecosystem. For clarity, subsequent references to "actors" by [Ts20] should be understood as pertaining to the specific roles that organizations or individuals fulfill within the ecosystem.

For the EPM, three constructs on ecosystem level are identified (1. value proposition, 2. user segments, 3. actors) and six on actor level (1. resources, 2. activities, 3. value creation, 4. value capture, 5. risk, 6. dependencies). The ID ecosystem was modeled, and the constructs analyzed for the research project that forms the framework of this study.[5] Due to the limited space available, the focus at this point must be on the construct of value capture on actor level. While the ecosystem creates value for the end users through interactions, each ecosystem actor strives for individual benefits and must be able to acquire them [Te86]. The value capture construct represents how, what type and how much value created by the ecosystem is captured by an actor – one of the main motivations for joining an ecosystem. Furthermore, actors expect to receive a fair share of the value created. The value does not necessarily have to be of a monetary nature if an ecosystem-external monetization (e.g. of reputation, growth, knowledge) is possible [Ta20]. The focus on the construct of value capture as one of the main motivations for joining an ecosystem therefore appears to be justifiable at this point.

In the subsequent analysis, we will consider the actors[6] within a prototypical ID ecosystem and their capacity for value capture. We have aligned this ecosystem with the EUDI Wallet Ecosystem as described in the Architecture and Reference Framework [Ei23]. The key

[5] ONCE, funded by the Federal Ministry for Economic Affairs and Climate Protection (www.once-identity.de).
[6] A particular organization in the ecosystem can simultaneously represent different actors.

participants identified (refer also to Figure 1) include: (1) *Issuers and data sources for digital IDs*, (2) *Service Providers* (Relying Parties) who use digital IDs for their service provision, (3) *Infrastructure Providers* who develop, provide and administer infrastructure services for the ID ecosystem (e.g. technical middlewares/gateways into the ID ecosystem, trusted lists, ID lifecycle services), or who control the entire ecosystem and represent it externally, (4) *End Users* or citizens and, finally, (5) *User Systems* (this work focuses on smartphone wallet applications (wallet instance) for end users), through which end users manage their digital IDs and access desired services via the ID ecosystem.

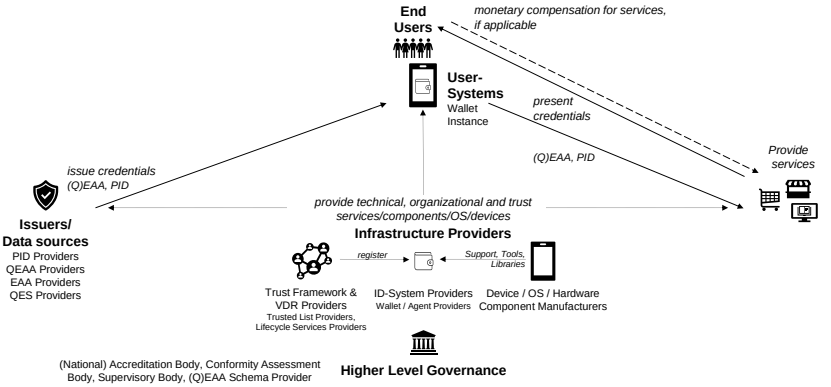


Fig. 1: Actors and relationships in an ID ecosystem (aligned with EUDI Wallet roles)

3.2 Empirical basis of the analysis

The following analyses are based on qualitative interviews conducted with 23 organizations with the potential to function as issuers and data sources, service providers, and, in some instances, as infrastructure service providers within ID ecosystems. The sample comprises eight municipalities, four other public administration entities (e.g., at state level), three IT service providers specializing in public administration, four organizations from the mobility sector, and four enterprises from the hotel and tourism sector. Participants were recruited in the context of a research project focusing on digital IDs; hence, the sample is not random but intentionally diverse. The goal was to encompass organizations across sectors that had at least some initial exposure to digital IDs and diverse security and trust requirements, to provide a broad perspective on the subject. The qualitative methodology was specifically chosen to facilitate in-depth discussion of the relatively novel and rapidly evolving concepts and allow detailed inquiry where needed.

The guided interviews were conducted via video calls from September to November 2021, with the recordings made after receiving the interviewees' consent. Recordings were transcribed, coded with the use of MAXQDA software, and analyzed. The interview guide covered topics such as the current state of digital ID usage within the organizations, the

driving factors and obstacles associated with the adoption of digital IDs, as well as key information about the organizations and the interviewees themselves. This article focuses solely on a single facet of the survey's findings. The interviews varied in length, lasting from 30 to 90 minutes. The interviewees held various positions within their respective organizations, ranging from CEO to identity management specialists. The critical criterion was not the specific title held, but rather that the interviewees had a deep understanding of their organizations and the relevant areas of operation, qualifying them as experts (key informants) for their organizations [Ho12]. An overview of the interviewees and their organizations is provided in Table 1 (available at https://gitlab.cc-asp.fraunhofer.de/mkubach/once/-/blob/master/EVIDE_Annex.pdf).

3.3 Results of the analysis

Concerning *issuers and data sources for digital IDs*, it is important to differentiate between state institutions, such as municipal administrations, and private sector issuers. Public administration actions, including the issuance of certain digital IDs, are primarily dictated by legislative requirements. Thus, the potential for value capture is a secondary consideration for their participation in the ID ecosystem. Consequently, a political mandate or legal framework could compel state issuers to distribute digital IDs for a specific ID ecosystem to foster digitization efforts. For private sector issuers and data sources, direct added value from joining the ecosystem is essential. This benefit must exceed the costs associated with joining (such as investments for process modifications, purchasing or updating software) and the issuance of digital IDs, including ensuring the accuracy of ID data. However, a clear value capture mechanism for the issuance of digital IDs is not evident. It is unlikely that end customers would be willing to pay for secure IDs [Ro14]. Compensating original issuers for the utilization of digital IDs by third-party service providers would necessitate tracking every ID use, which would infringe upon user privacy and contradict the core principles of Self-Sovereign Identity (SSI), rendering this unfeasible in an SSI ecosystem. Moreover, there seems to be minimal willingness from service providers to pay, making this avenue for monetization unsuitable. Thus, issuers and data sources are likely limited to capturing intangible values, such as leveraging ID data for their own service offerings, which could then enable value capture when they assume the role of a service provider (discussed in the following point).

Similar to issuers, a key distinction must be made between state and private sector entities when considering *service providers*. State service providers adhere to political and legal mandates. For them, the potential for efficiency gains through the digitalization of administrative procedures, which can be realized by an effective ID ecosystem, stands out as a value capture mechanism. However, some municipalities indicated in the interviews that their constrained budgets may prevent them from financially affording the necessary investments before such improvements can be realized. For private sector service providers, gains in efficiency due to process digitalization are also desirable. Additional benefits include the potential to increase revenue by introducing new digital products and services, cost reductions compared to alternative identification methods (such as video ID,

assuming an equivalent level of trust is necessary), diminished process costs by preventing fraud and misinformation, and increased sales through accelerated customer onboarding and higher conversion rates. Some service providers expressed a basic readiness to pay for these intangible benefits, depending on the specific use case, yet they also demonstrated significant price sensitivity. The use of platform-based identities (like login via Google or Facebook), which are free of charge and facilitate user data collection to enhance services, serves as a point of comparison in the interviews. Fees associated with credit card transactions or PayPal, which are often reluctantly accepted, continue to deter many businesses from offering these payment options. Moreover, in many scenarios, the mere receipt of payment suffices; further verified identity is unnecessary and thus unlikely to be compensated. However, if costly ID methods (i.e., Video ID) can be substituted, there is a discernible willingness to pay since the costs are apparent and can be justified.

Depending on the design of a specific ecosystem, various types of *Infrastructure Service Providers (ISSPs)* can be identified. For the sake of simplicity, we distinguish between two primary roles: the technical ISSP and the ecosystem orchestrator. A technical ISSP is responsible for developing technical components such as middleware and connectors that enable participation in the ecosystem. It may also operate these components on behalf of ecosystem participants. An ecosystem may host multiple technical ISSPs, in a coopetition dynamic or specializing in different tasks (e.g., developing connectors for state eIDs, creating SSI components, operating web agents, etc.). The ecosystem orchestrator, as described by [LHG22], takes on the overarching management of the ecosystem, including the responsibility for its rule set (trust framework), which encompasses common norms and standards, procedures, rules, and principles. Its key role is to establish and maintain trust within the ecosystem, which also entails managing the ecosystem's economic foundation. Moreover, the orchestrator may provide certain technical infrastructure services (e.g. directory services) and act as the external representative of the ecosystem, serving as the main point of contact for political entities or prospective new participants.

Regarding value capture, technical ISSPs can capitalize on their participation in the ecosystem through various revenue streams. They have the potential to generate license income from the components they develop, as well as fees for integrating these components into the systems of other ecosystem participants and for operating components (e.g., web agents). A range of pricing models can be applied, including a flat base fee, transaction-based fees, fees per user, etc. For the ecosystem orchestrator, value capture opportunities may arise through certification fees (for instance, for component manufacturers, or perhaps for service providers requiring ID data at an elevated security level) associated with joining the ecosystem, as well as ongoing membership fees. If the orchestrator also provides technical infrastructure services, usage fees could be levied. However, it is crucial to recognize that any fees imposed for joining or participating in the ecosystem can act as a deterrent to adoption. As a result, the suitability of these value capture strategies, particularly in the initial stages of ecosystem development, is debatable. High entry or usage costs could hinder its growth and widespread acceptance.

End users, or citizens, engage with digital services and utilize the ID ecosystem when they

perceive a benefit in doing so. The value capture for them, which is primarily intangible, can come from a variety of advantages such as broad applicability, a high level of trust in the ecosystem, increased convenience, or enhanced control over their personal data. Direct monetary gains are not a necessity for end users. On the flip side, it is equally improbable that end users will be prepared to pay for ID systems, meaning that financing an ID ecosystem through direct payments from end users is not a viable expectation [Ro14]. The organizations involved in the study do not foresee a willingness among end users to bear costs for such systems in the medium term either.

In the context of *user systems*, this analysis focuses on smartphone wallet applications. Since smartphones, their operating systems, and specialized secure hardware components are not custom-built for the ID ecosystems discussed in this article, the emphasis is on wallet manufacturers that develop and supply these applications specifically for a given ID ecosystem. The potential for value capture through fees charged to the end user for using the wallet is minimal due to the general reluctance to pay and induced barriers to adoption. Consistent with this, one wallet manufacturer interviewed for this article has chosen to offer its wallet to end customers free of charge. Consequently, alternative value capture methods must be explored. Wallet manufacturers may consider monetizing through branding or custom adaptations tailored for certain issuers, data sources, and service providers who could then compensate for these services. This could be coupled with consultancy services on a fee basis pertaining to the technology and ecosystem. However, the sustainability of a wallet application as a standalone product for the manufacturer is questionable, in particular if it has to compete with free wallets included in smartphone systems and free EUDI wallets provided by EU member states. Hence, it may be necessary for the manufacturer to also provide additional technical components as a technical ISSP, where the prospects for value capture might be more substantial.

Actor	**Value**	**Willingness to pay**
Issuer and data source	only as SP	none
Service Provider	medium	low - medium
Infrastructure Service Provider	unclear	none
End User	low - middle	low
User Systems	unclear	none

Table 2: Actors in the ID ecosystem: value capture and willingness to pay

The value chain analysis, focusing on value capture (summarized in Table 2), reveals that this aspect poses a significant challenge for ID ecosystems. Service providers appear to have the best chances of capturing value, even though it is predominantly intangible. End users can also reap clear benefits from a well-functioning ID ecosystem. If issuers and data sources also participate as service providers within the ecosystem, they can tap into similar value capture potentials. Otherwise, these actors must consider why they would invest in joining the ecosystem and potentially assume liability for the accuracy of the digital IDs they issue. For other participants in the ecosystem, the feasibility of the value capture methods proposed – primarily through various fees – is uncertain, especially

without erecting barriers to adoption that are too high in light of a limited willingness to pay. Consequently, the next chapter will explore potential strategies for fostering economically sustainable ID ecosystems, taking these considerations into account.

4 Strategies to develop economically viable ID ecosystems

As delineated in the preceding chapter, the scope for value capture and the potential for willingness to pay among ecosystem participants are notably constrained. The likelihood of value capture predominantly exists between service providers and their users. Other ecosystem participants, assuming they do not assume the role of service providers, need financial compensation. This, in turn, hinges on their willingness to pay, which, as identified in Chapter 3, is markedly scarce. In light of these factors, the emergence of a self-sustaining market for SSI driven purely by supply and demand seems implausible under the present conditions. Nonetheless, given the compelling impetus to forge such an ecosystem, as explicated in the introduction, it is prudent to examine a spectrum of viable approaches for its inception and augmentation, inclusive of public sector engagement. Drawing upon the insights of [OS06], we consider several strategic alternatives.

An only minimal intrusive strategy might involve launching an informational campaign to highlight the benefits of secure digital identities, thereby bolstering demand if users recognize these benefits as tangible added value. However, historical evidence suggests that the impact of such campaigns tends to be marginal. At the opposite end of the intervention spectrum lies forced adoption, representing the most aggressive form of market intervention. Forced adoption could be universally enforced, mandating the use of secure identities under penalty of fines, or it might target specific sectors or applications through regulatory mandates that stipulate a required trust level. In highly regulated industries, such as gambling and financial services, similar regulations have already led to the widespread adoption of video identification procedures. Should these no longer be considered sufficiently secure by regulation, or if more cost-effective alternatives become available, it could catalyze a shift in demand towards other secure identity solutions. Intermediate measures include the bundling of secure digital identities with complementary goods or subsidization. Bundling necessitates strong partnerships with service providers and their offerings. Subsidization, on the other hand, would depend on the public sector's readiness to support the operating costs of the identity ecosystem until a self-sufficient market materializes (or even permanently). The duration and certainty of such a market's development are currently unpredictable. Given the aforementioned challenges in value capture, it appears plausible that without (continuous) market intervention, the envisioned ID ecosystems may struggle to achieve economic viability.

5 Conclusion

The current development of new ID solutions predominantly concentrates on technical

dimensions such as data security, privacy, and usability. Nevertheless, for these solutions to deliver their full potential, it is imperative to also scrutinize the digital ID market and its complex economic interrelationships. This paper's analyses have elucidated the intricate nature of the multi-sided market for digital IDs, particularly the challenges associated with value capture by entities within the ID ecosystem. The potential for value capture is confined to a select group of stakeholders, and the imposition of fees for ecosystem participation may present prohibitive barriers due to the generally low willingness to pay. Our exploration of strategies to foster economically sustainable ID ecosystems reveals several avenues, yet uncertainty persists regarding the feasibility of creating and maintaining such ecosystems through market forces alone, absent significant government intervention or investment. If this assessment holds and is widely acknowledged, then further discourse is warranted. From a societal perspective, the ID ecosystem might be considered an essential piece of digital infrastructure. Consequently, it would be a policy decision to determine whether these digital infrastructures, if deemed socially beneficial, should receive ongoing state subsidies. In light of these considerations, it is crucial to engage in a measured discourse that not only addresses the technical design but also the economic dynamics and the state's role in the evolution of this infrastructure.

6 Literature

[CH02] Camenisch, J.; Van Herreweghen, E.: Design and Implementation of the Idemix Anonymous Credential System. In: Atluri, V. (ed.): Proceedings of the 9th ACM Conference on Computer and Communications Security, CCS 2002, Washington, DC, USA, November 18-22, 2002: ACM, 2002, pp. 21–30.

[Ei23] eIDAS Expert Group: The Common Union Toolbox for a Coordinated Approach Towards a European Digital Identity Framework - The European Digital Identity Wallet Architecture and Reference Framework, github.com/eu-digital-identity-wallet/eudi-doc-architecture-and-reference-framework/blob/main/docs/arf.md. accessed 2024-03-16. — Version 1.3.0.

[EC22] European Commission: European Digital Identity, ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-digital-identity_en. accessed 2022-08-19. — Digital Identity for all Europeans.

[Ev03] Evans, D. S.: Some Empirical Aspects of Multi-sided Platform Industries. In: Review of Network Economics 2/3, pp. 191–209, 2003.

[Gi23] Giannopoulou, A.: Digital Identity Infrastructures: a Critical Approach of Self-Sovereign Identity. In: Digital Society 2/2, pp. 18, 2023.

[Ho12] Homburg, C.; Klarmann, M.; Reimann, M.; Schilke, O.: What Drives Key Informant Accuracy? In: Journal of Marketing Research 49/4, pp. 594–608, 2012.

[KS21] Kubach, M.; Sellung, R.: On the Market for Self-Sovereign Identity: Structure and Stakeholders. In (Roßnagel, H.; Schunck, C. H.; Mödersheim, S., ed.): Open Identity Summit 2021. Bonn: Gesellschaft für Informatik e.V., 2021, pp. 143–154.

[LHG22] Lingens, B.; Huber, F.; Gassmann, O.: Loner or team player: How firms allocate orchestrator tasks amongst ecosystem actors. European Management Journal 40/4, pp. 559–571, 2022.

[Ma23] MarketsandMarkets: Digital Identity Solutions Market Size, Share and Global Market Forecast to 2028 | TC 7537, www.marketsandmarkets.com/Market-Reports/digital-identity-solutions-market-247527694.html. accessed 2024-02-13.

[Mi09] Microsoft: Windows CardSpace, learn.microsoft.com/en-us/previous-versions/dotnet/netframework-3.5/ms733090(v=vs.90). accessed 2023-01-09.

[Mi14] Microsoft: U-Prove, www.microsoft.com/en-us/research/project/u-prove/. accessed 2024-01-05.

[OS06] Ozment, A.; Schechter, S.E.: Bootstrapping the Adoption of Internet Security Protocols. In: WEIS Proceedings. Cambridge, 2006, pp. 1–19.

[PGH21] Pöhn, D.; Grabatin, M.; Hommel, W.: eID and Self-Sovereign Identity Usage: An Overview. In: Electronics 10/22, pp. 2811, 2021.

[RCS15] Rannenberg, K.; Camenisch, J.; Sabouri, A.: Attribute-based credentials for trust. In: Identity in the Information Society, Springer, 2015.

[Ro14] Roßnagel, H.; Zibuschka, J.; Hinz, O.; Muntermann, J.: Users' willingness to pay for web identity management systems. In: European Journal of Information Systems 23/1, pp. 36–50, 2014.

[SC22] Schardong, F.; Custódio, R.: Self-Sovereign Identity: A Systematic Review, Mapping and Taxonomy. In: Sensors 22/15, pp. 5641, 2022.

[SN21] Soltani, R.; Nguyen, U. T.; An, Aijun: A Survey of Self-Sovereign Identity Ecosystem. In (Galdi, C. ed.): Security and Communication Networks, pp. 1–26, 2021.

[TCL23] Tan, K. L.; Chi, C.; Lam, K.: Survey on Digital Sovereignty and Identity: From Digitization to Digitalization. In: ACM Computing Surveys 56/3, pp. 61:1-61:36, 2023.

[Te86] Teece, D.J.: Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. In: Research Policy 16/6, pp. 285–305, 1986.

[TA19] Toth, K. C.; Anderson-Priddy, A.: Self-Sovereign Digital Identity: A Paradigm Shift for Identity. In: IEEE Security & Privacy 17/3, pp. 17–27, 2019.

[Ts22] Tschirsich, M.: Praktischer Angriff auf Video-Ident, Version 1.2: Chaos Computer Club, 2022.

[Ta20] Talmar, M.; Walrave, B.; Podoynitsyna, K.; Holmström, J.; Romme, A. Georges L.: Mapping, analyzing and designing innovation ecosystems: The Ecosystem Pie Model. In: Long Range Planning 53/4, pp. 101850, 2020.

[Wh19] White, O.; Madgavkar, A.; Manyika, J.; Mahajan, D.; Bughin, J.; McCarthy, M.; Sperling, O.: Digital Identification: A Key to Inclusive Growth: McKinsey Global Institute, 2019.

[ZR12] Zibuschka, J.; Roßnagel, H.: Stakeholder Economics of Identity Management Infrastructures for the Web. In: Proceedings of the 17th Nordic Workshop on Secure IT Systems (NordSec 2012). Karlskrone, Sweden, 2012.

Evaluating the Good Practices of User Experience for Mobile and Electronic Governmental Services

Rachelle Sellung [1], Lennart Kiss [2]

Abstract: With digitalization in the focus of governmental services for citizens, mobile services need to provide users with a good user experience and usability to encourage greater user acceptance. [SHB22] defined ten Good Practices to support greater User Experience and Usability for Mobile Governmental Services. These Good Practices are evaluated and validated in this paper by a User Study that consisted of Qualitative and Quantitative results. Good practices of user experience can help impact and support the integration of the basic user experience needs into the technical development processes for future digitalization of mobile governmental services.

Keywords: User Experience, Mobile Services, Mobile Governmental Services, Electronic Governmental Services, Usability, Digitalization

1 Introduction

Governments have realized the importance of digitalization in the last decades. However, most are still facing key challenges regarding adoption and user experience of these digitalized services. When developing or digitalizing services, electronic and governmental services have fallen short on requirements in security, privacy, but also largely user friendliness [CS19]. [KR19] have pointed out that technical requirements take priority over user experience requirements throughout this process. In order to create services that will be adopted and used, it is important to ensure that users see the value in them and that it is as familiar and as easy as possible. Building on previous research, this paper will present a user study evaluation of the Good Practices of User Experience and Design Research described in [SHB22]. The ten Good Practices presented in [SHB22] were based on desk research of user experience and design research that was done on various mobile government and electronic government services around the world. It is implied that when considering these good practices that the user experience and usability of the services will improve. This paper attempts to test this hypothesis on whether the Good Practices defined in [SHB22] are correlated to a positively perceived user experience and usability.

[1] Fraunhofer IAO, Identity Management, Nobelstrasse 12, Stuttgart, 70569, rachelle.sellung@iao.fraunhofer.de , https://orcid.org/ 0000-0003-1235-030X
[2] University of Stuttgart IAT, Identity Management, Allmandring 35, Stuttgart, 70569, lennart.kiss@iat.uni-stuttgart.de, https://orcid.org/0009-0000-8839-5470

doi:10.18420/OID2024_03

The results of the user experience study presented consist of three electronic governmental (eGovernmental) and mobile governmental (mGovernmental) service pilots that were developed in the mGov4EU project, which is funded by the European Commission. By considering the data and feedback from this study, this will be considered in validating the significance of key user experience and design practices by investigating their impact on overall user satisfaction and comprehension. By analyzing the user feedback of the pilots, we test for a correlation between high scores on the Good Practices and improved user experience. The paper is structured in the following manner. The second and third section highlights related work and methodology. The fourth section presents the results of the user studies and the significance between the user experience, Good Practices and overall user satisfaction for the three pilots that were tested. The fifth section discusses the results and possible interpretations. The sixth and seventh sections summarizes limitations and potentials for future work and conclusions.

2 Related work

This paper evaluates the Good Practices of User Experience and Design that were established in [SHB22]. They described ten Good Practices that were based on desk research in the field of user experience and design research of mobile government and electronic government research. The following Table 1 summarizes the Good Practices listed in [SHB22].

Learnability	The ability to easily use, remember, and learn a service.
Minimalistic & Simple Design	Improves accessibility of user groups with varying capabilities.
Language	Simple Language that is understandable by a wide user group
User Readable Terminology	Terms used in buttons, labels, messages, should be simple, familiar, and understandable for users with little technical knowledge.
Help and Feedback	Resources for users to refer to when they have questions or need technical support.
Error Handling	Errors that disrupt the users experience should provide information on what happened and what the user can do next.
Search and Filter	Ability to search or filter through the server with key terms through the product.
Operability	Users can use the product or service with a high quality on any size or device type.

Placement of Information	Having a straightforward layout of information and clear instructions and functionality.
Use of Colors	The "look and feel" must be appealing to users to help impose trust, positivity, and consistency in the service.

Table 1 Good Practices listed in [SHB22]

These Good Practices are results of the desk research done by [SHB22]. [IW18] provided insights to learnability with their studies and research on how learnability can play a key role in improving user friendliness and eventually user acceptance. [KR19], [IW18], [CLH20], [SF20], [KMM18], [LUN16] supported [SHB22] second Good Practice on minimalistic and simple design, where their work contributed to understanding the challenge of providing a service to a user group with a wide range of needs. For "Help and Feedback" Good Practice, [SHB22] support it with examples from [SF20], [HB11] and their input on design patterns of how to communicate feedback to users in a user friendly way. [HB11] laid the foundation for the "Search and Filter" Good Practice with support of the findings in [SF20] for the convenience of the users in finding different functions. Research conducted by [IW18], [IW17] highlights operability and the use of mobile devices and user expectations. For "Placement of information", this was summarized by research conducted by [KR19], [IW18], [CLH20], [SF20], [IW17]. It highlights the importance of placing the information in a straightforward and functional arrangement. Lastly, the "Use of Color" is supported by research by [IW18], [CLH20], [KMM18], where it is emphasized how color can negatively impact users perception or user experience.

3 Methodology

The study from which the data originates was conducted as a part of the mGov4EU project [mG24]. A mixed-methods approach was employed to assess the user experience of three separate pilots. Although the pilots share the general technology used, they differ in their respective use cases. User stories and tasks were created for each pilot. Tab. 2 offers an overview of the pilots' use cases and the associated user stories.

Pilot Name	Use Case Description	User Story
I-Voting Pilot	Integrates mGov4EU identification mechanism and SDG-layer into the online voting system. Allows users to vote remotely and be authenticated.	User is a student voting on the usage of a donation for extracurricular activities. The user must prove their identity in the university voting portal, verify active enrollment by uploading a certificate, and cast a vote for the preferred use of a donation.

Smart Mobility Pilot	Applies mGov4EU infrastructure to state-subsidized mobility services. User uses the pilot to confirm eligibility for a discounted taxi ride.	The user, either a German citizen or an Austrian citizen in Germany, aims to receive a discounted taxi ride. Tasks involve selecting a region, proving identity, and, for Austrian citizens, obtaining proof of address. After completing these tasks, the user test concludes.
eSignature Pilot	Focuses on creating advanced and qualified electronic signatures meeting eIDAS-Regulations. Tests the eSignature building block in mGov4EU.	The user, a dual citizen (German/Austrian), has separate tasks: signing a code of conduct agreement for the Rugby Club using a German ID and uploading a business contract with signatures for the "GovAssist" project using an Austrian ID. The user needs to add partners and observers for the business contract and must complete these tasks to manage their commitments efficiently.

Tab. 2 Pilot use case and user story description [mG24]

The user study of the projects' pilots included key performance indicators (KPIs) (e.g., task success rate, time per task, ease of use), interview questions along their experience of different functions and tasks evaluated of the pilot and standardized questionnaires like the System Usability Scale (SUS) [Sy24] and the User Experience Questionnaire (UEQ) [LHS08]. The data originates from a set of three user tests that followed the same methodology. The pilots were tested with at least ten participants per pilot up to a maximum of 14 for one of them. Additionally, each pilot had a set of user tasks assigned to them. After task completion, participants were inquired about difficulties they encountered, their personal preferences and suggestions that might enhance user experience. After the user has completed their tasks, where they tested different functions or use cases from the pilot, they continued to the quantitative part of the user study. This included filling out the SUS, UEQ questionnaires and a set of post-test questions inquired about general feedback and impressions of the pilot.

In order to validate the Good Practices established in [SHB22], the following process was taken. Each participant's feedback from the user study was carefully examined and received a valence score for each Good Practice. The scoring system was based on the amount and severity of positive and negative feedback that fit a respective Good Practice. Scores were assigned on a five-point Likert scale ranging from "strongly negative (-2)" to "strongly positive (+2)". Therefore, if a user mentioned e.g., the mismatch of colors that were used or missing help structures, the score would be in the negatives, whereas if the feedback was positive, the respective Good Practice score would also be positive. However, when a Good Practice was not touched upon the score

would be zero indicating a neutral score. To validate the thesis the scores were checked for correlation with two established usability and user experiences questionnaires, namely the SUS and the UEQ. The primary objective of this validation was to investigate the relationship between adherence to Good Practices and perceived usability and user experience.

4 User Experience Study Comparison

This section presents the key findings and outcomes derived from the evaluation of the Good Practices and the results of the user study. The user study conducted within the mGov4EU project consisted of a sample of 34 participants, approximately evenly distributed between male and female participants. The participants were aged between 18 and 55 years, with almost 50 % having a master's degree and around 75 % describing themselves as tech-savvy.

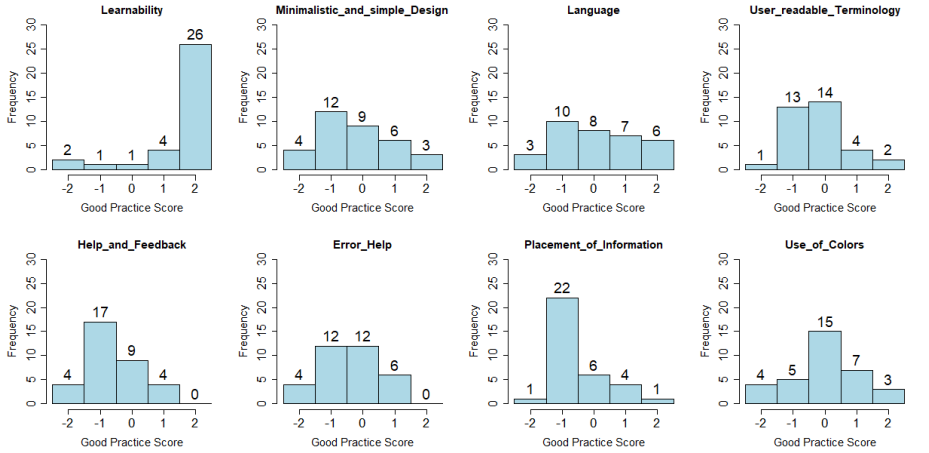


Fig. 1 Frequency of assigned scores per Good Practice (except Search and Filter, and Operability)

The qualitative data gathered from the evaluation was categorized and each participant's feedback received a rating per Good Practice following the scoring system. Fig. 1 summarizes the frequency of scores that were assigned to each Good Practice. The graphs indicate a prominent tendency towards higher scores only for the Good Practice of Learnability. At the same time, as no participant directly addressed the Good Practice of Search and Filter or Operability to an extent that would warrant assigning a weighted score, these two Good Practices were omitted from Fig. 1. However, the rest of the Good Practices were addressed extensively in the qualitative research, indicating a slight tendency towards lower scores.

4.1 Quantitative Insights on Good Practices

This case study focusses on the relationship between adherence to Good Practice and improved usability and UX. To validate this correlation, their relationship was analyzed. Fig. 2 presents the relationship between the participants' SUS score and their average Good Practice score resulting from the Good Practices rating system.

Fig. 2 Relationship between SUS Score and Good Practices Average

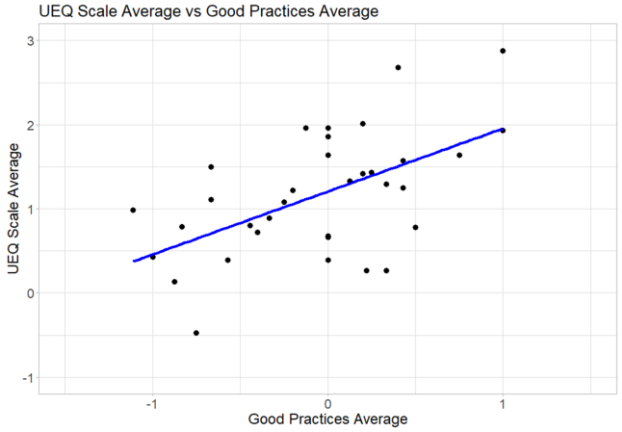


Fig. 3 Relationship between UEQ Scale Average and Good Practices Average

The correlation analysis revealed a statistically significant, strong, and positive linear relationship between the two variables ($r = .77$, $p < .001$). Therefore, lower ratings in the Good Practices category seem to correlate with diminished perceived Usability as indicated by lower SUS scores. Similar to the relationship between the SUS score and Good Practices, the average of the six UEQ subscales also indicate a positive

relationship with the Good Practice score ($r = .56$, $p < .001$) (see Fig. 3). These findings indicate a medium to strong correlation between adherence to Good Practices and improved usability and UX.

4.2 Qualitative Insights on the Good Practices

This section provides insights gathered from the qualitative feedback on which the categorization and Good Practice scores were based. Subsequently, the Good Practices will be addressed one by one to highlight their importance and further describe their impact on usability and UX.

The Learnability of a system poses an important aspect of the system's UX. It has a significant impact on how well users can grasp, execute, and recall actions within the system. In the usability evaluation, Learnability was mostly addressed by the users, when asked if they had problems executing a task and if they felt that they would be able to redo a task later on. A remarkable part of the study sample reported for almost all tasks, that they felt confident in repeating the tasks. Additionally, the straightforwardness of processes was mentioned multiple times in this context. However, users who were less confident in their comprehension of the task and the included steps often mentioned that certain breaks in their workflow threw them off. Therefore, to avoid user flow breaks that might impede Learnability, one should be aware of said breaks and either resolve them or if not applicable prepare the user for what is expected.

Adhering to the Good Practice of Minimalistic and Simple Design contributes towards the user's understanding and reduces cognitive load while interacting with a system. Users specifically addressed this Good Practice when commenting on the ease of recognizing the purpose of navigational UI elements or the clear association between icons and functionalities. However, when faced with a visually complex and unclear interface, users encountered challenges that hindered task completion. Therefore, when designing interfaces involving icons and other UI elements, one should not only keep cultural associations in mind but also carefully review the number of elements to prevent visual clutter. The Good Practice of Language relates to the use of widely understood languages for systems that cater to broad user groups. This involves abstaining from overly complicated texts or intricate wording. The goal is to offer a clear and simple way for users to understand the system, irrespective of their knowledge. It is important to note, that the adherence to this Good Practice varied tremendously between the three pilots. Nevertheless, even within a pilot group users reported different levels of understanding. This segmentation of users based on comprehension challenges highlights the critical role of clear language. At first User-Readable Terminology seems to head in a similar direction as the Good Practice of language but with some defining differences. Rather than assessing entire textual components, the focus lies on words that are typically stand-alone or with minimal context. This includes the labelling of buttons, tabs, functionalities, or simple elements like headlines. Users expressed a desire for alternative wording of certain use case specific terms, with abbreviations being

mentioned the most in this context. Difficulties arose in comprehending entire tasks and processes due to the terminology. This was especially prominent when users were faced with a decision that involved the use of a single digital gateway component (SDG-Component). The abbreviation itself was seen as the source of confusion; however, even after they were informed about its meaning, it did not contribute to their understanding. This highlights the need for terminology, that is universally understood by users irrespective of their demographics or technical proficiency. Aligning with this Good Practice, implementing features like tooltips or an info button for users to access explanations can contribute to user understanding while maintaining technical and legal integrity.

The second-lowest Good Practice score was reported for the category Help and Feedback. Help structures are responsible for assisting the user when they encounter challenges in achieving their goals. This need for assistance can be fulfilled by either means of direct interaction or software solutions that provide helpful insights. The primary goal should be to proactively prevent problems before they arise but in cases where challenges inevitably occur, users should be provided with suitable structures to resolve the situation. In the usability evaluation, the absence of feedback structures was pointed out on multiple occasions. Striking problems were found in the clarity of system status. At times it was unclear to the user whether they had completed tasks, whether the system was undergoing background processes, or in one of the worst scenarios, when users were not aware that an error had occurred. Features that were suggested by the users described the addition of progress bars into complex processes, as well as the option to obtain more information through user action (e.g., info buttons). Effective Error Handling is essential for preventing user frustration and maintaining trust and security within the application. Nevertheless, errors are often unavoidable. The way they are handled however strongly affects the users' chances to solve them. When designing error messages, it is important to include three aspects: First, the user should be made aware that an error occurred. Second, they should be informed about the nature of the error and provided with additional information that is relevant to them and the case. Third, they should receive clear and concise guidance on how to resolve the problem. Additionally, by reducing the time between the error cause and the error alert, the user can achieve a sense of agency that prominently highlights the importance of the error. The most frequent problems, that occurred in the usability evaluation can be summarized as user errors that stay undetected from the system, misunderstood, or overlooked error messages, and technical bugs without possible solutions for the user.

While the Good Practice of Search and Filter did not receive as much attention or detailed feedback as to warrant a specific valence score, users did express suggestions that fit this category. Generally, it is expected of a system that there is a way for users to search for specific information. This can be achieved by implementing either a search engine or employing filters. In one instance users reported the wish for a search function when they were asked to provide a contact email for one of the tasks. The search engine should replace the need to know the full address of known contacts. Whenever a

considerable amount of information is either presented or inquired search and filter options decrease the cognitive load a user has to deal with. Participants did encounter challenges related to the Good Practice of Operability, such as the scaling of the UI and unexpected swap of devices in one use case. However, the feedback they provided did not directly address operability. Nonetheless, developers should keep challenges related to adoption, interoperability, and various technical bugs in mind. Especially when the use case requires another device to be used, clearly informing the user about the specific processes becomes essential.

The average score for the Good Practice of Placement of Information turned out to be the lowest, indicating a higher prevalence of problems and criticism in this aspect. Strategical placement of information is critical to the user to comprehend the presented information and, consequently, to act in ways that work towards their goals. Employing recognizable formats, adding tooltips for unfamiliar words relevant to the user's actions, all while avoiding information overload, align with the principles of effective placement of information. Negative feedback additionally involved the prioritization of information in lists and not drawing the user’s attention through changes in typography. Typically, the provided text was first read when the user could not continue without further information, or an error occurred. Therefore, information should be presented in proximity to the elements it is relevant for. This enables easy association and positively impacts the user’s ability to navigate the application. The Use of Colors can influence user attention, contribute to visually appealing experiences, and support recognizability, but also negatively impact the user through inconsistent use or usage that contradicts their expectations. User comments expressed a desire for a more vibrant and engaging interface. Through highlighting and employing visual cues users are supported in understanding the significance of different elements. Furthermore, aligning the color scheme with the branding of the related provider contributes to a cohesive and branded experience, that can also spark familiarity with the service. To successfully use colors in UI design, one should keep user preferences, expectations, and general design principles in mind. The user feedback was fairly evenly distributed. Users who liked the use of colors either reported, that the use was generally a good fit for the application or that some visual cues were prominently designed due to the choice of color. However, users who disliked the use of colors reported some inconsistencies, a mismatch of colors with their expectations, or the general absence of colors in the UI.

5 Discussion

This paper aimed to test and validate the Good Practices of mobile services that were described by [SHB22]. The Good Practices described by [SHB22], were based on desk research of user experience and design research in the field of mobile and electronic governmental services. This paper took the ten Good Practices and conducted a comparison of the Good Practices and the results of a user study of governmental services.

In order to validate the Good Practices, this paper depicts how the values of the quantitative and qualitative data were transformed into a scoring system on a five-point Likert scale of the impression of the Good Practices. The quantitative data considered KPIs, Task Questions, and a General Post Questionnaire. The qualitative data was based on interview answers and comments users gave throughout the user test as they completed different tasks of the various functions in the pilots. This was then checked for correlation with two already established usability and user experience questionnaires, the SUS and the UEQ. By transforming the qualitative and quantitative data into a Good Practice score and then comparing it to the results of the UX and Usability questionnaires this allows to validate and investigate the relationship between adherence to the Good Practices and the perceived user experience and usability. As the user study was constructed and conducted independently of the consideration of Good Practices, two Good Practices (e.g., Search and Filter & Operability) could not be included in this comparison as the results regarding these two Good Practices were not extensive enough to warrant assigning a valence scoring. The Good Practice Search and Filter was not thematized by any user feedback. This is likely due to the short and straightforward design of the user tests' tasks. The same goes for the Good Practice of operability. While there were comments that highlighted issues with operability, the volume of such comments did not reach a significant level.

The results showed that there is a statistically significant correlation found between the Good Practices and the scoring system that considered the user study data with user experience and usability. This implies that if the Good Practice has a positive value, then the user experience is also positive. In turn, this validated the outcome that the Good Practices indeed align with greater perceived user experience and usability. In addition to the correlation, as shown in Fig. 1 it can be observed that Learnability was the only practice that had a dominant positive impression. One could assume that to some extent, the learnability of the pilots was successful and easy to learn for the users. However, it is important to also consider that the users were asked more directly about aspects of how they understood and learned from their experience in the use case. This provided more opportunities to give feedback, whether it be good or bad. The following Good Practices showed a more distributed impression, Language and Use of Color. The remaining Good Practices showed a more negative leaning impression; Minimalistic and Simple Design, User Readable Terminology, Help and Feedback, Error Handling, and Placement of Information. The qualitative data from the mGov4EU user study showed a slight tendency towards negatively phrased feedback. Additionally, this bias toward negative feedback is supported by the inherent salience of hindrances over positive aspects of UX, if they were not inquired directly. Participants often find it easier to articulate hindrances or elements that impede the system's UX over reporting what elements had a positive impact. Therefore, although users were asked about positive and negative aspects of each task, it could be that users found it easier to articulate their challenges rather than what they liked. Overall, the results presented in this paper show that by applying these Good Practices it could improve or lead to a higher user experience for mobile service applications. This was supported by a statistically significant correlation supported by a

user study conducted on mobile governmental services.

6 Limitations and Future Work

This study acknowledges the limitations of its methodology. As with every correlation, it is important to note that cause-and-effect relationships between identified usability issues and the absence of Good Practices are unclear. The central question of whether a subpar usability evaluation results from a lack of adherence to Good Practices remains open. However, the studies' results show a strong correlation between the two aspects and therefore warrant further investigation in this direction. Another potential limitation could be from the user study data, which consisted of three different pilots. The pilots generally used the same technology and had a great overlap on tasks that the users needed to conduct, regardless, each pilot had their own individual use case. The varying use cases could impact the overall familiarity and perception of user experience. In addition, participants often find it easier to articulate challenges or negative aspects in a user test, it is important to ensure that users are given the opportunity to balance this with direct questions of positive feedback. To gain a deeper understanding of the impact of Good Practices on usability and UX, future research could employ a survey-based approach. This could eliminate the need for extensive qualitative categorization and open the scope of the study to a larger sample size. Moreover, the scope could be extended to develop a predictive model for UX evaluations based on identified Good Practices. This would involve correlating survey responses with UEQ subscales to predict areas with the most room for improvement and offer tailored recommendations for action that correlate with better scores for the respective subscale. The results could serve as a foundation for evidence-based design choices and their impact on perceived usability and UX.

7 Conclusion

Overall, this paper aimed to look into the Good Practices that were defined by [SHB22]. These Good Practices were based on desk research of user experience and usability research done of various studies on mobile and electronic governmental services. The goal of this paper was to evaluate the Good Practices and seek if it is possible to validate them according to User Study Data, which consisted of Quantitative Data (General Questionnaires and Feedback, UEQ, SUS) and Qualitative Data (Interview questions throughout the User Test). In order to do that, the authors established a scoring system of the user study data to compare the Good Practices with the perceived user experience and usability. This led to testing and concluding with statistical significance that there is a correlation between the two. Leading to the conclusion, that these Good Practices indeed support greater perceived Usability and User Experience. The Good Practices can help support the future development and improvement of user experience and usability of mobile services, whether it be in a mobile governmental context or otherwise.

8 Bibliography

[Br95] Brooke, J.: SUS: A quick and dirty usability scale. Usability Eval. Ind., 189, 1995.

[CLH20] Chang, D.; Li, F.; Huang, L.A.: User-centered Evaluation and Redesign Approach for E-Government APP, 2020.

[CS19] Cahyono, T. A.; Susanto, T. D.: Acceptance Factors and User Design of Mobile e-Government Website (Study Case e-Government Website in Indonesia). Procedia Computer Science, 90–98, 2019.

[HB11] Hoober, S.; Berkman, E.: Designing mobile interfaces: patterns for interaction design, 2011.

[IW17] Isagah, T.; Wimmer, M. A.: Mobile Government Applications: Challenges and Needs for a Comprehensive Design Approach. In ICEGOV, New Delhi, India, 423–432, 2017.

[IW18] Isagah, T.; Wimmer, M. A.: Addressing Requirements of M-Government Services: Empirical Study from Designers' Perspectives. In ICEGOV '18, Galway, Ireland. doi: 10.1145/3209415.3209469, 2018.

[KMM18] Kö, A.; Molnar, T.; Matyus, B.: A User-centred Design Approach for Mobile-Government Systems for the Elderly. Hungary, 2018.

[KR19] Kureerung, P.; Ramingwong, L.: A Framework for Usability Design to Promote Awareness of Information Disseminated via Mobile Government Applications. Thailand, 2019.

[LHS08] Laugwitz, B.; Held, T.; Schrepp, M.: Construction and Evaluation of a User Experience Questionnaire. In USAB 2008, 63–76. doi: 10.1007/978-3-540-89350-9_6, 2008.

[LUN16] Lönn, C.-M.; Uppström, E.; Nilsson, A.: Designing an M-Government Solution: Enabling Collaboration through Citizen Sourcing. In ECIS 2016, Turkey, 68, 2016.

[mG24] mGov4EU. [Online]. Available: https://www.mgov4.eu/, accessed: Feb. 13, 2024

[SF20] da Silva, L. F.; Freire, A. P.: An Investigation on the Use of Interaction Design Patterns in Brazilian Government Mobile Information Systems. In Brazilian Symposium on Information Systems, 3–6, 2020.

[SHB22] Sellung, R.; Hölscher, M.; Burgstaller-Hochenwarter, L.: Good Practices of User Experience and Design Research for Mobile and Electronic Governmental Services. In Electronic Government and the Information Systems Perspective. Springer International Publishing, 138–149. doi: 10.1007/978-3-031-12673-4_10, 2022.

[Sy24] System Usability Scale (SUS) | Usability.gov. [Online]. Available: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html, accessed: Nov. 03, 2023.

Accountable Banking Transactions

Sebastian Mödersheim [1] and Siyu Chen[2]

Abstract: This paper shows how to apply the idea of *Three branches of Accountability* by Mödersheim and Cuellar to make banking transactions accountable, i.e., neither can the customer later deny to have placed the order, nor can the bank execute a transaction that the customer did not order. This is done in a general way that deliberately gives freedom to instantiate the system in several different ways, as long as it follows a few basic principles, and we show accountability holds in every instance.

Keywords: Accountability, Formal Verification, Security Protocols

1 Introduction

A security expert (whose identity we do not disclose here) once complained in a conversation with the authors about the following seemingly pointless security hurdle: after logging into his online banking system with a national Single-Sign On system (SSO), whenever he ordered a transfer, the bank would require another interaction with the SSO system. This nuisance, the expert argued, would only help if a user would leave open their machine unattended and this should better be prevented by automatic logout after short inactivity.

There is of course a different reason for the confirmation, and it is in fact very similar to the classic non-electronic banking. Here, a customer would in general need to show the bank clerk some identity document to authenticate themselves. Still, in order to make a bank transfer, the customer would have to fill out a form and *sign it*. The point of this signature is not authentication (the bank clerk already knows the identity of the customer), but gives the bank a proof that the customer indeed ordered this transaction. This proof is transferable, i.e., it can be shown to others. Thus, neither could a dishonest customer later deny a transaction they made, nor could a dishonest bank clerk (or dishonest bank) execute a transaction that had not been ordered without a substantial risk to themselves. In the electronic scenario, the second involvement of the SSO replaces the customer's signature, since typically the customers have no public/private key pairs with legal binding to their identity.

Accountability (see for instance [KTV10]) is a concept that protects security goals which cannot be enforced directly by cryptography, e.g., the goal that a customer never makes unfounded claims or a bank does not perform illegal transactions. Making the transactions accountable thus means that the actors cannot later deny what they did which can involve punishment in case of wrongdoing. We later discuss the relation with the common notion of *non-repudiation*.

1 DTU Compute, Kgs. Lyngby, Denmark, samo@dtu.dk, https://orcid.org/0000-0002-6901-8319
2 DTU Compute, Kgs. Lyngby, Denmark,

doi:10.18420/OID2024_04

We use in this paper the framework *Three Branches of Accountability* by Mödersheim and Cuellar [MC22]. A main idea is here not to study a particular protocol, but define through a legislative framework which messages have legal meaning and what actions are illegal. Agents are thus free to do anything that is not forbidden, and the legislation does not prescribe the protocols to be used, e.g., whether a bank has to use TLS to secure the connection with the customers. This considerably departs from standard paths of security protocol design and verification. Normally, we have a set of honest agents that follow the protocol and an intruder who can act as a participant, but who does not necessarily follow the protocol. What the intruder can do is defined by the common Dolev-Yao model: the intruder controls the communication medium and can apply encryption and decryption functions with known keys. In our accountability model, *all* agents are like such intruders communicating over a network where everybody can add messages and everybody can see (but possibly not decrypt) all messages. This gives us a transition system in which agents can truly "do whatever they want". We will make a restriction on the agents behavior, but essentially the goal is to verify that this large transition system has no attack state, i.e., a state that violates given security goals.

The legislative is the first branch of accountability, giving us a definition for every state in this transition system whether an agent has committed a crime (even though this is possibly not detectable for police and justice) and what legal terms hold in this state, e.g., that a public key PK is legally bound to agent A. The second branch is the executive branch which does not play any role in our case, and third is the judicial branch. It will be invoked in our case when a customer complains about a transaction, and should decide whether the customer or the bank is guilty of violating the law. The judicial branch is modeled as special transitions where an honest judge follows a protocol to interact with other participants. The participants can choose what to say to the judge (again within the Dolev-Yao model), and at the end of the protocol one or more participants are convicted.

There are now two things to prove: first, that this system is *lawful* in the sense that the innocent cannot be wrongly convicted, and that a violation of the security goal leads to some conviction.[3] We assume then that actors only commit a crime if it as *perfect* crime, i.e., where they are sure that they cannot be convicted. Under this assumption it then follows that the security goals are never violated.

The contributions of this paper are to formalize the accountability for banking transactions with SSO in a very generic way that can be instantiated by many concrete bank transaction systems and protocols. We then prove the lawfulness and accountability, i.e., violation of security goals leads to a conviction. We also show a minor variant that corresponds to a simple oversight where accountability would not hold.

3 The security goals in general do not require to prevent every illegal action: there may be illegal actions that the security goals are not directly concerned with (e.g., a user sharing their password with a friend). Moreover, we may have security goals that are not enforced by the three branches.

2 The Legislative

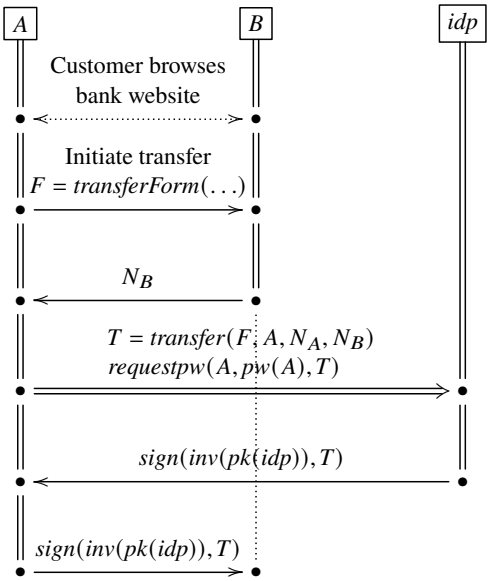


Fig. 1: Generic schema of a bank transaction between customer A, Bank B, and identity provider *idp*.

Figure 1 shows the generic schema of a transaction. It starts with a customer A who browses the website of their bank B, e.g., looking at their account balance. Typically, this entire communication is secured using a protocol like TLS and some authentication mechanism like a login with password. At some point A may decide to order a money transfer. For this purpose we assume there is an electronic form *transferForm* that contains a number of fields like the sender's and receiver's IBAN number, the amount (and currency) to be transfered. Such a form is some (non-cryptographic) way of structuring the information to be entered into the form, e.g., XML. Up to this point, we do not specify exactly how the interaction between A and B should work: this is up to the bank how to design their web interface, what authentication mechanism to use, and how to encrypt the transfer.[4] B will now generate a random number N_B and ask the user the authorize the transfer in an accountable way using N_B. To that end, since A does not have a legally registered public/private key pair, A turns to the identity provider *idp* in order to get the transfer signed by the *idp*. A first creates $T = transfer(F, A, N_A, N_B)$, another electronic form that contains the transfer form F, the name of A, a random number N_A created by A as well as N_B from before. A packs this into yet another form $requestpw(A, pw(A), T)$ that contains the name and password that A has with *idp*. This distinction of three forms is just for conceptual simplicity: the *transferForm*(...) contains the information that would also be present on a

4 In fact, for the purpose of the accountability proof below, the communication does not even have to be encrypted. However, for privacy this would be terrible: everybody on the network could observe A's interactions with the bank; also the authentication mechanism obviously cannot be password-based if messages were unencrypted.

non-electronic bank transfer form and is thus only banking-related without other technical aspects; the *transfer*(...) form is what should actually be signed by the *idp* containing the random numbers; and finally *requestpw*(...) is the message that A sends to the *idp* and that should not be observed by anybody else. This message indeed has to be encrypted in a way that only the *idp* can read, because it contains the user's password, and this will be part of the legal requirements below. We have in fact depicted this transmission in the figure using a double arrow to highlight that this transmission crucially needs confidentiality. Again, we do not require a particular way of achieving this: while one typically uses TLS, simply encryption with the public key of *idp* would be sufficient. If everything is fine, the *idp* signs the request and sends it back to the customer who can forward it to the bank, which then executes the transfer. Here $pk(idp)$ is the public key of *idp* and $inv(pk(idp))$ is the corresponding private key.

This protocol schema has some assumptions: the *idp* must be honest, otherwise there are trivial attacks against this. (In contrast, A and B may well be dishonest.) Further, nobody except A and *idp* knows $pw(A)$, and everybody knows that $pk(idp)$ is the public key of *idp* (and can thus verify the signatures produced here).

We call this a protocol schema because it leaves many details open and therefore gives banks the freedom to implement their online banking systems as they see fit, as long as they obey a few legal regulations. The three named formats *transferForm*, *transfer* and *requestpw* have distinct legal meaning, and thus the law must fix a particular way to implement these forms, e.g., as XML formats. Similarly, some cryptographic parameters must be fixed like permissable signature schemes and key sizes. We assume here that these details are fixed already.

The legal system now consists of the following laws; as is standard in legal text, some "commentary" will be given as footnotes:

§1 There is a public key legally bound to the identity provider, denoted $pk(idp)$ in the following. If any agent other than *idp* posses the corresponding private key $inv(pk(idp))$, then *idp* is punishable.[5]

§2 There is a set of actors who are registered as banks. To each bank B a public key denoted $pk(B)$ is legally bound. If any agent other than B posses the corresponding private key $inv(pk(B))$, then B is punishable.

§3 There is a set of actors who are registered as citizens, they each have a unique identifier and password registered with the *idp*. The password of citizen A is denoted $pw(A)$. If anybody other than A and the *idp* knows $pw(A)$, then A is punishable.[6]

5 As said before, we will assume that the *idp* is honest and therefore will never leak their private key, so they will never be punished.

6 The logistic requirements are fulfilled for instance in Denmark by the fact that every legal resident has a CPR number and authentication credentials with the national identity provider MitID. We do not consider here two-factor authentication for simplicity or mechanisms for changing the password for simplicity. The problem that an honest actor may lose their password is discussed below.

§4 When a client A wants to order a bank transfer, then they are required to generate a random number N_A and ask the bank for a random number N_B. Then they may issue the format $T = transfer(F, A, N_A, N_B)$ with F the fields of the transfer, and issue the format $requestpw(A, pw(A), T)$. Producing this message legally binds A to this transfer detailed by F, and A is punishable if N_A has been used in a different transfer. The client is also obliged to encrypt the message in such a way that only the *idp* can read it and that no agent can change the content of the message.[7]

§5 When the *idp* receives a message $requestpw(A, pw(A), T)$ that uses the correct password $pw(A)$ of A and T is a format *transfer* containing sender name A, then the *idp* may sign the message T with their private key $inv(pk(idp))$. If the *idp* signs with $inv(pk(idp))$ a *transfer* message other than according to this law, then the *idp* is punishable.[8]

§6 When a bank B receives *transfer* message T signed by the identity provider, it may execute the transfer requested in T, provided that: the random number N_B contained in T was freshly generated by B and has not been used in any other transaction, the name A corresponds to a legal account holder at B denoted by the IBAN number of the sender in T. A bank who performs a transfer without a signed transfer message T according to this law is punishable. The bank is obliged to save the signed T message and produce it when subpoenaed by a judge; if the bank fails to answer the subpoena to a transaction they have performed, they are also punishable.

§7 A customer can complain to a judge if a bank has performed a transfer without the customer having legally requested it. The client is punishable if they issue such a complaint while they did legally request that transfer.

3 The Judicial Branch

The second branch of The Three Branches of Accountability paper is the executive branch: the police who may discover some criminal activity, and provide the evidence to the judicial branch. This is not necessary in our case, so we directly come to the third branch, the judicial branch. This is about declaring how a judge should proceed when a customer A registers a complaint that a bank B made a transfer from A's account without A ordering that.

The judge has to follow a simple procedure: they subpoena the bank B to provide a signed transfer form that § 6 requires before the bank can make a transfer. If B does not produce a signed *transfer* form that matches the transaction, then B is convicted. Otherwise the customer A is convicted.

7 The requirement that no other agent than *idp* can learn it follows already from §3; here it also required that no other agent can alter the message. Note that the law leaves open how this is done, e.g., by public key encryption or TLS.

8 Note that the law only says *may*, because the *idp* shall not be punishable for instance for downtimes. Moreover the law does not forbid the *idp* to use its private key for other purposes, as long as the signed message is not of the *transfer* format.

Thus we have a procedure for the judge that will identify one participant as guilty, whenever a customer complains about a bank transfer they allegedly did not order. It may seem intuitively clear that this procedure is never convicting somebody who is innocent, i.e., somebody who abided the law. However, this is not the case as demonstrated by the following attack.

3.1 Breaking and Fixing the Transfer Protocol

Suppose an agent A issues a transfer request to the *idp* who signs it, and the client forwards this signed request to the bank B. Suppose now B is malicious and executes the exact same transfer two times, effectively doubling the amount that A transfers. Now A is intuitively right to complain about the second transfer (or alternatively about the first one). However, if they complain to the judge and the judge subpoenas B, B can show a valid signed request that matches the transaction, and thus, with the above judge procedure, the client would now be convicted without having broken any laws.

To fix the situation, recall that we have required that the transfer contains random numbers N_A and N_B, and that each side is obliged to create them freshly. Note that this is also subtle, because the bank cannot prevent that a nonce N_B they created could be used by a malicious customer in several different requests, so it is not so easy to tell in general who has failed to adhere to the protocol.

However, if the judge looks at two (or more) identical transfers that B has executed, and subpoenas for all of them, and B presents for each the *same* signed *transfer* form, including the same nonce N_B used in each of them, then B is actually punishable by § 6 which explicitly obliges B to check that N_B was generated by B and not used in another transaction. So even if for instance a malicious client has issued several transfers with the same N_B, it is the duty of B to check that each N_B can only be used once. It should also be noted that without the fresh random numbers N_A and N_B it would be impossible now to tell whether the bank or the client broke the law.

We thus modify both § 7 and the procedure of the judge (the change is underlined):

§ 7 A customer can complain to a judge if a bank has performed a transfer without the customer having legally requested it, or more often than the client has requested it. The client is punishable if they issue such a complaint while they did legally request that transfer at least as many times as the bank executed it.

The judge now follows this procedure: given a set of transfers that a client complains about, the judge subpoenas the bank B to provide signed *transfer* form for each transfer according to § 6. If B fails to produce a signed *transfer* forms with distinct values of N_B, then B is punishable. Otherwise the client is punishable.

4 Security Goals and Proof

4.1 Lawfulness

First, we prove that this system is *lawful*, i.e., that no innocent (i.e., law-abiding) agent can ever be convicted (which actually could happen in the first version as demonstrated by the attack).

To that end, consider again the procedure of the judge: when a customer A complains about a set of n transfers, and the judge thus subpoenas the bank to produce n signed *transfer* forms with distinct numbers N_B, then we have two cases. First, if B does not comply with this subpoena, then B must indeed have broken the law: they were obliged by § 6 only to perform transfers for which they have a corresponding signed *transfer* form, which must all have distinct numbers N_B according to § 6 as each N_B can only be used once. Finally, by § 6, B is also obliged to store each signed *transfer* form and produce them upon a subpoena.[9] Thus, if B does not answer the subpoena correctly, B is rightfully convicted.

Second case, if the bank does correctly answer the subpoena by producing n matching signed *transfer* forms $sign(inv(pk(idp)), T_1), \ldots, sign(inv(pk(idp)), T_n)$ with distinct numbers for the N_B-field in the T_i. Then we have to show that the customer is now rightfully convicted. (This is the case where the first flawed version could potentially lead to a wrong conviction.) Recall that we have assumed that the *idp* is honest.[10] This implies that the *idp* has not leaked their private key $inv(pk(idp))$ and thus the signed transfer requests $sign(inv(pk(idp)), T_i)$ have been made by *idp* where, again by honesty of *idp*, we can conclude that the *idp* has followed § 5 when they they signed the transfer forms, i.e., they must have received the $requestpw(A, pw(A), T_i)$ first. Again since the *idp* is honest, one of two things must be the case. First possibility: A has produced all these requests, then A is punishable by § 7 for issuing n request and claiming that they did not. Second possibility: A has leaked $pw(A)$ to somebody else who then produced some of these requests (maybe unbeknownst to A); then A is punishable according to § 3. (Of course, if *idp* were dishonest, it could have leaked $pw(A)$, but we assumed *idp* to be honest.) Thus either way, A is rightfully convicted also in this case and we have established that we are in a lawful system.

9 One may ask whether it is impractical that a bank can always check that the N_B in a request is different from every N_B they have ever accepted, as this seems to imply that the bank always have to consult their entire transaction history. However, following the schema in Fig. 1 is a simple solution: they create a *fresh* random number N_B (that is with overwhelming probability different from all previous such random numbers). The bank remembers N_B *for this session* and accept at most one incoming signed transfer with this N_B. If this does not arrive within a certain time window, the bank simply closes the session and forgets N_B. This is legal, since the bank is not obliged to eventually perform the transfer.

10 This is not an unproblematic assumption as we discuss below, but if we think of a national identity provider, it is at least reasonable that a judge would value their statements (and thus signatures) as trustworthy.

4.2 The Security Goals

The security goals we would like to ensure in this system are:

1. A bank never performs a transfer more often than the customer has ordered it. Observe that this formulation also includes the case that the bank performed a transfer that the customer did not order at all.
2. A customer never complains about a set of bank transfers that they have ordered at least as many times.

The Three Branches of Accountability approach defines a *perfect crime* as an illegal action of an agent where the agent knows they will never be convicted for this action. For instance, in our example, the customer may illegally reuse the same nonce N_A for more than one transaction. As nothing in the described system triggers on that, the agent will never be convicted for that. However, the agent also does not have any advantage from this crime—it is irrelevant.

It is now easy to see that under the assumption that agents will only commit perfect crimes, the security goals hold. Suppose we could reach a state where the first goal is violated, i.e., where a bank has performed a transfer more often than the customer ordered it. Then the client will complain with the judge, which will convict either the client or the bank.[11] Either way, since we have shown the system is lawful, the convicted party has indeed broken the law. The fact that they have been convicted for it shows that it was not a perfect crime, contradicting the assumption. Suppose we could reach a state where the second goal is violated, i.e., a customer dishonestly complains about a set of transfers they have indeed ordered. Again the procedure of the judge will lead to the rightful conviction of either the customer or the bank, which by the perfect crime assumption is absurd. Thus, no state violating the security goals is reachable.

5 Conclusions

Essentially, this paper shows how transactions, like a bank transfer, can be made accountable: at the core the bank needs a transferable proof that the customer indeed ordered the transfer in question. For such a transferable proof, a pure authentication of the customer to the bank is not sufficient, because this would give the bank nothing to convince a third party like a judge. The only option is that we have a signature that could either be produced by the customer themselves or via a trusted third party as in our case study. This solution has the disadvantage that in every transaction, the trusted third party has to be involved, and if it should be hacked for instance, the security guarantees of the entire system are void (as it crucially needs to be an honest party). On the other hand it has the advantage that it can be deployed based on existing identity management infrastructures such as the Danish MitID.

11 Observe here that the client might still be punishable here, even though they did not order all the transfers: they may have given out their password and somebody else did order them.

There are several points one can criticize about the system sketched here. First of all, one may wonder if customers can be blamed if their password gets leaked: of course, not all these leaks are because a customer intentionally gave it to somebody else (like their spouse) but they may have been observed entering it or their computer or phone may have been hacked. This is in fact an old problem that we have for instance for leaked PINs for credit cards: to our knowledge, when a thief withdraws money from an ATM using a stolen card with the correct PIN, the default assumption (unless there is contrary evidence) of the courts is that the card owner must have been careless with their PIN (e.g., kept it written down along with the card). We believe that this is a general dilemma as part of digitalization namely that digital secrets (like PINs, passwords, private keys) have legal meaning, and that losing them can have substantial legal and financial consequences for an individual. There are several mitigations such as multi-factor authentication that at least make it harder for an attack to take over one's identity. One of the anonymous reviewers of this paper suggested that a customer who became victim to a cyber attack could inform the police, and use the police report to get (at least temporarily) reimbursed by the bank. In fact, it is common that the total transfer amounts per day are limited and that banks are ensured against the limited fraudulent transfer that can occur within that time window.

Note, however, that this is a more general problem of digitalization. For our example, suppose the bank loses their database of transfers due to a cyberattack (or simply a software bug) and suddenly become liable in court for all transfers they have executed and cannot justify if subpoenaed for it. This indeed shows that the systems like the one described here cannot be absolute: we have to have human judges who evaluate other evidences like a forensic investigation of a customer's phone for instance. The accountability proof thus provides a strong argument that the system is reasonably well-designed to deter criminals, but it is not an absolute that cannot be overridden when new evidence appears.

In fact, the absolute trust in the *idp* is a second serious problem of this system. A standard approach to replace the single trusted *idp* by a consortium of identity providers of which only a majority needs to be honest is not very promising as this requires a consensus amongst the consortium and is thus not the light-weight solution we have sketched here.

But let us end this list on a third problem for which there is actually a good solution: suppose a customer has obtained via the *idp* a valid signed *transfer* form and send it to the bank. If the bank is malicious, they might hold it off, not executing the transfer but they have the right do so at any moment. The customer is thus in limbo: the transaction has not gone through but could be accepted at any moment. This is a fair exchange problem that in general also requires a trusted third party to resolve; however, one can make this *optimistic* in the sense that the trusted third party would only need to be involved if customer and bank do not come to a consensus on their own [ASW00].

There are several research articles on accountability, most importantly Küsters et al. [KTV10], Künnemann et al. [KGB21], as well as Mödersheim and Cuellar [MC22] that we have used. While the former two are based on computational adversaries, we follow the third

approach and consider a Dolev Yao-style intruder who cannot break the cryptography. We have chosen the approach of [MC22] because the concept of the legal system offers the flexibility to support a wide range of systems rather than fixing a particular protocol (like TLS) that is largely irrelevant to the accountability question.

Many works rather use the term *non-repudiation* instead of *accountability*; for an overview, see e.g. [KTV10, MC22]. While the term *non-repudiation* puts the emphasis on ensuring that actors cannot deny actions they have performed, *accountability* goes further: a bank who has (undeniably) performed a transfer may also be required by a judge to *justify* their action, showing that they acted legally. Depending on the protocol, the answer from an honest actor to a subpoena may give the judge further evidence to investigate [MC22].

As future work, we plan to follow the idea of Bruni et al. [BGS17] to investigate if and how protocol verification tools like Tamarin [Me13], ProVerif [Bl01], or PSPSP [He21] could be employed, and possibly adapted, to verify accountability questions in such an open scenario as the legal system in this paper.

Acknowledgements. This work was supported by the Horizon Europe project TaRDIS (Trustworthy And Resilient Decentralised Intelligence For Edge Systems).

Bibliography

[ASW00] Asokan, N.; Shoup, Victor; Waidner, Michael: Optimistic fair exchange of digital signatures. IEEE J. Sel. Areas Commun., 18(4):593–610, 2000.

[BGS17] Bruni, Alessandro; Giustolisi, Rosario; Schürmann, Carsten: Automated Analysis of Accountability. In (Nguyen, Phong Q.; Zhou, Jianying, eds): ISC 2017. volume 10599. Springer, pp. 417–434, 2017.

[Bl01] Blanchet, Bruno: An Efficient Cryptographic Protocol Verifier Based on Prolog Rules. In: Computer Security Foundations Workshop. pp. 82–96, 2001.

[He21] Hess, Andreas V.; Mödersheim, Sebastian; Brucker, Achim D.; Schlichtkrull, Anders: Performing Security Proofs of Stateful Protocols. In: 34th IEEE Computer Security Foundations Symposium (CSF). volume 1. IEEE, pp. 143–158, 2021.

[KGB21] Künnemann, Robert; Garg, Deepak; Backes, Michael: Accountability in the Decentralised-Adversary Setting. In: 2021 IEEE 34th Computer Security Foundations Symposium (CSF). IEEE Computer Society, pp. 95–110, 2021.

[KTV10] Küsters, Ralf; Truderung, Tomasz; Vogt, Andreas: Accountability: definition and relationship to verifiability. In: Proceedings of the 17th ACM conference on Computer and Communications Security. pp. 526–535, 2010.

[MC22] Mödersheim, Sebastian; Cuellar, Jorge: Three Branches of Accountability. In: In Festschrift for Joshua Guttman, LNCS 13066, 2021. . Springer, 2022.

[Me13] Meier, Simon; Schmidt, Benedikt; Cremers, Cas; Basin, David A.: The TAMARIN Prover for the Symbolic Analysis of Security Protocols. In: Computer Aided Verification. pp. 696–701, 2013.

Gaining Back the Control Over Identity Attributes: Access Management Systems Based on Self-Sovereign Identity

Kenneth-Raphael Keil[2], Ricardo Bochnia[1], Ivan Gudymenko[2], Stefan Köpsell[3], and Jürgen Anke[1]

Abstract: Digital employee cards used for door access control offer benefits, but concerns about traceability, profiling and performance monitoring have led to opposition from workers' councils and employees. However, the emerging identity management approach, Self-Sovereign Identity (SSI), can address these concerns by giving control over disclosed identity attributes back to the end user. This paper analyzes a real-world access management scenario in a hospital building and applies the SSI paradigm to address the identified issues. The analysis assumes a semi-honest observing attacker sniffing on the payload and the transport layer. The SSI-based proof of concept is shown to have a high potential to protect against traceability and profiling. However, in addition to the careful technical implementation of SSI, it is important to consider non-technical factors such as governance for a holistic solution. We propose potential strategies to further minimize privacy risks associated with SSI-based employee identity management using mediators.

Keywords: Self-Sovereign Identity, Traceability, Privacy, Access Control, Profiling, Architecture

1 Introduction

Privacy is a crucial concern in the workplace where there is often a need to balance the employees' privacy and the employer's needs [BTD20]. Digital employee ID cards enable door access control use cases but also introduce privacy risks, such as traceability and profiling. The possibility that employee data could be used to track and analyze movement patterns within the company building and for employee productivity monitoring could lead to undesirable consequences such as stereotyping and discrimination. This may be an obstacle to adoption of such systems and creates a demand for privacy-preserving solutions.

To explore how privacy-preserving digital employee ID cards could be designed, we examine a hospital door access management use case from the SEMECO project cluster[4]. SEMECO researches secure medical microsystems and communications. In this use case, the hospital wants to introduce a new digital employee ID card for door access control. However, the

1 HTW Dresden, Digital Service Systems Group, Friedrich-List-Platz 1, 01187 Dresden, Germany, ricardo.bochnia@htw-dresden.de, https://orcid.org/0009-0007-4317-1810; juergen.anke@htw-dresden.de, https://orcid.org/0000-0002-9324-9387
2 Deutsche Telekom MMS GmbH, , Riesaer Straße 5, 01129 Dresden, Germany, kenneth-raphael.keil@telekom.de; ivan.gudymenko@telekom.de
3 Barkhausen Institut, , Schweriner Str. 1, 01067 Dresden, Germany, stefan.koepsell@barkhauseninstitut.org, https://orcid.org/0000-0002-0466-562X
4 https://semeco.info/

doi:10.18420/OID2024_05

workers' council demands that the employed solution does not lead to privacy violations, e.g., profiling, and does not pave the way for employee performance monitoring.

Many access management solutions are inherently prone to profiling and tracing attacks. However, new paradigms have emerged to address this issue, such as Self-Sovereign Identity (SSI). SSI is a user-centric identity management approach that gives users more control over personal identity attributes than traditional identity systems. Privacy and data-minimization are core principles of SSI [Al16; Ču22; Se22]. In SSI, users (holders) receive Verifiable Credentials (VCs) from a trusted issuer, which they can present to a verifier (relying party). SSI represents a paradigm shift by allowing the user to decide if and which attributes of the VCs he wants to share with a third party. While SSI still has challenges, e.g., interoperability, it offers promising benefits for organizations [BRA24].

However, the issue of traceability and profiling with digital employee ID cards is not limited to the healthcare sector. Therefore, our research aims to generalize its applicability beyond the healthcare context by focusing on the following research question:

> To what extent can SSI prevent employee tracing and profiling in the access management system of a building?

To answer this question, we created an access management system design based on SSI and validated it by implementing a proof of concept (PoC). Rooms and specific roles in the hospital are deliberately abstracted into generic entities. Traceability and profiling mitigation are evaluated by logging and analyzing all transactions between the employee's SSI wallet and the door lock (verifier). The focus is on avoiding linkability, i.e., limiting re-identifying characteristics, and identifying possible attack points for quasi-identifiers within the system architecture. Note that we have only evaluated the direct privacy implications of the SSI-based system, excluding contextual data generated by e.g., cameras or rosters. While such contextual information needs to be considered in designing an overall solution, it is a prerequisite that the individual building blocks of an overall architecture are as privacy-friendly as possible by design.

The remainder of the paper is organized as follows: First, the related work is reviewed and analyzed. Next, we present the requirements for the attacker model and the protection goals. Next, we explain the design and architecture of the proposed SSI-based access control system and justify our design choices. We then evaluate the ability of our PoC to track and profile employees using the access control system by running and logging test transactions. The logs are analyzed for potential tracking and profiling attacks. In the discussion, we examine the implications and limitations of our solution and provide an outlook for future work.

2 Related Work

There have been several studies on the use of SSI in healthcare. However, most of these studies focus on patient data. A notable exception is the NHS Digital Staff Passport, formerly known as the COVID-19 Digital Staff Passport, which is a digital employee ID [LC22; NH24]. Although it is an employee ID, the focus is on reducing onboarding time for staff moving between hospitals, not on access control.

Watanabe et al. [Wa24] also developed a PoC for an anonymous door-unlocking system based on SSI. They focused on removing anonymity in the event of an incident rather than tracing and profiling risks. They introduce a new role called *opener*, which acts as a trustee. In essence, each holder has a unique *uid* attribute in their VCs that is part of the presentation to the verifier. However, the *uid* is encrypted using the *opener's* key. Thus, only the *opener* can re-identify the holder if an incident occurs. While this method allows re-identification for incidents, it requires that *opener* and verifier do not collude. This ultimately shifts the risk of tracing and profiling from the verifier to the *opener*.

3 Attacker Model and Protection Goals

The overall goal is to design a privacy-preserving door access control system that preserves the security and privacy features of a physical key-based system while providing the flexibility of an IT-based access control management system.

We assume the following attacker model: an observing, computationally limited attacker with access to all communication messages. In this use case, the attacker can be represented by door terminals and the backend system(s) can monitor all messages. We assume the semi-honest attacker, which implies that no malicious intervention into the communication protocol, altering of messages, etc., takes place. This type of attacker is also known as "honest but curious". We decided to choose this attacker model since in the context of our system, it represents the minimum capability set an attacker would have in a real life scenario against which the respective countermeasures must be provided.

As mentioned above, the potential for attackers to correlate user door interactions with external information is out of scope in our model because our main focus is on designing a hospital door access IT system that avoids user profiling.

Given our attacker model, we explicitly address two types of attacks:

1. Observing attacks on the message payload
2. Observing attacks on the transport layer

Considering the classical privacy goals, we focus primarily on *unlinkability*. According to [PH10], *"Unlinkability of two or more items of interest (IOIs, e.g., subjects, messages, actions, ...) from an attacker's perspective means that within the system (comprising these*

and possibly other items), the attacker cannot sufficiently distinguish whether these IOIs are related or not." We also assume that the existence of a sufficient anonymity set in the system, i.e., the number of active users producing transactions while interacting with the access management system, is large enough to prevent trivial user profiling. However, we do not pursue any further formal analysis of this in this paper.

4 Design of the Privacy-Preserving Door Access Control System

Our SSI-based PoC is supported by an anonymous credential system. An anonymous credential system [Ch85] allows the holder of a given credential to prove that the credential satisfies a logical predicate specified by the verifier. The verifier learns nothing more than whether the proof was successful. Furthermore, the verifier cannot link two runs of the proof protocol, and thus will not learn if the two protocol runs involve the same holder. This is true even if the verifier and the credential issuer collude.

The only case where this strong privacy and security property does not hold is when a revocation mechanism is applied. However, this must be explicitly addressed during system setup and cannot be added by a malicious entity after the fact. Revocation is beyond the scope of this paper.

For our door access control use case, the holder is an employee, and the credentials issued by the hospital express the authorization to open a given door. The verifier is a given door, and the logical predicate expresses "has the right to open the door". The overall design and its concrete implementation are illustrated in Fig. 1 and 2.

Fig. 1: Overview of the privacy-preserving door access control systems utilizing anonymous credentials.

For implementing our design, we choose Hyperledger Aries[5] as our SSI toolkit, along with DIDComm v1[6] as the communication protocol and AnonCreds[7] as the credential format. The reason for this choice was their widespread use and popularity in the SSI domain. AnonCreds was chosen as the credential format because it supports selective disclosure and zero-knowledge proofs (ZKP) to ensure the privacy and unlinkability of the holder. Although the format has drawbacks, it is widely used [Yo22]. In addition, other popular credential formats do not provide the same privacy-preserving features, especially with respect to ZKP [YY23].

Each entity – issuer, verifier, holder – is represented by an Aries Cloud Agent - Python (ACA-Py) hosted on a dedicated virtual machine with a public IP address. The wallet is

5 https://www.hyperledger.org/projects/aries
6 https://github.com/hyperledger/aries-rfcs/blob/main/concepts/0005-didcomm/README.md
7 https://www.hyperledger.org/projects/anoncreds

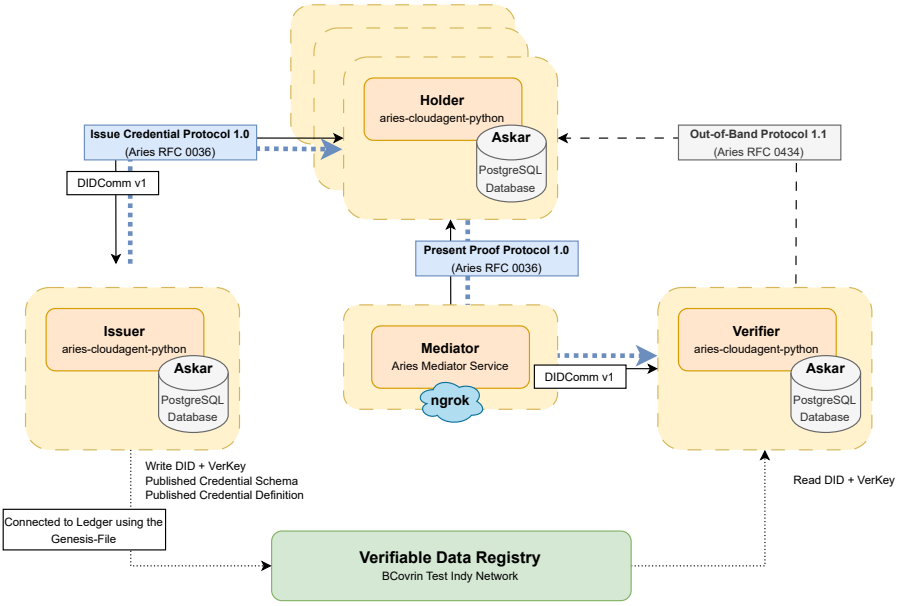


Fig. 2: The architecture overview of the PoC.

Aries Askar, with a PostgreSQL database. While the issuer and holder communicate directly, the communication between the holder and verifier relies on an Aries Mediator Service. The mediator uses Ngrok to provide a public URL for communication. The mediator's job is to prevent the verifier from knowing the holder's IP address. The BCovrin Test Indy Network serves as a verifiable data registry to host the schema and credential definition. The key factors in choosing this network were the constant availability of the network and the fact that there was no need to host our own network.

5 Anonymization and Profiling Evaluation

To evaluate our system, we constructed a fictive scenario where three employees entered the door five times, as illustrated in Fig. 3. Based on our semi-honest observational attacker model, we analyze both the payload and transport layers to identify potential vulnerabilities. Thus, our analysis is twofold: First, we examine the content of the presentation that the holder makes to the verifier. Then we look at the routing of the presentation. For each of these, we examine potential traceability and profiling risks. In addition to manual investigation, we also used Python scripts that searched all log files for correlated data to identify patterns that might indicate traceability risks.

The first script compares JSON objects and identifies shared key-value pairs between them, even within nested structures. It generates a table of duplicated values grouped by key and listed with the corresponding file names. The other script analyzes the log files of the mediator or verifier for shared keys and values. It uses regular expressions to extract certain attributes – such as did, connection_id, or serviceEndpoint – from each line, and then checks whether these attributes have the same values in different files. The output is a report that lists the shared values and their corresponding files.

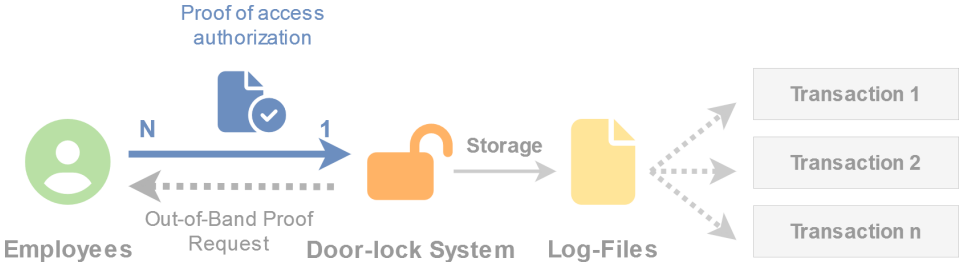


Fig. 3: The evaluation consists of conducting five transactions for each of the three employees.

Five test transactions are run for each of the three test employees to generate data for analysis. To query an access authorization from an employee, the door-lock system must initiate a query that is carried out via an out-of-band invitation in combination with a proof request attachment, as specified in Aries RFC 0037: Present Proof Protocol 1.0 [Kh19]. The proof request is designed to grant access only if the employee has a value of 10 ≥ for the *access_room* attribute in their employee ID. This simulates simple hierarchical permissions (e.g., ≥ 10 for access to the main building, ≥ 20 for access to the ward, etc.).

The ACA-Py configuration setting *–log-level "debug"* logs the steps of the AnonCreds presentation dataflow. In addition, the command *sudo tcpdump -i any 'port 8000'* logs at the network level of the virtual machine. In addition, the following items are logged as a JSON object for each transaction: the DIDComm connection established and the REST API responses to the creation of a proof request (verifier), an out-of-band invitation (verifier), the receipt of the out-of-band invitation (holder), the sending of the presentation (holder) and the presentation received by the holder (verifier). Parts of the proof are shown in Fig. 4.

```
{
  "revealed_attrs": {},
  "a_prime": "699...399",
  "e": "135...461",
  "v": "584...918",
  "m": {
    "access_room": "115...454",
    "master_secret": "152...048"
  },
  "m2": "308...508"
}
```

```
{
  "c_hash": "993...755",
  "c_list": [
    [ 2,41, ... ],
    [ 249, 87, ... ],
    [ 42, 212, ... ],
    [ 205, 234, ... ],
    [ 2, 25, ... ],
    [ 171, 141, ... ]
  ]
}
```

```
{
  "revealed_attrs": {},
  "self_attested_attrs": {},
  "unrevealed_attrs": {},
  "predicates": {
    "additionalProp1": {
      "sub_proof_index": 0
    }
  }
}
```

Fig. 4: The eq_proof (equal proof), aggregated_proof and the requested_proof.

5.1 Payload

To analyze whether each proof is unique and does not contain a recurring identifier, a direct comparison of all proofs received in the presentations of the test transactions was performed. The comparison shows that each proof is unique and, therefore, does not contain any potentially recurring identifiers or identifiers that would enable direct personal traceability or profiling. However, the schema ID and the credential definition represent a recurring, identifiable identifier (see Fig. 5) that could be used for group-based profiling.


```
{
  "schema_id": "4jeMLJGSfmAAyHtFXBTyiG:2:access-card:1.0",
  "cred_def_id": "4jeMLJGSfmAAyHtFXBTyiG:3:CL:241769:access-card",
  "rev_reg_id": null,
  "timestamp": null
}
```


Fig. 5: The following attributes were the same across transactions.

In a hospital environment, conclusions could be drawn about those who issued the credentials. On the one hand, this is intended to ensure the authenticity of a credential. On the other hand, if there are several central authorities in the hospital that issue credentials, a classification could be made at this point (e.g., department A issued credential B to employees X, Y, and Z). However, if the anonymity set is substantial enough, it does not negatively impact the achievement of the profiling and unlinkability goals.

5.2 Transport

The analysis of transport logs and paths has shown that the verifier does not generate personal tracking or profiling. However, the frequency and timing of transactions can be used to identify general behavioral patterns for accessing a particular location (e.g., X people open the door at time interval Y). However, these behavioral patterns can only lead to the re-identification of a specific person by adding other data sources (e.g., hospital duty rosters or information from video surveillance cameras, which is outside the scope of our adopted attacker model).


```
ngrok_1               | t=2023-12-28T16:20:16+0000 lvl=info msg="join
connections" obj=join id=32b7b39e55c7 l=172.21.0.3:2015 r=20.203.139
mediator_1            | 2023-12-28 16:20:16,925
aries_cloudagent.messaging.base_handler INFO Received forward for:
E8DbAjd4e4eWTzDYx8V1AFhoECcTgpQRaeyAfuTjn4pm
aries_cloudagent.messaging.base_handler INFO Forwarding message to connection:
d52a2d91-1573-4d8a-a190-d8c1da7dfa45
ngrok_1               | t=2023-12-28T16:20:18+0000 lvl=info msg="join
connections" obj=join id=43291c212b83 l=172.21.0.3:2015 r=20.100.196
mediator_1            | 2023-12-28 16:20:18,620
aries_cloudagent.messaging.base_handler INFO Received forward for:
E8DbAjd4e4eWTzDYx8V1AFhoECcTgpQRaeyAfuTjn4pm
mediator_1            | 2023-12-28 16:20:18,634
aries_cloudagent.messaging.base_handler INFO Forwarding message to connection:
d189bb56-b704-42a7-805b-46be3d149d0d
```


Fig. 6: The mediator log (truncated). The mediator could create a behavior profile.

The traceability and profiling risk on the part of the verifier is therefore considered low, since the transport data transmitted and the logged logs do not reveal anything about a specific person without additional sources of information. As mentioned above, our system uses anonymous credentials for credential generation by design, which makes the user's transactions indistinguishable from those of other users and therefore unlinkable within the system.

In contrast, the risk of traceability and profiling is considered high when using a mediator. Although the mediator hides the holder's endpoints from the verifier, thus preventing traceability and profiling on the part of the verifier, it may itself capture certain behavioral patterns and unique identifiers such as connection IDs or IP addresses from its logging (see Fig. 6).

6 Discussion

Our results show that SSI can provide employee privacy and prevent employee profiling, but the traceability and profiling risks depend on the specific implementation. It is naive to believe that using an SSI solution does not involve privacy issues, even if selective disclosure and zero-knowledge proofs are implemented. The main risk for our PoC is not in the content of the VP, but in the routing of the VP, since the holder's IP address itself may be a correlatable identifier. In the following sections, we will first discuss the implications of our findings and possible strategies to further minimize the remaining risks. We will then discuss the limitations, and finally, we will address areas for future work.

6.1 Implications

While our PoC makes it hard for the door-lock system (verifier) to trace and profile employees, the risk is only shifted to the mediator. This issue is similar to Watanabe et al. [Wa24] where the risk is shifted to the opener which we have already discussed in the related work section. The mediator can recognize from whom and where a message is sent. Therefore, additional actions are necessary. In the following, we present several strategies to mitigate the risk of traceability and profiling further. Fig. 7 and 8 illustrate several strategies. It is possible and advisable to combine multiple strategies.

Independent Mediator: Mediators may be hosted by independent, non-colluding entities in the SSI ecosystem (separation of concerns). For example, independent and trustworthy operators could take over the provision of the mediators. This further complicates data consolidation due to the virtual, physical, and legal separation of the operator and verifier. However, the problem of determining which entities are trustworthy enough to operate mediators remains. To reduce the risk of collusion, multiple independent mediators can be chained together.

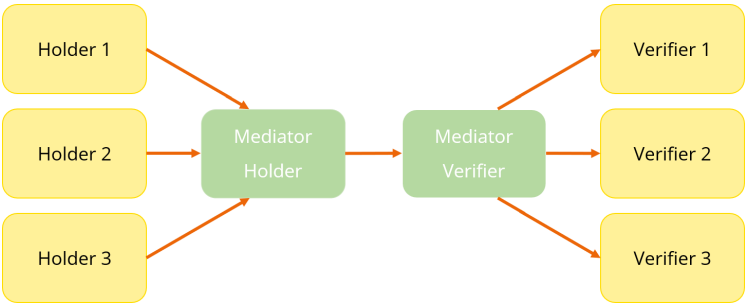


Fig. 7: Communication involves a chain of two mediators and each mediator also groups several holders or verifiers. The two mediators should be operated by independent, non-colluding entities.

Mediator Chaining: Both the verifier and the holder can use separate mediators, creating a chain of mediators between the verifier and the holder. Additionally, this chain can be extended by adding more mediators. However, consolidating the mediator data would again create the problem described above. Therefore, mediator chaining should be used in conjunction with independent mediators and group mediators.

Mediator Grouping: By further consolidating the endpoints of the holders through a mediator and the endpoints of the verifiers as well (also known as agency or herd privacy), the mediator of the holder does not know which verifier the message ultimately goes to, and the mediator of the verifier does not know which holder the message originated from. However, this implies that different mediator groups of verifiers and holders do not collude.

Mediator Rotation: In the PoC, a default mediator was used that routes all incoming and outgoing messages by default. Aries also allows a specific mediator to be defined for a specific message, and an agent can have multiple mediators [Ha21]. This would allow each message to be routed through a randomly selected mediator, as shown in Fig. 8. Instead of using a random mediator for each message, it may be sufficient to rotate the mediator that transmits the messages regularly.

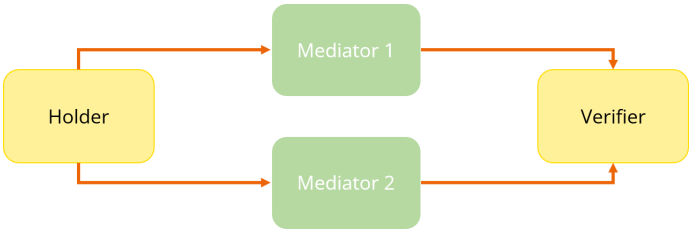


Fig. 8: The holder rotates between two or mediators for communication with the verifier.

Governance: Proper governance is an important non-technical measure to achieve desired privacy properties. Technical solutions support and sometimes technically enforce governance, but not everything can be supported by technical means. In practice, a significant

portion of the requirements are specified by specific policies and enforced by processes that involve human interaction and control. Examples of such policies include how long logs are retained, who has access to them, what attributes are required for employee IDs, external audits, etc.

6.2 Limitations

We only examined DIDComm v1 and AnonCreds. Although there are other communication protocols, such as OID4VC, and credential formats, such as SD-JWT, they were not included in this analysis. It is important to note that OID4VC does not offer a mediator service and always uses synchronous communication. However, if OID4VC is used during the presentation of AnonCreds, the findings – that the transport layer is the main concern – of this study should still apply. In such cases, the authorization request and response must be designed to ensure that no correlatable identifier is present. Most other credentials, such as JWT-SD, do not offer ZKP, only selective disclosure. However, if selective disclosure is used in a way that the presentation object does not contain any unique or correlatable identifier, then our findings should also hold. Nevertheless, this depends on the actual implementation, and future research may be required to confirm these assumptions. However, even then, there may still be a conflict between unlinkability and cloneability, which will be discussed below.

An important issue is the problem of clonability when using unlinkability. Although the credential itself is blinded and can only be read by the holder's link secret with the actual cryptographic signature, this does not prevent cloning of the generated proof and the holder's cryptographic key material. The CL signatures used for anti-cloning are not supported by common hardware security modules required for high assurance use cases [Yo22; YY23]. The issue of simultaneously achieving unlinkability and unclonability while using ZKP is still an open issue [AL19; Ku20; YY23].

6.3 Future Work

Future research could investigate the unlinkability revocation of previously used and randomized authorization proofs due to legal reasons, e.g., if an incident occurs. This presents a challenging task: Striking a balance between the main objective of preventing traceability while also considering the requirement for unlinkability revocation on demand. The recent work by Watanabe et al. [Wa24] may offer a potential starting point. As we showed not only the content of the presentation must be privacy-preserving but also the routing strategy. Mediators do offer privacy-preserving routing by utilizing herd privacy. However, if the mediators are not trustworthy, potential privacy issue arises. Thus, further research on privacy-preserving routing is required, e.g., by expanding upon the outlined strategies. Another way to prevent profiling is establishing rules and laws prohibiting it. In

our case, the governance of the mediator must be technically enforceable, e.g., machine-readable governance (as proposed by Hardman [Ha20]). Thus, a further area of research is privacy-preserving machine-readable governance. Another potential research topic is investigating the issue of revocation and traceability, which we did not examine. Although there has been research on the issue of traceability and revocation, it remains unsolved due to the trade-off between privacy and scalability.

7 Conclusion

To address the research question: *To what extent can SSI prevent employee tracing and profiling in the access management system of a building?* we designed and evaluated a PoC based on Hyperledger Aries, DIDComm v1, and AnonCreds. Our findings suggest that the problem of traceability and profiling in SSI is essentially a transport and not a payload load problem when using these technologies. Various strategies can mitigate this issue, such as using at least two chained mediators that are independently operated or regularly switching mediators. Thus, SSI is a promising tool to enable hospital access management without compromising the staff's privacy while complying with regulations and addressing the concerns of workers' councils and employees alike. The presented system design is a first step in this direction. However, several challenges, such as unlinkability revocation, remain.

Acknowledgement

This work is partially supported by projects from the Germany Federal Ministry of Education and Research (SEMECO-Q1, Grant no. 03ZU1210AA) and the German Federal Ministry of Economics and Climate Protection (Grant no. 01MN21001A). The author from Barkhausen Institute is also financed based on the budget passed by the Saxonian State Parliament.

References

[Al16] Allen, C.: The Path to Self-Sovereign Identity, 2016, URL: https://www.lifewithalacrity.com/2016/04/the-path-to-self-soverereign-identity.html.

[AL19] Arnold, R.; Longley, D.: Zero-Knowledge Proofs Do Not Solve the Privacy-Trust Problem of Attribute-Based Credentials: What if Alice Is Evil? IEEE Communications Standards Magazine 3/4, pp. 26–31, 2019.

[BRA24] Bochnia, R.; Richter, D.; Anke, J.: Self-Sovereign Identity for Organizations: Requirements for Enterprise Software. IEEE Access 12/, pp. 7637–7660, 2024.

[BTD20] Bhave, D. P.; Teo, L. H.; Dalal, R. S.: Privacy at Work: A Review and a Research Agenda for a Contested Terrain. Journal of Management 46/1, pp. 127–164, 2020.

[Ch85] Chaum, D.: Security without identification: transaction systems to make big brother obsolete. Communications of the ACM 28/10, pp. 1030–1044, 1985.

[Ču22] Čučko, Š.; Bećirović, Š.; Kamišalić, A.; Mrdović, S.; Turkanović, M.: Towards the Classification of Self-Sovereign Identity Properties. IEEE Access 10/, pp. 88306–88329, 2022.

[Ha20] Hardman, D.: Aries RFC 0430: Machine-Readable Governance Frameworks, 2020, URL: https://github.com/hyperledger/aries-rfcs/blob/main/concepts/0430-machine-readable-governance-frameworks/README.md.

[Ha21] Hardman, D.: Aries RFC 0046: Mediators and Relays, 2021, URL: https://github.com/hyperledger/aries-rfcs/blob/main/concepts/0046-mediators-and-relays/README.md.

[Kh19] Khateev, N.: Aries RFC 0037: Present Proof Protocol 1.0, 2019, URL: https://github.com/hyperledger/aries-rfcs/blob/main/features/0037-present-proof/README.md.

[Ku20] Kubach, M.; Schunck, C. H.; Sellung, R.; Roßnagel, H.: Self-sovereign and Decentralized identity as the future of identity management? In: Open Identity Summit 2020. Gesellschaft für Informatik e.V., Bonn, pp. 35–47, 2020.

[LC22] Lacity, M.; Carmel, E.: Implementing Self-Sovereign Identity (SSI) for a digital staff passport at UK NHS, University of Arkansas, 2022.

[NH24] NHS: NHS Digital Staff Passport, 2024, URL: https://beta.staffpassports.nhs.uk/, visited on: 01/26/2024.

[PH10] Pfitzmann, A.; Hansen, M.: A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management, 2010, URL: https://api.semanticscholar.org/CorpusID:150929990.

[Se22] Sedlmeir, J.; Barbereau, T.; Huber, J.; Weigl, L.; Roth, T.: Transition Pathways towards Design Principles of Self-Sovereign Identity. In: ICIS 2022 Proceedings. International Conference on Information Systems, Copenhagen, Denmark. P. 4, 2022.

[Wa24] Watanabe, K.; Kimura, Haruto, Morita, Kohei; Tsutsumi, O.; Muzahid, Z.: An Anonymous Door Unlocking with Anonymity Revocation, 2024, URL: https://devpost.com/software/an-anonymous-door-unlocking-with-anonymity-revocation, visited on: 02/01/2024.

[Yo22] Young, K.: Being "Real" about Hyperledger Indy & Aries / Anoncreds - Identity Woman, 2022, URL: https://identitywoman.net/being-real-about-hyperledger-indy-aries-anoncreds/.

[YY23] Young, K.; Yang, L.: Standards-Based Digital Credentials: Flavors Explained, 2023, URL: https://drive.google.com/file/d/1JwWLayXdJRxhGvO56G1UKEGt0R-6uuVm/view.

Evaluating the evaluation criteria for account-recovery procedures in passwordless authentication

Manuel Keil [1,3] and Alf Zugenmaier[1,2]

Abstract: Passwordless authentication avoids the weaknesses of password based authentication such as guessable passwords and password reuse. However, when passwordless authentication becomes impossible for the user, e.g. due to loss of the security token, an account recovery method has to be used. Kunke et al. [Ku21] analysed these recovery mechanisms in respect of criteria they extracted from the literature. However, these criteria in the literature were based on researchers' opinions and were not grounded in practical experience.

To achieve this grounding, semi-structured interviews were conducted with practitioners in various industries. These experts were asked to rate the existing criteria and contribute additional criteria if required. The result is a weighted list of criteria that can be used in future to evaluate account recovery procedures.

Keywords: passwordless authentication, account recovery, requirements evaluation

1 Introduction

Password-based authentication is vulnerable to a multitude of attacks, including phishing and dictionary attacks [PMA22; Ra12]. Initially password managers were deployed to mitigate these vulnerabilities[Bo12]. Password managers pose problems when they are to be used on different devices by the same user, as a synchronization of the password manager must be implemented or a device with the password manager needs to be carried at all times [Bo12]. The centralization of online password managers introduces additional security risks [Ar23]. Two-factor or multi-factor authentication making use of the factors knowledge, ownership and biometrics also mitigates the vulnerabilities of password based authentication, but still has the same usability issues. The logical next step is to move to passwordless authentication, which eliminates the factor knowledge. However, passwordless authentication shares a problem with password based authentication: it still requires a process for recovering account access in case of lost hardware tokens or biometric feature changes, e.g due to injury. Processes for account recovery already exist, and a few of them have been specifically designed for passwordless authentication, e.g. the recommended practices published by the FIDO alliance, like using multiple authenticators per account [GLS19]. This is necessary because the security advantages of passwordless authentication should not be undermined by the recovery process [Ku21]. According to Kunke et al. [Ku21]

1 Munich University of Applied Sciences, Munich, Germany,
2 Concordia University, Montreal, Canada,
3 ARCA-Consult GmbH, Pfaffenhofen a. d. Ilm, Germany,

doi:10.18420/OID2024_06

none of the established methods fulfill all their requirements, either in terms of usability or in terms of security. Kunke et al. draw their requirements and evaluation criteria from prior publications, which are mostly based on those defined by Bonneau et al. [Bo12], i.e. defined by researchers. As passwordless authentication is gaining traction in business contexts, the question to be addressed here is whether the requirements set out more than 10 years ago are still relevant and complete for business use cases.

To address these issues, a mixed methods study was conducted with experts from various industries to verify known criteria, identify new ones, and examine existing criteria definitions. The wide range of interviewees ensures applicability to a broad spectrum of industries. The results of these interviews were analyzed using a qualitative, category-guided text analysis [Ma10] to determine which of the already known criteria are no longer relevant for use in business environments, which criteria are still considered in principle but no longer fall under the definition of Kunke et al. [Ku21], and which additional, not yet documented criteria are of relevance. This paper also proposes further categorization for the catalog of criteria into general and optional criteria.

2 Theoretical Background

2.1 Definitions

The definitions used in this paper for passwordless authentication and account recovery are analogous to the ones given by Kunke et al. [Ku21]. For convenience of the reader, these are given in Appendix A. Likewise, the definitions of the three categories into which criteria are grouped, namely usability, deployability, and security, agree with those used by Bonneau et al. [Bo12] and Kunke et al. [Ku21], and are also given in Appendix A.

As this work is based on the evaluation criteria defined by Kunke et al. [Ku21], the reader is referred to that paper or to Appendix A, where these are summarized.

2.2 Related Work

The most important studies on this topic are the aforementioned ones by Kunke et al. [Ku21] and Bonneau et al. [Bo12], which defined criteria for evaluating authentication and recovery methods as well as proposed practices for evaluation. The criteria mentioned by those studies seem plausible, but are solely defined by the authors of those papers and not backed by input from practitioners. Nevertheless, their results and methodology for evaluation as well as the definitions for criteria are the basis for this study. Kunke et al. [Ku21] already collected possible criteria from multiple sources and reduced previously mentioned criteria to those related to account recovery with passwordless authentication. Other earlier papers like those by Saltzer and Schröder [SS75] Nielsen [Ni94], and Stajano [St11] proposed and evaluated criteria for security and usability of processes.

Gerlitz et al. [Ge23] measured deployment of various recovery mechanisms. They used their experience from these measurements to suggest enhancements to existing recovery mechanisms. However they stop short of defining actual criteria. This paper and another by Amft et al. [Am23], published almost concurrently, are very similar, as they both measured the deployment of recovery mechanisms and provided recommendations. Li et al. [Li22] are proposing a new recovery scheme, and are evaluating it in respect of security and performance. Because account recovery should be a procedure that is only rarely invoked, the performance evaluation seems superfluous. Wahab et al. [Wa21]) also propose inclusion of keystroke dynamics into the recovery procedure, and analyse it in respect of false positive and negative rate.

The closest work to this work is probably the thesis by Tiller [Ti20], which performs an online user study with questionnaires to determine user preferences of recovery methods for two-factor authentication, and in which correlations between preferences and various demographic factors are explored. The preferences in this work are structured slightly differently, and are often specific to individual recovery authentication methods, but are also covered by the criteria used in this paper. One novel preference from the two open questions is to not be intrusive to an emergency contact. This requirement is also specific to a particular recovery method and wouldn't be applicable in an enterprise setting.

It should be noted that in all the above work, the focus is on recovery for individual accounts. Recovery from mass compromise as analysed by Fritsch [Fr23] is not considered here.

3 Methodology

This study aimed to validate the criteria for the evaluation of account recovery by conducting interviews with experts in industry responsible for the fields of IT administration, digitalization, and controlling. These semi-structured interviews were set up according to guidance given in Helfferich [He22]. The experts were in positions that would be involved in procurement or deployment decisions on authentication solutions. The selection of the interview partners was trying to cover the three fields as well as trying to cover a variety of industries. The study aimed to avoid bias, by searching on Google for companies in relevant industry and size and then addressing the relevant person, and by searching in LinkedIn for potential interview partners, aiming at covering the different fields and industries. There may still be some self selection bias, as the response rate for interview requests was about 10%.

The interviews were conducted between 27 October 2023 and 06 December 2023 with seven experts[4] from three different fields in six companies in Germany. The interviews were conducted in different industries, including education, consulting, manufacturing, and banking. The size of companies ranged from 20 to over 100 000 employees. The interviews

4 This is on the low end of interview partners according to Creswell [Cr98], but still within the given range of 5-50 participants for qualitative research.

were conducted online as well as in person, with the majority of the interviews conducted via phone conference. With consent of the interview partners, the interviews were recorded and transcribed using an audio recorder and transcription software. The transcripts were later manually corrected and formatted.

The interviews serve to answer the research question of which criteria are relevant for the evaluation of account recovery processes for passwordless authentication in companies. The analysis is conducted using a mixed methods approach. The qualitative part is analyzed using category-oriented text analysis, focusing on the criteria that decision-makers in companies consider relevant, and can influence the final decision for a new procedure. The interview partners' responses are coded to provide additional information and a simple evaluation. The coding was performed by one person, with a second person performing spot checks to validate the coding. In the quantitative part the experts evaluated the relevance of the existing criteria on a 5 point Likert scale.

The interview process was structured according to a questionnaire which was made available to the participants ahead of time, so they could understand the definitions of the existing criteria. In part one of the interview, the interview partners were asked to describe their own experiences with account recovery, and to describe the current process for account recovery in their company. The interview partners were then asked if they have any experience with or whether the company is using or is planning to switch to passwordless authentication. The interviewees were also asked about optimizations that can be implemented to improve the process.

In part two of the interview, the criteria selected by Kunke et al. [Ku21] were evaluated and criteria are expanded upon. Each of the criteria was rated and the rating was usually given with justification. The scale for evaluation was from 0 to 4, with 0 indicating the criteria is irrelevant, and 4 indicating high relevance. As interview partners also gave a reason for their rating it was possible to determine whether the rating would be specific to the company or industry or whether they would need to be redefined to fit within the current requirements for authentication and account recovery.

The questionnaire was pre-tested to gauge the duration and whether there are issues with the questions which led to improvements before the actual interviews. The results from the pre-test were not included in the final analysis.

4 Results

4.1 Interview part 1: experience

The first part of the interview, in which experts were asked about current processes and their experiences with passwordless authentication and account recovery showed that:

1. The most commonly used form of account recovery at the moment is helpdesk-based, meaning the user is not able to recover their accounts on their own. In one case, the procedure for authenticating to help desk personnel is highly complex, even requiring governmental documents.
2. Most of the respondents would prefer a form of self-service for their account recovery procedures.
3. Passwordless authentication is rarely deployed in corporate environments. Where it is used the most common forms are platform authenticators like mobile phones or Windows Hello. However, most of the respondents have personal experience with passwordless authentication.
4. Companies, that do not offer passwordless authentication themselves but are using it externally usually do not have account recovery procedures defined for accounts with passwordless authentication.
5. Where security is of the essence, password managers, Single-Sign-On and multi-factor authentication are implemented as a more secure alternative to passwordless authentication.

For example, a statement supporting item 2 given by an interviewee when talking about his experience with passwordless authentication would be:

"[self-service] would of course be more practical, because as I said, the [help desk] resources would be wasted, which in the end, [...] will cost something"(Translated from German)

4.2 Interview part 2: relevance of criteria

In the second part of the interview, the experts evaluated the criteria defined by Kunke et al. [Ku21] and have a chance to propose what they believe to be missing. Table 1 provides the average importance of the individual criteria together with the responses standard deviation, which indicates how well the experts and the environments for which they would evaluate such a recovery mechanism are aligned. The scale to rate the criteria was: 0: irrelevant, 1 low relevance, 2: medium relevance, 3: relevant, and 4: very relevant.

To keep the length of this paper in check, we will only discuss the most and least important criteria further, as well as some honorable mentions.

- *Resilient-to-Phishing* was rated as very relevant by all interviewees, saying that it should certainly be able to withstand a phishing attack, attack surface in general should be minimized and that users don't always have the experience to recognize such an attack.
- *Resilient-to-Targeted-Impersonation* like the following two criteria was rated very relevant, calling targeted impersonation one of the greatest threats currently and

Criterion	Avg. rating	Std. dev.
Resilient-to-Phishing	4,00	0,000
Resilient-to-Targeted -Impersonation	3,86	0,378
Easy-to-Learn	3,86	0,378
Resilient-to-Leaks-from-other-Verifiers	3,86	0,378
Memorywise-Effortless	3,86	0,378
Negligible-Cost-per-User	3,71	0,756
Resilient-Theft	3,71	0,488
Resilient-to-Physical-Observation	3,57	0,535
Complete-Mediation	3,57	0,787
Resilient-to-Internal-Observation	3,57	0,787
Accessible	3,57	0,787
Scalable-for-User	3,57	0,787
Implemented	3,43	0,976
Work-Factor	3,43	0,787
No-trusted-Third-Party	3,43	1,134
Unlinkable	3,29	1,253
Physically-Effortless	3,14	1,464
Nothing-to-Carry	3,00	1,155
Requiring-Explicit-Consent	2,71	1,603
Browser-Compatible	2,57	1,272
Non-Proprietary	2,43	1,134
Match-System-to-Real-World	2,29	1,496
Open	2,00	1,527

Tab. 1: Average rating and std. dev. according to the experts

relating it to cyber-bullying. One expert said it is only relevant, justifying it by mentioning that not everything can be the most important.

- *Easy-to-Learn* was rated highly because multiple experts explained that a process that is easy to learn can reduce training costs and does not stress users with complexity. One expert rated the criterion as just relevant, but gave no explanation. Therefore it is likely just personal preference.
- *Resilient-to-Leaks-from-other-Verifiers* was rated just as highly with only one expert categorizing it as only relevant calling the criterion unattainable because attackers have access to highly sophisticated attacks and regularly leak data to the darknet. Others called it essential or repeated that it should not be possible to get user data from other service providers.
- *Non-Proprietary* is closely related to *Open* but considers the roll-out effort and costs more than security. With ratings ranging between 1 and 4 on the scale, there is no clear answer for the relevance of this criterion. For some, cost is of higher value than

security. Others say licensing costs are expected, and with free offers, the developer bears no liability.

- *Match-System-to-Real-World* even received ratings between 0 and 4, covering the whole scale. Generally the ratings were more on lower side, though two experts stated it would be very relevant. One said, an unknown process could unsettle the users and therefore rated it highly while the lowest ratings came with the explanation that it would not be necessary, and that the real world would change so often that processes would constantly need to be adjusted.
- *Open* received the lowest average rating of all criteria, with some experts giving similar reasoning to *Non-Proprietary*. High ratings were not accompanied by explanations while experts who gave lower ratings reasoned, that a proprietary, but certified system would be chosen just as likely. One even mentioned that open-source software would carry a negative impact since their cyber-insurance would not cover damages caused by open-source software.
- *Implemented* was viewed as relevant by most of the experts, with especially experts in large corporations reasoning, that software development is not their business and that self-developed software could also cause problems regarding the cyber-insurance. Rather surprisingly, experts from smaller companies expressed that hiring someone to develop this process would be acceptable as well.

For example, one expert from a large company gave the following statement in regards to the criterion *Implemented*:

"[...] we are not a software company [...]. I think there are other people who are better equipped for this and [we] are not [...] hiring someone here who only takes care for this topic."(Translated from German)

5 Discussion

The quantitative analysis of the expert opinions suggests that there are some mandatory requirements that a account recovery procedure needs to fulfill, while others could be considered optional. This is corroborated by the interviews. For this paper, we defined the requirements with a standard deviation of > 1 as optional, and with an average rating of < 2.5 as not relevant, which is reflected in Table 1.

The inductive analysis also allows to infer new criteria, which are defined below:

1. *Reliable (Usability):* The user can expect successful execution of the recovery procedure if followed properly. A procedure should receive a good rating if it is already widely used. (Mentioned by 1 expert)
2. *Ease-of-Use (Usability):* The procedure is designed not to complicate the user-experience and does not require additional devices. (Mentioned by 1 expert)

3. *Trusted-Vendor (Security):* The manufacturer of a possession factor like a hardware token can be trusted. Its adherence to security standards has been proven. (Mentioned by 1 expert)
4. *Cost-to-Benefit (Deployability):* The procedure offers more benefits for the same price or the same benefits at a lower price than a chosen reference procedure. (Mentioned by 3 experts)
5. *Resilient-to-Failure (Deployability):* The mechanism offers built-in redundancy as a measure against system failure or can be rolled out redundantly.(Indirectly mentioned by 1 expert)
6. *Regulatory-Compliant (Deployability):* The process meets the company's regulatory requirements. This includes, for example, security standards such as NIS2 [EU22], IT-GS (in Germany) [BS17] or the ISO27000 series [IS18a]. (Indirectly mentioned by 1 expert)

The criteria "Resilient-to-Failure" and "Regulatory-Compliant" were not defined explicitly by the interviewed experts. The above definitions for these two criteria were therefore generalized from the conversations. Newly added criteria are categorized as optional, since other experts could not confirm their relevance. This was due to the difficulty of acquiring experts willing to donate time for research, and due to the small number of new criteria.

6 Conclusion

The paper grounds evaluation criteria for account recovery mechanisms for enterprise use in requirements coming from practitioners in industry. It was possible to uncover several new requirements, as well as evaluate the importance of the requirements given in existing literature. The first part of the interview gave a snapshot of state of deployment of passwordless authentication and account recovery procedures at the time the interviews took place.

As might be expected, some requirements in literature are not considered to be as important as others. Interestingly, the criteria "open" seems to have a negative impact on insurability, or is at least perceived as such. This shows that the criteria in the category "deployability" may need to be extended further to include even more business related aspects as experience and interdependence with other aspects increase.

Future work in this area may include surveying how the relevance of criteria changes as companies collect more experience with passwordless authentication and the associated processes. A re-run of the interviews with a larger sample size and including the newly discovered criteria may result in a more conclusive ranking of the criteria. Additionally, collecting non-experts' views, i.e. users' views, may provide a different perspective.

References

[Am23] Amft, S.; Höltervennhoff, S.; Huaman, N.; Krause, A.; Simko, L.; Acar, Y.; Fahl, S.: "We've Disabled MFA for You": An Evaluation of the Security and Usability of Multi-Factor Authentication Recovery Deployments. In: Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security. CCS '23, Association for Computing Machinery, Copenhagen, Denmark, pp. 3138–3152, 2023, URL: https://doi.org/10.1145/3576915.3623180.

[Ar23] Arntz, P.: 1Password reports security incident after breach at Okta, 2023, URL: https://www.malwarebytes.com/blog/news/2023/10/1password-reports-security-incident-after-breach-at-okta.

[Bo12] Bonneau, J.; Herley, C.; Oorschot, P. C. v.; Stajano, F.: The Quest to Replace Passwords: A Framework for Comparative Evaluation of Web Authentication Schemes. In: 2012 IEEE Symposium on Security and Privacy. Pp. 553–567, 2012.

[BS17] BSI: IT-Grundschutz Methodik, BSI-Standard 200-2, Bundesamt für Sicherheit in der Informationstechnik, 2017, URL: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschutz/BSI_Standards/standard_200_2.pdf?__blob=publicationFile&v=2.

[Cr98] Creswell, J. W.: Qualitative inquiry and research design: choosing among five traditions. Sage Publications, Inc., 1998.

[EU22] EUROPEAN PARLIAMENT: Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive), Directive 2022/2555 OJ L333/80, 2022, URL: https://eur-lex.europa.eu/eli/dir/2022/2555/oj.

[Fr23] Fritsch, L.: Electronic identity mass compromize: Options for recovery. In: Open Identity Summit 2023. Gesellschaft für Informatik e.V., Bonn, pp. 141–146, 2023, ISBN: 978-3-88579-729-6.

[Ge23] Gerlitz, E.; Häring, M.; Mädler, C. T.; Smith, M.; Tiefenau, C.: Adventures in Recovery Land: Testing the Account Recovery of Popular Websites When the Second Factor is Lost. In. USENIX Association, pp. 227–243, 2023, ISBN: 978-1-939133-36-6, URL: https://www.usenix.org/conference/soups2023/presentation/gerlitz-adventures.

[GLS19] Gomi, H.; Leddy, B.; Saxe, D. H.: Recommended Account Recovery Practices for FIDO Relying Parties, FIDO Alliance, 2019.

[He22] Helfferich, C.: Leitfaden- und Experteninterviews. In (Nina, J. B.; Blasius, eds.): Handbuch Methoden der empirischen Sozialforschung. Springer Fachmedien Wiesbaden, pp. 875–892, 2022, ISBN: 978-3-658-37985-8, URL: https://doi.org/10.1007/978-3-658-37985-8_55.

[IS18a] ISO: Information technology — Security techniques — Information security management systems — Overview and vocabulary, Standard ISO 27000:2018, International Organization for Standardization, 2018, URL: https://www.iso.org/obp/ui/en/#iso:std:iso-iec:27000:ed-5:v1:en.

[IS18b] ISO: Ergonomics of human-system interaction, Standard ISO 9241-11:2018, International Organization for Standardization, 2018, URL: https://www.iso.org/obp/ui#iso:std:iso:9241:-11:ed-2:v1:en.

[Ku21] Kunke, J.; Wiefling, S.; Ullmann, M.; Lo Iacono, L.: Evaluation of Account Recovery Strategies with FIDO2-based Passwordless Authentication. In: Open Identity Summit 2021. Gesellschaft für Informatik e.V., Bonn, pp. 59–70, 2021, ISBN: 978-3-88579-706-7.

[Li22] Li, Y.; Chen, Z.; Wang, H.; Sun, K.; Jajodia, S.: Understanding Account Recovery in the Wild and its Security Implications. IEEE Transactions on Dependable and Secure Computing 19/1, pp. 620–634, 2022.

[LWS18] Li, Y.; Wang, H.; Sun, K.: Email as a Master Key: Analyzing Account Recovery in the Wild. In: IEEE INFOCOM 2018 - IEEE Conference on Computer Communications. Pp. 1646–1654, 2018.

[Ma10] Mayring, P.: Qualitative Inhaltsanalyse. In (Mey, G.; Mruck, K., eds.): Handbuch Qualitative Forschung in der Psychologie. VS Verlag für Sozialwissenschaften, pp. 601–613, 2010, ISBN: 978-3-531-92052-8, URL: https://doi.org/10.1007/978-3-531-92052-8_42.

[Ni94] Nielsen, J.: Enhancing the Explanatory Power of Usability Heuristics. In. Association for Computing Machinery, pp. 152–158, 1994, ISBN: 0897916506, URL: https://doi.org/10.1145/191666.191729.

[PMA22] Papathanasaki, M.; Maglaras, L.; Ayres, N.: Modern Authentication Methods: A Comprehensive Survey. AI, Computer Science and Robotics Technology 1/June, 2022, URL: https://doi.org/10.5772/acrt.08.

[Ra12] Raza, M.; Iqbal, M.; Sharif, M.; Haider, W.: A survey of password attacks and comparative analysis on methods for secure authentication. World Applied Sciences Journal 19/4, pp. 439–444, 2012.

[SS75] Saltzer, J.; Schroeder, M.: The protection of information in computer systems. Proceedings of the IEEE 63/9, pp. 1278–1308, 1975.

[St11] Stajano, F.: Pico: No More Passwords! In (Christianson, B.; Crispo, B.; Malcolm, J. A.; Stajano, F., eds.): Security Protocols XIX - 19th International Workshop, Cambridge, UK, March 28-30, 2011, Revised Selected Papers. Vol. 7114. Lecture Notes in Computer Science, Springer, pp. 49–81, 2011, URL: https://doi.org/10.1007/978-3-642-25867-1%5C_6.

[Ti20] Tiller, L. N.: Account Recovery Methods for Two-Factor Authentication (2FA): An Exploratory Study, Master Thesis in Psychology, Old Dominion University, 2020.

[Wa21] Wahab, A. A.; Hou, D.; Schuckers, S.; Barbir, A.: Utilizing Keystroke Dynamics as Additional Security Measure to Protect Account Recovery Mechanism. In: Proceedings of the 7th International Conference on Information Systems Security and Privacy - ICISSP. INSTICC, SciTePress, pp. 33–42, 2021, ISBN: 978-989-758-491-6.

A Appendix - Definitions from literature

Passwordless authentication is a primary authentication method that does not rely on knowledge. Security tokens, smartcards or biometrics are examples [PMA22]. Multi-factor authentication that use knowledge (e.g. password) as one factor are not considered passwordless.

Account Recovery is a process to recover access through an alternative authentication method (also called secondary or fallback authentication) when the primary user authentication method can't be used, e.g. due to lost security token or forgotten password, [Ge23; Ku21; LWS18]. Since Account Recovery is also a method of authentication on a system, it should be noted that the procedure needs to provide at least equivalent security compared to the primary authentication method.

Usability criteria represent improvements in efficiency and user satisfaction. [IS18b]. It considers the simplicity of the mechanism to provide users with a smooth, positive experience without additional support.

Deployability criteria consider the use of corporate resources for the deployment and maintenance of the resources required for the mechanism. The term "deployability" was coined by Bonneau et al. [Bo12]. The aim is to keep the necessary resources as low as possible so that companies of all sizes are able to use a considered process.

Security and its criteria address the method in terms of its resilience to various attacks as well as potential security risks that may arise from third parties or a closed system. In addition, this category also lists criteria dealing with privacy and data protection [Bo12].

Kunke et al. [Ku21] defined the following criteria, which are taken as the basis for this work. The suffixes indicate which category the criteria fall under: *U* for Usability, *D* for Deployability, and *S* for Security.

(1U) Memorywise-Effortless: The user does not have to remember an authentication secret.
(2U) Scalable-for-User: No additional burden is introduced when using the mechanism with hundreds of services.
(3U) Nothing-to-Carry: The user does not need to carry any additional physical item to use the recovery mechanism at any time.
(4U) Physically-Effortless: Users do not need to perform any physical activities during the process beyond pressing a button.

(5U) Easy-to-Learn: The mechanism is intuitively designed and thus easy to learn.
(6U) Match-System-to-Real-World: The access recovery mechanism is based on real world concepts. The user can operate it intuitively because it is based on real world operations.
(7D) Accessible: Users must be able to use this mechanism, even with physical limitations.
(8D) Negligible-Cost-per-User: The financial cost per user must be very low.
(9D) Browser-Compatible: Mechanism can be used with any standard web browser without installing additional plugins or other software.
(10D) Non-Proprietary: The mechanism can be used at no additional cost for royalties and are not protected by patents or other trade secrets.
(11D) Implemented: The mechanism must be implemented as a practical application. It must not exist only as a theoretical concept.
(12S) Resilient-to-Physical-Observation: Despite observing the user while using the mechanism, attackers fail to successfully legitimize themselves as the user.
(13S) Resilient-to-Targeted-Impersonation: The attacker can not impersonate the user to the mechanism with background knowledge, which he may be able to obtain, e.g., via social networks.
(14S) Resillient-to-internal-Observation: Despite intercepting user input at participating devices, e.g., smartphone or desktop PC, it is impossible for an attacker to imitate the user.
(15S) Resilient-to-Leaks-from-Other-Verifiers: A user uses other services that use the same or similar mechanism but whose data is made public. The attacker cannot impersonate the user with the obtained data at that service.
(16S) Resilient-to-Phishing: Attacker is able to fake a legitimate mechanism and convince the user to use the faked version, but cannot successfully impersonate the user to the service with the resulting data.
(17S) Resilient-to-Theft: Refers to mechanisms that require the factor possession in the form of an object as proof of legitimacy. If attackers gain possession of a user's object, they must not succeed in legitimizing themselves as the user to the mechanism.
(18S) No-trusted-Third-Party: The mechanism for checking the authorization of the access recovery process is not based on a third trusted party, which could have been taken over or manipulated by an attacker to become an untrusted party.
(19S) Requiring-Explicit-Consent: The access recovery mechanism must only be performed with the user's conscious consent. It must never be started accidentally or automatically.
(20S) Unlinkable: The information processed by this mechanism cannot be used to draw conclusions about what other services a user is using.
(21S) Open: The code or at least the functionality of the mechanism must be openly accessible to everyone.
(22S) Work-Factor: The mechanism should be designed in such a way that an attacker has to invest many resources to falsely successfully legitimize against the mechanism.
(23S) Complete-Mediation: The authorization to use the mechanism must be verified every time. It is not enough to assume that the person that operates the mechanism during an open session is the legitimate user.

GRAIN: Truly Privacy-friendly and Self-sovereign Trust Establishment with GNS and TRAIN

Martin Schanzenbach[1], Sebastian Nadler[2], Isaac Henderson Johnson Jeyakumar[3]

Abstract: Robust and secure trust establishment is an open problem in the domain of self-sovereign identities (SSI). The TRAIN [KR21] concept proposes to leverage the security guarantees and trust anchor of the DNS to publish and resolve pointers to trust lists from DNS. While the DNS is a corner stone of the Internet, its continued use is primarily a consequence of inertia due to its crucial function as the address discovery system for existing Internet services. Research and development in the area of SSI is — for the most part — green field. The choice of DNS as a core building block appears fainthearted given its open security issues. Recently, the IETF paved the way to experiment with alternative name systems in real world deployments by reserving the special-use top-level domain ".alt" in the domain name space [KH23]. This allows us to use alternative name systems such as the GNU Name System (GNS) [SGF23a] without intruding into the domain name space reserved for DNS. In this paper, we show how we can use the GNS as a drop-in replacement for DNS in TRAIN. We show how TRAIN-over-GNS (GRAIN) can deliver security and privacy improvements the security concept of TRAIN-over DNS and show that it is practically feasible with limited modifications of existing software stacks.

Keywords: SSI; Name System; Trust; Decentralization

1 Motivation

Almost since the dawn of the Internet, from SPKI [Yl99] to X.509 [Bo08], experts have been grappling with the concepts of identity and trust. A recent development is the "lightweight trust management infrastructure for self-sovereign identity" (TRAIN) [KR21] which allows participants to manage and securely resolve trust lists through DNS with security extensions (DNSSEC) [Ho23]. DNS is a mature corner stone of the Internet and consequently an obvious choice for a mechanism that provides a distributed directory that binds names to records.

[1] Fraunhofer AISEC, Lichtenbergstraße 11, 85748 Garching bei München, Germany
schanzen@aisec.fraunhofer.de
[2] Technische Universität München, Boltzmannstraße 3, 85748 Garching bei München, Germany
sebastian.nadler@tum.de
[3] Fraunhofer IAO, Nobelstraße 12, Stuttgart, Germany
isaac-henderson.johnson-jeyakumar@iao.fraunhofer.de

doi:10.18420/OID2024_07

TRAIN is an evolution of the LIGHTest project[4] and extends it with concepts from the world of SSI. It is currently finding its way into large-scale cloud platform designs such as the European GAIA-X Federation Services (GXFS)[5]. The GXFS TRAIN specifications mention the use of alternative name systems as "advanced concepts" to be researched. Considering that TRAIN is a mostly green field technology, this implies that it builds on DNS out of convenience rather than necessity. As assessed in RFC 8324 [Kl18], even after 40 years of patching, attempts to secure DNS fail to address critical security issues such as the lack of end-to-end security and query privacy, censorship, and centralization of root zone governance. In particular, adoption of DNSSEC is still a concern: The unavailability of DNSSEC in some name spaces is considered in TRAIN deployments such as GAIA-X where "[. . .] if DNSSEC is not available, it will use standard DNS."[6] Consequently, we want to pick up on this idea of using alternative name systems for TRAIN and investigate how the GNU Name System (GNS) [SGF23a] can alleviate current shortcomings.

For TRAIN, we can identify three security and privacy issues that arise from the use of DNS: a lack of (query) privacy, reliance on externally controlled trust anchors and disjoint trust domains.

Query leakage: Resolution of trust through DNS has potentially severe privacy implications: Neither DNS nor DNS with DNSSEC provides any kind of *query privacy* [Kl18]. This can lead to the exposure of trust relationships and service usage patters of users and with it, social patterns. A passive observer or DNS server administrator (or state actor [GWE15]) may be able to track users across services and interactions by analyzing the queried list of trusted entities. Protocol extensions for DNS such as DNS-over-HTTPS (DoH) attempt to alleviate this issue, but it each come with their own set of privacy issues in practice[7].

Externally managed root of trust: Trust verifiers in TRAIN may curate lists of trusted domains in order to specify trusted organizations and entities themselves, but the root of trust the allows them to securely resolve those names remains to be the DNSSEC trust anchor. One of the key value offerings claimed in the TRAIN concept is that it provides a decentralized and flexible trust model. However, the central root of the DNS stops this idea in its tracks as the DNS inherently relies on a centrally controlled trust anchor, which is an established security concern [No23, Kl18, Gr18b]. This trust anchor is managed not by the verifiers, but by the Internet Corporation for Assigned Names and Numbers (ICANN)[8]. In other words, verifiers have no ultimate say in which names are resolvable and thus available for trust establishment. Names that are not or no longer available for solution, cannot be trusted.

Disjoint trust domains: One aspect in which TRAIN extends LIGHTest is by allowing

[4] https://www.lightest.eu

[5] https://www.gxfs.eu/specification-phase-2/, https://www.gxfs.eu/download/10499

[6] https://www.gxfs.eu/download/10499

[7] https://labs.ripe.net/author/bert_hubert/centralised-doh-is-bad-for-privacy-in-2019-and-beyond/

[8] https://icannwiki.org/Root_Zone

pointers from the DNS to trust lists via decentralized identifiers (DIDs) [St22]. DIDs and their myriad of DID methods with varying security properties [SS23] also come with distinct trust models and, with exceptions, trust anchors. At the same time, DNSSEC brings its own, disjoint trust anchor, further diminishing the flexibility envisioned by TRAIN. In the end, regardless of the trust policies resolved using TRAIN through DNS and DIDs, the root of trust is always the DNSSEC root trust anchor. This situation can result in deployments of TRAIN that rely on a variety of disjoint trust domains, not to mention the significant technical debt incurred through the flexibly of supporting arbitrary DID methods. Consequently, in the process of trying to establish trust into an entity through TRAIN, at least two independent trust anchors are required to do so, raising the question of why one of the required trust anchors and systems cannot be used in its stead.

2 TRAIN-over-GNS

In this paper, we propose TRAIN-over-GNS (GRAIN): Using the GNU Name System (GNS) as defined in RFC 9498 [SGF23a] as a drop-in replacement for DNSSEC and show that it aligns better with the goals of TRAIN and provides better privacy guarantees, in particular query privacy for trust list resolution and truly sovereign and decentralized trust management and governance. GNS is a censorship-resistant name resolution protocol following a security and privacy by design approach. It provides also aims to provide some DNS compatibility in order to minimize migration efforts. In the following, we show how GRAIN can alleviate the issues that we identified in the previous section:

Query privacy: GNS is designed from the start with security and privacy in mind. As a consequence, a GNS zone is created by generating a zone key pair using any of the supported cryptographic schemes. At the time of writing, this includes ECDSA and EdDSA keys. Records in GNS are encrypted and – ideally – stored in a remote, decentralized key-value storage. The particularities of this storage are not defined by RFC 9498, but the GNS reference implementation uses an efficient and resilient Distributed Hash Table (DHT) [SGF23b] as record storage. The key used to store and find a record set in the remote storage is derived from a label and the zone public key. Similarly, the encryption key used to encrypt the record set is derived from the same information. This implies that both the knowledge of the name (or label) *and* the zone of a record set is necessary to find it in the remote storage and decrypt it. Privacy is further enhanced using signatures created from blinded private/public key pairs (cf. [ESS21]) as to not trivially disclose the zone that signed the encrypted record set. This applies to the provider or providers (in the case of a DHT) of the remote storage as well as to any passive observers of queries and responses in the network, realizing the property of query privacy.

Locally managed trust: In GNS, the concept corresponding to the DNS root zone is called “Start Zones”. The Start Zones in GNS map name suffixes to zone keys. This provide entry points for name resolution to a resolver implementation similarly to how DNS root servers do. An initial set of Start Zones is shipped as part of a GNS implementation and is

freely configurable and extensible by each user on the local host. This is a significant and important deviation from the strictly external governance of the DNS root zone. Both the contents of the DNS root zone itself, but also the cryptographic trust anchor in the case of DNSSEC cannot be modified by the user of the system. This is not a technical restriction but a restriction through governance and protocol standards. Consequently, the user, and in particular the verifier in the TRAIN context, is given no choice and no trust agility, it is a "take it or leave it" approach.

The GNS governance model primarily revolves around the "Start Zones" enhanced with domain- and application-specific enhancements or modifications. For example, it explicitly allows administrators (or verifiers) to harden their GNS resolvers by only having highly trusted zone keys in their Start Zone configurations. It also allows the configuration of completely private mappings, using zones that are not public. Even if a governance body like ICANN emerged for the default "Start Zones" shipped with GNS implementations they can always be overridden by the the users in the spirit of self-sovereignty. We argue that this approach is much more in line with the flexible trust model envisioned in TRAIN when comparing it to the governance and trust model of DNSSEC [Ho23].

Coherent trust domain: Establishment of a coherent trust domain in TRAIN is not so much achieved as a direct consequence from the use of GNS, but is a more general usage and deployment insight. The issue of disjoint trust domains results from the flexibilities offered by TRAIN: For example, a commonly chosen DID method due to its reliance on readily-available, tried and tested protocol stacks (DNS and HTTPS) is "did:web"[9]. In "did:web" the DID Document corresponding to any given DID is retrieved by contacting a HTTPS server after resolving a DNS domain name using DNS (without requiring DNSSEC). This implies that a PKIX trust establishment [Bo08] must take place to verify the identity of the web server and the associated domain name in DNS. Note how the root of trust for PKIX is not the same as the root of trust in DNSSEC.

In comparison, the use of GNS for both trust list resolution as well as storage allows verifiers and Trust Scheme Organizations (TSO) in the TRAIN design to use a homogeneous technology stack by employing DIDs that match the existing technology used for trust list resolution in a secure and trusted fashion. To facilitate this, a DID method specification for GNS [SS23, SSB23] already exists.

3 Implementation

GNS and its reference implementation are designed to be used as a drop-in replacement of DNS. As a result, almost all DNS resource records are supported in GNS. A study has shown [Gr18a] that users do not even notice if DNS or GNS is used for name resolution in browsers, demonstrating this feature.

[9] `https://w3c-ccg.github.io/did-method-web/`

However, while GNS records are binary compatible with DNS, some extensions and modifications to GNS and RFC 9498, were necessary: The peculiar handling in GNS of so-called *underscore* labels that also receive special treatment in DNS [Cr19] was insufficient. In GNS, a special record type called "BOX" is used to indicate underscore-prefixes in the format "_<port>._<protocol>.example.com". GNS resolvers, upon encountering a BOX record, try to validate and parse the name suffix accordingly and expect port and protocol descriptions mapping to the respective service and port numbers as defined in their respective IANA number registries[10]. However, as clarified in RFC 8552 [Cr19], underscore labels may have application-specific meaning not related to Internet protocols and numbers. This particular aspect is also used in TRAIN and consequently requires a minor update to the GNS protocol. Specifically, we introduced a new record type to support the flexible definition of underscore labels in GNS by specifying a new resource record type called "SBOX" for GNS[11].

We have implemented TRAIN-over-GNS by extending the reference implementation of GNS[12] and TRAIN[13]. The resulting limited changes in the TRAIN architecture are illustrated in Figure 1.

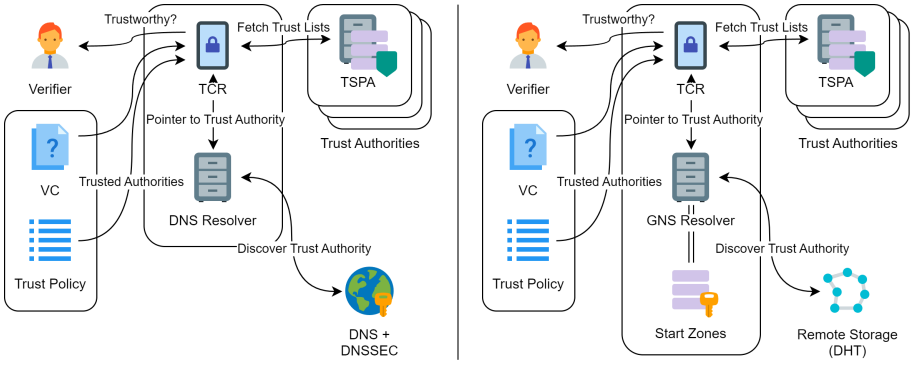


Fig. 1: Left: The original TRAIN overview with trusted content resolver (TCR) querying the DNS secured with DNSSEC and retrieving trust lists from the Trust Scheme Publication Authority (TSPA). Right: The modified TRAIN architecture where GNS replaces DNS.

The TRAIN "ZoneManager" component, which acts as a proxy layer between the components that publish trust-related information in the underlying name system, is modified to allow the administration of locally managed GNS zones instead of DNS zones. GNS is designed as a drop-in replacement for DNS which allowed us to reuse most of the semantics and processes that are also used for DNS zones. In particular, this eliminates the necessity of altering the components utilized by the Trust Scheme Organization (TSO) to handle their trust-related data, as they communicate with the name system through the ZoneManager.

[10] Example: `https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml`
[11] `https://lsd.gnunet.org/lsd0008`
[12] `https://git.gnunet.org/gnunet.git/`
[13] `https://gitlab.eclipse.org/snadler/grain`

The Trusted Content Resolver (TCR) is modified to optionally use the GNS resolver instead of a validating DNS resolver for name resolution and trust scheme discovery. Consequently, DNSSEC verification of the resource records is no longer required as GNS already ensures data integrity and authenticity. Our modifications are publicly available.

4 Related Work

The two currently used trust attestation, management and establishment technologies which most notably include X.509 [Bo08], OpenPGP [Fi07] identities and message formats. Both notably lack an efficient distribution mechanism.

Distribution and resolution of OpenPGP public keys and the associated management of trust levels through the Web of Trust is infamously cumbersome and user-unfriendly resulting in security-impacting skewed topologies [Ul11]. Generally, OpenPGP keys may be distributed through the DNS as well using the CERT record type [Jo06], but it is rarely used.

Commonly, X.509 trust anchors are distributed out of band in trust stores of browsers or HTTPS clients. Given the sorry state the SSL landscape is in [Ho11], the often difficult to prohibitively difficult to modify trust stores pose a significant threat to secure trust establishment. An alternative and alleviating strategy, which is arguably much less used at the time of writing, is a distribution of X.509 certificates through DNSSEC with DANE and the respective TLSA records [DH15] or in the CERT record type [Jo06].

Historically, SDSI/SPKI [Yl99] tried to address the issues identified in X.509 and PGP-based trust establishment but it never gained significant traction. The core ideas behind SDSI/SPKI lie in its reliance on local names which are not globally unique: Local starting points for trust resolution rooted in the user's social graph. Those ideas live on in the GNU Name System (GNS) [SGF23a], which we use in this paper to enhance the TRAIN concepts with the same notions of user self-sovereignty and self-governance that is found in SDSI/SPKI. Notably, SDSI/SPKI certificates themselves are also supported in the CERT record type [Jo06] of the DNS and, due to its compatibility, in GNS.

5 Summary and Future Work

We have demonstrated how GNS can replace DNS with DNSSEC as a secure, decentralized directory in TRAIN-over-GNS (GRAIN). The combination of GNS and TRAIN is also particularly advantageous with SSI support systems such as re:claimID [SBS18] that already make use of GNS. In the future, we plan to integrate GRAIN trust establishment for re:claimID [SBS18] and demonstrate its viability. We also believe that it is necessary to evaluate and categorize suitable DID methods for high security and privacy use cases that also call for GRAIN. In addition, we want to explore the possibility of managing trust lists directly in the directory structure of the name system, mitigating and bypassing the requirement of external storage systems and protocols such as web servers or DIDs.

References

[Bo08] Boeyen, Sharon; Santesson, Stefan; Polk, Tim; Housley, Russ; Farrell, Stephen; Cooper, David: , Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile. RFC 5280, May 2008.

[Cr19] Crocker, Dave: , Scoped Interpretation of DNS Resource Records through Ünder-scoredN̈aming of Attribute Leaves. RFC 8552, March 2019.

[DH15] Dukhovni, Viktor; Hardaker, Wes: , The DNS-Based Authentication of Named Entities (DANE) Protocol: Updates and Operational Guidance. RFC 7671, October 2015.

[ESS21] Eaton, Edward; Stebila, Douglas; Stracovsky, Roy: Post-quantum key-blinding for authentication in anonymity networks. In: LATINCRYPT 2021. Springer, pp. 67–87, 2021.

[Fi07] Finney, Hal; Donnerhacke, Lutz; Callas, Jon; Thayer, Rodney L.; Shaw, Daphne: , OpenPGP Message Format. RFC 4880, November 2007.

[Gr18a] Grothoff, Christian; Schanzenbach, Martin; Laube, Annett; Benoist, Emmanuel; Mainini, Pascal: , Decentralized Authentication for Self-Sovereign Identities using Name Systems, 10/2018 2018.

[Gr18b] Grothoff, Christian; Wachs, Matthias; Ermert, Monika; Appelbaum, Jacob: Towards Secure Name Resolution on the Internet. Computers & Security, 77:694–708, August 2018.

[GWE15] Grothoff, Christian; Wachs, Matthias; Ermert, Monika: NSA's MORECOWBELL: Knell for DNS. 2015.

[Ho11] Holz, Ralph; Braun, Lothar; Kammenhuber, Nils; Carle, Georg: The SSL landscape: a thorough analysis of the x. 509 PKI using active and passive measurements. In: Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference. pp. 427–444, 2011.

[Ho23] Hoffman, Paul E.: , DNS Security Extensions (DNSSEC). RFC 9364, February 2023.

[Jo06] Josefsson, Simon: , Storing Certificates in the Domain Name System (DNS). RFC 4398, March 2006.

[KH23] Kumari, Warren "Ace"; Hoffman, Paul E.: , The .alt Special-Use Top-Level Domain. RFC 9476, September 2023.

[Kl18] Klensin, Dr. John C.: , DNS Privacy, Authorization, Special Uses, Encoding, Characters, Matching, and Root Structure: Time for Another Look? RFC 8324, February 2018.

[KR21] Kubach, Michael; Roßnagel, Heiko: A lightweight trust management infrastructure for self-sovereign identity. Open Identity Summit 2021, 2021.

[No23] Nottingham, Mark: , Centralization, Decentralization, and Internet Standards. RFC 9518, December 2023.

[SBS18] Schanzenbach, Martin; Bramm, Georg; Schütte, Julian: reclaimID: Secure, Self-Sovereign Identities Using Name Systems and Attribute-Based Encryption. In: 2018 17th IEEE TrustCom/BigDataSE. pp. 946–957, 2018.

[SGF23a] Schanzenbach, Martin; Grothoff, Christian; Fix, Bernd: , The GNU Name System. RFC 9498, November 2023.

[SGF23b] Schanzenbach, Martin; Grothoff, Christian; Fix, Bernd: The R5N Distributed Hash Table. Internet-Draft draft-schanzen-r5n-03, Internet Engineering Task Force, August 2023. Work in Progress.

[SS23] Steele, Orie; Sporny, Manu: , DID Methods, August 2023.

[SSB23] Schanzenbach, Martin; Schwieren, Tristan; Bellebaum, Thomas: , The GNU Name System DID Method. LSD 0005, August 2023.

[St22] Steele, Orie; Sporny, Manu; Longley, Dave; Sabadello, Markus; Reed, Drummond; Allen, Christopher: , Decentralized Identifiers (DIDs) v1.0, July 2022.

[Ul11] Ulrich, Alexander; Holz, Ralph; Hauck, Peter; Carle, Georg: Investigating the openpgp web of trust. In: Computer Security–ESORICS 2011: 16th European Symposium on Research in Computer Security, Leuven, Belgium, September 12-14, 2011. Proceedings 16. Springer, pp. 489–507, 2011.

[Yl99] Ylonen, Tatu; Thomas, Brian; Lampson, Butler; Ellison, Carl; Rivest, Ronald L.; Frantz, William S.: , SPKI Certificate Theory. RFC 2693, September 1999.

Towards Building GDPR-Friendly Consent Management Systems on Top of Self-Sovereign Identity Ecosystems

Julia Schramm [1] and Tobias Eichinger [1]

Abstract: Consent is a legal basis that legitimizes the processing of personal data under the General Data Protection Regulation (GDPR). Implementing consent management systems in a GDPR-compliant fashion has proven difficult. A major pain point of current implementations is that users only have insufficient means to prove that they withdrew consent. Controllers can, therefore, plausibly deny having received a notification of consent withdrawal and it is thus at their discretion to continue the processing of personal data against the user's will. As a remedy, it has been proposed to log consent withdrawal events in blockchains to make them non-repudiable by controllers. This approach is typically at odds with the GDPR's fundamental principle of Storage Limitation. The issue is that a consent withdrawal event has to permit identification of the user who submitted it, yet only until the controller has received it. However, if they are logged in a blockchain, identification is possible indefinitely, as blockchains are append-only databases that do not facilitate deletion. In the paper at hand, we alleviate this issue and present work in progress on a consent management system in which users (i) give consent by issuing a verifiable credential to a controller and (ii) withdraw consent by revoking it. These two functions are natively provided in Self-Sovereign Identity (SSI) ecosystems.

Keywords: Consent Management System, User-centric, Self-Sovereign Identity, GDPR, Identity Management System, Storage Limitation

1 Introduction

Consent is one of six legal bases that legitimize the processing of personal data under the General Data Protection Regulation (GDPR) [Un16] (see Article 6(1) therein). Users give consent to a controller to legitimize the processing of their personal data by that controller and withdraw consent to delegitimize it. Note that no processing of personal data can be legitimized by consent. An empirical analysis indicates that out of all purposes for which website providers process personal data, only about 13% are based on consent [Ro20].

Consent is the only legal basis that users have control over. If personal data are processed under any of the remaining five legal bases, the user cannot counterfeit the processing of their data. A very prominent example of this is that banks base the processing of personal data for background checks on their clients on the legal basis of *legal obligation* in order to circumvent, for instance, the provision of a bank account to a malicious user.

Since consent is the only legal basis over which users do have control, it is all the more important to make sure that users can actually control, in particular on a technical level,

1 Technische Universität Berlin, Service-centric Networking, Ernst-Reuter-Platz 7, 10587 Berlin, Germany, julia.marie.schramm@campus.tu-berlin.de, https://orcid.org/0009-0003-5180-5212; tobias.eichinger@tu-berlin.de, https://orcid.org/0000-0002-8351-2823

doi:10.18420/OID2024_08

to whom and for what purposes they give and withdraw their consent. It is, therefore, that Article 4(11) GDPR mandates that consent be "*freely given, specific, informed and unambiguous*".

Current consent management systems undermine user control. To date, consent management systems typically do not provide sufficiently easy-to-understand and easy-to-use control mechanisms. Users on the Web for instance are nudged into giving consent or giving consent to more purposes than necessary on a day-to-day basis (see for instance [No20; So20]). Once given, withdrawing consent typically requires considerable cognitive effort by the user and renders consent withdrawal impossible in practice.

The difficulties in withdrawing consent become immediately evident in the example of withdrawing consent for an online newsletter. Users have to be able to (a) prove that they own an email address, (b) contact the controller of the online newsletter, and (c) withdraw their previously given consent in a non-repudiable manner. All three issues can immediately be addressed in so-called *Self-Sovereign Identity* (SSI) ecosystems.

In the paper at hand, we propose a consent management system that is built on top of an SSI ecosystem and centers around the control of users over their consent. The benefit of our proposed system in comparison with existing systems is that it better adheres to the fundamental principle of Storage Limitation as defined in the GDPR, which mandates that personal data shall be stored in a form that permits identification of the user to whom they relate no longer than is necessary (see Article 5(1)(e) GDPR). Before we describe our proposed system, we first recapitulate on existing consent management systems.

2 Related Work

Consent management systems collect and store consent from users in order to legtimize the processing of their personal data by a controller. Traditional consent management systems store consent in a database of the controller. It is clear that, in this case, users do not have immediate control over their own consent [Al19]. As a remedy, so-called "*user-centric*" consent management systems have been proposed [Ag20; Al20; Me21].

User-centric consent management systems center around the idea of giving users control over their own consent by controllers exposing an interface through which users can withdraw their consent stored in the database of the controller. User control is, however, limited by the trustworthiness of the controller, as users cannot make sure that their consent has been deleted after having withdrawn their consent. This is the inherent trust issue between users and controllers in consent management systems.

In order to allow users to make sure that controllers delete their consent when they withdraw it, it has been proposed to use blockchains to log the giving and withdrawing of consent. The idea is that users cannot deny having given consent, and, vice versa, controllers cannot deny that users have withdrawn their consent since blockchains are immutable append-only

databases. In other words, blockchains render the giving and withdrawing non-repudiable for both users and controllers alike.

It is not permissive to persist any personal data in a blockchain under the GDPR whatsoever, even if the personal data are hashed [Eb21]. Therefore, giving and withdrawing consent by means of logging events in a blockchain can only be lawful if the logs do not represent personal data. We see that logging consent events in a blockchain solves the trust issue between users and controllers since blockchains provide non-repudiation. Such logging is, however, unlawful under the GDPR since logs represent personal data.

We see that the main problem that user-centric consent management systems face is how to establish non-repudiation of consent giving and withdrawing without logging personal data in a blockchain. We describe such a system in the following.

3 Concept

We propose a user-centric consent management system that does not persist personal data in a blockchain. We build it on top of an SSI ecosystem as shown in Figure 1. Since a general description of SSI and SSI ecosystems goes beyond the scope of the paper, we assume that the gentle reader is familiar with the underlying concepts and kindly refer to [Al16; Ca08; Ma12; TR16] for conceptualizations of SSI and to [GMM19; Mü18; Su21] for characterizations of SSI ecosystems. We begin by mapping the roles of consent management systems to those of SSI ecosystems.

3.1 Mapping Roles from Consent Management Systems to SSI systems

Consent management systems feature the following four distinct roles:

- **Data Subject**: The individual to whom the data to be processed relate and who gives and withdraws consent for the processing thereof.
- **Controller**: The organization that manages data collection, storage, and processing. Controllers do not necessarily process personal data themselves.
- **Processor**: The organization that collects consent from users and processes data about that user. Processing activities are typically recorded for auditing purposes.
- **Auditor**: Verifies whether processing is only performed in the presence of consent and only under the terms specified in that consent. Auditors can be internal auditors of the controller or external (typically data protection authorities).

For simplicity, we refer to data subjects as users and assume that personal data are processed by the controller. We then map the three roles of *user*, *controller*, and *processor* to the usual roles of *issuer*, *verifier*, and *holder* in SSI ecosystems.

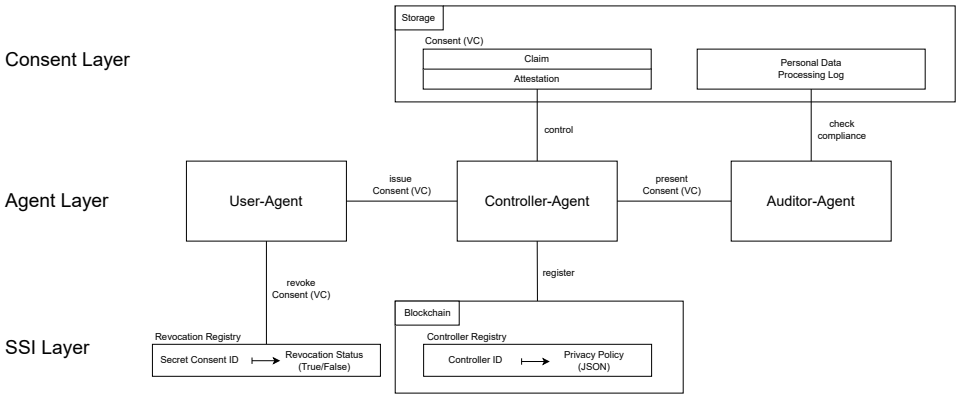


Fig. 1: The architecture of our proposed consent management system, which is built on top of an SSI ecosystem. The diagram is an extended adaptation of Mühle et al. [Mü18]'s SSI reference architecture.

- **Data Subject** $\mapsto$ **Issuer**: Creates and issues Verifiable Credentials (VCs).
- **Controller** $\mapsto$ **Holder**: Stores and presents VCs to the auditor.
- **Auditor** $\mapsto$ **Verifier**: Verifies the validity of VCs.

From the two role mappings above, it is clear that we want to represent the consent of a user as a VC. In order to do so, we can make use of domain-specific ontologies [Ku21] or goal-oriented modeling languages [PS18]. Note that by digitally signing the VC for issuance, the user digitally testifies to the controller that consent was "*freely given, specific, informed and unambiguous*". It remains to describe the workflows.

3.2 Mapping Workflows from Consent Management Systems to SSI systems

We describe the workflows for consent collection, withdrawal, and audit on the basis of the architecture shown in Figure 1.

- **Consent Collection**: The controller agent requests consent from the user agent with respect to his privacy policy persisted in the controller registry. The user agent issues a verifiable credential with respect to the credential schema defined by the privacy policy to the controller agent who persists the VC in his database. The user agent also assigns a secret consent identifier specific to the issued VC and forwards it to the controller agent. The secret consent identifier can be used to check the revocation registry to see whether the given consent has been revoked.
- **Consent Withdrawal**: The user agent revokes a previously issued VC in the revocation registry such that the controller agent can look up whether it has been revoked with the help of the secret consent identifier.

- **Audit**: The auditor agent accesses the controller's personal data processing log to understand whose personal data have been processed on the basis of consent, when and how. For each processing event, the auditor checks whether the user to whom the processed personal data relate has given consent by means of verifying the VC that has been issued by the user and, if so, checks via the revocation registry whether that consent is still valid.

A fundamental design property of our proposed architecture, which distinguishes it from prior work, is that it pushes the burden of proof away from data subjects to controllers. In other words, it is no longer the data subjects who must prove to the auditor that they have approached a controller and withdrawn their consent, and the processing of their personal data is thus no longer lawful. Instead, it is now the controllers who must prove to the auditor that the data subjects have not withdrawn their consent through the revocation registry and that the processing of personal data by them is thus still lawful. We see that the revocation registry is at the heart of our proposed architecture.

The revocation registry must:

- (R1) be highly available for both users and controllers,
- (R2) be append-only,
- (R3) circumvent linking the logged revocation of a VC and the user who issued it.

In order to fulfill Requirements (R1) and (R2), we make use of an append-only distributed hash table. Users revoke their consent by writing a key-value pair onto the distributed hash table, where keys are hash values of the secret consent identifier that has been chosen by the user when issuing the VC and values as the hash values thereof. Users make use of the secret consent identifier, as every user should only be able to revoke their own consent. Note that controllers are also technically able to withdraw their users' consent, which we assume not to be the case as there typically is no incentive to do so.

In order to fulfill Requirement (R3), we allow users to write in the distributed hash table under pseudonyms. We further make sure that users choose their secret consent identifiers independently from any personal information in order to mitigate the chances of linking the withdrawal of consent to the user.

So far, we have only described the bare bones of our proposed consent management system. There remain several issues with its implementation that stand in the way of broad adoption and use by both users and controllers, which we discuss in the following.

4 Discussion

At the time of writing, we have only just begun to implement our proposed user-centered consent management system. In this section, we first discuss some of the technical issues

we face in implementing our proposed system. We then discuss some technical aspects of data protection of our proposed system.

4.1 Implementation Issues

In order to make our proposed consent management system available for a wide range of users and controllers, we require a standardized interface for describing and communicating consent. Semantic ontologies that describe consent form the basis for the creation of such a standard interface [Ku21]. Notably, the *Data Protection Vocabulary* (DPV) is a semantic ontology that allows to serialize consent in the JavaScript Object Notation (JSON) [Pa19]. The description of consent in JSON format can immediately be used to describe consent as a VC. However, since there exist multiple alternative semantic ontologies, it remains to be seen which will find adoption and arise to an industry standard.

As of late, a multitude of agent software and SSI ecosystems have emerged that support the issuance and verification of VCs. Among the most popular SSI frameworks are *Sovrin* [Fo19; Fo20], *uPort* [uP20], *Civic* [Mo18], *IDunion* [ID23] and *Gaia-X* [GA22; GA23]. Both agent software and SSI ecosystems will only be adopted by the majority of controllers if large controllers invest in them and start realizing use cases. Furthermore, committee work will be required to make agents support the consent interfaces and standards mentioned in the previous paragraph.

Last but not least, user acceptance is crucial for the widespread usage of user-centered consent management systems [MTC21]. This realization is, in particular, one of the findings of *EnCoRe* (Ensuring Consent and Revocation), a European proposal architecture developed by *HP Laboratories* in 2011, which aimed to model consent and revocation management in a way that users have control over their consent and can be leveraged by controllers [CSP12]. They report on the necessity to onboard and educate users about how user-centric consent management can be used and how it impacts their privacy.

4.2 Towards Better Technical Data Protection

Article 5(1)(e) GDPR characterizes Storage Limitation as: "*personal data shall be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed*". This implies that any system must either allow personal data to be deleted or transformed in such a way that they no longer facilitate identification of the user to whom they relate. For system developers, this raises the question of whether and to what extent append-only databases such as blockchains can be used to enable use cases in which personal data are used.

Our proposed user-centric consent management system encapsulates the personal data that are necessary to describe consent given by a user to a controller into a VC. The issuance of

VCs does not require the logging of personal data in a blockchain, as the digital signature of the VC by the users ensures non-repudiation, that is users cannot refute that they gave consent after they have applied their digital signature. The verification of VCs in our proposed system, however, foresees the logging of personal data in a revocation registry.

The key-value pairs that represent consent withdrawal technically represent personal data since they are linked to VCs that hold personal data. In order to see this, recall that the controller to whom a VC is issued knows about the link between the secret consent identifier and the VC that contains personal information, such as for which purposes the user has given consent. Now, if the controller sees that the secret consent identifier has been used in the revocation registry to revoke the VC, the controller understands which user withdrew their consent. This would not be possible if the controller had deleted the VC.

Controllers have to delete all personal data associated with consent, including the VC, as soon as consent is withdrawn. In the context of the previous paragraph, we see that after users withdraw their consent by writing a key-value pair in the revocation registry, the key-value pair no longer facilitates identification of the user as it no longer points to anything. We emphasize that the key-value pair still represents personal data yet does not facilitate identification anymore and, in this sense, caters to the Storage Limitation principle. This is the benefit that our proposed system provides compared to current user-centric consent management systems in terms of technical data protection.

Data protection hardliners are likely to disfavor any form of revocation registry in which personal data are stored, even if the can be transformed to no facilitate identification as described in the previous paragraph. An alternative to using a revocation registry represents the use of expiring consent. Users would not need to log the revocation of VCs since they lose validity after a given date or period. Expiring consent is, to the best of our knowledge, never used in practice. Even if they were used, Custers concludes that expiry dates for consent would not be applicable in the age of Big Data [Cu16].

5 Conclusion

We describe a user-centric consent management system that, in contrast to current user-centric consent management systems, caters to the GDPR's principle of Storage Limitation. Current systems are at odds with the Storage Limitation principle since they log the giving and withdrawing of consent in blockchains in order to make the giving of consent non-repudiable by users and the withdrawing of consent non-repudiable by controllers. Logging such consent events goes against the GDPR's Storage Limitation principle since consent events represent personal data that can neither be deleted nor transformed in a way that no longer facilitates identification of the user who withdrew consent.

Our proposed user-centric consent management system is built on top of an SSI ecosystem and circumvents logging consent events on blockchains. Users give consent by means of

issuing a verifiable credential to a controller and withdraw their consent by revoking the issued verifiable credential through a revocation registry. Although our proposed revocation registry makes use of personal data, our proposed system can make it such that the personal data on the revocation registry to no facilitate identification of the user who withdrew consent. It is due to this property that our proposed system caters to the Storage Limitation principle, while current systems do not.

6 Acknowledgement

This research was partially funded by the German Federal Ministry of Economics and Climate Protection (BMWK) in the frame of the IDunion showcase project. The authors would like to thank Wassim al Shami and Muhammad Mohsin Nisar for their help in collecting and reviewing the papers that formed the basis of our paper.

References

[Ag20] Agarwal, R. R.; Kumar, D.; Golab, L.; Keshav, S.: Consentio: Managing consent to data access using permissioned blockchains. In: 2020 ieee international conference on blockchain and cryptocurrency (icbc). IEEE, pp. 1–9, 2020.

[Al16] Allen, C.: The path to self-sovereign identity. Life with Alacrity/, 2016.

[Al19] Aldred, N.; Baal, L.; Broda, G.; Trumble, S.; Mahmoud, Q. H.: Design and implementation of a blockchain-based consent management system. arXiv preprint arXiv:1912.09882/, 2019.

[Al20] Albanese, G.; Calbimonte, J.-P.; Schumacher, M.; Calvaresi, D.: Dynamic consent management for clinical trials via private blockchain technology. Journal of Ambient Intelligence and Humanized Computing 11/, pp. 4909–4926, 2020.

[Ca08] Cameron, K.: A user-centric identity metasystem. Microsoft Corp/, 2008.

[CSP12] Casassa Mont, M.; Sharma, V.; Pearson, S.: EnCoRe: Dynamic Consent, Policy Enforcement and Accountable Information Sharing within and across Organisations. HP Laboratories Technical Report/, 2012.

[Cu16] Custers, B.: Click here to consent forever: Expiry dates for informed consent. Big Data & Society 3/1, p. 2053951715624935, 2016.

[Eb21] Ebbers, F.; Karaboga, M.; Bremert, B.; Bile, T.; Ochs, C.; Meier, Y.; Weiler, S.: Datenschutz in der Blockchain. White Paper./, ed. by Friedewald, M., 2021.

[Fo19] Foundation, T. S.: Sovrin Glossary V2./, Accessed on December 10, 2023, 2019, URL: %5Curl%7Bhttps://sovrin.org/wp-content/uploads/Sovrin-Glossary-V2.pdf%7D.

[Fo20] Foundation, T. S.: Innovation Meets Compliance Data Privacy Regulation and Distributed Ledger Technology./, 2020, URL: %5Curl%7Bhttps://sovrin.org/wp-content/uploads/GDPR-Paper_V1.pdf%7D.

[GA22] GAIA-X: GAIA-X Secure and Trustworthy Ecosystems With Self Sovereign Identity./, 2022, URL: %5Curl%7Bhttps://gaia-x.eu/wp-content/uploads/2022/06/SSI_White_Paper_Design_Final_EN.pdf%7D.

[GA23] GAIA-X: Gaia-X European Association for Data and Cloud AISBL, Accessed on December 10, 2023, 2023, URL: %5Curl%7Bhttps://gaia-x.eu/who-we-are/association/%7D.

[GMM19] Grüner, A.; Mühle, A.; Meinel, C.: An integration architecture to enable service providers for self-sovereign identity. In: 2019 IEEE 18th International Symposium on Network Computing and Applications (NCA). IEEE, pp. 1–5, 2019.

[ID23] IDUnion: IDUnion - About the Project./, Accessed on December 10, 2023, 2023, URL: %5Curl%7Bhttps://idunion.org/projekt/?lang=en%7D.

[Ku21] Kurteva, A.; Chhetri, T. R.; Pandit, H. J.; Fensel, A.: Consent through the lens of semantics: State of the art survey and best practices. Semantic Web/Preprint, pp. 1–27, 2021.

[Ma12] Marlinspike, M.: What is Sovereign Source Authority./, Accessed on December 10, 2023, 2012, URL: %5Curl%7Bhttp://www.moxytongue.com/%202012/02/what-is-sovereign-source-authority.html%7D.

[Me21] Merlec, M. M.; Lee, Y. K.; Hong, S.-P.; In, H. P.: A smart contract-based dynamic consent management system for personal data usage under GDPR. Sensors 21/23, p. 7994, 2021.

[Mo18] Mountain, P.: Working together for better self-sovereign identity: civic, selfkey, and peer mountain./, Accessed on December 10, 2023, 2018, URL: %5Curl%7Bhttps://medium.com/peermountain/working-together-for-better-self-sovereign-identity-civic-selfkey-and-peer-mountain-282bca9a8e4a%7D.

[MTC21] Mahula, S.; Tan, E.; Crompvoets, J.: With blockchain or not? Opportunities and challenges of self-sovereign identity implementation in public administration: Lessons from the Belgian case. In: DG.O2021: The 22nd Annual International Conference on Digital Government Research. DG.O'21, Association for Computing Machinery, Omaha, NE, USA, pp. 495–504, 2021, ISBN: 9781450384926, URL: https://doi.org/10.1145/3463677.3463705.

[Mü18] Mühle, A.; Grüner, A.; Gayvoronskaya, T.; Meinel, C.: A survey on essential components of a self-sovereign identity. Computer Science Review 30/, pp. 80–86, 2018, ISSN: 1574-0137, URL: %5Curl%7Bhttps://www.sciencedirect.com/science/article/pii/S1574013718301217%7D.

[No20] Nouwens, M.; Liccardi, I.; Veale, M.; Karger, D.; Kagal, L.: Dark patterns after the GDPR: Scraping consent pop-ups and demonstrating their influence. In: Proceedings of the 2020 CHI conference on human factors in computing systems. Pp. 1–13, 2020.

[Pa19] Pandit, H. J.; Debruyne, C.; O'Sullivan, D.; Lewis, D.: GConsent-a consent ontology based on the GDPR. In: The Semantic Web: 16th International Conference, ESWC 2019, Portorož, Slovenia, June 2–6, 2019, Proceedings 16. Springer, pp. 270–282, 2019.

[PS18] Peixoto, M.; Silva, C.: Specifying privacy requirements with goal-oriented modeling languages. In. Pp. 112–121, 2018.

[Ro20] Roßnagel, A.; Bile, T.; Nebel, M.; Geminn, C.; Karaboga, M.; Ebbers, F.; Bremert, B.; Stapf, I.; Teebken, M.; Thürmel, V., et al.: White Paper Einwilligung./, 2020, URL: %5Curl%7Bhttps://epub.ub.uni-muenchen.de/95296/1/Whitepaper-Einwilligung.pdf%7D.

[So20] Soe, T. H.; Nordberg, O. E.; Guribye, F.; Slavkovik, M.: Circumvention by design - dark patterns in cookie consent for online news outlets. In: Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society. NordiCHI '20, Association for Computing Machinery, Tallinn, Estonia, 2020.

[Su21] Sury, U.: Die datenschutzrechtlichen Aspekte der Self-Sovereign Identity, 2021.

[TR16] Tobin, A.; Reed, D.: The inevitable rise of self-sovereign identity. The Sovrin Foundation 29/2016, p. 18, 2016.

[Un16] Union, E.: REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016, URL: %5Curl%7Bhttps://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%5C%3A32016R0679%7D.

[uP20] uPort.me: uPort identity System./, Accessed on December 10, 2023, 2020, URL: %5Curl%7Bhttps://www.uport.me/%7D.

A Trust Registries Enrollment Tool Supporting Decentralized Ecosystem Governance: Use Case Healthcare

Isaac Henderson Johnson Jeyakumar [1], Michael Kubach[1], Juan Vargas[2], and John Walker[3]

Abstract: Decentralized governance models have gained prominence in business ecosystems. These require trust, transparency, and collaboration among diverse stakeholders. Trust registries play a pivotal role in ensuring the integrity and authenticity of participants within these decentralized networks. However, the enrollment process presents challenges such as identity verification and reputation assessment. This paper introduces a Trust Registries Enrollment Tool (TRET) to facilitate the process. It simplifies procedures, strengthens trust, and enables secure and efficient participation in an ecosystem. This paper outlines its architecture, technical implementation, and potential impact. The practical use case is COVID19 certificate providers, highlighting its transformative potential for decentralized governance in healthcare and beyond.

Keywords: Decentralized Governance, Trust Registry, Enrollment, COVID-19 certificates, Trust frameworks, Trust ecosystems

1 Introduction

Facilitated by information technology that has increasingly enabled interorganizational cooperation, organizational networks have emerged. Such business ecosystems comprise suppliers, customers, competitors, and other stakeholders. Moreover, the use of information technology, Internet of Things, and other developments, have enabled decentralization of these ecosystems [Wi19], which can also be a result of an equal power structure among its participants. Central coordination in ecosystems isn't always necessary nor optimal. According to [Sc18], governance is central to coordinating the interactions between participants in an ecosystem and based on their analysis characterized by mechanisms such as governance structure, resources & documentation, accessibility & control, trust & perceived risk, pricing, and external relationships. In this context we focus further on trust registries. Trust registries facilitate the determination of the authenticity and authorization of entities in an ecosystem and are an integral component for the practical implementation of governance mechanisms. They enable governing authorities to specify actions authorized for governed parties (e.g., issuers of credentials, trusted service providers), and to validate whether entities are authorized to be part of the ecosystem as well as to perform defined actions. Technical implementations of trust

[1] Fraunhofer IAO, Nobelstraße 12, 70569 Stuttgart, Germany firstname.lastname@iao.fraunhofer.de
[2] University of Stuttgart, Institute of Human Factors and Technology Management IAT, Allmandring 35, Stuttgart, 70569, juan.vargas @iat.uni-stuttgart.de
[3] SemanticClarity, 44 Bonnie Lane, CA 94708 Berkeley, USA jwalker@semanticclarity.com

doi:10.18420/OID2024_09

registries can come in different forms, for example as signed .xml trust lists [EC23] or in ledger-based smart contract registries [EB23].

For this paper, healthcare provides an ideal environment to showcase the challenges and our approach to supporting decentralized ecosystem governance. Effective governance is essential for ensuring quality, safety, and efficiency in healthcare delivery. However, the complex nature of the healthcare ecosystem with healthcare providers, regulators, insurers, researchers, and patients, presents challenges to governance processes. The evolution of smart healthcare has allowed for the implementation of intelligent diagnostic and clinical support systems, data management systems, telemedicine, digital vaccination certificates and many other tools [GL19]. Despite all these advances, the processes for accessing health services, payment and reimbursement, and interaction with health professionals have largely remain unchanged. The constraints of traditional means of accessing health services as well as inflexible and siloed governance rules has prompted a lack of trust, and promoted access inequalities to health services [Bl23] [MS23]. Finally, the COVID-19 pandemic highlighted the need for efficient coordination and collaboration among various actors. Enrolling and managing these entities quickly and effectively with decentralized governance is essential for an efficient response [Gr22].

This paper contributes to the research and development on decentralized governance of business ecosystems by proposing a novel approach that leverages an enrollment tool for trust registries called TRET. We provide an architecture with implementation of the enrollment tool for trust registries exemplary in the healthcare ecosystem. For this, the remainder of the paper is structured as follows. Section 2 introduces decentralized governance practices in the healthcare ecosystem and different enrollment tools. Section 3 provides the architecture of the enrollment tool. Subsequently, in Section 4 we briefly describe how the solution could be applied for COVID-19 Vaccine Service Providers as the specific use case, and discuss the implementation. In section 5 we evaluate the TRET and other tools and then conclude the paper.

2 State of the Art

This section addresses the state of the art of decentralized governance practices in the healthcare ecosystem, initiatives for healthcare ecosystem governance, and enrollment tools for trust registries. This serves as a foundation for presenting the trust registries enrollment tool in the subsequent section.

2.1 Decentralized Governance Practices in the Healthcare Ecosystem

The following overview of several studies shows how decentralized governance practices in the healthcare ecosystem are transforming the way healthcare systems operate and are managed. Traditional healthcare models often rely on centralized structures, which can lead to inefficiencies, lack of transparency, and limited patient involvement. In contrast,

decentralized governance models aim to distribute decision-making power, foster collaboration, and empower various ecosystem stakeholders.

[SS19] provides a detailed study of the effect of decentralization or centralization of governance of health services on access to health care, utilization of health services, population health and other outcomes of interest. The study points out that health system decentralization varies in Europe, East Asia and North America depending on the political and public administrative structure of countries and the organization of the health system itself. While the paper focuses on finance, service delivery and organization, the digitization and interoperability aspects between different countries are not covered. [ABB19] present a systematic analysis of mechanisms and contextual factors for policymakers and implementers to pay attention to in their efforts to maximize the positive impact of decentralized governance in the healthcare ecosystem. The context–mechanism–outcome (CMO) configuration findings are documented, and it is explained how decentralization influences health system equity, efficiency, and resilience. Particularly, this paper points out how socio-economic context-based factors play a role in the efficiency of the execution of decentralized governance. [Ra23] present a decentralized platform for COVID-19 vaccinations in Germany. The proposed EU-GDPR-compliant platform model connects various actors and enables them to involve, conduct, and track the vaccination process. The open and decentralized platform model illustrates the potential for facilitating international interconnectivity and therefore the management of future global pandemics or another global health-related crisis. But it is also mentioned that the platform might require additional actors with functionalities depending on the socio-economic region when deployed internationally.

2.2 Related Projects and Initiatives in the Use Case Context

In this section, projects and initiatives that are covering aspects of decentralized governance based on trust registries in the healthcare ecosystem are presented. Their development status and shortcomings are highlighted.

TRAIN (TRust mAnagement INfrastructure) provides components for a flexible and cross-domain trust infrastructure to sovereignly manage trust anchors with Domain Name System (DNS) and verify the inclusion of entities (e.g., issuers of self-sovereign identity credentials) in trust frameworks [Ju21]. TRAIN uses the global, well-established, and trusted infrastructure of the Internet Domain Name System DNS as its root of trust, leveraging DNS's ubiquitous use and recognition. Trust lists store trusted service provider information in a format based on ETSI standard TS 119 612. TRAIN is a generic approach originating in use cases around digital identities and credentials. Its value for a healthcare use case has been demonstrated in [Ju21]. Lately, it has been integrated for the provision of decentralized trust into the Gaia-X Federation Services - GXFS [GX23] that serve the Gaia-X ecosystem. However, adding entities into the trust registry in TRAIN is still a manual process as it currently does not provide an easy-to-use enrollment tool supporting users in this task. This is a hurdle for the wider adoption of this approach.

Digital TRUST Infrastructure for Discovery and Validation (Regi-TRUST) is an infrastructure project sponsored and hosted at the United Nations Development Programme (UNDP) [Re23]. It intends to develop and provide a suite of tools to enable the discovery and validation of trusted services by leveraging the infrastructure of the DNS and its security extensions. Application areas are broad, with e-government, banking, but also healthcare. Implementers can leverage the core trust registry and enrollment tools to create purpose-sized networks or ecosystems for trusted digital services or service networks. It provides service providers with a mechanism to publish and expose trusted service definitions and provides users, via the public internet, the ability to discover needed services and access relevant service information through trusted endpoints, thus enabling informed decisions about whether to trust and use a service. Most importantly, Regi-TRUST aims to enable networks of networks at scale, using a decentralized, cloud-agnostic architecture. Developing practical tools for decentralized governance, in particular to manage the trust registry, is one part of the work currently pursued.

The World Health Organization (WHO) has established the **Global Digital Health Certification Network (GDHCN)** [WH23]. It builds on regional networks for COVID-19 certificates and takes up the infrastructure and experiences with the digital European Union Digital COVID Certificate (EU DCC) system, which adopted in the EU and 51 other countries and territories. It has been designed to be interoperable with other existing regional and sectoral networks (e.g., ICAO VSD-NC, DIVOC, LACPass, SMART Health Cards) specifications. Participating governments, private sector entities and consortia shall be able to contribute standards-based content and authentication mechanisms (e.g., public keys) in a decentralized manner, linking health record issuance, possession, and verification to authorized individuals and institutions. Healthcare providers will more easily be able to verify health records to support continuity of care. However, architecture and implementation approach are not yet published in detail.

2.3 Enrollment Tools in the Use Case Context and Beyond

This section examines the latest advancements in tools for registering trusted service providers (TSPs). The EU DCC and COWIN platforms, instrumental during the COVID-19 pandemic, facilitated the registration of vaccine providers in nearly 100 countries. Additionally, there is a broader strategy that extends outside of healthcare, employed by eIDAS 1.0. This approach offers qualified certificates through TSPs within the EU.

The **EU DCC Enrollment Tool** [EU21], established by the European Union, is a platform to support the enrollment process for service providers in issuing and verifying EU Digital COVID Certificates. The EU DCC is a standardized digital document that provides proof of vaccination, a negative test result, or recovery from COVID-19. Service providers, such as healthcare providers, testing laboratories, and authorized bodies, play a crucial role in issuing and verifying these certificates. The tool provides an interface for service providers to register and enroll for becoming authorized to issue and verify EU DCCs. The tool ensures compliance with the technical and security standards set by the EU enabling

service providers to manage their enrollment process, track the progress of their applications, and receive updates and notifications regarding their status. Still, there are few options for individual states to enact their own policies with respect to what kind of data they obtain from service providers and what kind of data they can make public. While the public keys of service providers are public and can be used to validate certificates, data on compliance and service provider status (i.e., revocation/suspension) is not public.

Co-WIN [Co23] is a platform used for vaccine provider enrollment in India's DIVOC (Digital Infrastructure for Vaccination, Open and Co-WIN) system. It is an integral part the country's COVID-19 vaccination program. The tool has a dashboard that enables healthcare facilities, such as hospitals and vaccination centers, to register themselves as vaccine providers and participate in the vaccination drive by providing necessary details and documentation. Additionally, it helps to verify data by verifying the credentials and eligibility of healthcare facilities before granting them the status of authorized vaccine providers. However, the whole enrollment approach is centralized and managed by the Ministry of Health and Family Welfare, Government of India as central authority. Similar to the EU DCC there is little room for federal states in India to enact their own policies with respect to data related to service providers and its use for other regional outbreaks is currently not discussed in detail.

The **eIDAS 1.0 Trust Registry enrollment tool** [EC23] provided by the EU Commission is a dashboard platform designed to facilitate the enrollment process for trust service providers (TSPs) under the eIDAS Regulation. eIDAS sets out the legal framework for electronic identification and trust services. The Trust Registry Enrollment Tool serves as a federated system where TSPs can register and submit their information to become recognized providers of trusted services. It streamlines the enrollment process, allowing TSPs to manage their applications, documentation, and relevant data required for compliance with eIDAS requirements. The federated approach is not fully decentralized. Like the other tools it is limited to a specific region (the EU) and lacks global interoperability and flexibility to accommodate different identity ecosystems.

3 Trust Registries Enrollment Tool

The current enrollment tools as presented before in section 2.3 come with several shortcomings. Most are managed by a centralized or a federated authority (for example: a national health authority of a country or member state) and information on service providers is difficult to obtain. Most of the tools make only the public keys of service providers publicly available in centralized registries. Other compliance-relevant data or revocation status are not available. Moreover, most approaches provide limited flexibility regarding the identity technology (i.e., x.509, DID). Additionally, with current implementations it is not possible to implement a governance structure (e.g., a specific reviewing process of service providers) with particular policies on the regional level. To address these shortcomings, this chapter presents the implementation details and

architecture of the Trust Registries Enrollment Tool (TRET) that we developed to specifically support decentralized governance of ecosystems. Beyond the architecture, an overview of the data model of the trust registry will be presented as well.

3.1 Architecture of the TRET

The architecture of the trust registries enrollment tool is shown in Figure 1. To make decentralized governance possible with flexibility and interoperability, the enrollment tool was built as an extension of the TRAIN infrastructure mentioned in section 2.2. The core components of the tool are the Trust Scheme Publication Authority (TSPA), DNS Zone Manager, Client UI Web Application, Keycloak Server and Trust Registry.

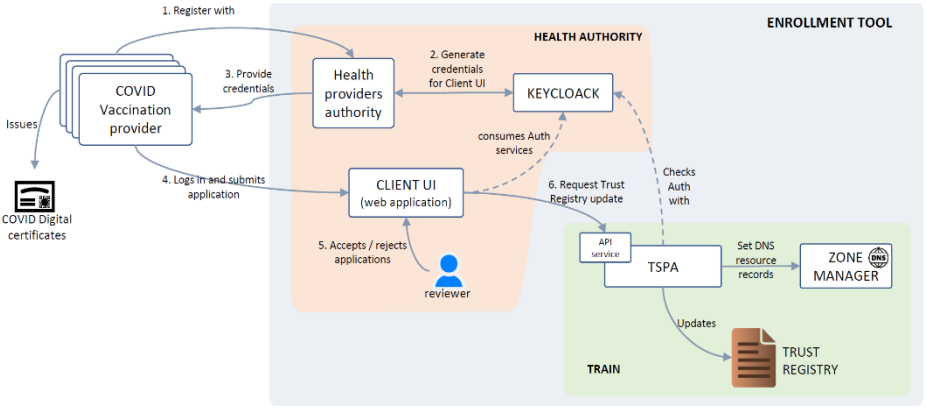


Fig. 1: Architecture for Trust Registries enrollment tool (TRET)

The **TSPA** is a backend component of the Trust Registry enrollment tool, responsible for exposing API endpoints that can be used by the front-end Client UI to update the Trust Registry. The TSPA provides a configuration file to connect with the Keycloak server or any other OpenID compliant identity provider to establish a role-based authorization allowing services providers the publication of trust services in a trust registry. The TSPA must be configured with the DNS Zone Manager by which the URL of the Trust Registry can be anchored in the DNS zone file, allowing interoperability and global discovery.

The **DNS Zone Manager** component is primarily a server that anchors Trust Registry updates into the DNS system for global discovery and interoperability. Any organization or entity controlling a Domain Name Server can set up their own instance of the enrollment tool. A DNS Zone Manager instance has the flexibility to host multiple Trust Registries with a unique *Trust Scheme Name.* This is a DNS name and several use cases or registries can be set up using different Trust Scheme Names. For example: the vaccine registry can be hosted under the Trust Scheme Name vaccine.region1.com, whereas the registry for applications related to hospitals in different regions can be hosted under hospitals.region1.com, hospitals.region2.com etc. This will be used later for API lookups

intended for validation of certificates.

The front-end **Client UI** is exposed to participants and reviewers. The participants can use it to enroll themselves into a trust registry. The reviewer(s) will be able to view different submissions created by the participants and the reviewers' approval will publish the data of the participant to the Trust Registry.

The **Keycloak Server** is used for Open ID connect role-based authentication and authorization. Three roles are defined: (1) The admin, who is responsible for setting up the infrastructure and can issue client UI credentials to Reviewers and Participants. (2) Participants, who are service providers wanting to enroll themselves in the registry. They do so by submitting a request to the Trust Registry Admin. (3) The reviewer, responsible for reviewing the information provided by the participant. If the submission complies with the governance policies, the submitter will be enrolled as trusted service provider in the Trust Registry and becomes a Participant. There is no restriction of the architecture of the enrollment tool to the three roles as described. Further delegation and new roles can be included depending on the specific governance structure.

The **Trust Registry** contains a list of trusted service providers for the trust domain along with their services. The hosting location of the Trust Registry is anchored as URI record via the DNS Zone Manager in a Zone File for global discovery and interoperability. A trust registry model implemented for the use case of COVID-19 certificates as presented in the following section. Different formats can be used, for example XML and JSON, depending on the respective requirements. In the COVID-19 implementation, although an xml-based approach is pursued for the registry, special API endpoints have been created for exposing the trust registry in JSON.

3.2 Data Model of the Trust Registry

The Trust Registry for enrolling trusted service providers is composed of three parts. (1) *TSPInformation* includes the service provider's business, legal and certification details. The trust registry can also accommodate different identifiers and certification details based on the governance. An example could be a governance norm requiring the service provider to provide a vLEI (verifiable Legal Entity Identifier) to enroll in the trust registry. (2) *SubmitterInfo* contains details of the person or organization who is responsible for submitting the entry for enrollment in the trust registry. (3) *TSPServices* contains information regarding the service offered by the provider. This contains details regarding which type of identifier is used for which service. Moreover, it lists information regarding the digital identity (public key) required to verify the credential issued by the provider. Both, PKI (x.509 certificates) and Decentralized Identifiers (for example: DIDs [Sp23]) can be accommodated. Additionally, this contains detailed information pertaining to the governance rules for the particular service, which the verifier can use to validate the authenticity of the provider's service. The JSON format data model of the trust registry can be obtained at https://t.ly/7z3-2.

4 Healthcare Use Case: COVID-19 Certificate Service Providers

Healthcare ecosystems serve as an ideal source for a use case allowing to showcase our approach. Here, ensuring quality, safety, and efficiency in healthcare delivery are pivotal. The basis for this is an efficient governance that cannot be fully centralized.

4.1 Use Case Scenario

Our tool provides means for decentralized governance at different levels, for example at global, regional, country, and organizational level. Our use case spans from global to regional level where health providers authorities like ministries of health of different countries allow COVID vaccination providers to enroll themselves into every country's health-related trust framework. This allows COVID-19 vaccination certificates issued by health providers in a specific country to be verified in other countries inside the same trust domain. Three main actors are covered: (1) the healthcare providers' authority which is present at the country level, (2) the reviewer, a role assigned by the healthcare providers' authority with the ability to verify providers' compliance with the policies and rules set by the authority, and (3) the COVID vaccination providers as participants.

The healthcare providers' authority assigns a system administrator for the Client UI application and at least one reviewer. As a first step, a COVID vaccination provider registers with the corresponding authority in its country through an online or offline process. The authority then issues credentials to the provider which allows it to access the enrollment tool. Then, the provider initiates the enrollment process by logging into the Client UI application and submitting an application providing the necessary information for the trust registry (see section 3.2). The reviewer, can then approve or reject this submission through the Client UI. In case of approval, the Client UI uses the TRAIN infrastructure, specifically the API endpoints provided by the TSPA to start the trust registry update process. The TSPA not only updates the trust registry but also creates and sets the necessary DNS resource records to make the trust registry discoverable via the widely adopted DNS infrastructure. Upon completion, the existence of a trusted service provider in a trust registry ca now be globally verified via the DNS system.

4.2 Implementation of the Enrollment Tool in the Use Case

The tool implementation is available as open source in [Tr23]. This includes a set of endpoints provided by the TSPA that allow the publication of trust registries as well as the fetching of full trust registry and trusted service provider general information in JSON format. Detailed information on individual TSPs can also be queried. Moreover, a TSP update mechanism is provided, allowing for the dynamic management of the provided list of services, for example, for additional vaccination services. The implemented Client UI web application provides features for service providers to enroll their services in the trust registry by using the *Submit Network Entry* menu (see Figure 2).

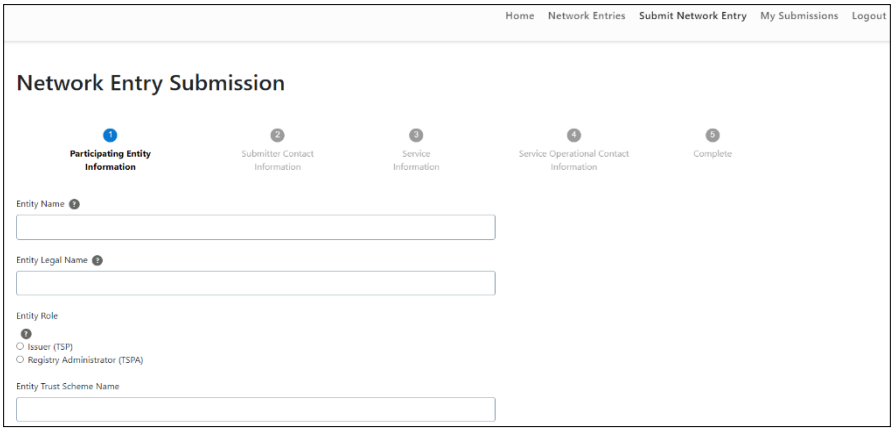


Fig. 2: WebView for enrolling service providers in the Trust Registry

The Client UI (Figure 3) provides role-based access for reviewers. It provides the ability to check accepted submissions and the pending status of applications to be reviewed.

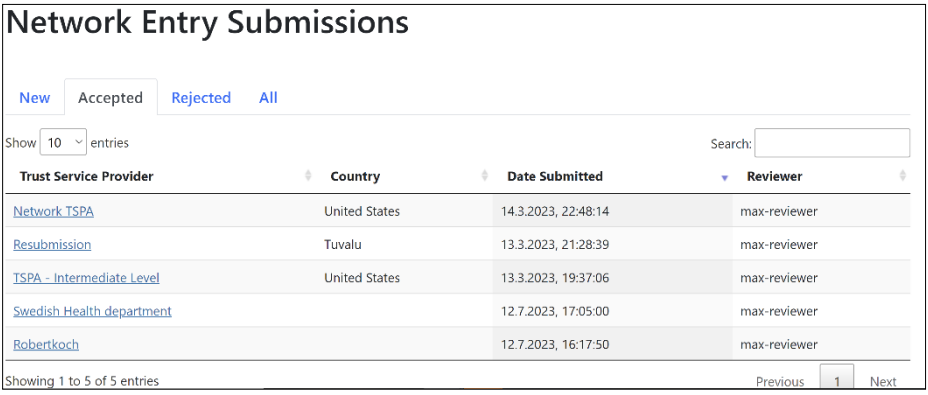


Fig. 3: WebView of Reviewer User Interface

The service feature in the trust registry allows vaccination service providers to add other vaccines (i.e., Malaria) and types of certificates to their services in the trust registry, not limiting the use of the tool to COVID-19 certificates. This illustrates the flexibility to accommodate other services without building a new infrastructure from scratch or having to duplicate and redeploy the current one. Additionally, the tool can be used by governments and the private sector to possibly digitize the vaccine service provider records in a decentralized way. Once enrolled, service providers are able to view the entries in the *My Submissions* tab. An example is shown in Figure 4.

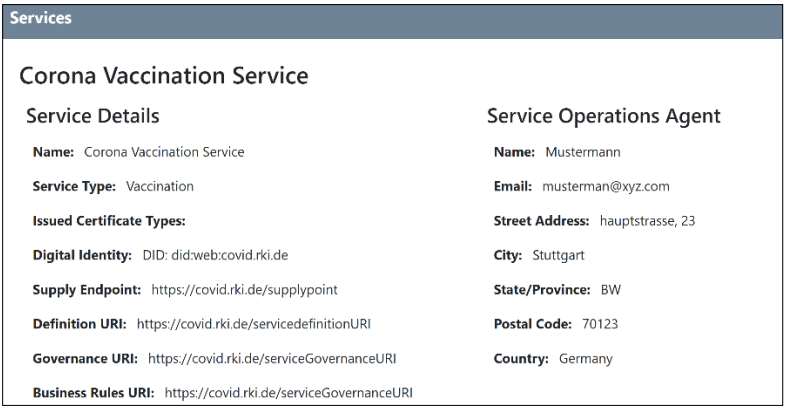


Fig. 4: WebView of Service Details offered Vaccination Service Provider

5 Discussion

The Trust Registries Enrollment Tool (TRET) stands out in its ability to satisfy several key features that support decentralization, including interoperability, portability, transparency of user/organizational control, data minimization, and open-source licensing, setting it apart from other tools in the healthcare sector. Interoperability is paramount in healthcare systems, and the TRET excels in this aspect by adhering to open standards and employing a modular architecture. This ensures seamless data exchange and collaboration across diverse healthcare platforms and providers, overcoming interoperability challenges commonly faced by proprietary tools. Portability is essential for universal access to healthcare services, and the TRET achieves this through its web-based design and API endpoints compatibility with various devices and operating systems. This enables users to access the platform conveniently from any location, fostering inclusivity and accessibility. Data minimization is a core principle in privacy protection, and the TRET upholds this by limiting the collection and storage of sensitive information by distributing the enrollment process so that the single trust list enrollment office will not be responsible for holding and hosting all service provider information as in the case of different existing enrollment tools mentioned above. The EU DCC and the eIDAS 1.0 enrollment tool offer federated enrollment approaches. However, their data collection leans towards centralization, and they are not designed to be used outside the European context. Finally, Co-WIN offers a centralized enrollment approach with single point data collection from all service providers. Clearly, both Co-WIN and EU-DCC were specifically designed for COVID and are not operational anymore. Whereas TRET due to its decentralized nature and its portability can maintain its operational nature and can be used beyond COVID use cases.

6 Conclusion

Decentralized governance of business ecosystem remains challenging. It requires trust and transparency among diverse stakeholders that must collaborate and co-create value. Our paper contributes to this field by presenting an enrollment tool for trust registries for such ecosystems. Our approach, as developed in a healthcare scenario, is a building block for an effective and scalable governance of decentralized ecosystems, ensuring trust and security among participants. We showcase that the tool could represent an innovative and reliable solution to the current challenges faced in healthcare ecosystems and beyond. Although this paper demonstrated only the COVID-19 use case, the tool possesses the potential and flexibility to serve in other scenarios covering diverse regions or continents and including a wide range of digital certificates. Leveraging the TRAIN infrastructure to anchor the trust registry provides a method to make all kinds of data that requires a certain level of trust and/or access control interoperable across the globe. Moreover, the integration of TRAIN into the GXFS enriches our tool's relevance, offering a robust mechanism for enrolling entities within Gaia-X and related ecosystems. This development enhances the tool's utility, facilitating broader adoption and interoperability in fostering European digital sovereignty and secure data spaces. Some limitations apply to the current version of the tool as presented. So far, we have only piloted it in the use case that was presented here. Further development and evaluations for other scenarios, in particular beyond healthcare, are certainly required. Moreover, for the current demonstration deployment of the use case, a single Keycloak server was used for role-based authorization with 3 different roles. When deployed in real-world application scenarios, there might be multiple roles required for managing the whole infrastructure, which has not yet been implemented and evaluated. These shortcomings will be addressed in future development iterations of the enrollment tool.

7 References

[ABB19] Abimbola, S.; Baatiema, L.; Bigdeli, M.: The impacts of decentralization on health system equity, efficiency and resilience: a realist synthesis of the evidence, Health Policy Plan., vol. 34, no. 8, pp. 605–617, Oct. 2019.

[Bl23] Blanchette: Decentralized autonomous organizations: Health care in the metaverse, https://insurancenewsnet.com/innarticle/decentralized-autonomous-organizations-health-care-in-the-metaverse, accessed: 13.07.2023.

[Co23] Co-WIN: CoWIN, https://www.cowin.gov.in/, accessed: 14.07.023.

[EB23] EBSI: Issuers trust model - Accreditation of Issuers, EBSI Specifications, https://ec.europa.eu/digital-building-blocks/wikis/display/EBSIDOC/Issuers+trust+model+-+Accreditation+of+Issuers, accessed: 20.07.2023.

[EC23] European Commission: eIDAS Dashboard, https://eidas.ec.europa.eu/efda/tl-browser/#/screen/home, accessed: 14.07.2023.

[EU21] EU-DCCG : eHealth Network Guidelines on Technical Specifications for EU Digital COVID Certificates Volume 2, 2021, https://health.ec.europa.eu/document/download/5c88867a-c4e5-4956-8461-9c5441797dda_en

[GL19] Garbuio, M.; Lin, N.: Artificial Intelligence as a Growth Engine for Health Care Startups: Emerging Business Models, Calif. Manage. Rev., vol. 61, no. 2, pp. 59–83, Feb. 2019, doi: 10.1177/0008125618811931.

[Gr22] Greer, S. L.; Rozenblum, S.; Falkenbach, M.; Löblová, O.; Jarman, H.; Williams, N.; Wismar, M.: Centralizing and decentralizing governance in the COVID-19 pandemic: The politics of credit and blame, Health Policy Amst. Neth., vol. 126, no. 5, pp. 408–417, May 2022, doi: 10.1016/j.healthpol.2022.03.004.

[GX23] GXFS: Gaia-X Federation Services (GXFS) werden erweitert, GXFS.eu, https://www.gxfs.eu/de/ausschreibung-identity-trust/, accessed: 12.07.2023.

[Ju21] Jurado V.M.; Vila, X.; Kubach, M.; Johnson, I. H.; Solana, A.; Marangoni, M.: Applying assurance levels when issuing and verifying credentials using Trust Frameworks, in Open Identity Summit 2021, H. Roßnagel, C. Schunck, and S. Mödersheim, Eds., in Lecture Notes in Informatics (LNI). Bonn: Köllen Druck + Verlag, 2021, pp. 167–178.

[MS23] Mateus, S.; Sarkar S.: Can Decentralized Autonomous Organizations (DAOs) Revolutionize Healthcare? Calif. Manag. Rev. Insights, Jan. 2023, https://cmr.berkeley.edu/2023/01/can-decentralized-autonomous-organizations-daos-revolutionize-healthcare/, accessed: 13.07.2023.

[Ra23] Radonjic-Simic: Decentralized Open Platform for Vaccination—A German Example: COVID-19-Vacc., https://www.mdpi.com/2199-8531/7/3/186, accessed: 12.07. 2023.

[Re23] Regi-TRUST: Governance Consultation: How to Operate and Govern a Global Trust Network of Networks?, https://www.sparkblue.org/Regi-TRUST, accessed: 12.07.2023.

[Sc18] Schreieck, M.; Hein, A.; Wiesche, M.; Krcmar, H: The Challenge of Governing Digital Platform Ecosystems, in Digital Marketplaces Unleashed, C. Linnhoff-Popien, R. Schneider, and M. Zaddach, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2018, pp. 527–538. doi: 10.1007/978-3-662-49275-8_47.

[Sp23] Sporny: Decentralized Identifiers (DIDs) v1.0, https://www.w3.org/TR/did-core/, accessed: 15.07.2023.

[SS19] Sreeramareddy, C. T.; Sathyanarayana, T.: Decentralised versus centralised governance of health services, Cochrane Database Syst. Rev., no. 9, 2019, doi: 10.1002/14651858.CD010830.pub2.

[Tr23] Trust Registries Enrollment Tool Github Code, 20. Jun.2023, https://github.com/undp/Regi-TRUST, accessed: 12.07.2023.

[WH23] WHO: Global Digital Health Certification Network, https://www.who.int/initiatives/global-digital-health-certification-network, accessed: 12.07.2023.

[Wi19] Wieringa, R. J.; Engelsman, W.; Gordijn, J.; Ionita, D.: A Business Ecosystem Architecture Modeling Framework, 2019 IEEE 21st Conference on Business Informatics (CBI), Jul. 2019, pp. 147–156.

Learnings from a Guided Method for Experience Design: Psychological Needs in the Context of the Privacy Value

Anne Elisabeth Krueger [1] and Stefan Brandenburg [2]

Abstract: This position paper introduces a guided method for experience design that addresses the importance and challenge of considering the rather abstract psychological (user) needs and values as input for creative ideation processes of interactive systems. We present exemplary empirical results from the application of the method concerning the value of privacy, revealing how needs and values can become tangible for user experience designers. Also, interdependencies between the value of privacy and psychological needs, and between the concepts of psychological needs and values in general, were identified by applying the guided experience design method. Learning about the connections of needs and values provide valuable insights for experience design, which are discussed in the paper and should be further explored.

Keywords: Experience Design, Psychological Needs, Values, Needs Persona, Needs Empathy Map

1 Introduction

In the past, there has been considerable discussion and research on the notions of values and psychological needs in various fields. This discussion has led to a myriad of conceptualizations and ideas about what values and needs are. However, there is no universally accepted understanding of these concepts across different research communities. We respect the diversity of perspectives and do not claim to encompass them. The aim of this position paper is to introduce a method for Human-Computer Interaction (HCI) that considers both values and psychological needs. Moreover, we seek to investigate how to incorporate privacy into the realm of HCI, acknowledging the importance of values and psychological needs in this context.

1.1 The Concept of Values and Psychological Needs

Values encompass deeply rooted and meaningful beliefs, attitudes, ideals, and needs shared by members of a society. They play a substantial role in shaping an individual's character, identity, and cultural context. Serving as a fundamental lens, values influence how individuals perceive and assess the world around them. Academic research on values has extended

1 Fraunhofer Institute for Industrial Engineering , 70569 Stuttgart , Germany, anne-elisabeth.krueger@iao.fraunhofer.de, https://orcid.org/0000-0002-3472-1186
2 Technical University Chemnitz, Cognitive Psychology and Human Factors, Wilhelm-Raabe-Straße 43, 09120 Chemnitz, Germany, stefan.brandenburg@psychologie.tu-chemnitz.de, https://orcid.org/0000-0002-7628-5247

doi:10.18420/OID2024_10

for more than 25 years, yielding results such as Friedman's [Fr13] compilation of values with ethical implications. These include next to privacy considerations for human welfare, ownership and property, freedom from bias, universal usability, trust, autonomy, informed consent, accountability, courtesy, identity, calmness, and environmental sustainability. Over time, the significance of ethical considerations in the human-centered development of digital products and services has been steadily rising. This mirrors an increasing acknowledgment of the role values play in shaping technology and its impact on individuals and society [KK18; Kr23]. In addition to values, it appears crucial in the field of HCI to consider the psychological needs of users to craft a positive user experience. If the psychological needs of the users are fulfilled in the context of an interactive product or service, a positive user experience might arise [Ha03; Ha13]. Desmet and Fokkinga [DF20a] propose a needs-typology, which embraces thirteen fundamental needs like autonomy, beauty, community, comfort, competence, impact, morality, purpose, and recognition. It can already be seen here - e.g., with autonomy - that there is some overlap between the two concepts. Thus, we operate in this paper under the assumption that values and psychological needs, while not entirely distinct, are nevertheless independent yet interdependent concepts.

1.2 Experience Design based on Values and Psychological Needs

The human-centered design process, as defined in [IS19], is a widely utilized framework for developing interactive products and services through iterative and user-centered practices (EN ISO 9241-210). Despite being commonly employed and prioritizing user experience, the human-centred design process still lacks substantial integration of values, psychological needs, and ethical considerations across its four stages. Nevertheless, these elements significantly influence the user's experience with a product or service. As a result, certain studies have concentrated on specific values, such as privacy, to devise and assess methods that aid in understanding the role of people's values in their interactions with digital products and services [HIB22]. However, there is still a lack of methods that effectively bring together people's psychological needs and values, considering the apparent overlap between these two concepts. What psychological needs and values share is their abstract nature, situated in a psychological context. Consequently, they are often not readily accessible for designers. To achieve a successful design process, it seems crucial for technology managers and designers to possess a thorough understanding of the psychological needs and values of their users. Thus, there is a necessity to make these somewhat implicit concepts tangible and comprehensible for them (cf. [KFP15; Kr17]), allowing them to use these as a foundation for their creative process.

1.3 Objectives

The present paper has two objectives. First we want to introduce the guided experience design method that can help designers to understand their users psychological needs and

values, and the interdependencies between them. Therefore, we present exemplary empirical findings concerning the value of privacy, identifying the impact that individual psychological needs have on the design of positively perceived privacy experiences. Second, we want to learn about the interaction of psychological needs and values in designing interactive products and services. In this regard, our focus is on making psychological needs and values tangible, igniting creative inspiration for the ideation processes of designing interactive systems.

2 Related Research: Designing for Values and Psychological Needs

2.1 Value Sensitive Design

Value Sensitive Design is a methodologically sound approach to technology design, incorporating human values in a systematic and thorough manner across the entire design process. This approach proposes a tripartite methodology involving three interconnected investigations: conceptual, empirical, and technical [FKB02]. Conceptual investigations address issues such as what values are, whose values should be prioritized, and how technological designs impact values. These inquiries provide carefully crafted conceptualizations of specific values. Empirical investigations complement conceptual inquiries by observing, measuring, and documenting human activities related to the technical artifact. These investigations explore how stakeholders apprehend values, prioritize them in design trade-offs, and consider both espoused and actual practices, examining the impact of technology on groups and individuals. Finally, technical investigations recognize that technologies inherently support certain activities and values. The first form assesses how existing technological properties hinder or support human values. The second form involves proactively designing systems to support values identified in conceptual investigations. Technical investigations focus on the technology itself, differentiating from empirical investigations that concentrate on people or larger social systems affected by the technology. The method presented in this paper combines aspects of the empirical and technical investigation [FKB02]. Moreover, the guided experience design method helps to make abstract needs and values more tangible by providing specific definitions of values. Defining what values mean in the context of designing technological systems has been difficult in the past (cf. [Um20]), especially for designers that are not ethical experts (cf. [Br17]).

2.2 Experience Design based on Psychological Needs

In experience design, several approaches exist, and one of them already incorporates ethical considerations. In this context, the available resources introduce and sensitize designers to basic human needs and possible resulting requirements for the design. Notable examples include needs cards [Ha10; Ha13; HD12; Sh01] wellbeing determinant cards [RD15],

and recent materials on psychological needs by Desmet and Fokkinga [DF20b; DF21; DF22]. However, these materials are primarily designed to offer insights into specific needs and their implications for design. The designers are not explicitly aided in a systematic reflection process concerning the needs. Additionally, there are only a limited number of methods in experience design (e.g. [KLH18; PA20]) for systematically designing based on psychological needs (cf. [PAC20].

In the context of this position paper, we refer to the "Needs Profiles"method [Kr17], which can be used to actively sensitize designers. It also allows them to systematically access and gather their (partly implicit) shared knowledge about psychological needs in the realm of experience design (cf. [KFP15]). Thus, it leverages the designer's existing knowledge of psychological needs and expands it through personal, systematic shared reflection-in-action [Sc83] processes with other designers. Moreover, the resulting "Needs Personas"– a personification of psychological needs - can serve as a solid foundation and creative input for ideation processes [KFP15; Kr22].

3 Guided Experience Design Method

The guided experience design method incorporates psychological needs and ethical considerations. Within the scope of this paper, the emphasis will be on the conceptualization (see section 3.1) and materialization (see section 3.2) of psychological needs and the value of privacy. Overall, method sections 1 and 2 are intended to provide inspiration for the experience design of privacy-related aspects and interactive solutions.

3.1 Method Section 1: Conceptualization

One main objective of this method section is to explore the partially overlapping concepts - needs and values - and to establish a shared understanding of those. Furthermore, participants have the opportunity to specify which needs they associate with and how they link them to the value of privacy, and which of them they would like to address in the next Method Section.

Step 1: Understanding Values and Psychological Needs: To prepare the participants for the topic of values and psychological needs in the context of experience design, to understand their attitudes towards the concepts - whether, how and where they see a link, the modified Warm UP Object Presentation (cf. [KM22]) was used. Participants were asked to build their understanding of the terms *psychological need* and *value* individually using a small set of Lego® bricks. The results were then shared and discussed with all participants.

Step 2: Connecting Privacy with Psychological Needs: The focus is then placed on the value of privacy and a brief theoretical introduction was given. Then the needs typology according to Desmet and Fokkinga [DF20a] is presented with the help of the materials

from [Bu23]. The participants are asked to listen actively - they were asked to reflect while listening and put on sticky notes which psychological needs they associate with the value of privacy and how. These findings are clustered and evaluated. The participants are asked to pick out the four most relevant psychological needs for them.

3.2 Method Section 2: Materialization

In the second method section, the initial focus is for the participants to generally bring together their various perspectives on the defined needs and, in doing so, become aware and visualize unconscious elements of knowledge or beliefs. Subsequently, the participants are tasked with putting themselves in the perspective of users who have the needs pronounced in an extreme sense in the context of the privacy value, and deriving concrete strategies for fulfilling those needs, i.e., requirements for the design of interactive systems.

Step 1: Constructing Psychological Needs: The Needs Profiles [Kr17] are applied to make the selected, rather abstract psychological needs and the value tangible and concrete - thereby, the first step is to metaphorically built the previously defined needs with Lego®-bricks.

Step 2: Designing Needs Personas: The Needs Profiles define that the insights gained during the former step are then transferred to the Needs Empathy Map [Kr22] and used for the development of the Needs Personas [KFP15] in the context of the value of privacy. Thereby, the participants creatively work out how individuals with the previously defined needs would act and behave and what their motivations, thoughts and feelings are – in the context of the value of privacy. This means that the participants have dealt during the method with how the Needs Personas would live out the value and what influence the needs might have on privacy behaviour.

4 Results

The findings presented are drawn from the first implementation of the guided experience design method with nine participants at a scientific HCI-conference (4 female, 5 male). The participants had a background in a wide variety of research fields, none of them had a focus on experience or value-sensitive design.

4.1 Conceptualization - Step 1: Values and Psychological Needs as Related Concepts

The detailed presentation of the individual Lego®-models (cf. Method Section 1 - *Understanding Values and Psychological Needs*) and the subsequent in-depth discussion showed that the participants were well supported in dealing intensively with the topic of psychological needs and values. All participants were able to build a Lego®-model that

included their perspective on the issue. In the subsequent presentation, discussion, and debate of the buildings, it became clear that, for most of the participants, the two concepts were fundamentally linked and mutually dependent. It also became apparent that the two concepts were mutually dependent for the participants. For some of them, an internalized understanding of values could trigger psychological needs, while for other participants, a certain predominant set of needs was a prerequisite for living out a value. Overall, the exercise proved to be well suited for activating the participants' prior knowledge and opinions as well as making these aspects openly communicable as part of the method implementation. In this way, we were able to pick up the participants' convictions and positions on the topic and prepare them accordingly for the subsequent creative process and the implementation of the following exercises.

4.2 Conceptualization - Step 2: Linking Psychological Needs to the Value of Privacy

For the participants, eleven out of the 13 psychological needs (cf. [DF20a]) were directly linked to the value of privacy (cf. *Connecting Privacy with Psychological Needs* in Method Section 1). Furthermore, they associated concrete strategies for fulfilling those needs and the influence of different needs on how the value of privacy should be designed for (see Figure 1). However, no linkages were made for the needs fitness and stimulation. Moreover, it was discovered that certain overlaps within the typology of needs exist. For instance, distinguishing between the needs for connectedness and community did not appear as straightforward for the participants. Furthermore, two participants expressed that data protection was for them mostly associated with the need for security, but that they had now realized that it can also be linked to so many other needs. Therefore, the participants decided not to go on with the obvious need for security in the context of privacy but to proceed with the four needs of comfort, community, recognition, and autonomy to inspire the creative process.

4.3 Materialization - Exploring Needs in the Context of the Value of Privacy

First, the participants were split into four groups (cf. *Constructing Psychological Needs* in Method Section 2), with each group building two of the selected needs using Lego® bricks. In doing so, they were encouraged to take a closer look at the psychological need and its characteristics - i.e., not yet to address the value of privacy. In doing so, they were supported to access the psychological need in general, first. Exemplary results can be found in [Kr17; Kr23].

Then, two out of the four groups were formed with each of them developing a Needs Persona (cf. *Designing Needs Personas* in Method Section 2). Due to space constraints, within the scope of this paper, only the Needs Persona "Alex" (see Figure 1) is presented. Alex is a progressive activist who lives in a commune and has several children. He likes to lie outside

Tab. 1: Linking the Psychological Needs to and the mutual dependence on the Value of Privacy.

	Psychological Need in the context of the Value of Privacy
Autonomy	Autonomy was connected to self-determination, control over the disclosure of information and individualism; the freedom to live out personal preferences, making decisions about the disclosure of information, preserving the right to privacy as a personal choice.
Beauty	Aesthetics was seen as important factor for the efficient communication of privacy matters. To be perceived as beautiful (currently it is not), one should feel safe, calm, not have to worry.
Comfort	Deemed to prevent tension and difficulty, privacy seems essential for creating a comfortable environment. Ease in adapting settings, e. g. non-compulsory defaults, security measures seem crucial.
Community	Community requires general openness; necessitates the support of intentional engagement with consent, simultaneously. Reliable, standardized privacy settings to balance the exposure to the community as chose were claimed.
Competence	The freedom of choice seemed crucial; it was found necessary to provide information and (privacy) settings that can be changed. The ability to enact private protection was also emphasized.
Impact	Proof that a stated preference (e.g. regarding data-handling) has been complied with. To see if the abandonment of data protection.
Morality	The violation of privacy was generally estimated as ethically wrong. The willingness to pay for a service which uses sensitive data was associated with moral considerations.
Purpose	It seems important that it is clearly and comprehensibly explained what the requested information is used for.
Recognition	Recognition involves sharing data for appreciation. Protecting privacy is a delicate balance between openness and personal space, especially in the realm of social media where recognition and stimulation abound but pose a challenge to privacy.
Relatedness	Sharing information and protecting privacy seem crucial for building relationships - involving contribution, knowledge sharing, and achievements within a trusted connection; a secure and meaningful relation requires rather gradual trust-building.
Security	If privacy is neglected, people lack general security. Systematic enforcement of data protection seemed essential to protect both data and the people behind it. Thus, it seemed crucial to them that information and data be stored securely.

Fig. 1: Needs Persona addressing the Psychological Needs Comfort and Community within the Context of the Value of Privacy.

on a blanket, is generally rather cosy and has little cosy dog as a pet. He has a university education and is active in both offline and online communities. He needs easy, convenient access to his online and offline community. He wants to feel secure that his data is protected - but doesn't want to go to a lot of trouble to do so. Therefore, he prefers private servers and likes secure passwords, but finds them rather inconvenient because they are easy to forget. Moreover, additional insights were elaborated regarding Alex's behaviour and emotional life, derived from his two defined psychological needs. This information illustrates how he navigates the value of privacy in his daily and personal life, both online and offline.

5 Discussion and Future Work

The first objective of this paper was to introduce the guided experience design method that may help to understand and design for the psychological needs and values of the users. Second, we wanted to learn about the interdependencies of needs and values in general by applying the method to the value of privacy. The results revealed that the conceptualization step of the method, which involved a systematical introduction to needs and values, an initial engagement with the concepts through a playful exercise [KM22], and an exploration of chosen needs, facilitated a profound understanding of the concepts and allowed participants to make preliminary assessments of the needs in the context of the value of privacy. Thereby, it was revealed that various psychological needs have diverse relationships to the value of privacy, influencing how privacy is experienced and how this experience can be modified by addressing user needs. It is noteworthy that the psychological needs for fitness and stimulation were not linked to privacy. While fitness may play a role in safeguarding privacy in the analogue domain, its relevance in the digital realm was not as apparent to the

participants. The missing relationship of stimulation and privacy may be attributed to the peoples' perception that privacy applications are mostly considered unappealing.

When it came to finding ideas for interactive solutions, the Needs Profiles [Kr17] were particularly relevant in addition to the designers' own decision in favour of certain needs (Method Section 2: Materialization). Here, the designers could gain a detailed view and more concrete insights into their personal and potential user perspectives regarding the psychological needs in the context of the value of privacy. After building the needs with the Lego®-Bricks to access possible latent knowledge about and visualize the needs first, the Needs Persona - a personification of the psychological needs – provided the participants with even more detailed and tangible insights into possible design requirements a user might have for privacy-related matters. Simultaneously, creating a Needs Persona using a Needs Empathy Map served to efficiently establish a common ground for the subsequent ideation process and integrate (the particpants) various perspectives on the resulting requirements of the needs for an interactive solution.

Furthermore, through the application of the method, we explored the relationship between the two concepts – values and psychological needs – in the context of the value of privacy. We found that the value of privacy and psychological needs have some distinct but also some interconnected characteristics. Therefore, designing for the needs that are related to privacy may help to address the value of privacy and, therefore, a positive (user) experience in the context of interactive products and services. Thereby, value appear to be more universal and offering a broader perspective on the individual's attitude towards life in general and the perceptions of interactive systems in particular than individual psychological needs. In general, assessing peoples' values helps to understand more universal personal motivations and life goals. When designing interactive systems, they can be regarded as a rather indirect influence on specific interaction situations. However, an individual value set may determine the importance of individual psychological needs. In contrast, psychological needs appear as a rather direct influence an interaction situation, with users trying to accomplish needs fulfilment when interacting with interactive products or services. Different psychological needs seem to have the capacity to generate different demands regarding the value of privacy. Throughout interactions, the value of privacy consistently remains in the background, acting as a canvas or target variable that can influence the ongoing need situation. Although, values seem to remain rather stable over time, they might still be subject to change. Thus, specific psychological needs, such as community and comfort, may potentially shift the importance of the value somewhat into the background, necessitating a different, less rigid privacy-oriented design requirement.

It can be concluded that the unique features of values and psychological needs suggest that values act as a foundation that must be considered before addressing needs effectively. Simultaneously, needs can influence how a value is expressed, defining how individuals experience that value in diverse situations or the needs that emerge from it.

Based on the insights gained from this study, we seek to aid designers by applying the

guided method for experience design in understanding the wide range of design possibilities for privacy based on different needs. We encourage integrating the awareness concerning the relevance of needs and values as well as the corresponding methods into the design of interactive systems, always with the goal of ensuring a positive user experience next to usability issues. These insights lead us to envision the potential development of a framework integrating values and needs for the experience-oriented design of interactive products and services - and thus, the design of privacy. However, to create a thorough ethical experience design, it seems essential to evaluate various values, acknowledge potential conflicts between them, and consider alternative design approaches. Due to limitations in space and the paper's focus, this aspect is not discussed here - see [Ho23; KHB23] for more details.

Based on these presented results, further empirical investigations are planned to deepen our understanding of the interplay of needs and values and how both concepts can be addressed with a sound methodology. Furthermore, we want to gather more empirical data. Here, a comparison of the awareness and knowledge of designers prior the application of the experience design method and after its application would help to determine, how much they have learned through the application. A more comprehensive exploration of the theoretical foundations will also be necessary in the future.

Acknowledgments. This work is part of our activities in the "Mittelstand-Digital Zentrum Fokus Mensch" (MDZ-FM). The MDZ-FM is part of the Mittelstand-Digital initiative funded by the German Federal Ministry for Economic Affairs and Climate Action, funding number: 01MF23003E.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

[Br17] Brandenburg, S.; Minge, M.; Cymek, D.; Zeidler, L.: Ethische Aspekte in der menschzentrierten Erforschung und Entwicklung technischer Geräte-Erfahrungen einer Ethikkommission. Forschung 10/, pp. 101–106, 2017.

[Bu23] Burmester, M.; Janssen, D.; Krüger, A. E.; Laib, M.: Bedürfnismaterialien, 2023, URL: https://www.digitalzentrum-fokus-mensch.de/wissen/beduerfnismaterialien, visited on: 01/05/2024.

[DF20a] Desmet, P.; Fokkinga, S.: Beyond Maslow's pyramid: Introducing a typology of thirteen fundamental needs for human-centered design. Multimodal technologies and interaction 4/3, p. 38, 2020.

[DF20b] Desmet, P.; Fokkinga, S.: Thirteen chairs, thirteen fundamental needs./, 2020.

[DF21] Desmet, P.; Fokkinga, S.: The House of Happiness. Inspiration Poster, 2021, URL: https://diopd.org/wp-content/uploads/2021/05/House-of-happiness-screen-scaled.jpg, visited on: 10/17/2022.

[DF22] Desmet, P.; Fokkinga, S.: Thirteen fundamental needs, 2022, URL: https://needtypology.com/.

[FKB02] Friedman, B.; Kahn, P.; Borning, A.: Value sensitive design: Theory and methods. University of Washington technical report 2/8, 2002.

[Fr13] Friedman, B.; Kahn, P. H.; Borning, A.; Huldtgren, A.: Value sensitive design and information systems. Early engagement and new technologies: Opening up the laboratory/, Publisher: Springer, pp. 55–95, 2013.

[Ha03] Hassenzahl, M.: The thing and I: understanding the relationship between user and product. In: Funology. Springer, pp. 31–42, 2003.

[Ha10] Hassenzahl, M.: Experience design: Technology for all the right reasons. Synthesis lectures on human-centered informatics 3/1, pp. 1–95, 2010.

[Ha13] Hassenzahl, M.; Eckoldt, K.; Diefenbach, S.; Laschke, M.; Len, E.; Kim, J.: Designing moments of meaning and pleasure. Experience design and happiness. International journal of design 7/3, 2013.

[HD12] Hassenzahl, M.; Diefenbach, S.: Well-being, need fulfillment, and Experience Design. In: Designing Well-being Workshop. Retrieved August. Vol. 25, p. 2013, 2012.

[HIB22] Hoth, V.; Ivanova, M.; Brandenburg, S.: UX Design Pattern für Datenschutz und Vertrauen./, Publisher: Gesellschaft für Informatik eV, 2022.

[Ho23] Hoth, V.; Krueger, A. E.; Langner, M.; Brandenburg, S.: Ethical Experience Design for the Value of Privacy based on Psychological Needs. In: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems. CHI '23: CHI Conference on Human Factors in Computing Systems. ACM, Hamburg Germany, pp. 1–4, 2023, ISBN: 978-1-4503-9422-2, URL: https://dl.acm.org/doi/10.1145/3544549.3574166, visited on: 12/11/2023.

[IS19] ISO: Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems, Norm ISO 9241-210:2019, Genf, Schweiz: ISO, 2019.

[KFP15] Krüger, A. E.; Fronemann, N.; Peissner, M.: Das kreative Potential der Ingenieure–menschzentrierte Ingenieurskunst. In: Stuttgarter Symposium für Produktentwicklung, SSP 2015 Entwicklung smarter Produkte für die Zukunft. Pp. 1–10, 2015.

[KHB23] Krueger, A. E.; Hoth, V.; Brandenburg, S.: How to Design for Ethical Experiences: Introduction of a Guided Method based on Psychological Needs for the Value of Privacy./, Publisher: Gesellschaft für Informatik e.V., 2023, URL: https://dl.gi.de/handle/20.500.12116/42412, visited on: 12/11/2023.

[KK18] Kakar, A. K.; Kakar, A.: IS THE TIME RIPE FOR BRANDING OF SOFTWARE PRODUCTS. SAIS 2018 Proceedings 21/, 2018.

[KLH18] Klapperich, H.; Laschke, M.; Hassenzahl, M.: The positive practice canvas: gathering inspiration for wellbeing-driven design. In: Proceedings of the 10th Nordic Conference on Human-Computer Interaction. Pp. 74–81, 2018.

[KM22] Krueger, A. E.; Minet, S.: Designing Positive Experiences in Creative Workshops at Work Using a Warm UP Set Based on Psychological Needs. Multimodal Technologies and Interaction 6/10, 2022, ISSN: 2414-4088, URL: https://www.mdpi.com/2414-4088/6/10/90.

[Kr17] Krueger, A. E.; Kurowski, S.; Pollmann, K.; Fronemann, N.; Peissner, M.: Needs profile: sensitising approach for user experience research. In: Proceedings of the 29th Australian Conference on Computer-Human Interaction - OZCHI '17. ACM Press, Brisbane, Queensland, Australia, pp. 41–48, 2017, ISBN: 978-1-4503-5379-3, URL: http://dl.acm.org/citation.cfm?doid=3152771.3152776, visited on: 11/30/2019.

[Kr22] Krueger, A. E.: Two methods for Experience Design based on the Needs Empathy Map: Persona with Needs and Needs Persona. In (Marky, K.; Grünefeld, U.; Kosch, T., eds.): Mensch und Computer 2022 - Workshopband. Gesellschaft für Informatik e.V., Bonn, 2022.

[Kr23] Krüger, A. E.; Ivanova, M.; Sattink Rath, D.; Brandenburg, S.: Experience Design Based on Values and Psychological Needs in a Corporate Context. In: International Conference on Human-Computer Interaction. Springer, pp. 146–163, 2023.

[PA20] Peters, D.; Ahmadpour, N.: Digital wellbeing through design: Evaluation of a professional development workshop on wellbeing-supportive design. In: 32nd Australian Conference on Human-Computer Interaction. Pp. 148–157, 2020.

[PAC20] Peters, D.; Ahmadpour, N.; Calvo, R. A.: Tools for Wellbeing-Supportive Design: Features, Characteristics, and Prototypes. en, MTI 4/3, p. 40, 2020, ISSN: 2414-4088, URL: https://www.mdpi.com/2414-4088/4/3/40, visited on: 12/01/2021.

[RD15] Rafael Calvo; Dorian Peters: Wellbeing Determinant Cards, Publication Title: Positive Computing, 2015, URL: http://www.positivecomputing.org/p/were-pleased-to-share-some-of-tools-and.html, visited on: 07/15/2020.

[Sc83] Schön, D. A.: The reflective practitioner: how professionals think in action. Basic Books, New York, 1983, ISBN: 978-0-465-06878-4 978-0-465-06874-6.

[Sh01] Sheldon, K. M.; Elliot, A. J.; Kim, Y.; Kasser, T.: What is satisfying about satisfying events? Testing 10 candidate psychological needs. Journal of personality and social psychology 80/2, Publisher: American Psychological Association, p. 325, 2001.

[Um20] Umbrello, S.: Imaginative value sensitive design: Using moral imagination theory to inform responsible technology design. Science and Engineering Ethics 26/2, pp. 575–595, 2020.

Hyperledger Indy Besu as a permissioned ledger in Self-sovereign Identity

Alexander Shcherbakov[1]

Abstract: Self-sovereign Identity (SSI) represents an approach to digital identity that prioritizes privacy and empowers individuals to maintain control over the information associated with their identity. This approach aligns with GDPR and similar regulations and is gaining adoption across various governments, non-profit organizations, and commercial entities worldwide. A foundational element in SSI is a Verifiable Data Registry (VDR), which serves as a trusted repository for registering and accessing public keys, schemas, identifiers, and other data. A natural choice for a VDR is a distributed ledger or blockchain. Among the most stable and popular frameworks for SSI is Hyperledger Indy. Indy includes a custom implementation of a public permissioned ledger as a VDR. The Indy community has been developing a new experimental approach for a VDR in Indy: a permissioned ledger based on Hyperledger Besu. In this paper, we are going to discuss the importance, benefits, and technical details of this initiative.

Keywords: SSI, Self-sovereign Identity, Decentralized Identity, Verifiable Credentials, DID, W3C VC, VDR, AnonCreds, Hyperledger Indy, Hyperledger Aries, Hyperledger Besu, Distributed Ledger Technologies, Blockchain, Permissioned Ledger, Ethereum

1 About Self-Sovereign Identity and Verifiable Data Registry

Self-sovereign Identity (SSI), or Decentralized Identity, is an approach to digital identity focusing on privacy and enabling individuals to maintain control over the information associated with their identity [PR21]. SSI is not tied to a specific framework or library; instead, it encompasses multiple specifications, standards, frameworks, and tools that implement SSI principles. Three key concepts are fundamental in SSI: *Verifiable Credentials* (VC), *Decentralized Identifiers* (DID), and *Verifiable Data Registries* (VDR).

A *credential* is a set of one or more claims made by an issuer. Generally, these claims describe certain properties of the credential holder. A *verifiable credential (VC)* [VC24] is a tamper-evident credential with authorship that can be cryptographically verified. Verifiable credentials can be used to build verifiable presentations, which can also be cryptographically verified.

A *Decentralized Identifier (DID)* [DI24] refers to any subject (e.g., a person, organization, thing, data model, abstract entity, etc.). In contrast to typical, federated

[1] DSR Corporation Europe, Rua Sa da Bandeira 819, 2 DTO, Porto, 4000-438, alexander.sherbakov@dsr-corporation.com

doi:10.18420/OID2024_11

identifiers, DIDs may be decoupled from centralized registries, identity providers, and certificate authorities. DIDs are URIs that associate a DID subject with a *DID document* allowing trustable interactions associated with that subject. DID document is a set of data describing the DID subject, including mechanisms, such as cryptographic public keys, that the DID subject or a DID delegate can use to authenticate itself and prove its association with the DID.

Finally, a *verifiable data registry (VDR)* [VC24][DI24] is a system that facilitates the creation, verification, updating, and/or deactivation of decentralized identifiers, DID documents, keys, and other relevant data, such as verifiable credential schemas, revocation registries, issuer public keys, and so on. In other words, VDR can be considered as a trusted place used to publish and access public keys for verification of verifiable credential signatures.

Example verifiable data registries include trusted databases, decentralized databases, government ID databases, and distributed ledgers. Distributed ledgers, in particular, are popular options for VDRs due to their effective combination of decentralization and trust. While the use of public permissionless ledgers and blockchains (such as Ethereum Main Net) as a VDR can be a reasonable option for many use cases, there are scenarios where a permissioned ledger becomes the only viable choice. This is particularly evident in government-driven ledgers or situations where the use of crypto tokens is deemed undesirable.

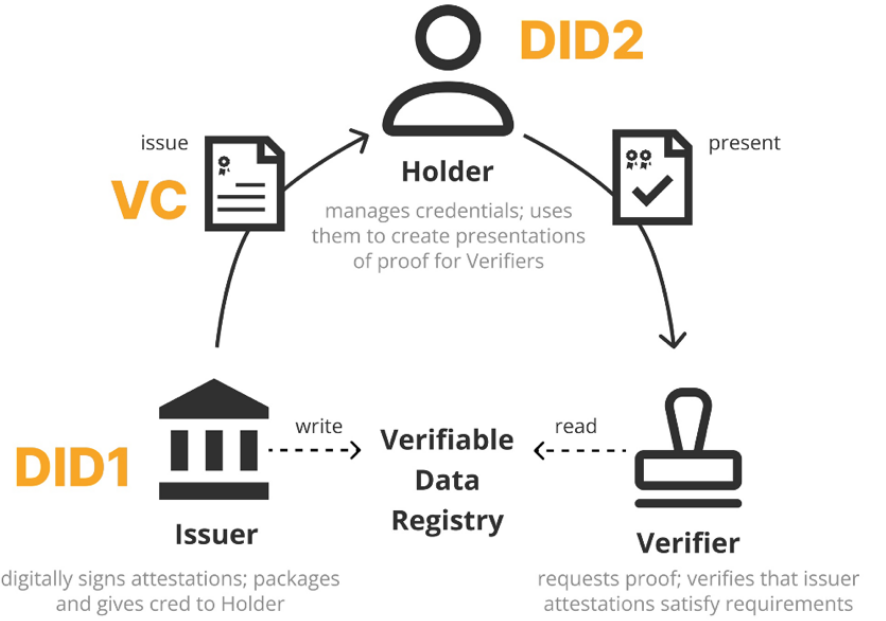


Fig. 1: SSI Workflow

Let's explore a standard SSI-based workflow (Fig. 1) by integrating all three concepts [VC24], [DI24]:

1. *Pre-requisites*. The Issuer possesses a private-public key pair used to sign verifiable credentials. The key is linked to a DID (DID1) by including the public part in the corresponding DID Document. The DID and DID Document are then published to a Verifiable Data Registry (VDR), such as a distributed ledger.

2. *Issuance*. The Issuer issues a credential for a holder identified by a DID2. The credential is signed by a key associated with DID1. Both DID1 and DID2 are included in the credential, and hence signed. The Issuer may also require the Holder to prove ownership of DID2 keys before issuance. The issued verifiable credential (claims and issuer's signature) is sent to and stored by the Holder, usually in the Holder's mobile or web wallet.

3. *Presentation*. When the Verifier requests proof from the Holder, the Holder examines the wallet and selects verifiable credential(s) that can satisfy the proof request. The Holder then creates a verifiable presentation comprising required information from the selected verifiable credential(s), as well as the corresponding Issuer's signature(s) (such as DID1's signatures) and the Holder DID2's signature (a proof of VC ownership by the Holder). The Verifier resolves Issuer's public keys associated with the Issuer's DID (DID1) via the VDR and verifies the signatures.

2 Hyperledger Indy Besu

The Indy community has been actively working on an experimental approach for a Verifiable Data Registry (VDR) within Hyperledger Indy— a permissioned ledger based on Hyperledger Besu [IB24]. To provide a comprehensive understanding, we will begin with an overview of both Hyperledger Indy and Hyperledger Besu. Subsequently, we will introduce the Indy Besu approach, delving into its importance, benefits, and technical details.

2.1 About Hyperledger Indy

Hyperledger Indy [HI24] is an open-source project under the Linux Foundation's Hyperledger umbrella. It is a decentralized and self-sovereign identity (SSI) management platform designed to give individuals, organizations and devices control over their digital identities. Indy provides the tools and protocols necessary to create, manage and verify digital identities in a secure and privacy-preserving manner. Many

Hyperledger Aries frameworks depend on Indy or some API from it under their hood.

Indy consists of two main components, ledger and client SDKs:

- Indy Ledger [IN24] is a public permissioned distributed ledger providing a decentralized, secure and tamper-evident infrastructure for managing identities. It's based on RBFT consensus protocol implemented as part of the Indy Plenum [IP24] project. Indy ledger can be used as a verifiable data registry (VDR) with did:indy [ID24] and did:sov [SD24] methods. It can also be used as a registry for CL AnonCreds [AC24] verifiable credentials to publish credential schemas and issuer's credential definition, public keys, revocation registries, etc.

- Indy SDK [IS24] is a collection of software libraries, tools and APIs that developers can use to build applications and systems that incorporate SSI features based on Indy. It includes the components for communication with the Indy Ledger, managing CL AnonCreds [AC24], verifiable credentials, establishing pairwise connections, wallet functionality, CLI, etc. The code is written in Rust and contains wrappers for all popular programming languages and platforms, including mobile.

Indy is a graduated Hyperledger project since 2019. Indy Ledger has successfully run in production for many years without significant issues as part of Sovrin [So24] and other networks.

2.2 About Hyperledger Besu

Hyperledger Besu [HB24] is a Java-based Ethereum client that has been an active and graduated Hyperledger project since 2020. This versatile framework serves two main groups of use cases:

1. *Public Networks***:** These include public Ethereum networks.

2. *Private Permissioned Networks***:** These are often associated with enterprise or supply chain ledgers.

For each category, Hyperledger Besu adopts different approaches concerning consensus protocols, supported features, components, etc.

Hyperledger Besu features a pluggable architecture and encompasses multiple implementations of consensus algorithms, including proof-of-stake (PoS), proof-of-work (PoW), and proof-of-authority (PoA). In private permissioned networks, consensus algorithms such as IBFT 2.0, QBFT, or Clique, which fall under the proof-of-authority category, are commonly employed. For implementing application-specific business logic and transactions, developers can use Solidity smart contracts.

2.3 Why New Indy Ledger

The Indy Ledger [HI24] stands out as one of the most stable and widely adopted frameworks for decentralized and self-sovereign identity. It has been successfully deployed in production systems and played a pioneering role in the early days of SSI projects, significantly contributing to the adoption and popularization of the self-sovereign identity concept. Despite its reputation for 'just working,' it is worth noting that there has been a lack of ongoing maintenance and implementation of new features.

The Indy Ledger project was initiated in 2016. At that time, stable frameworks for a public permissioned ledger were not readily available. Consequently, Indy Ledger not only incorporates the business logic of SSI-specific transactions but also encompasses the implementation of the auxiliary blockchain framework itself, including components such as the consensus protocol, permissioned logic, ledger, storage, etc. This dual role contributes to a sizable and intricate codebase, making maintenance a non-trivial task.

When the Indy project commenced, it offered a functional implementation of SSI principles at a time when modern SSI standards, such as W3C VC [VC24] and W3C DID [DI24], had not yet been established and finalized. To maintain its role as a driving force in the SSI space, Indy now requires the implementation of new features and support of recent specifications.

While there are various options for CL AnonCreds registries beyond Indy, such as cheqd or Cardano [AM24], it's noteworthy that all of these alternatives are built on permissionless proof-of-stake ledgers. While a permissionless ledger can be a reasonable choice in many cases, there are specific scenarios where a permissioned ledger, like the existing Indy Ledger, serves as a more suitable alternative.

The current Indy Ledger has limitations concerning performance (throughput, latency). Typically, performance is not a critical factor for SSI cases, as the primary actor who needs to write to the ledger is the verifiable credentials Issuer. In many cases, writing just a couple of transactions (e.g., publishing issuer’s public keys, credential schemas, etc.) is sufficient. However, if support for the revocation of verifiable credentials is required [AC24], issuers may need to write to the Indy ledger frequently (e.g., on each VC revocation). This increased frequency places higher demands on the ledger's throughput.

Decentralization is another crucial aspect. While the number of nodes in a permissioned network is typically limited, having more nodes participate in the validation of new transactions is advantageous for trust. The current Indy Ledger network often operates under the assumption of a limit of 25 nodes [So24], beyond which performance experiences a significant drop.

In response to the challenges observed in the current implementation of the Indy Ledger and the recognized need for a public permissioned ledger as an option for VDR, a new Indy Besu [IB24] initiative was proposed, and the first MVP was implemented.

The primary objective of this initiative is to preserve the favorable characteristics of the current public permissioned Indy Ledger while simultaneously reducing complexity, simplifying maintenance, enhancing performance and scalability, accelerating the development of new features, improving the end-user experience, and lowering the operational costs of Indy nodes.

A key enhancement involves replacing the custom consensus protocol implementation with Hyperledger Besu [HB24], a stable and maintained framework. This shift allows Indy Ledger to concentrate on SSI-specific business logic.

2.4 Indy Besu Benefits

In contrast to the current Indy Ledger [IN24][IP24], the new Indy Besu [IB24] has a more compact and simpler codebase. This is achieved by encapsulating blockchain complexity within the Hyperledger Besu framework, upon which the new Indy Ledger is built.

It has the following benefits and advantages for the Indy community, maintainers and users:

- Business logic (transactions) is implemented in Solidity, one of the most popular and adopted languages for smart contracts. Solidity smart contracts are easy and understandable. This will attract new developers, make it much easier to support the code, and add new technologies that are common on the market.
- The new consensus protocol significantly increases network throughput (up to 10 times) [FLK22], which can be especially beneficial for revocation features where verifiable credential issuers may publish quite significant number of transactions to the ledger.
- Hyperledger Besu allows to increase the number of validators in the network, which improves decentralization properties and trust of the network [FLK22].
- Permissioned mode (proof-of-authority consensus) follows the same principles as the current Indy ledger implementation.
- Hyperledger Besu is part of the Hyperledger family, so it is a logical bridge between two graduated Hyperledger projects.
- The new Solidity-based Indy contracts can be run on a public Ethereum Main Net as an alternative to the permissioned case.
- There is a possibility to deploy SSI/Indy logic into existing private permissioned deployments based on Hyperledger Besu (for example, to extend supply chain cases).

- A possibility to implement light client solutions [LC24].
- Indy Besu has lower hardware requirements [BD24][So24][IN24], contributing to a reduction in the overall cost of operating and maintaining Indy nodes.

Moreover, the new implementation is compatible with the old one (did:sov [SD24] and did:indy [DI24] methods), and have clear migration guides for existing deployments.

2.5 Indy Besu Technical Details

Indy Besu ledger [IB24] serves as a Verifiable Data Registry (VDR) for verifiable credentials, supporting both *W3C VC* format [VC24] and *Hyperledger AnonCreds* format [AC24]. Indy Besu, like the current Indy Ledger, can be utilized as a VDR for Hyperledger CL AnonCreds.

Additionally, similar to the existing Indy SDK and Indy VDR libraries, Indy Besu features a client SDK written in Rust, complemented by wrappers for popular languages and seamless integration with Hyperledger Aries projects.

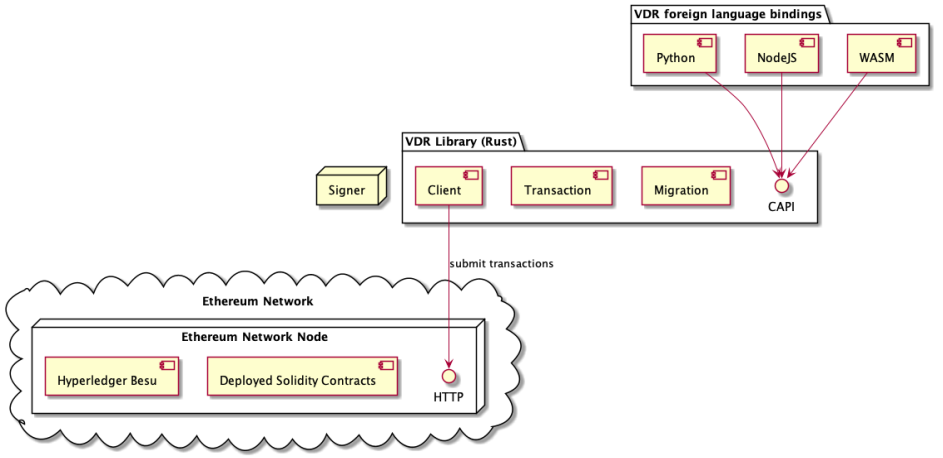


Fig. 2: Indy Besu Components

Indy Besu supports both the *did:ethr* and an extension of *did:indy* (*did:indy:besu)* methods. The choice of an appropriate DID method depends on particular deployment specifics and other requirements, such as permissionless or permissioned ledger, the need for data migration, etc. We expect that one or another method will be selected as the main one as a result of adoption.

The *did:ethr* DID method [ED24] is a well-known, standardized approach for DID management on Ethereum networks. This method adheres to *ERC-1056* [ERC24] and is designed to utilize Ethereum addresses as fully self-managed DIDs. The method is

preferrable for permissionless deployments, as it requires lower gas costs.

The new *did:indy:besu* (extension of *did:indy*) method [IB24] is designed to be compatible with the legacy *did:indy* and *did:sov* by supporting DID aliases, allowing legacy identifiers to be mapped to and associated with the new identifiers. Additionally, it utilizes Ethereum addresses as fully self-managed DIDs and provides support for light client approaches similar to the existing Indy Ledger. This method is recommended for permissioned deployments.

A compatibility layer is in place to facilitate the migration of existing Indy Ledger deployments (with *did:indy* [ID24] or *did:sov* [SD24] identifiers) to the new *did:ethr* or *did:indy:besu* based networks.

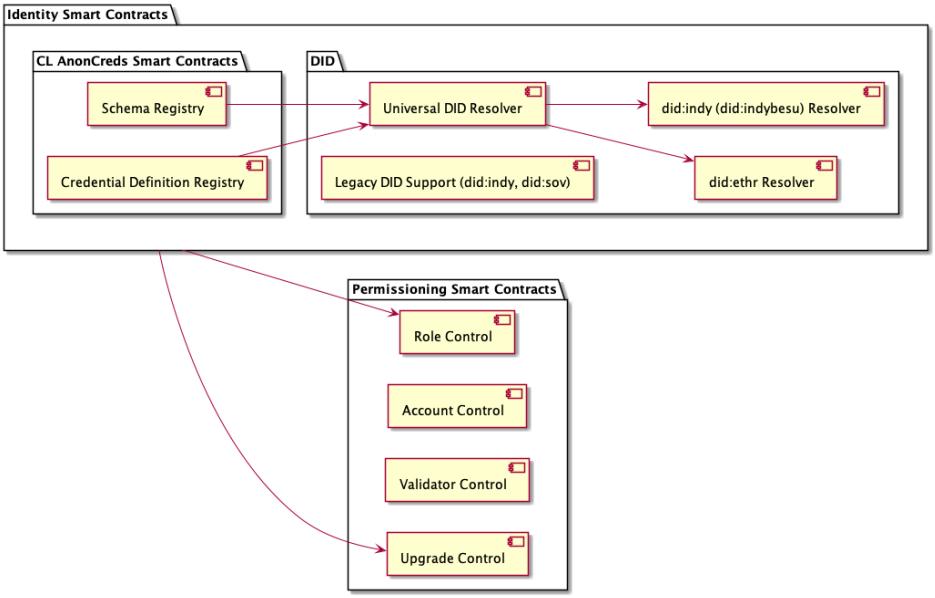


Fig. 3: Indy Besu Smart Contracts

Much like the existing Indy Ledger, Indy Besu operates as a public permissioned ledger, allowing public access to read requests while restricting write requests and the setup of new validator nodes. Leveraging the permissioned capabilities of Hyperledger Besu, Indy Besu extends these functionalities with roles-based authorization for accounts.

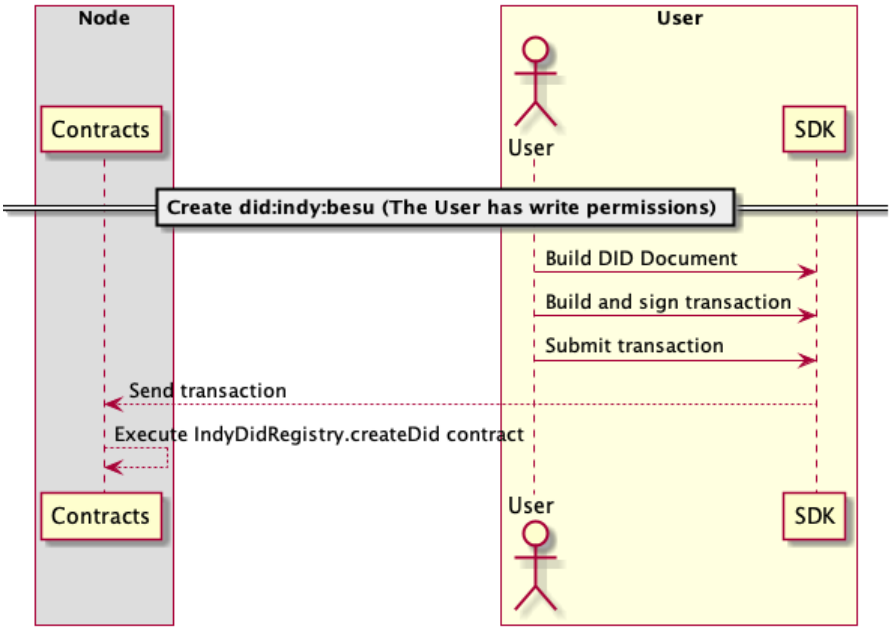


Fig. 4: Create did:indy:besu sequence diagram

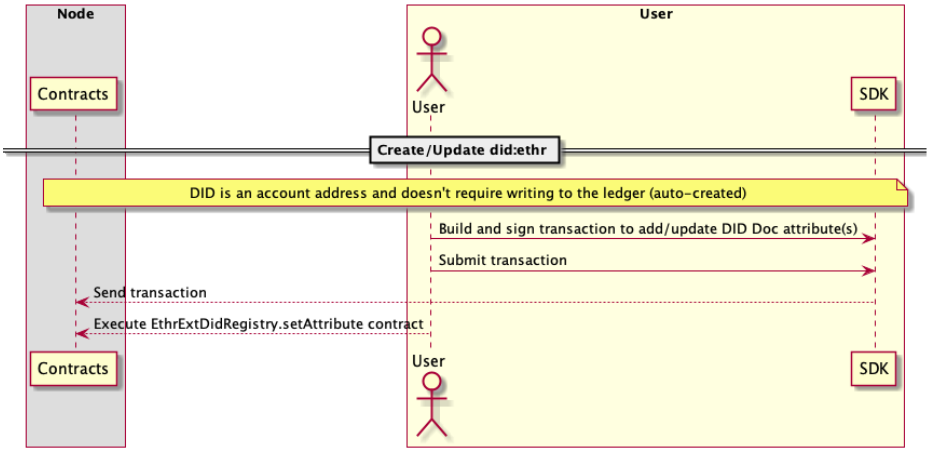


Fig. 5: Create did:ethr sequence diagram

Not all identity owners may have permissions to write transactions to a permissioned ledger. Therefore, similar to the existing Indy Ledger, Indy Besu has transaction endorsement support. The transaction endorsement is a mechanism for executing transaction writes to the ledger by a special party with an Endorser role while preserving

the original transaction author as the entity owner. This approach is applicable to all transaction types, including Hyperledger AnonCreds, did:ethr and did:indy / did:indy:besu transactions.

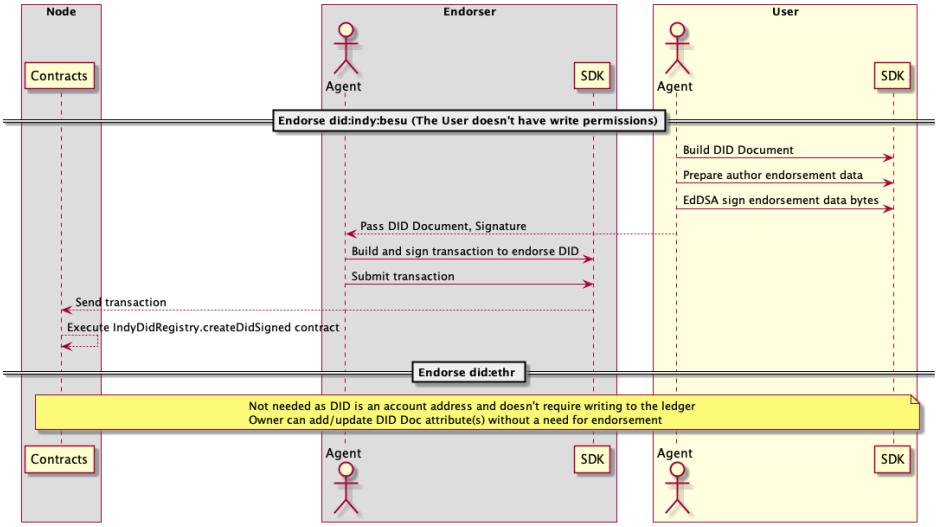


Fig. 6: DID transaction endorsement sequence diagram

3 Summary

The Indy Besu initiative is currently in an experimental state and resides in a separate Indy repository [IB24]. A minimal valuable product (MVP) has been already implemented. This project is attracting considerable attention from the Indy community and existing Indy deployments. There are promising indications that it has the potential to gradually replace the legacy Indy Ledger implementation and contribute to Self-Sovereign Identity adoption.

Bibliography

[PR21] Preukschat, A.; Reed, D.: Self-Sovereign Identity: Decentralized digital identity and verifiable credentials. Manning, 2021.

[IB24] Hyperledger Indy Besu, github.com/hyperledger/indy-besu, accessed: 31/01/2024.

[VC24] W3C Verifiable Credentials Data Model v1.1, www.w3.org/TR/vc-data-model, accessed: 31/01/2024.

[DI24] W3C Decentralized Identifiers (DID) v1.0, www.w3.org/TR/did-core, accessed: 31/01/2024.

[HI24] Hyperledger Indy, www.hyperledger.org/projects/hyperledger-indy, accessed: 31/01/2024.

[IN24] Hyperledger Indy Node, github.com/hyperledger/indy-node, accessed: 31/01/2024.

[IP24] Hyperledger Indy Plenum, github.com/hyperledger/indy-plenum, accessed: 31/01/2024.

[IS24] Hyperledger Indy SDK, github.com/hyperledger/indy-sdk, accessed: 31/01/2024.

[ID24] Indy DID Method Specification (did:indy), hyperledger.github.io/indy-did-method, accessed: 31/01/2024.

[SD24] Sovrin DID Method Specification (did:sov), sovrin-foundation.github.io/sovrin/spec/did-method-spec-template.html, accessed: 31/01/2024.

[AC24] CL AnonCreds Specification, hyperledger.github.io/anoncreds-spec, accessed: 31/01/2024.

[AM24] CL AnonCreds Methods Registry, hyperledger.github.io/anoncreds-methods-registry/, accessed: 31/01/2024.

[So24] Sovrin, sovrin.org, accessed: 31/01/2024.

[HB24] Hyperledger Besu, www.hyperledger.org/projects/besu, accessed: 31/01/2024.

[BD24] Hyperledger Besu Documentation, besu.hyperledger.org/private-networks, accessed: 31/01/2024.

[ED24] ETHR DID Method Specification (did:ethr), github.com/decentralized-identity/ethr-did-resolver/blob/master/doc/did-method-spec.md, accessed: 31/01/2024.

[ERC24] ERC-1056, github.com/ethereum/EIPs/issues/1056, accessed: 31/01/2024.

[LC24] Light Clients, ethereum.org/developers/docs/nodes-and-clients/light-clients, accessed: 31/01/2024.

[FLK22] Fan, C.; Lin, C.; Khazaei, H.; Musilek, P.: Performance Analysis of Hyperledger Besu in Private Blockchain: 2022 IEEE International Conference on Decentralized Applications and Infrastructure (DAPPS), 2022. Noah & Sons, San Francisco, pp. 64-73, 2022.

Strengthen Digital Sovereignty of Smartphone Users: Evaluation Results of a Tailored Analysis Tool for App Behavior

Susen Döbelt[1] and Dominik Lange[2]

Abstract: A usable analysis tool that provides information on risky app behavior and offers options for action, can contribute to strengthen the digital sovereignty of smartphone app users. To this end, it should be tailored and meet the requirements of a human-centered design. Therefore, we conducted a lab test with $N = 38$ participants. They evaluated a prototype of our analysis tool in terms of its usability, transparency and potential to increase self-efficacy for data protection and privacy preservation. Furthermore, we investigated the effects of the tailoring by providing a congruent and an incongruent variant for behavioral stages. Both, usability and transparency evaluations differed positively from the average. Moreover, the interaction with the tool significantly increased the participants' self-efficacy and thus strengthened their digital sovereignty. Our tailoring of texts had a positive impact at least on the efficiency evaluation. This could be further developed by extended tailoring of e.g., the GUI.

Keywords: Usable Privacy, User Research, Human-Centered Design, Tailoring, Digital Sovereignty, Smartphone Apps.

1 Introduction

The right to informational self-determination [FLR17] also applies in the digital space, and not just since the introduction of the European GDPR. However, realizing and establishing digital sovereignty [Go17] is difficult for any individual in everyday life. Smartphones and apps, for example, are daily companions that on the one hand provide easily accessible information, but on the other hand also access users' data and pass it on to third parties. There is little transparency here, but this is a prerequisite for the valued preservation of one's privacy [TM15] [Ca09] and taking appropriate - often demanding - measures. Therefore, the goal of our research project was to create a usable tool that provides transparent information on risky app behavior and offers options for action to strengthen the digital sovereignty of smartphone app users. The purpose of this study was to review the human-centered and –tailored design of this tool.

[1] Professorship of Cognitive Psychology and Human Factors, Chemnitz University of Technology, Wilhelm-Raabe-Str. 43, 09120 Chemnitz, Germany, susen.doebelt@psychologie.tu-chemnitz.de
[2] Professorship of Cognitive Psychology and Human Factors, Chemnitz University of Technology, Wilhelm-Raabe-Str. 43, 09120 Chemnitz, Germany

doi:10.18420/OID2024_12

Several analysis methods that create transparency on the behavior of smartphone apps and provide information regarding the associated risk have been available for some years [En14]. Results can be requested by interested users from specific platforms (e.g., App-Checker [In22], Appicator [Fr24]) or provided via an app directly to their smartphone (e.g., Androlyzer [Te20]). Very few of these tools have been developed systematically in terms of user- or human-centeredness [In19], empirically evaluated concerning usability, user experience (UX), or even behavioral implications.

In the field of research, Gerber et al. [Ge18] investigated the effects of the app *FoxIT*, which entailed a static permission analysis method and feedback about the risk of installed apps. Results of a field trial showed that privacy awareness and knowledge of participants increased. Furthermore, changes in smartphone settings were reported. In a lab study with a provided smartphone, Bal et al. [BRH14] were able to show that their tool *Styx* based on dynamic monitoring of information flows via *TaintDroid* [En14], was perceived as user-friendly. The usage contributed to increasing user confidence due to the transparent information on app information flows. Van Kleek et al. [Kl17] suggested using Data Controller Indicators to disclose the transfer of information to third parties. The user evaluation revealed that more transparent information leads to decisions for apps with fewer organizations receiving data and could support users to make confident decisions.

In addition, Döbelt et al. [Dö20] were also able to highlight the importance of transparency and usability based on two user studies (online and lab) and five guidelines for tools that analyze app behavior. However, transparency played a specific role here on two dimensions: 1.) In terms of information delivery about apps, and 2.) In terms of the design of the tool itself: it should meet the requirements of transparency on its data handling [Dö20]. If tools are to be developed to further strengthen the digital sovereignty of users, additional options for action should be offered [Dö20]. Here, however, a 'one size fits all' solution is not sufficient [DH23] to encourage behavior change. Instead, a comprehensive alignment between the users and the tool may further increase its effectiveness [Li16]. Therefore, a tailored approach has been proposed [Kn15]. The *User-Tailored Privacy by Design* model [Wi17] suggests personalized nudges. In this context, the classification of behavioral stages and tailored interventions, like those for changing pro-environmental behavior [Ba13b], could serve as a template [DG21].

Based on this literature, we developed a human-centered prototype of an analysis tool with behavioral stage-specific tailoring. The tool was supposed to make the handling of app data accessible, visible and assessable. In addition to usability and transparency, we also investigated the impact on perceived self-efficacy. This describes the individual's assessment of being able to learn an action, even if they have not currently been able to master it [SJ02]. It is an important factor in predicting whether a person will consider or perform an action [Ba00]. Therefore, it presents an interesting variable to study the impact of our prototype on the digital sovereignty of smartphone app users.

3 Research questions and hypotheses

In our laby study, we focused on two research questions for our prototype: **(RQ1)** How usable and transparent is it evaluated? and **(RQ2)** Can it increase self-efficacy regarding data protection and privacy preservation of smartphone app users? Based on the presented content as well as the underlying user-centered design guidelines, we assume:

H1: The prototype is evaluated as above-average usable and transparent.

H2: Self-efficacy regarding data protection and privacy preservation is higher after interaction with the prototype than before.

In addition, a third exploratory question arose regarding the tailoring of the prototype: **(RQ3)** On which aspects of UX does the congruent tailoring to the behavioral stage have a positive impact?

4 Method

4.1 Study Design

Independent variables
For RQ1, the independent variable represents the use of our prototype; for RQ2, the time of measurement before (T1) and after (T2) the use. For RQ3, a distinction was made about the variant of the prototype. Based on the behavioral stage assessed during the recruitment, the congruent or incongruent variant was presented randomly and counterbalanced. This self-assessment of behavioral stage was adapted from the domain of pro-environmental behavior [Ba13b]. Participants who assigned themselves to the stage *predecision* or *preaction* were grouped into the early behavioral stage group, all others (*action* or *postaction*) into the late one.

Dependent variables
To evaluate our prototype, several established questionnaires were used: For testing H1, the System Usability Scale (SUS; [Br96] [RRR13]). It captures the usability of a system using 10 items answered via a 5-point rating scale ("*strongly disagree*"-"*strongly agree*"). Answers are aggregated to the SUS score (0-100), which allows results to be classified in grades from "*A+*" to "*F*" [LS18]. We chose 68.00 ("*C*") as a benchmark, which is at the center of the curved grading scale. Additionally, the SIPAS questionnaire [Sc21] was used to assess transparency. It measures the ability to perceive, understand and predict the information processing of a system [Sc21]. The 6 items are answered using a 6-point rating scale ("*not at all*"-"*completely*" true). The suggested unidimensional scale [SF22] is used here. Since the questionnaire was developed for a different field of application, the mean value (3.5) of the response scale serves as a benchmark.

was adapted to our application context (privacy preservation and data protection). The 10 items are answered on a 4-point rating scale (from "*not agree*" to "*agree exactly*").

Further, the User Experience Questionnaire (UEQ; [LHS08]) was used to investigate the exploratory RQ3. It captures UX by 6 subscales [Sc23]: Attractiveness (valence and overall impression), Perspicuity (easy to learn), Efficiency (without unnecessary effort), Dependability (control of interaction), Stimulation (exciting and motivating use), and Novelty (innovative and creative design). The UEQ contains 26 items and is designed as a 7-level semantic differential (from "*-3*" to "*+3*") with contrasting adjectives. The benchmarks referenced in [STH17] were used to interpret mean evaluations.

4.2 Sample

A total of $N = 38$ (26 female) subjects participated in the lab study. They were 24 years old on average ($M = 23.95$, $SD = 5.03$, min = 18.00, max = 41.00), and 71% reported a high school diploma as their highest educational qualification. The majority (92%) were third-semester students ($M = 2.57$, $SD = 2.33$) of psychology (69%). Participants rated themselves as significantly less tech-savvy ($M = 3.76$, $SD = .93$; $t(37) = -2.51$; $p = .017$, $d = -0.41$) than a comparable sample ($N = 300$; $M = 4.14$; [FAW19]). With regard to smartphone competence, the ratings ($M = 3.82$, $SD = .70$) were significantly ($t(37) = 3.13$, $p = .003$, $d = 0.51$) higher than those of a 2009 comparison sample ($N = 460$, $M = 3.47$; [Ka09]). This is only marginally true for their competence with apps ($M = 3.64$, $SD = .53$; $t(37) = 1.96$, $p = .057$, $d = .32$). The participants mainly (53%) used a smartphone with an Android operating system (OS) v12.

4.3 Procedure

The study started at the end of 2022 in Chemnitz, Germany, was conducted under pandemic-related restrictions, and approved in advance by the university's ethics committee. The procedure here is shortened and originally included additional tasks, variables, and qualitative data collection, which are omitted in favor of a reduced presentation.

Recruitment and selection of participants
The study announcement contained the rough goal (understand apps in more detail and evaluate a prototype) as well as the different remuneration options. It was distributed via various channels online and offline. Interested persons accessed the recruitment questionnaire via a QR code or link, implemented with *LimeSurvey* (v3.28.29+220920). After a brief welcome, respondents were first asked to generate an individual subject code for a pseudonymization of the data. This was followed by demographic information (age, gender, education, employment) and on the smartphone in use, i.e. OS and version. In the following, the behavioral stage was assessed (adapted from [Ba13a]; incl. n/a option). Lastly, contact details and compensation preferences were entered separately. To complete the questionnaire took a mean of 13 min ($SD = 5.60$). Participants selected

for the lab test used a smartphone with an Android OS > 6.0, to ensure seamless handling of the test smartphone. Furthermore, as of this version, runtime modification of permissions was possible. Individuals for whom behavioral stages were also available received an invitation.

Equipment and laboratory study
The laboratory room was equipped with two tables one for the participant and one for the experimenter. They were aligned at right angles to each other to avoid direct observation. Additional hardware (*ASUS* UX32V notebook, *LG* Flatron E2411 monitor, *Logitech* B110 mouse, *Cherry* G230 keyboard and the test smartphone (*Samsung Galaxy A33 5G*, OS v12) was placed on the participants' table to conduct the study.

After the room and equipment had been prepared in accordance with the hygiene policy, the participants were welcomed, signed the information on participation and data protection, and were presented with the experimental materials. They were informed that they could ask the experimenter questions at any time and that the app was not yet fully developed. The test started with the collection of self-descriptive variables (technology affinity, smartphone competence). Participants were then asked to rate their self-efficacy expectations in advance. They then had 10 minutes to freely explore the prototype. This was followed by two tasks of 5 minutes each: 1.) Inform yourself about the app *FitnessPro* with the help of the prototype; 2.) Take action using the prototype to minimize the risk of the app *FitnessPro*. Next, the participants were asked to fill out the evaluation questionnaires and the repeated self-efficacy assessments. In the end, the participants were remunerated. The test lasted a mean of 74 minutes (*SD* = 7.23).

4.4 Material: Prototype user interface

The prototype was implemented as smartphone app on the device it provides information about. The content was presented in German, but anonymized and translated here. The user interface (UI) was structured into three main areas: 1.) App risk, 2.) Device security, and 3.) Third-party providers. The examined area is described in Figure 1 and below (for more information see [Ch24]).

The start page of the app displayed a pie chart that included the number as well as the average risk score of the apps analyzed (only one for the lab test). The app list entries below were color- and numerically-coded according to their risk value. When selected, a drop-down at the very top of the subsequent screen explained this value. Below, the user could: a) Access analysis details: app permissions and third-party providers, and b) Take options for action: Change permissions, install alternatives or delete the app. The *Analyzed Permissions* page contained results from dynamic app analysis and highlighted whether a permission is currently withdrawn (green hand) or granted (orange exclamation mark). The *Analyzed Third-Party Providers* section entailed an overview of those providers to which data is forwarded other than the app provider. Here, also their risk value and the corresponding description were displayed. In the lower part *Options*

installed from a store to be defined. The final option was to uninstall the app.

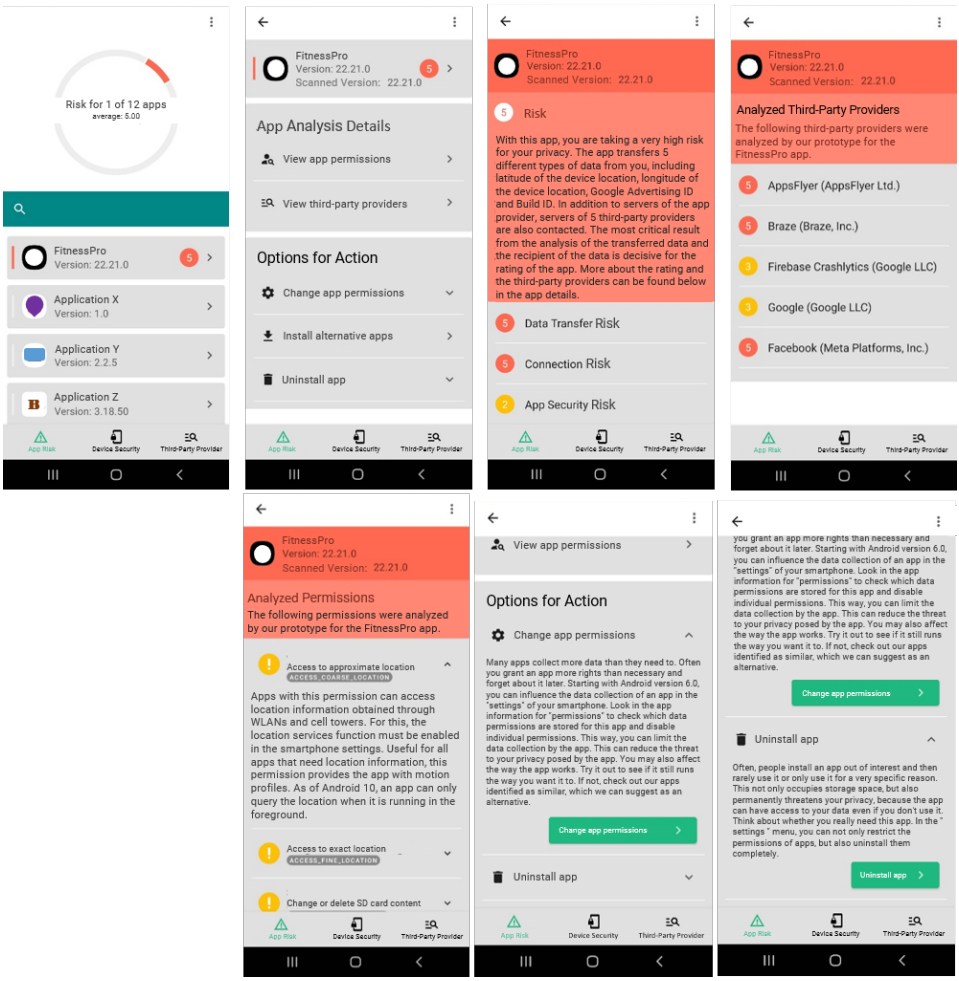


Fig. 1: UI of the prototypes risk area at the time of the laboratory test, texts here are translated into English and were originally presented in German, names and icons of apps have been anonymized

Prototype tailoring: As theoretical work (see section 2) implies that the adaption of a system to the user offers advantages, two prototype variants were implemented. These are tailored to the early or late behavioral stages. Different texts have been applied in the three areas under *Options for Action* to provide specific guidance to users. For the early behavioral stage, the texts began with a problem description to build problem awareness. The description of the options for action is also more detailed than that of the late ones. Additional feedback for behavior was added here instead.

5 Results

The raw data was first recoded according to the respective analysis instructions, and mean values or subscales were calculated. These are first reported descriptively and analyzed with respect to their distribution. Afterwards, a non- or parametric test was applied to examine the hypotheses.

5.1 **Usability and transparency** (RQ1)

Descriptive data revealed that the subjects rated the prototype with a mean SUS score of $M = 80.00$ ($SD = 13.90$). This corresponds to the grade "*A*-", i.e. a very good rating. The items assessing transparency were rated with "*somewhat agree*" ($M = 4.39$, $SD = 0.71$). The Shapiro-Wilk test for the SUS score showed a significant difference ($W(38) = .884$, $p < .001$) from the normal distribution. Therefore, we tested nonparametric and according to H1 one-tailed versus the mean SUS score of 68.00. The Wilcoxon test revealed a significantly higher SUS score with a high effect size compared to the selected benchmark ($W_s = 660.00$, $z = 4.02$, $p < .001$, $d = 1.86$). To further examine H1, the normal distributed transparency score ($W(38) = .946$, $p = .066$) was tested by a one-tailed t-test against the mean of the response scale (3.5). This also revealed a significant difference ($t(37) = 7.643$, $p < .001$, $d = 1.24$) with a large effect size.

5.2 **Self-efficacy** (RQ2)

Descriptive analysis of self-efficacy ratings revealed that at T1 the mean assessment was $M_{T1} = 2.38$ ($SD_{T1} = 0.54$, "*hardly true*") and at T2 $M_{T2} = 2.52$ ($SD_{T2} = 0.50$, "*somewhat true*"). At both times of measurement, the data was normally distributed ($W_{T1}(38) = .971$, $p_{T1} = .423$; $W_{T2}(38) = .984$, $p_{T2} = .852$). A one-tailed t-test for dependent samples was used to test the H2 and revealed a significant increase in self-efficacy ($t(37) = - 2.71$, $p = .005$, $d = - 0.44$; see Figure 2) with a small effect size.

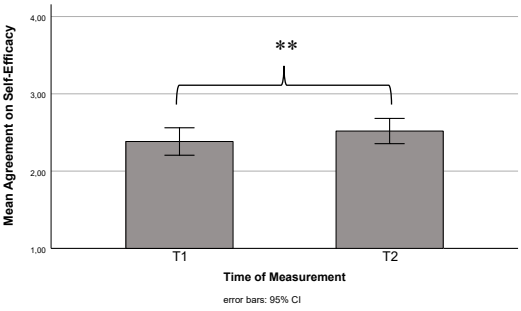


Fig. 2: Mean agreement of $N = 38$ participants for on self-efficacy, pre (T1) and post (T2) interaction with the prototype, ** marks a significant difference $p < .01$

The descriptive results of the UEQ (Table 1) showed that the mean ratings in the incongruent condition could be categorized largely as "*above average*" or "*good*", while in the congruent variant were mostly "*excellent*".

	Incongruent					Congruent				
UEQ Scale	*M*	*SD*	Min	Max	Category	*M*	*SD*	Min	Max	Category
Attr	1.41	0.78	-0.83	3.00	Above Average	1.57	0.67	0.50	2.83	Good
Persp	1.70	1.22	-2.25	3.00	Good	2.05	0.75	-0.25	3.00	Excellent
Eff	1.47	0.63	-0.25	2.25	Good	1.97	0.58	1.00	3.00	Excellent
Dep	1.78	0.56	0.75	2.75	Excellent	1.82	0.56	0.50	2.75	Excellent
Stim	0.96	0.92	-0.75	2.50	Below Average	1.29	0.55	0.00	2.25	Above Average
Nov	0.95	0.90	-1.50	2.50	Above Average	1.20	0.89	-0.50	2.75	Good

Tab. 1: UEQ evaluation for incongruent (n = 19) and congruent (n = 19) prototype variant, Attr = attractiveness, Persp = perspicuity, Eff = efficiency, Dep = dependability, Stim = stimulation, Nov = novelty; benchmark category for mean evaluation follows [STH17]

The normal distribution was violated for perspicuity ($W_{incon}(19) = .819$, $p_{incon} = .002$; $W_{con}(19) = .836$, $p_{con} = .004$) and was therefore analyzed nonparametrically. A one-tailed t-test for independent samples was calculated for all other subscales and a Bonferroni correction ($p = .010$) was applied. Only one statistically significant difference with a strong effect size emerged for the efficiency ($t(36) = 2.543$, $p = .008$, $d = 0.83$; Figure 3).

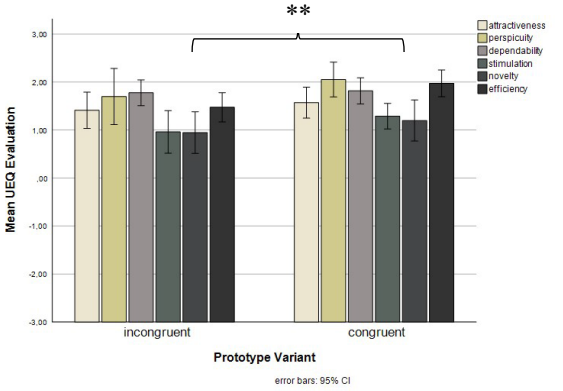


Fig. 3: Mean UEQ evaluations for subscales in the incongruent (n = 19) and congruent (n = 19) prototype condition, ** marks a significant difference $p < .01$

6 Discussion and limitations

The results of our laboratory test showed that our prototype was appreciated for its usability and transparency (H1 confirmed). Thus, the human-centered development based on specific design guidelines resulted in a positive evaluation of the tool. However, the evaluation could be positively biased, due to the notice that the app was a prototype developed as part of a research project. The participants may therefore have rated our app more leniently than they would have done with mature products. Additionally, the first positive impression might wear off long term. Least is opposed by the fact that usability evaluations seem to increase over time [Ku11]. In addition, the interaction with our prototype was predetermined by the tasks set. Thus, all participants got to know it equally and comprehensively. It remains unclear whether all the information would have been noticed in a self-determined interaction.

The H2 was also confirmed, as interaction with our prototype significantly increased participants' self-efficacy. Thus, our tool can contribute to strengthening users' digital sovereignty. Even though the prototype's objective was not described in advance, it was revealed during the lab test. A positive biasing of the evaluation due to social desirability is also possible here. In order to address this, it was pointed out that there are no right or wrong answers.

For our third exploratory question, the results show that our tailoring to the behavioral stages had a positive effect on the evaluation of efficiency. Our tool was perceived as more pragmatic and less cluttered in the congruent condition. However, differences remained limited to this aspect, probably because the prototype variants differed only in the texts. Different graphical designs might have been more crucial for further aspects.

The limitations of our lab study relate mainly to its external validity and transferability to real world settings. We provided a prepared smartphone to our participating sample of students, which differ in many respects from the overall population [An16]. Therefore, some results may be over- or underrepresented. Regarding the internal validity of the study, it can be noted that established instruments were slightly reformulated to adapt them to the present context. These changes could have an impact on construct validity, which cannot be followed up here.

6.1 Future work

Our tool will be further developed during the research project [Ch24]. The scope of smartphone security in particular will be expanded and evaluated again. This will include a follow-up survey examining the impact on behavioral change beyond the lab. Further research is also needed on the tailoring of analysis tools like ours. There are still many possibilities in the design of the graphical user interface (GUI) of information and options for action that can contribute, e.g., to increase awareness of the problem and the motivation to exercise digital sovereignty.

The usability of our prototype was rated as "*very good*" and its transparency was "*somewhat approved*". Both evaluations differ statistically significantly and positively from the average. Our prototype can therefore be considered as usable and transparent (RQ1). Additionally, it was encouraging that the interaction with our prototype significantly increased the self-efficacy of the participants (RQ2). Our prototype thus can contribute to the purpose of strengthening the digital sovereignty of smartphone app users. In addition, we found that our tailoring for the behavioral stages had a positive impact on the efficiency evaluation (RQ3). The behaviorally congruent app variant thus contributes to decreasing at least the resource investment of app users interested in protecting their privacy. This first result could be further developed by extended tailoring of other e.g., GUI elements.

Acknowledgments

This study was conducted as a part of the research project "*PANDERAM*" (funding code: 16SV8521, [Ch24]), funded by the German Federal Ministry of Education and Research. We thank our project partners Fraunhofer SIT, the Institut für Technik und Journalismus e.V., the DAI-Laboratory of TU Berlin and secuvera GmbH for their support in the prototype development. Furthermore, we want to thank Anna Graue, who was in charge of the experiments in the lab and supported the analysis of the data.

Bibliography

[An16] Andone, I. et.al.: How age and gender affect smartphone usage. In: Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing: adjunct, pp. 9-12. 2016.

[BRH14] Bal, G.; Rannenberg, K.; Hong, J.: Styx: Design and Evaluation of a New Privacy Risk Communication Method for Smartphones. In (Cuppens-Boulahia, N., Cuppens, F., Jajodia, S., El Kalam, A. A., Sans, T. eds.): ICT Systems Security and Privacy Protection. IFIP Advances in Information and Communication Technology, 428. Berlin, Heidelberg: Springer. 2014.

[Ba13a] Bamberg, S.: Applying the stage model of self-regulated behavioral change in a car use reduction intervention. Journal of Environmental Psychology, 33, 68–75. 2013.

[Ba13b] Bamberg, S.: Changing environmentally harmful behaviors: A stage model of self-regulated behavioral change. Journal of Environmental Psychology, 34, 151–159. 2013.

[Ba00] Bandura, A.: Social Cognitive Theory: An Agentic Perspective. Annual Review of Psychology, 52, pp. 1–26. 2000.

[Br96] Brooke, J.: SUS: A "Quick and Dirty" Usability Scale. Usability Evaluation. In Industry, 189(194), pp. 4-7. 1996.

[Ca09] Cavoukian, A.: Privacy by design: The 7 foundational principles. Information and

Privacy Commissioner of Ontario, Canada, 5. 2009.

[Ch24] TU Chemnitz, Professorship Cognitive Psychology and Human Factors: PANDERAM: https://www.tu-chemnitz.de/hsw/psychologie/professuren/allpsy1/forschung/panderam/index.html.en, accessed: 15/01/2024.

[Dö20] Döbelt, S. et.al.: Clearing the Hurdles: How to Design Privacy Nudges for Mobile Application Users. In (Moallem, A., ed.): Proceedings of the Second International Conference on HCI for Cybersecurity, Privacy and Trust, Held as Part of the 22nd HCI International Conference, Copenhagen, Denmark, pp. 326-353, Springer, Cham. 2020.

[DG21] Döbelt, S.; Günther, M.: Two Values Work Alike: Linking Proenvironmental and Privacy Preserving Behavior. In (Huckauf, A., Baumann, M., Ernst, M., Herbert, C., Kiefer, M., Sauter, M., eds.): TeaP 2021. Lengerich: Pabst Science Publishers. 2021.

[DH23] Döbelt, S.; Halama, J.: Clean up my Phone: Field Trial Results of a Privacy Tool Increasing Transparency about App Behavior. International Symposium on Technikpsychologie, Darmstadt, Germany. 2023.

[En14] Enck, W. et.al.: TaintDroid: an information-flow tracking system for realtime privacy monitoring on smartphones. ACM Transactions on Computer Systems, 32(2). 2014.

[FAW19] Franke, T.; Attig, C.; Wessel, D.: A personal resource for technology interaction: development and validation of the affinity for technology interaction (ATI) scale. International Journal of Human–Computer Interaction, 35(6), pp. 456-467. 2019.

[Fr24] Fraunhofer SIT. Appicator – Framework for App Securtiy Tests. https://www.sit.fraunhofer.de/en/appicaptor/?MP=iwllzb, accessed: 15/01/2024.

[FLR17] Friedewald, M.; Lamla, J.; Roßnagel, A.: Informationelle Selbstbestimmung im digitalen Wandel, Springer Vieweg, ISBN 978-3-658-17661-7. 2017.

[Ge18] Gerber, N. et.al.: FoxIT: enhancing mobile users' privacy behavior by increasing knowledge and awareness. In: Proceedings of the 7th Workshop on Socio-Technical Aspects in Security and Trust, pp. 53-63. ACM. 2018.

[Go17] Goldacker, G.: Digitale Souveränität. https://publica-rest.fraunhofer.de/server/api/core/bitstreams/71c726ab-133b-4cd9-8e4c-3d259453fcf8/content. 2017. accessed: 15/01/2024.

[In19] International Organization for Standardization: Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. (ISO 9241-210:2019). https://www.iso.org/standard/77520.html. 2019. accessed: 15/01/2024.

[In22] Institut für Technik und Journalismus e.V.: Mobil Sicher – App Checker. https://appcheck.mobilsicher.de/. 2022. accessed: 15/01/2024

[Ka09] Karrer, K. et.al.: Technikaffinität erfassen–der Fragebogen TA-EG. Der Mensch im Mittelpunkt technischer Systeme, 8, pp. 196-201. 2009.

[Kl17] Kleek, M. v. et.al.: Better the devil you know: Exposing the data sharing practices of smartphone apps. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 5208-5220. ACM. 2017.

[Kn15] Knijnenburg, B. P.: A user-tailored approach to privacy decision support, Ph.D. Thesis, University of California, Irvine, CA. 2015.

[LHS08] Laugwitz, B.; Held, T.; Schrepp, M.: Construction and Evaluation of a User Experience Questionnaire. In (Holzinger, A., ed.): HCI and Usability for Education and Work, pp. 63–76. Springer. 2008.

[LS18] Lewis, J. R.; Sauro, J.: Item benchmarks for the system usability scale. Journal of Usability Studies, 13(3). 2018.

[Li16] Liu, B. et.al.: Follow my recommendations: A personalized privacy assistant for mobile app permissions. In: SOUPS'16 Proceedings of the Twelfth USENIX Conference on Usable Privacy and Security, pp. 27-41. 2016.

[RRR13] Rummel, B.; Ruegenhagen, E.; Reinhardt, W.: System Usability Scale (Translated into German). https://blogs.sap.com/2016/02/01/system-usability-scale-jetzt-auch-auf-deutsch/. 2013. accessed: 15/01/2024.

[Sc23] Schrepp, M.: User ExperiencevQuestionnaire Handbook. https://www.ueq-online.org/Material/Handbook.pdf. 2023. Accessed: 05/05/2024

[STH17] Schrepp, M.; Thomaschewski, J.; Hinderks, A.: Construction of a Benchmark for the User Experience Questionnaire (UEQ). In: International Journal of Interactive Multimedia and Artificial Intelligence, 4(4), 40. Universidad Internacional de La Rioja. 2017.

[Sc21] Schrills, T. et.al.: Are Users in the Loop? Development of the Subjective Information Processing Awareness Scale to Assess XAI. 2021.

[SF22] Schrills, T.; Franke, T.: How do Users Experience Traceability of AI Systems? Examining Subjective Information Processing Awareness in Automated Insulin Delivery Systems. 2022.

[SJ02] Schwarzer, R.; Jerusalem, M.: Das Konzept der Selbstwirksamkeit. In (Jerusalem, M., Hopf, D., eds.): Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen, pp. 28–53. Beltz. https://www.pedocs.de/frontdoor.php?source_opus=3930. 2002. accessed: 15/01/2024.

[SJ03] Schwarzer, R.; Jerusalem, M.: SWE. Skala zur Allgemeinen Selbstwirksamkeitserwartung. In (Leibniz-Institut für Psychologie, ed.): Open Test Archive. Trier: ZPID. 2003.

[Te20] TU Berlin, DAI-Labor.: Androlyzer – Privacy Scanner for Android: https://download.cnet.com/Androlyzer-Privacy-Scanner/3000-20432_4-78524108.html. 2020, accessed: 15/01/2024.

[TM15] Trepte, S.; Masur, P. K.: Privatheit im Wandel. Eine repräsentative Umfrage zur Wahrnehmung und Beurteilung von Privatheit. Stuttgart: Universität Hohenheim. https://www.uni-hohenheim.de/fileadmin/einrichtungen/psych/Team_MP/Berichte/Bericht_-_Privatheit_im_Wandel_2014-06-18.pdf. 2015.

[Wi17] Wilkinson, D. et.al.: Work in Progress: User-Tailored Privacy by Design. Paper presented at 24th Annual Network and Distributed System Security Symposium, San Diego. 2017.

Qualified Electronic Signatures with the EU Digital Identity Wallet

Tobias Wich[1], Detlef Hühnlein[1], Florian Otto[1], and Mike Prechtl[1]

Abstract: Art. 5a of the amended eIDAS-Regulation (EU) 2024/1183 establishes the European Digital Identity Framework and introduces the European Digitial Identity Wallet (EUDIW), which will meet the requirements of assurance level "high" for identity proofing and authentication (see Art. 5a Nr. 11) and is envisioned to be able to create Qualified Electronic Signatures (QES) free of charge for non-professional purposes (see Art. 5a Nr. 4 (e) and Nr. 5 (g)). As it will not be feasible in practice to certify the secure elements of all smartphones in the market as Qualified Signature Creation Device (QSCD), one needs to look at remote signature solutions along the lines of ETSI TS 119 432 and the specification developed within the Cloud Signature Consortium (CSC) . The Architecture and Reference Framework (ARF) makes it clear that the EUDIW will support Verifiable Credentials (VCs) for the purpose of strong identification and authentication and the only missing step to enable QES in the EUDIW seems to be the integration of Verifiable Credentials and Verifiable Presentations according to W3C with the remote signature protocol of the CSC-API. The present paper shows how to integrate the two worlds to enable QES in the EUDIW using emerging standards, such as Selective Disclosure JSON Web Tokens (SD-JWT) and OpenID for Verifiable Presentations (OID4VP).

Keywords: QES, EUDIW, eIDAS, OID4VP, SD-JWT, Verifiable Credential (VC)

1 Introduction

Based on the existing remote signature standard ETSI TS 119 432 [ET20] produced by the technical committee for Electronic Signatures and Infrastructures (ESI) within ETSI, which refers to version 1.3 of the CSC-API, which by now has been updated to version 2.0.0.2 [CSC22], and the specification of the main functionality and interfaces of the EU Digital Identity Wallet (EUDIW) within the Architecture and Reference Framework [ARF23] it is straightforward to come up with a high level architecture for the realisation of QES with the EUDIW, as depicted in Figure 1.

In this architecture the EUDIW is both used for the identification step according to Art. 24 Nr. 1a (a) of (EU) 2024/1183 [EU24], which is necessary to create a qualified certificate, and for the subsequent Signature Activation Protocol (SAP), whereas the EUDIW is serving as Signer Interaction Component (SIC). While based on the [ARF23] it is clear that the protocol, which is to be used for the identification step is OID4VP, it is not that obvious which protocol should be used to realise the SAP.

Lastly, it is worth mentioning that four large-scale pilot projects have been initiated to test the EUDIW. One of these initiatives, the POTENTIAL consortium, is known to have

1 ecsec GmbH, Sudetenstr. 16, 96247 Michelau, Germany, tobias.wich@ecsec.de; detlef.huehnlein@ecsec.de; florian.otto@ecsec.de; mike.prechtl@ecsec.de

doi:10.18420/OID2024_13

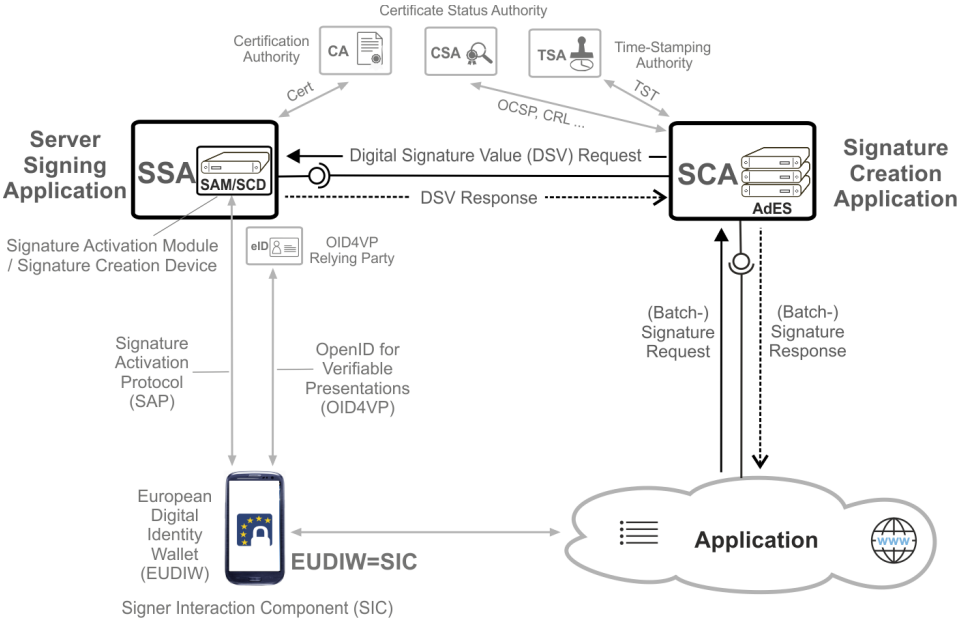


Fig. 1: Solution outline for realising QES with the EUDIW

produced a draft for a QES specification [PO24]. This draft is expected to be published as POTENTIAL D3.1.

2 Standards for defining and using Verifiable Credentials

The technology developed in the vicinity of VCs is still in constant flux and seems to be changing rapidly. At first sight it may even seem that the rate of new specifications in the field is still increasing. A closer look however shows that

a) some of the new specifications replace other unfinished ones without changing the fundamental concepts, or b) provide new and simpler approaches.

Furthermore, it needs to be noted, that there is a strong commitment of the European Union to the OID4VP, OID4VCI and SD-JWT specifications and as Art. 5a (1) and Art. 5c (1) of (EU) 2024/1183 [EU24] requires that each European Member state issues or endorses at least one certified EUDIW until November 21 2026 one may expect that the relevant technical standards will be finalised rather soon.

VC Formats form the core of the VC ecosystem. They must be differentiated between the definition of the VC as an abstract verifiable entity and the definition of the use case

specific data. While the foundational specifications of VCs are developed by work groups in W3C, IETF, and other standardisation bodies, the use case specific data models are usually in the hands of organisations related to the systems issuing the actual credentials. They are therefore out of scope here.

On the structural definition side, a noteworthy variant are AnonCreds [An23], which are centered around the Hyperledger blockchain system. Having this strong binding to a specific blockchain system, these credentials have some major limitations regarding offline capability and the aspirated self sovereignty.

A variant trying to overcome these limitations has been created with the W3C VC Data Model Version 1.1 [W3C22b], which is currently being renewed to create the W3C VC Data Model Version 2.0 [W3C24a]. The W3C model is based on JSON-LD [W3C20] in order to provide flexible mechanisms to extend and transform the credential data in the receiving software. While these are desirable properties, it comes at the cost of a more complicated model and less mature libraries and tools to process the data. Furthermore, W3C credentials are completed by numerous specifications describing the proof algorithms, which range from the classical signature schemes based on ECDSA [W3C24b] to zero knowledge proofs using a BBS signature algorithm based on pairing-based cryptography [W3C24c]. It is important to note, that these algorithms are designed to provide a selective disclosure functionality.

Considering that the W3C credentials are quite heavy with respect to the complexity and high number of required additional specifications, SD-JWT [SD24a; SD24b] strives to provide a simpler alternative based on the established technology of JSON Web Tokens, while still providing selective disclosure. The basic principle of SD-JWTs is that the JWT itself contains only a list of hashes instead of the actual claims. The claims that should be revealed are then simply attached to the serialised form, when revealed to the verifier.

Issuance and proof protocols form the bridge between wallet, issuer and verifier. These protocols enable to exchange credentials and proofs of credentials between these parties.

DIDComm [DID23] is a generic communication protocol based on the use of Decentral Identifiers (DIDs) [W3C22a]. It is completed by numerous specifications for issuance, proof conveyance, and further use cases[2]. A major problem with DIDComm is the fact, that it comes with its own messaging paradigm (routing, transport types, etc.), which is not a very good fit for the classical web.

OID4VCI and OID4VP aim to reduce this complexity by only focusing on issuance and verification in the style of the well established methodology used within the OAuth 2.0 Authorization Framework and OpenID Connect. Additionally these protocols are extensible with respect to the used VC and key proof types.

2 https://didcomm.org/search/

3 Signature Release within the CSC-API using VCs

Signatures are released within the CSC-API by performing a Credential Authorization. Credential hereby refers to the signing credential, not the method of the authorisation. The CSC-API provides two variants to perform the Credential Authorization:

1) *Explicit Credential Authorization* embeds the authorisation directly into the CSC-API by providing respective endpoints.
2) *OAuth 2 Authorization* uses an external OAuth 2 Authorization Server in order to obtain an Access Token for use within the CSC-API.

The following two sections go into detail how OID4VP can be included into these authorisation methods in a minimally intrusive way with respect to the CSC-API specification. Thanks to the extensibility of OID4VP, this makes it possible to use arbitrary VCs to authorise the credential usage and thereby also the release of a signature.

3.1 Explicit Credential Authorization using OID4VP

Explicit Credential Authorization (ECA) provides a means to use a method for authorisation by explicitly providing the authentication parameters in the credential authorisation request. The anticipated cases include providing a password or OTP code, starting an external process such as an app-based validation, or using a challenge response protocol. The challenge response based method can also be used to model a flow using OID4VP.

In order to use OID4VP in the Explicit Credential Authorization, a new authentication type with a respective Authentication Object must be specified. The communication pattern defined by a so called “ChallengeResponse, out of band response” [CSC22, Sec. 8.3.1.6] is thereby suitable for integrating OID4VP. A detailed illustration of the Explicit Credential Authorization using OID4VP flow is presented in Figure 2.

In the `credentials/getChallenge` method, the Authorization Request (VP-Req) in the form of a Request URI, is returned to the signature application. Before starting the wallet with the VP-Req, the `credentials/authorize` method needs to be invoked, to conclude the challenge. As it can also seen in Figure 2, the `credentials/authorize` method defines additional parameters such as `credentialID`, which is necessary for the authentication process [CSC22, Sec. 11.6].

The VP-Req is subsequently returned to the user, who forwards it to the own wallet. The wallet then generates a VP, which is transmitted to the Cloud Signature Consortium API (CSC-API). Upon receiving the verifiable presentation, the CSC-API verifies the credential ID. If all verifications succeed, the signature can be released. The status of the authorisation can be monitored using the `credentials/authorizeCheck` method, with different status codes indicating various states [CSC22, Sec. 11.7]. For instance, a status code of 200 signifies the retrieval of the Signature Activation Data (SAD), which is necessary for signing

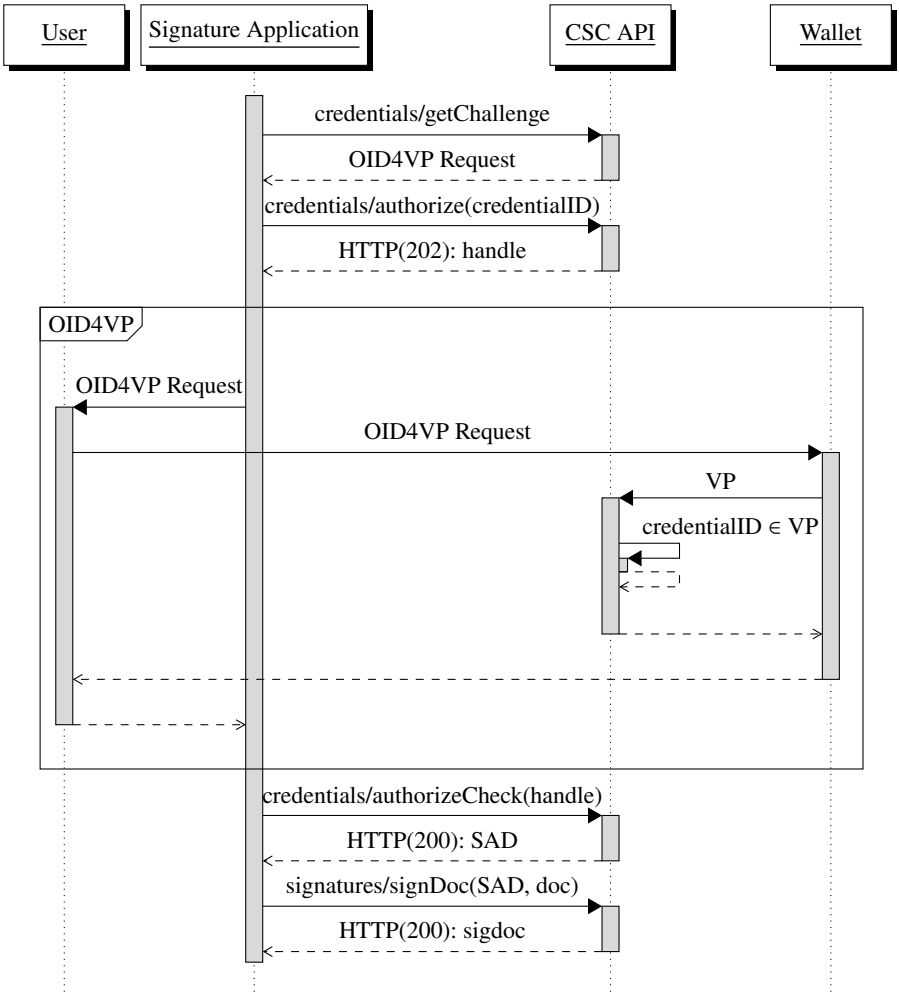


Fig. 2: Authorisation Process using Explicit Credential Authorization and OID4VP

the actual document. Meanwhile, HTTP status code 202 indicates that the authorisation process is still ongoing. Once the SAD is obtained, the signing process can commence. The CSC-API provides different endpoints for this purpose, such as `signatures/signHash` and `signatures/signDocument` [CSC22, Sec. 11.10-11.11].

3.2 Nested OAuth 2 based Credential Authorization

The CSC-API allows to use an attached OAuth 2 Authorization Server (AS) in order to release a signature [CSC22, Sec. 8.4]. The process commences with the client redirecting the user to the Authorization Server where the user performs the authentication. Thanks to the abstraction the OAuth 2 process provides, any authentication means can be used. This includes the possibility to present a VC using OID4VP. Upon successful authentication, the AS transmits an Access Token (AT) back to the client. This AT, along with the corresponding credentialID (matching the chosen VC), is then forwarded to the CSC-API. The CSC-API validates both the AT and credentialID. If all verifications are successful, the signature can be released. A detailed illustration of this nested OAuth 2 flow is presented in Figure 3.

The "nested" OAuth 2 Credential Authorization flow differs from Explicit Credential Authorization in the methods it employs. The `oauth/authorize` method is designed to handle OAuth 2 Authorization Requests using the Authorization Code Flow. Analogous to the `credentials/authorize` method, the `oauth/authorize` method expects certain parameters, including the credential ID, which are essential for the authentication process [CSC22, Sec. 8.4.2].

Upon processing, the `oauth/authorize` method returns an VP-Req, which is then passed back to the user. The user forwards the VP-Req to the own wallet, which generates a VP. This presentation is then transmitted to the AS. The AS conducts verification checks on the credential ID.

If all verifications are successful, a code is issued. This code serves as a means to obtain an AT from the AS via the `oauth/token` method [CSC22, Sec. 8.4.4]. Finally, the signature can be released by invoking either the `signatures/signHash` or `signatures/signDocument` method [CSC22, Sec. 11.10-11.11].

4 Document Proof Binding in OID4VP

Document Proof Binding describes a method trying to combine the proof of a VC and the associated authorisation process with the document that is about to be signed. This enables to establish a strong cryptographic binding between the VC and the signature process, which is a necessity when issuing Qualified Electronic Signatures. This section shows how this binding can be implemented with W3C VCs and SD-JWTs. It also shows how the document can be made accessible to the wallet when OID4VP is used.

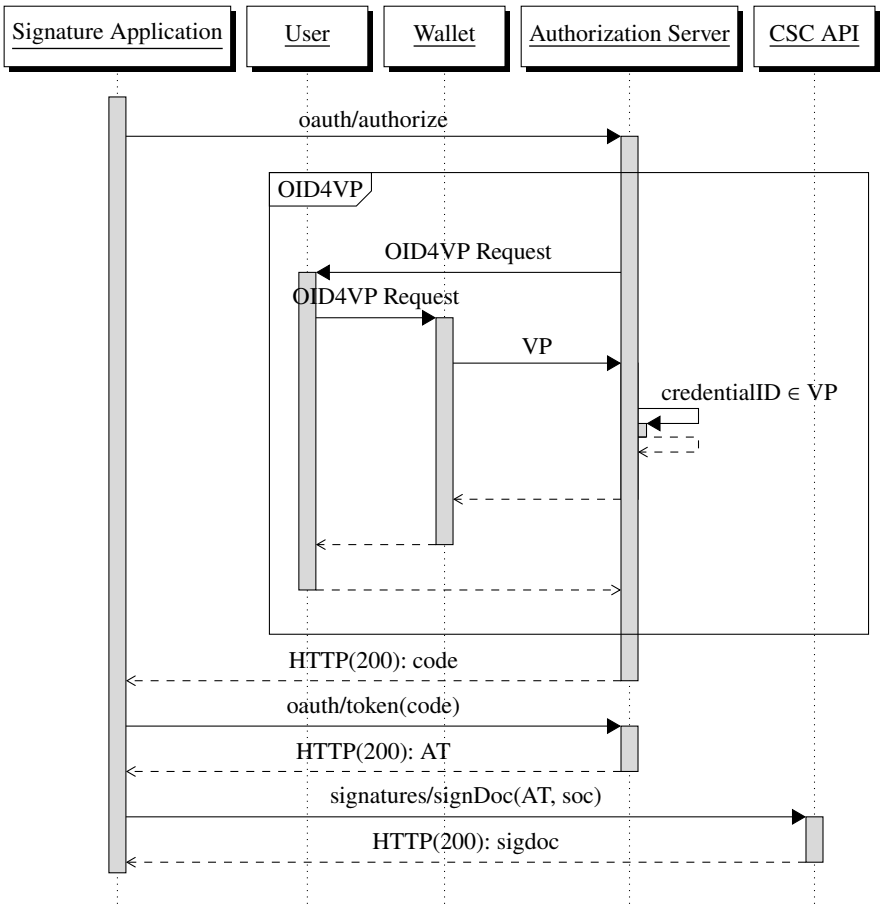


Fig. 3: Authorisation Process using OAuth 2 and OID4VP

4.1 JSON-LD Model

To prove the control of the document to be signed, the holder can create a VP containing a) the VC used for authentication and b) a VC containing the document hashes similar to a bearer credential. Note, that the holder is also the issuer of the latter, which is why one could argue that the creation of a complete VC containing a proof for the document hashes is unnecessary. This approach however ensures a correct validation of the VP, if strict schematic rules are enforced. This approach also requires providing a JSON-LD context defining the meaning of "documentHashes". Since the proof of the VP is used to validate all contained credentials the AS can then assume the control of the document by the holder, given the hashes are identical.

```
{
    "verifiablePresentation": {
        [...]
        "holder": "HOLDER_ID",
        "verifiableCredential": [
            {
               [...]
                "issuer": "ISSUER_ID",
                "credentialSubject": {[...]},
                "proof": {[...]}
            },
                {
                 "@context": [
                    "https://www.w3.org/2018/credentials/v1",
                    "https://...ContextForDocumentHashes..."
                ],
                "issuer": "HOLDER_ID",
                "credentialSubject": {
                    "documentHashes": ["OThlYTZlNG...", [...]]
                },
                "proof": {[...]}
            },
        ],
        "proof": {[...]}
    }
}
```

List. 1: Shortened example of a VP containing document hashes

4.2 SD-JWT Model

Within the SD-JWT specification [SD24a] there are two possible ways to add the document hashes to VPs. The document hashes can be part of a Key Binding JWT, which can be concatenated after the Issuer signed JWT and the disclosures as a last element. An example Key Binding JWT containing document hashes, can look like the following:

```
{
  "nonce": "4242424242",
  "aud": "https://verifier.example.org",
  "iat": 1702316015,
  "sd_hash": "nYcOXyP43v9szKryn_k_4GkRr_j3STHhNSS-i1Duauo",
  "documentHashes: ["OThlYTZlNG...", [...]]
}
```

List. 2: Shortened example of a Key Binding JWT containing document hashes

If the SD-JWT [SD24a] is enveloped in an outer JWT signed by the holder, the latter could also contain arbitrary properties, which could be used to add the document hashes, as shown in the following example:

```
{
  "aud": "https://verifier.example.org",
  "iat": 1580000000,
  "nonce": "iRnRdKuu1AtLM4ltc16by2XF0accSeutUescRw6BWC14",
  "_sd_jwt": "eyJhbGci...emhlaUJhZzBZ~eyJhb...dYALCGg~"
  "documentHashes: ["OThlYTZlNG...", [...]]
}
```

List. 3: Shortened example of a JWT containing document hashes and enveloping an SD-JWT

Since the holder signs the Key Binding JWT and/or the outer JWT, it is ensured that the hashes are known by the holder during creation of the presentation, enabling a strong association of the hashes to the holders key.

4.3 Wallet invocation with OID4VP

In order for the document proof extension, discussed in Sections 4.1 and 4.2, to be able to be created in the wallet, it is necessary to convey the document information, namely the document hashes, to the wallet. Once the hashes are available there, the wallet can include them in further checks and a suitable user consent before building the proofs and authorizing the signature.

In OID4VP, the invocation of the wallet is performed by sending the VP-Req to the wallet. Given that the invoking party has access to the hashes, it can simply add the hashes to the VP-Req. This additional data would represent an extension to OID4VP, but given the closed ecosystem of the EUDIW, this does not seem to be a major problem. While the exact way to send the VP-Req to the wallet is intentionally left unspecified in OID4VP, in most cases it will be encoded as a URI or displayed as a QR-Code. This however imposes some limitations on the size of the VP-Req. Typically that will not affect signature processes with only one hash, but might become problematic when performing batch signatures and signing multiple signatures.

Including the document hashes into the proof seems to provide a sufficiently strong method to relate the signature with the documents. This is also the anticipated method, as the hashes are present as a parameter in either the `credentials/authorize` and `oauth/authorize` calls. For the user this is however a problem, as a hash value does not allow the user in the context of the wallet to check that the hash belongs to the document, or let the user inspect the document itself which is about to be signed. Given the limitations described beforehand, sending the document via the VP-Req is not a viable option. This problem can be solved by providing an indirection to the document, by including a link to the document into the VP-Req. The wallet can retrieve the document and include its hash in the proof. As the document is however not in the CSC-API authorize calls, this would require an update of the specification.

5 Conclusion and Outlook

The EUDIW promises to become the core element of a multitude of digitalisation efforts in the European Union. It not only aims at reducing the gap between the different eID Systems in Europe, but also enables the use of Qualified Electronic Signatures for all citizens with access to a wallet containing respective VCs.

Taking the currently available and emerging major standards concerning VCs into account, the present paper presented different possible approaches for the secure integration of OpenID4VP [IE23] with the CSC-API [CSC22]. The proposal can be implemented with only small changes in the extension points of the CSC-API and OID4VP. One exception is the inclusion of the document itself in the validation process of the request in the wallet, which will require further discussions with experts within pertinent projects and standardisation bodies, if this stronger binding between proof and document is desirable.

We thankfully acknowledge, that the present paper benefits from fruitful discussion with members of the eIDAS expert group and experts in pertinent standardisation bodies, such as ETSI ESI and CSC for example. We plan to continue this fruitful discussion and contribute to the update of ETSI TS 119 432, while extending the present work to somewhat more complex use cases, which also require the attestation of additional professional qualifications and mandates according to Annex VI of (EU) 2024/1183 [EU24].

References

[An23] AnonCreds Specification, tech. rep. Draft, Hyperledger, 2023, URL: https://hyperledger.github.io/anoncreds-spec/.

[ARF23] European Commission: The European Digital Identity Wallet Architecture and Reference Framework, Version 1.1.0, 20.04.2023, URL: https://github.com/eu-digital-identity-wallet/eudi-doc-architecture-and-reference-framework/blob/main/docs/arf.md.

[CSC22] Architectures and protocols for remote signature applicaons, tech. rep. v2.0.0.2, Cloud Signature Consortium, 2022, URL: https://cloudsignatureconsortium.org/wp-content/uploads/2023/04/csc-api-v2.0.0.2.pdf.

[DID23] DIDComm Messaging v2.1, tech. rep. Draft 2.1, Identity Foundation (DIF), 2023, URL: https://identity.foundation/didcomm-messaging/spec/v2.1/.

[ET20] ETSI TS 119 432: Electronic Signatures and Infrastructures (ESI); Protocols for remote digital signature creation, tech. rep. V1.2.1, ETSI, 2020, URL: https://www.etsi.org/deliver/etsi_ts/119400_119499/119432/01.02.01_60/ts_119432v010201p.pdf.

[EU24] Regulation (EU) 2024/1183 of the European Parliament and of the Council of 11 April 2024 amending Regulation (EU) No 910/2014 as regards establishing the European Digital Identity Framework, 2024, URL: http://data.europa.eu/eli/reg/2024/1183/oj.

[IE23] OpenID for Verifiable Presentations, tech. rep. Draft 20, IETF, 2023, URL: https://openid.net/specs/openid-4-verifiable-presentations-1_0.html.

[PO24] D3.1-UC5-draft: QES Specification, tech. rep., Unpublished draft, a quality-controlled version will be published as POTENTIAL D3.1, POTENTIAL consortium, 2024.

[SD24a] Selective Disclosure for JWTs (SD-JWT), tech. rep. Draft 08, IETF, 2024, URL: https://www.ietf.org/archive/id/draft-ietf-oauth-selective-disclosure-jwt-08.html.

[SD24b] SD-JWT-based Verifiable Credentials (SD-JWT VC), tech. rep. Draft 03, IETF, 2024, URL: https://www.ietf.org/archive/id/draft-ietf-oauth-sd-jwt-vc-03.html.

[W3C20] JSON-LD v1.1, tech. rep. Recommendation, W3C, 2020, URL: https://www.w3.org/TR/json-ld/.

[W3C22a] Decentralized Identifiers (DIDs) v1.0, tech. rep. Recommendation, W3C, 2022, URL: https://www.w3.org/TR/did-core/.

[W3C22b] Verifiable Credentials Data Model v1.1, tech. rep. Recommendation, W3C, 2022, URL: https://www.w3.org/TR/vc-data-model/.

[W3C24a] Verifiable Credentials Data Model v2.0, tech. rep. Draft, W3C, 2024, URL: https://www.w3.org/TR/2024/CRD-vc-data-model-2.0-20240218/.

[W3C24b] Data Integrity ECDSA Cryptosuites v1.0, tech. rep. Recommendation Draft, W3C, 2024, URL: https://www.w3.org/TR/2024/CRD-vc-di-ecdsa-20240209/.

[W3C24c] Data Integrity BBS Cryptosuites v1.0, Achieving Unlinkable Data Integrity with Pairing-based Cryptography, tech. rep. W3C Candidate Recommendation Draft 28 April 2024, W3C, 2024, URL: https://www.w3.org/TR/vc-di-bbs/.

Further Conference Contributions

Secure Industrial Device Wallet

Ankita Kumari [1], Anita Aghaie [1] [2], Anne Passarelli [1], Niranjana Papagudi Subrahmanyam [1], Aliza Maftun [1]

Abstract: Industry 4.0 integrates technologies such as blockchain, self-sovereign identities, digital twins etc. within industrial environments. A key feature within such industrial advancements is the use of wallets that can facilitate the secure digitalization of industrial operations and the expansion of systems. This paper puts forth the idea of a secure industrial device wallet, that is robustly bound to the device and serves as a hardware trust anchor within self-sovereign identity architectures. The paper introduces and evaluates different approaches to achieve this binding and provides a proof of concept to prevent device counterfeit attacks.

Keywords: Industrial wallet, Generic Trust Anchor API, cryptographic agility, industrial ecosystem, anti-counterfeiting, self-sovereign identity

1 Introduction

Industrial devices require different credentials for different use cases like identification, authentication, authorization, etc. Currently, PKI X.509 certificates are commonly used for these use cases. Other scenarios like secure device onboarding, bootstrapping, etc. also require device identity management. Self-Sovereign Identity (SSI) [Pa21] is a concept where every entity manages its identity information by itself and decides when to share this information and with whom. An SSI system typically consists of an issuer, a holder, and a verifier. The issuer issues a verifiable credential (VC) [Pa21] for the holder by attesting claims made by the holder and signing it with its private key. The holder presents the credential proof to the verifier who verifies its authenticity by verifying the signature using issuer's corresponding public key obtained from a trusted registry. A holder can be a natural person, an organization (legal entity) or a thing (IoT device).

In industrial scenarios, VCs can be leveraged to depict device-specific claims and properties, such as enrollment tokens, process information, regulatory conformance certificates, etc. The VCs are managed by digital containers called wallets. Wallets can receive and store VCs, present verifiable proofs to the requesting party and manage entity relationships. Wallets are designed with protocols for secure communication such as DIDComm [Di20] for both issuance and presentation of VCs.

[1] Equal contribution, Siemens AG, Germany

[2] https://orcid.org/0000-0003-2470-3408

doi:10.18420/OID2024_14

Currently known SSI-based wallets like the Lissi [Li23] and the Esatus [Es23], manage identity and credentials of a natural person and legal entities. Furthermore, the European Digital Identity group is focused on the implementation of an EU-wide European Digital Identity (EUDI) wallet solution [Eu13] based on VCs. These wallets contain VCs that attest the attributes of only the natural person alone and not the device where the wallet is installed. This implies that a natural person can have the same credentials across multiple devices such as laptop, smartphones, cloud, etc. In contrast, VCs held in industrial wallets attest the attributes of the device itself such as those for proving device properties, device authentication, device reconfiguration etc. The credentials contained in the wallet should attest the authentic device and are not valid when exported to other devices.

This paper addresses this serious threat to the industry arising through device counterfeiting by proposing a solution to keep the industrial digital wallet, that holds the device credentials, tightly bound to the original device hardware. The paper first evaluates multiple approaches to achieve this device binding and further presents a proof of concept of one of the approaches, showcasing how device counterfeiting is prevented by binding the wallet to its hardware.

2 Methodology

A wallet can be realized on a device as an encrypted database with APIs for querying and writing contents into it. The wallet database is created and encrypted using a secret called wallet master key (WMK). The WMK prevents unauthorized users from accessing the wallet contents, even if they manage to export the encrypted database. This approach is adopted in SSI wallets like [Li23] and [Hy18]. The term wallet in the rest of this paper indicates only this database and not the API for data handling. The term wallet secrets denote device-specific confidential data, e.g. private keys. The wallet also stores other device data that is not necessarily confidential and therefore not considered in this paper. The term device-specific key denotes a secret key that is stored in a secure storage, e.g. secure element (SE) on the device. In a different variant, a device-specific key could also be a secret that is generated using some intrinsic hardware and/or software properties of the device, like a device hardware fingerprint, physical unclonable functions, encrypted serial numbers etc. This section explains three approaches for binding the wallet to device and evaluates their advantages and disadvantages.

Approach 1: Bind access to the wallet to the device by wrapping WMK with a device-specific key.

In this approach, the device binding of a wallet is achieved by binding its access (the WMK) to a device-specific key (Figure 1). The WMK is wrapped with the device-specific key, so that only the device in possession of the correct device-specific key can unwrap the WMK and access the wallet contents, including the wallet secrets.

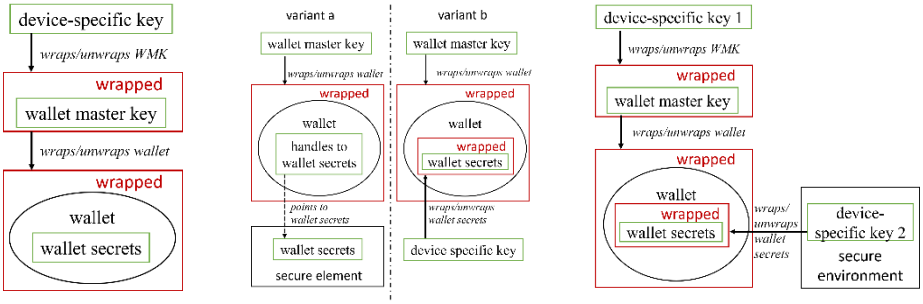
Approach 2: Bind wallet secrets to the device.

This approach can be realized in two different variants. In the first (variant a), the wallet secrets are stored in a hardware trust anchor of the device (e.g. SE). Storing wallet secrets in this secure hardware ensures that these secrets are tied to the specific device and cannot be easily extracted. The wallet stores only key handles pointing to the wallet secrets and all cryptographic operations with the wallet secrets take place inside the SE. This approach requires that the SE supports these cryptographic operations. In a second variant (variant b), instead of storing the wallet secrets directly in the expensive secure memory of the SE, they are wrapped with a device-specific key using the SE and stored as wrapped key blobs in the wallet which is placed in the file system of the device. If a cryptographic operation with a wallet secret should be performed, the wrapped wallet secret is passed through the wallet API to the SE, where it is unwrapped and then used for the requested operation. (Figure 1)

Approach 3: Bind both WMK and wallet secrets to the device with device-specific keys.

This approach is a combination of the previously explained approaches and provides a higher security level by increasing the dependency on hardware binding with multiple device-specific keys. Such device-specific keys can be in a SE or other trusted components of a trusted code base.

All the proposed approaches enhance security of the wallet by offering better protection for the WMK and wallet secrets. Both WMK and wallet secrets, when stored in file system are prone to attacks including unauthorized access or memory scraping. Secure environments offer isolation, anti-tamper mechanisms and protection, against side-channel attacks, unauthorized access, and prevent misuse of sensitive information. By safeguarding wallet secrets, either by confining operations within the SE or by securely wrapping them and storing them on the wallet, it is ensured that these secrets are never exposed in plain text. Moreover, by allowing access to wallet through a combination of user and device authentication, these approaches strengthen security measures against unauthorized access.



Approach 1 Approach 2 Approach 3

Figure 1: Different approaches to achieve device binding of the wallet.

Approach 2 (variant b) and approach 3 introduce flexibility and scalability by enabling the storage of wallet secrets outside the costly and limited SE memory and by supporting binding using different device specific aspects. This diversification in security mechanisms allows for a higher security level, mitigating risks if one binding measure is compromised. Approach 1 is a good solution to secure data at-rest, given that the attack potential for memory eavesdropping and interception of the WMK in plain text during communication with the SE, is low.

A significant constraint across these approaches is the dependency of devices on SEs, which may not be universally available. Secondly, the secure memory within a SE is also notably scarce and managing wallet secrets within this limited space poses a challenge. Lastly, the requirement for SEs to support cryptographic operations for handling wallet secrets can complicate their implementation. It is therefore essential to compare the costs of implementing such solutions against the benefit of avoiding costs incurred from data breaches from not having such secure solutions. The different approaches proposed above are summarized in the form of a table below.

Table 1: Summary of proposed approaches

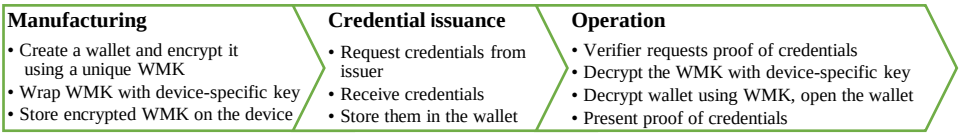
	Methodology	Targeted Attacks [Security level]	Implementation Overhead
1	Bind the wallet access (WMK) to a device-specific key	Anti- Counterfeiting [Lower]	Lower
2	Bind the wallet secrets to the device with a device-specific key	Anti-Counterfeit, Man in Middle, RAM attacks [Medium]	Medium
3	Bind both WMK and wallet secrets with different device-specific keys	Stronger Anti-Counterfeit, Man in Middle, RAM attacks [Higher]	Higher

3 Proof of concept

As discussed in the previous section, the evaluated approaches for binding a wallet on an industrial device can prevent replication of the wallet secrets on illegitimate devices. Device counterfeit attacks are prevalent in the industrial domain and cause, amongst others, financial and reputational losses. The proof of concept implements binding the access to the wallet (i.e. WMK) to an industrial device as proposed in approach 1 and elaborates how this binding is used to prevent device counterfeiting attacks.

3.1 Implementation scenario

Industrial device life cycle consists of three different phases:



For the scope of this paper, both manufacturing and credential issuance phases are considered as a pre-step and the device is assumed to be in operational phase with the issued credentials in the wallet. The proof of concept uses an industrial device with a TPM as the trust anchor and the open-source Indy SDK [Hy18] for the implementation of the wallet. To achieve the device binding, an agent application is implemented with an interface to an abstraction layer [Ge24] that is used to bind the WMK to the SE (TPM) and handles other wallet functionalities. This abstraction layer provides independence as to which SE is being used by the device, thus facilitating crypto agility. The abstraction layer is a technological standard called Generic Trust Anchor API (GTA API) [Ge24].

Figure 2 shows the software stack of the implemented approach that binds the WMK to the industrial device. During the manufacturing step, the WMK is encrypted using the device-specific key in the TPM (Figure 2) and the encrypted WMK is stored on the device. Every time the wallet needs to be accessed, the encrypted WMK is decrypted using the device-specific key in the TPM and the retrieved WMK is used to access the wallet.

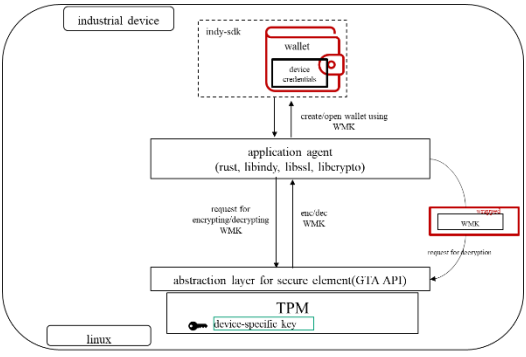


Figure 2: Workflow of enc/dec the WMK using the SE to access device credentials in the wallet.

3.2 Preventing the device counterfeiting attack

In this attack scenario, the attacker counterfeits the original device and copies all the data including the wallet containing the device credentials and the encrypted wallet master key. During the operation phase, the attacker tries to present the device credentials to the verifier. In the proof of concept, the WMK is bound to the SE present on the original device. To present the device credentials, the attacker tries to decrypt the WMK to access the wallet. As this is a different device, the device-specific key present in the SE of this device is different, leading to failed decryption. The attacker cannot open the wallet and thus cannot access and present the device credentials stored in the wallet (Figure 3). This proof of concept shows that by binding the access of the wallet to the device, a device counterfeiting attack can be prevented.

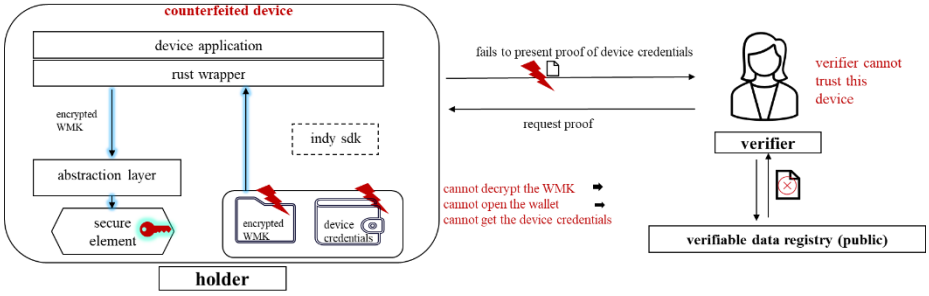


Figure 3: Counterfeited device trying to access the original device's copied credentials stored in the wallet but fails due to device binding of the WMK.

Conclusion

Device wallets find their application as a secure container storing device-related credentials and cryptographic keys for authenticating the device in different industrial scenarios like device onboarding, secure updates etc. Here misuse of the wallet and its credentials poses a serious threat and ensuring security of the industrial wallet is imperative and of paramount importance. The presented approaches in this paper propose a strong hardware binding to solve this threat. Further, the discussed proof of concept introduced a generic abstraction layer for performing cryptographic operations to enable crypto agility. Implementation of binding the wallet master key and the wallet secrets to multiple device-specific keys as mentioned in approach 3 is planned in future.

Bibliography

[Pa21] Preukschat, A. (2021). Self-Sovereign Identity. New York: Manning Publications.

[Di20] DIDComm Messaging v2.x Editor's Draft. (2020, January 20). Decentralized Identity Foundation: https://identity.foundation/didcomm-messaging/spec/.

[Li23] Lissi Wallets. (2023). Lissi: https://www.lissi.id/.

[Es23] Esatus wallet (2023): https://esatus.com/en/digital-identity/.

[Eu13] European Commission. (2013, January 3). The Common Union Toolbox for a coordinated Approach Towards a European Digital Identity Framework.

[Hy18] Hyperledger. (2018, December 18). Indy SDK. Hyperledger Indy: https://hyperledger-indy.readthedocs.io/projects/sdk/en/latest/docs/index.html.

[Ge24] ISO/IEC CD TS 30168: Generic trust anchor application programming interface for industrial IoT devices." (2024). https://www.iso.org/standard/53288.html.

MINERVA: Secure Collaborative Machine Tool Data Utilization Leveraging Confidentiality-Protecting Technologies

Andy Ludwig[1,2], Michael P. Heinl[1,2], Alexander Giehl[1,2]

Abstract: The digitization of shop floors opens up opportunities for innovative applications and business models due to the vast amount of generated data. However, a lot of this potential is currently not utilized because companies consider the risk of data sharing as too high compared to the corresponding benefit. Focusing on the machine tool sector, the research project MINERVA addresses these concerns by experimentally repurposing privacy-enhancing technologies as confidentiality-protecting technologies and applying them to the use case of condition monitoring to protect intellectual property and other information deemed critical by machine tool operators. Thereby, MINERVA's goal is to reduce the risk of data sharing and support the establishment of data-driven business models in the machine tool sector in the long term.

Keywords: Machine Tool Data; Confidentiality-Protecting Technologies; Privacy-Enhancing Technologies; Industrial Internet of Things; Defense-in-Depth; Supply Chain Security

1 Introduction

Industry 4.0, the digitization and interconnection of shop floors, leads to a steadily increasing amount of production-relevant data. This data has a large potential which the machine tool industry leverages only in parts [Ve21]. Applications such as condition monitoring or predictive maintenance are currently used mainly for single or few machines on the local shopfloor. In order to fully harness the data's potential and to build a broad basis for data-driven business models, the data has to be aggregated and analyzed across company borders. However, this data-driven innovation is hindered by the concerns of many machine tool operators who worry to suffer a competitive disadvantage due to data sharing.

2 Objectives

MINERVA addresses this conflict by creating a secure and transparent edge and cloud architecture for the Industrial Internet of Things (IIoT). In contrast to existing architectures,

[1] Fraunhofer AISEC, Department Product Protection and Industrial Security, Garching bei München, Germany, andy.ludwig@aisec.fraunhofer.de, michael.heinl@aisec.fraunhofer.de, alexander.giehl@aisec.fraunhofer.de
[2] Technical University of Munich, TUM School of Computation, Information and Technology, Department of Computer Engineering, Chair of IT Security, Germany

doi:10.18420/OID2024_15

MINERVA technologically underpins the machine tool operators' data sovereignty by applying Privacy-Enhancing Technologies (PETs) as an additional layer of defense. This corresponds to the widespread paradigm of defense-in-depth promoted by the established ISA/IEC 62443 series of standards [IE09] covering industrial security. Condition monitoring, which is highly relevant to the industry partners, serves as a tangible use case. For this, suitable Machine Learning (ML) / Federated Learning (FL) algorithms are used to train models in the cloud using "anonymized" data protected by PETs. PET parameters will be selected in accordance with the use case, also considering potential for automation, such as in the fine-tuning of Differential Privacy (DP). Subsequently, these models are transferred back to the shop floor, where they are used to assess the machines' condition. In the medium term, the industry partners can use collaboratively trained ML models without having to disclose sensitive data. The long term goal is to strengthen the trust of the entire machine tool industry in data-driven business models. Since MINERVA does not primarily use PETs to protect privacy due to the lack of Personally Identifiable Information (PII) but rather to protect confidentiality of industrial data, the term Confidentiality-Protecting Technology (CPT) is proposed and from now on used instead of PET in order to semantically match its purpose. For the same reason, the term *additionally protected* will be used instead of *anonymized*. Currently, MINERVA is in an early phase of development and the analysis of potential CPTs beyond the few explicitly mentioned during the course of this paper is part of the ongoing work.

3 Related Work

In security research, data privacy is an important area due to the dilemma between data-driven functions and the risks associated with data disclosure [CGL20]. CPTs are possible countermeasures that have been applied in different use cases [DB20; Th23; TM21]. Ensuring privacy in the context of big data is currently one of the main challenges [Cu21]. There are few published research items in the specific field of data sovereignty in Industry 4.0. For example, the industrial data space concept describes an implemented platform but does not incorporate common CPTs [Ot16]. There are also platforms for collaborative data usage, but not in the area of Industry 4.0 [Kh21]. However, future research directions have been suggested in this context [Ca19]. The importance of anonymization models in Industry 4.0 research is increasing [Fu10; Ji21]. DP has gained attention in this regard [HRC19; Hu21] and there are different developments due to the necessary adaptation for each use case [Fu10]. Advantages of DP include the ability to set an explicit privacy level due to its mathematical definition as well as its real-time capability [DR14; Dw08] which significantly increases the acceptance of DP [GHB21; Gi19]. Failure to apply data privacy models to machine data may lead to the reconstruction of product geometry, resulting in a loss of Intellectual Property (IP) [Gi19].

Summarized, there are basic approaches in the area of collaborative data utilization in Industry 4.0. However, there are still various unsolved issues that prevent companies from sharing their data. This is where MINERVA starts.

4 Methodology

The concept of trustworthy collaborative IIoT-as-a-service can facilitate data-driven business models. This can be achieved by implementing CPTs in a secure system architecture to prevent the loss of IP. The following sections outline the concept of trustworthy collaborative usage of machine tool data. The concept will be applied to the use case of condition monitoring for machine tools, in particular milling machines. Figure 1 schematically illustrates the planned workflow which is also described in the following sections.

4.1 Data Collection from Machine Tool

First, the local data of machine tools is collected and processed unprotected in the corresponding data lake of the edge device. For instance, this data contains the coordinates of the tool or specific forces over time. This can be enriched using sound sensors attached to the machine. Structure-borne noise data is an additional input for condition monitoring algorithms. The larger quantity of data allows a more detailed analysis of processes within the machine. The local data needs to be labelled for later usage of ML algorithms like condition monitoring. For this purpose, a human-machine interface at the edge device is necessary, allowing employees to assess the machine conditions. In general, the labels do not need to be determined throughout the entire life cycle or operational use but only during the training process at the beginning. Automated support for the labelling decision may be possible but will be conducted later, which further simplifies the collection of machine states. In addition to the machine condition, the corresponding data criticality must be determined. The criticality level is essential for applying CPTs with suitable parameters to the data in order to reach an appropriate level of protection.

4.2 Application of CPTs

Depending on the associated data criticality level, a suitable CPT is applied to the data in the edge device of the machine tool within the company boundaries. This additional protection strengthens the owner's data sovereignty and reduces the risk of losing IP. The impact of the CPT must be reliable and comprehensible. In cases where a company utilizes multiple machine tools, a local server can collect all the data and apply the CPT in an aggregated manner. The selection of appropriate CPTs depends on various factors, including the level of data criticality, basic suitability of data, its layout, and finally, the compatibility with subsequently applied ML algorithms. Furthermore, a synchronisation layer must be established between all participants. If each company can set data criticality independently, data could be preprocessed in different ways. Either the cloud service can handle these differences in data quality and format or a policy has to restrict the pool of possible CPTs.

4.3 Cloud Communication and Collaboration

After the preprocessing, the data can be transmitted to a central entity, where all data from all participating machine tool operators is aggregated. This comprehensive data set can be used to train ML models for specific use cases. A server collecting all data represents an attractive target for attackers. Although the usage of CPTs to protect data is an additional layer of defense, the infrastructure must also be hardened against attacks. State-of-the-art security mechanisms are necessary to protect not only the integrity and confidentiality of the data itself but also the services ingesting and processing the data. These objectives can be in parts accomplished through the utilization of other types of CPTs, such as Trusted Execution Environments (TEEs) or group signatures.

For instance, the secure infrastructure must include an authentication process, allowing only known and trustworthy partners to participate and contribute data for training purposes because data poisoning attacks are a potential risk for the system, leading to erroneous model results. However, what seems to be a straightforward task to be solved by established technology such as Public Key Infrastructure (PKI) can involve a couple of challenges. First of all, there has to be a PKI which is trusted by all participants. Since the operator of such a PKI represents a single point of failure, utmost care during its selection [He19] and compliance with special requirements [He23] are crucial. In order to establish an additional layer of defense against compromises involving information extraction attacks, IP should be separated from the participants' identities or the latter be protected by the usage of anonymous credentials. At the same time, this measure increases the risk of poisoning attacks conducted by untrustworthy or unnoticedly compromised participants. Therefore, a ring signature-like mechanism might be employed to establish a secure channel with a certain degree of anonymity which can, however, be revoked in case of misuse.

4.4 Leveraging the Model in Companies

Using the same secure channel, the model which has been collaboratively trained in the cloud can be transmitted back to each participating company. They can use the model without accessing the input data which enables each participant to benefit from sharing their data while at the same time protecting it from competitors or adversaries. Local data stored within company boundaries can be classified using the model without establishing a connection to external partners. Depending on the use case, a repeated transfer of new training data as well as the generated model is possible. Additionally, feedback loops must be implemented to increase model performance.

5 Conclusion

A major challenge of sharing data across company borders is the loss of IP. A possible threat is the competitors' capability to reconstruct specifics about a work piece. As a result,

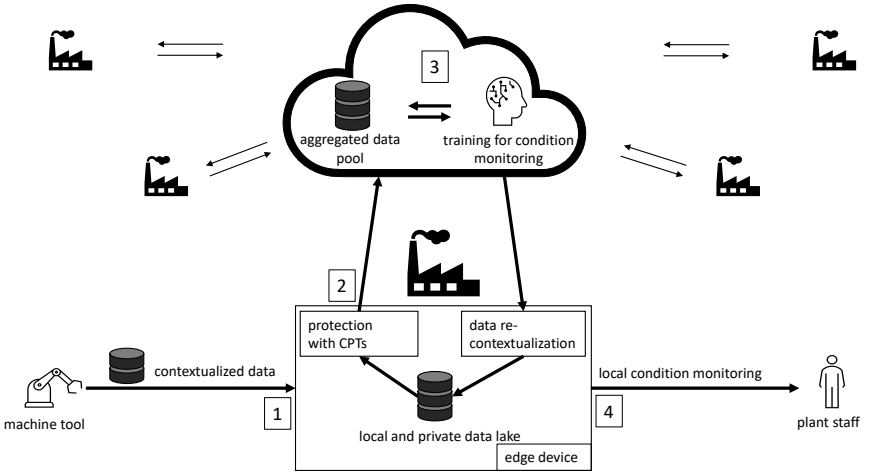


Fig. 1: Planned workflow: (1) Data Collection from Machine Tool; (2) Application of CPTs; (3) Cloud Communications and Training of Condition Monitoring Algorithm; (4) Using the Model in Companies.

companies are torn between optimizing their production flow and the fear of disclosing internal knowledge. Additionally, aggregated data at a central point, such as a cloud server, represents an attractive target for adversaries because one successful attack can result in the compromise of data from a plethora of companies (so-called *supply chain attack*). Generally, two main attacker models can be distinguished: an attack on intellectual property, which could also be conducted by an honest but curious participant. Additionally, the data can be compromised in its integrity, leading to a distorted model.

CPTs may provide a comprehensive solution. If each company maintains data sovereignty, the risk of IP loss by sharing data can be reduced. CPTs can be directly applied to data within the company boundaries. Afterwards, only additionally protected data is transferred to the cloud. This approach will be combined with a secure, transparent, and reliable data infrastructure between each company and the cloud. Possible application areas include improved energy consumption or reduced machine tool down times through the use of algorithms for condition monitoring or predictive maintenance. Detecting and eliminating product weaknesses through the implementation of these algorithms can lead to optimization for component suppliers or machine manufacturers.

The overall goal of reducing the machine tool operators' concerns regarding data sharing are evaluated by a survey taking place at the beginning and once again at the end of the project duration after presenting the project results to the survey participants. Finally, potentials to transfer the results to other areas will be identified to generally increase the willingness to collaboratively use data for the good of all.

Acknowledgement

This work was supported by the German Federal Ministry of Education and Research (BMBF) under grant number 16KIS1803K.

References

[Ca19] Capiello, C. et al.: Data Ecosystems: Sovereign Data Exchange among Organizations. Dagstuhl Seminar 19391, 2019.

[CGL20] Christen, M. et al.: The Ethics of Cybersecurity. 2020.

[Cu21] Curry, E. et al., eds.: Elements of Big Data Value. 2021.

[DB20] Domingo-Ferrer, J. et al.: Privacy-Preserving Technologies. In: The Ethics of Cybersecurity. Vol. 21, The International Library of Ethics, Law and Technology, 2020.

[DR14] Dwork, C. et al.: The Algorithmic Foundations of Differential Privacy. Foundations and Trends in Theoretical Computer Science 9 (3-4), 2014.

[Dw08] Dwork, C.: Differential Privacy: A Survey of Results. In: Theory and applications of models of computation. Lecture Notes in Computer Science, 2008.

[Fu10] Fung, B. C. M. et al.: Privacy-preserving data publishing. ACM CSUR 42 (4), 2010.

[GHB21] Giehl, A. et al.: Leveraging Edge Computing and Differential Privacy to Securely Enable Industrial Cloud Collaboration Along the Value Chain. In: IEEE CASE. 2021.

[Gi19] Giehl, A. et al.: Edge-computing enhanced privacy protection for industrial ecosystems in the context of SMEs. In: 12th CMI Conference on Cybersecurity and Privacy. 2019.

[He19] Heinl, M. P. et al.: MERCAT: A Metric for the Evaluation and Reconsideration of Certificate Authority Trustworthiness. In: 2019 ACM SIGSAC CCSW. 2019.

[He23] Heinl, M. P. et al.: From Standard to Practice: Towards ISA/IEC 62443-Conform Public Key Infrastructures. In: Computer Safety, Reliability, and Security. 2023.

[HRC19] Hassan, M. U. et al.: Differential Privacy Techniques for Cyber Physical Systems: A Survey. IEEE Commun. Surv. Tutor. 22 (1), 2019.

[Hu21] Husnoo, M. A. et al.: Differential Privacy for IoT-Enabled Critical Infrastructure: A Comprehensive Survey. IEEE Access 9, 2021.

[IE09] IEC: IEC/TS 62443-1-1:2009, Industrial communication networks – Network and system security – Part 1-1: Terminology, concepts and models. 2009.

[Ji21] Jiang, B. et al.: Differential Privacy for Industrial Internet of Things: Opportunities, Applications, and Challenges. IEEE Internet of Things Journal 8 (13), 2021.

[Kh21] Khokhar, R. H. et al.: Enabling Secure Trustworthiness Assessment and Privacy Protection in Integrating Data for Trading Person-Specific Information. IEEE Transactions on Engineering Management 68 (1), 2021.

[Ot16] Otto, B. et al.: Industrial Data Space: Digital Souvereignity Over Data. 2016.

[Th23] The Royal Society: From privacy to partnership: The role of privacy enhancing technologies in data governance and collaborative analysis. 2023.

[TM21] Timan, T. et al.: Data Protection in the Era of Artificial Intelligence: Trends, Existing Solutions and Recommendations for Privacy-Preserving Technologies. In: Elements of Big Data Value. 2021.

[Ve21] Verein Deutscher Werkzeugmaschinenfabriken e.V., Fachverband Werkzeugmaschinen und Fertigungssysteme im VDMA: Jahresbericht 2021. 2021.

Fulfilling Principles of Self-Sovereign Identity: Towards a Conformity Assessment Approach for Human Wallets

Dustin Doege[1], Ricardo Bochnia[2], and Jürgen Anke[2]

Abstract: Self-Sovereign Identity (SSI) represents a paradigm shift toward user-centric digital identity management by emphasizing principles such as user control and privacy. However, there is a notable gap in assessing how these principles are implemented within existing SSI products despite the ongoing research interest in the theoretical principles of SSI. Our research introduces a structured conformity assessment approach to bridge the gap between theoretical ideals and practical implementation. This approach enables the assessment of SSI products based on fulfilling requirements derived from SSI principles. This provides developers and policymakers with a tool to assess the adherence of SSI products to the fundamental principles. Thus, it may serve developers as a design guideline and policymakers as a basis for certification processes.

Keywords: Self-Sovereign Identity, Principles, Requirements, Conformity Assessment, Test, Wallet

1 Introduction

Self-Sovereign Identity (SSI) is not just a set of technologies but has been guided by overarching principles since its inception. Principles are fundamental, abstract guidelines aiming to capture the core objectives of SSI. When Allen coined the term Self-Sovereign Identity, he also proposed ten principles [Al16]. As SSI gained popularity, the initial set of principles was refined and expanded [Ču22; FCA19; Se22; So23; TA19].

Two notable recent works were done by Čučko et al. [Ču22] and Sedlmeir et al. [Se22]. Both employ a rigorous, iterative method using literature reviews and expert interviews. While [Ču22] derived a broad set of SSI principles with ranking and categorization, [Se22] focused on a core set with associated characteristics. A comparison can be found in Tab. 1. Furthermore, as [Ču22] showed in their literature review, a set of principles is common among the above-mentioned proposals, such as putting the user (holder) in control of their data and allowing them to maintain their privacy by using minimal disclosure.

Despite the research interest in the theoretical principles of SSI, it is not well-understood whether SSI products also adhere to these principles. This gap highlights a critical research need: An assessment approach for testing the conformity of SSI products to these principles.

To address this gap, we seek to answer the following research question: *How can SSI products be assessed regarding their fulfillment of SSI principles?* We propose an approach to assess

1 Deutsche Telekom MMS GmbH, , Riesaer Straße 5, 01129 Dresden, Germany, dustin.doege@telekom.de
2 HTW Dresden, Digital Service Systems Group, Friedrich-List-Platz 1, 01187 Dresden, Germany, ricardo.bochnia@htw-dresden.de, https://orcid.org/0009-0007-4317-1810; juergen.anke@htw-dresden.de, https://orcid.org/0000-0002-9324-9387

doi:10.18420/OID2024_16

Tab. 1: SSI Principles of Allen [Al16], Čučko et al. [Ču22], and Sedlmeir et al. [Se22]

Category[1]	Allen	Čučko et al.	Sedlmeir et al.[2]
Controllability	Control	*Ownership and Control*[3]	Control
	Existence	Existence and Representation	Representation
	-[4]	*Decentralization and Autonomy*	Reliability
Privacy	Consent	*Consent*	Control
	Minimalization	*Privacy and Minimal Disclosure*	Privacy
	-	Single source	Control
Security	Protection	*Security and Protection*	Security
	-	*Verifiability and Authenticity*	Verifiability
	-	*Verifiability and Authenticity*	Authenticity
Usability	Access	Accessibility and Availability	Control
	-	Usability and User Experience	Usability
	-	Recoverability	Reliability
Adoption	Transparency	Transparency	Flexibility
	Persistence	Persistence	Representation
	Portability	Portability	Flexibility
	Interoperability	*Interoperability*	Flexibility
	-	*Standard*	Flexibility
	-	Compatibility with legacy systems	-
	-	Cost	-

[1] Categories are based on [Ču22] [2] Each principle of [Se22] contains several characteristics that match to [Ču22] principles in most cases [3] Italized principles were ranked as most important [4] [Al16] views decentralization more as a prerequisite for his principles

the fulfillment of these principles in SSI products by deriving software requirements from SSI principles and matching them with features from existing products. This approach can serve as a design guideline for developers and as a foundation for certification processes for policymakers. For example, it could be used to certify SSI solutions that offer advanced privacy features. The goal is to ensure that SSI principles are not just theoretical ideals, but that products labeled as SSI also adhere to these principles.

Although bridging the gap between SSI principles and their practical implementation is crucial, it is also a sensitive topic for the reputation of wallet vendors. Conducting this assessment requires a robust and time-consuming effort. To avoid premature conclusions, we refrained from publishing preliminary results. However, our preliminary assessment already revealed discrepancies between the principles and existing products, and uncovering the gap between principles and existing products can lead to fruitful insights.

2 Assessing the Fulfillment of SSI Principles in Products

The development process of our approach is outlined in Fig. 1, which shows the steps for developing (green) and applying (gray) the approach. We focus on describing the methodology behind the development of the approach, covering the initial five steps. Due to space limitations, our explanations of the choices made in the development steps are brief. The remaining two steps related to the application of our approach are left for future work.

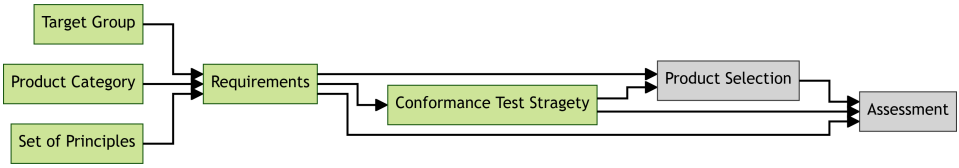


Fig. 1: The flowchart outlines the development process: Green steps depict completed steps in our approach development, while gray steps indicate future work to apply the approach.

Target Group and Product Category: Our focus is on individuals and their wallets because of the prevalence of these products. Given the variety of wallets available, this category provides a rich field for examining how SSI principles are applied and where improvements may be necessary. Wallets are the primary interface for end users and are the primary point of contact between users and principles such as control, privacy, and security.

Set of Principles: Due to the limited space, we only present a condensed set of principles and derived requirements (explained in the next section) based on [Ču22] and [Se22], already introduced in Tab. 1. We used the following criteria: Considered most important by experts [Ču22] but excluding decentralization as it can be considered a prerequisite of other principles [Se22]. Furthermore, consent is considered as a part of control [Se22] and standard as a part of interoperability [Se22]. Although [Se22] mention flexibility as a principle, most of its proposed characteristics are related to interoperability. Additionally, [Ču22] show that experts do not strongly associate any principle with their flexibility category. Furthermore, it is worth noting that several principles correspond to software product quality criteria defined in ISO/IEC 25010, and these criteria may be used to derive additional requirements.

Requirements: By analyzing the principles, their definitions, characteristics, etc., we specify requirements based on them. Some requirements are not based on a single principle but on several principles. Tab. 2 presents a set of requirements based on these principles, focusing on the requirements of wallets for individuals. An important condition for each requirement is that it is measurable. In an earlier version of our approach, each requirement was rated as fulfilled or unfilled. However, a preliminary assessment showed that in some cases, a requirement is only partially fulfilled. In some cases, the reason may be that the requirement is too broad and should be split into two or more. However, splitting requirements too much leads to too many requirements, which are difficult to manage. So we introduced a third possible value, and our possible values for the degree of fulfillment are: unfulfilled, partially

fulfilled, and fulfilled, which are illustrated by the Harvey balls ○, ◑, and ●. It is also possible to specify the degree of fulfillment as a blank if it cannot be determined. Finally, a more detailed scale may be applicable, but it is not provided in the current approach.

Tab. 2: The requirements must be examined during the assessment by determining the degree of fulfillment for each product. This table shows a fictive example assessment of a single product.

Principle[1]	No.	Derived Requirements: The system must ...	Source[2]	MyWallet[3]
Security	SEC-1	use state-of-the-art cryptography.	C, M	○
	SEC-2	encrypt interactions end-to-end.	C, M	◑
	SEC-3	use hardware-based cryptography.	C	●
Verifiability	VER-1	ensure that credentials and presentations are tamper-proof.	C, P	○
	VER-2	support revocable credentials.	C, P	○
Authenticity	AUT-1	support mechanisms for device binding.	C, M	●
	AUT-2	support mechanisms for holder binding.	C, M	◑
Privacy	PRI-1	support zero-knowledge proofs.	P, M	○
	PRI-2	support selective disclosure.	P, M	●
	PRI-3	comply with the holder's right to be forgotten.	P	○
Control	CON-1	require the user's consent for important actions.	P	●
	CON-2	ensure that the holder is the single source of truth about their identity.	P	◑
Interoperability	INT-1	use common, standardized credential formats.	C, P	●
	INT-2	use common, standardized communication protocols.	C, M	●
	INT-3	support a migration to another wallet.	P	◑
Usability	USA-1	follow accessibility guidelines.	P	○
	USA-2	be able to recover data in case of loss.	P	●

[1] Order based on the ranking of [Ču22].[2] P = Product, C = Code, M = Montoring [3] Exemplary Values.

Conformance Test Criteria: It is useful to define criteria for each requirement and to specify what it means to satisfy a requirement as part of the conformance test strategy. They help make the requirements more measurable and keep the number more manageable by allowing multiple criteria to be grouped into a requirement. For example, in addition to the splitting mentioned above, some requirements were combined and used as criteria. For example, for the requirement, *INT-1 Common, standardized credential formats* potential criteria are: *SD-JWT Support*, *JSON-LD Support*, *mDL Support*, etc. The first question is more important for comparing products in terms of their adherence to the SSI principles, as the latter are too detailed. However, they can serve as a basis for answering the first question.

Conformance Test Sources: Another part of the test strategy is the sources: Documentation (D) can be used as a baseline, and sometimes is the only source available. However, the documentation may differ from the actual product, so other sources are preferable. An obvious approach is to use the product (P). However, some requirements may require a

deeper level of analysis. Examining relevant parts of the source code (C) for open-source products is another effective but potentially time-consuming approach. Monitoring (M) product behavior, such as traffic capture or log analysis, can provide additional insight.

Requirement: To illustrate the application of our approach, we explain it using the example requirement *INT-1* and its criteria, which was introduced in a previous example. It was derived from the interoperability principle, as there is a need to ensure that credentials are comprehensible across systems. The requirements and their criteria could be assessed using the documentation, e.g., whether SD-JWT or JSON-LD is supported. However, a better option would be to examine the source code. Another option would be to test for interoperability using an interoperability conformance suite (if available) or other wallets that support SD-JWT. Monitoring may also be an option, as the credential could be captured during a transaction and analyzed for compliance with a particular specification.

Product Selection: The requirements and their required conformance test sources can also be used to guide product selection. Ideally, each conformance test source should apply to each product. This requires proper documentation, access to the product and source code, and the ability to monitor. In particular, access to source code may be difficult in some cases. However, depending on the scope of the assessment (small vs. large number of products), this may not be a disqualifier. For some product requirements, their degree of fulfillment may be unknown if the required conformance test source is not available. Once the products have been selected, the actual assessment can begin by determining the degree of fulfillment for each requirement with the appropriate conformance test source.

3 Discussion

As our current approach is based on existing literature and previous experience with SSI products, it is important to refine and validate it. The approach has already been refined from an earlier iteration following a preliminary assessment, but further work is required. Although there is some consensus on certain core principles, more consensus is needed. However, the more significant issue is to derive requirements from these principles based on consensus from a wide range of stakeholders, including those from outside academia, such as certification bodies. This missing stakeholder involvement can be considered a limitation of our research approach. Therefore, we plan to extend our approach by consulting experts and using methods such as the Delphi technique to reach a consensus on the necessary principles and requirements. These requirements can then serve as a solid basis for carrying out the assessment.

The approach outlines conformance test sources without specifying detailed conformance test methods. Therefore, further development in this area is necessary. For instance, while our approach centers on wallets, we recognize that wallets cannot function independently and require an underlying infrastructure for interaction. Thus, certain requirements may necessitate the examination of the wallet and relevant components of the overall infrastructure.

Besides assessing the relationship between principles and products, the applicability of current SSI principles to non-human holders, such as organizations or devices, must be considered. For instance, while an organization does not have privacy (only its members), confidentiality could serve as an analogous principle. It is important to investigate the applicability of these principles to non-human holders, especially organizations. Moreover, the requirements of organizations also differ from humans [BRA24], thus, our approach would likely need some minor adjustments (choosing a different target group, principles, and requirements) to be usable for, e.g., organizational wallets or to evaluate different kinds of verifiable data registries that are likely to be operated by organizations.

4 Conclusion

We emphasized the importance of assessing the implementation of SSI principles in current products and proposed an approach to address the question *How can SSI products be assessed regarding their fulfillment of SSI principles?* Using this approach in future research will allow us to determine how well SSI principles are being implemented. Because SSI products can be used for different use cases, the principles that guide them may sometimes require compromises. For example, cloud wallets sacrifice control for ubiquitous access and potentially improved usability compared to mobile wallets. It is important to balance the trade-off between adhering to SSI principles as guiding ideals and allowing for flexibility. Too much compromise in implementation could ultimately result in a failure to achieve the goals of SSI. Therefore, observing and maintaining the relationship between principles and products is key to preserving the core of SSI in practice.

References

[Al16] Allen, C.: The Path to Self-Sovereign Identity, 2016, http://www.lifewithalacrity.com/2016/04/the-path-to-self-soverereign-identity.html, visited on: 04/15/2020.

[BRA24] Bochnia, R.; Richter, D.; Anke, J.: Self-Sovereign Identity for Organizations: Requirements for Enterprise Software. IEEE Access 12, 2024, ISSN: 2169-3536.

[Ču22] Čučko, Š.; Bećirović, Š.; Kamišalić, A.; Mrdović, S.; Turkanović, M.: Towards the Classification of Self-Sovereign Identity Properties. IEEE Access 10, pp. 88306–88329, 2022, ISSN: 2169-3536.

[FCA19] Ferdous, M. S.; Chowdhury, F.; Alassafi, M. O.: In Search of Self-Sovereign Identity Leveraging Blockchain Technology. IEEE Access 7, 2019, ISSN: 2169-3536.

[Se22] Sedlmeir, J.; Barbereau, T.; Huber, J.; Weigl, L.; Roth, T.: Transition Pathways towards Design Principles of Self-Sovereign Identity. In: ICIS 2022 Proceedings, Copenhagen, Denmark. P. 4, 2022, https://aisel.aisnet.org/icis2022/is_implement/is_implement/4.

[So23] Sovrin Foundation: Principles of SSI V3, 2023, https://sovrin.org/principles-of-ssi/, visited on: 01/19/2024.

[TA19] Toth, K. C.; Anderson-Priddy, A.: Self-Sovereign Digital Identity: A Paradigm Shift for Identity. IEEE Security & Privacy 17 (3), pp. 17–27, 2019, ISSN: 1540-7993.

Determining the Efficiency of Mitigations Based on Covered Threats

Matthias Winterstetter [1]

Abstract: Prioritization of threats is an important skill for experts working in the cybersecurity field. With daily new discovered threats and a variety of tools providing information, warnings, and alerts, it is essential for experts working in cybersecurity to identify the most important warnings and threats and handle them efficiently to stay ahead of the growing competence, organization, and size of threat groups. To assist cybersecurity experts with these tasks, this paper provides an approach covering six steps that can be used to determine the efficiency of mitigations for a system under consideration. To this end, this paper describes a straightforward approach and provides an example in which it has already been used.

Keywords: Cybersecurity, IT-Security, Mitigations, Threat Actors, Threat Groups, Mitigation Efficiency.

1 Introduction

The resources a company is willing to invest in security tends to be limited financially and in terms of human resources. This stems from the fact that security does not add any inherent value to an organization but instead serves to ensure that the existing value of the organization is not decreased by accidents or malicious actors. This tends to lead to a very strict cost benefit analysis by the management. Additionally, experts in cybersecurity are, like all experts in the tech industry, very sought-after, further limiting the resources an organizations cybersecurity team has available. Furthermore, in the event of a cybersecurity incident occurs or a relevant new vulnerability is made public, a fast and efficient response is necessary.

This limit on resources necessitates a prioritization of vulnerabilities to determine which vulnerability to handle first. This necessity is covered by a large number of risk management tools that also highlight critical vulnerabilities. With the ability to determine which vulnerability has to be handled first comes the ability to choose which mitigations to implement first. However, mitigation techniques can be vastly different from each other and can require large amounts of resources. For instance, changing a registry key or uninstalling software can be done much faster and is much cheaper than buying and configuring a firewall. At the same time, a properly configured firewall can mitigate more vulnerabilities than a change in the registry or uninstalled software.

[1] University of Stuttgart, Insitut für Arbeitswissenschaft und Technologiemanagement, Nobelstraße 12, Stuttgart, 70569, Matthias.Winterstetter@iat.uni-stuttgart.de https://orcid.org/0000-0001-9093-4381

doi:10.18420/OID2024_17

While the management of risks and vulnerabilities has been given a lot of attention by the research community, the same cannot be said for managing and prioritizing mitigations. Simply knowing which mitigations can handle the given vulnerability does not enable one to know which mitigation is best for a given scenario. To make an informed decision in such a situation, knowledge about the time a mitigation takes to implement, the required cost for implementing the mitigation, how many relevant vulnerabilities a mitigation can cover and what mitigations are already in use is required. To assist security experts with this decision-making, this paper presents an approach that can be used to determine the most efficient mitigations with respect to the number of threats covered.

The next chapters of this paper will cover the related work in chapter 2, the used methodology in chapter 3, a discussion about the results and the methodology in chapter 4 and an outlook in chapter 5.

2 Related Work

There are many articles, scientific papers, frameworks, and methodologies that concern themselves with risk management and the handling of vulnerabilities. For instance, [Le21], [Ek23], [Za23] and [Ku20] are all relatively recent contributions to the scientific community and introduce new methods for risk management. There are also articles going over the top choices for risk management frameworks [Wi00]. On the vulnerability side, there are a number of globally accessible knowledge bases, among them is MITRE ATT&CK [Mi19] which, among other things, provides a taxonomy of tactics, techniques, and procedures (TTP).

While there are many methodologies for managing risks and prioritizing vulnerabilities, the prioritization of mitigations has not been given as much attention. For instance, [GMP20] states that most vulnerability assessment tools, while performing exceptional at identifying vulnerabilities, do not have the capability to perform prioritized mitigation of said vulnerabilities. It goes on to introduce the cybersecurity vulnerability mitigation framework (CyFEr) to fill this gap. An alternative method was presented by [Be20] which also focuses on prioritizing mitigations on the basis of potential damage.

When it comes to prioritizing mitigations regarding available resources, the landscape becomes scarcer still. However, there are some proposed solutions, like [Sa13] which presents a method that can determine the most resource efficient set of mitigations given a set of mitigations and threats based on the effectiveness at blocking threats, implementation costs and probability of attacks. Another alternative is presented in [OTK08] by introducing a method through which a pareto optimal selection of mitigations can be calculated. To this end, this method uses a graded security model, representing the degree of protection, security groups, representing a set of security measures and a fitness function that represents the confidence of achieving the given security goals for a set of countermeasures. Given the scarcity of resources that

cybersecurity professionals have to deal with, the topic of optimizing mitigation usage still requires more attention in the opinion of the author.

3 Approach

The approach presented in this chapter can be used to determine the efficiency of mitigations for a system under consideration. The approach can be split into the following six Steps:

Step 1: Determine the relevant threat groups based on the threat groups motivation and their capability to target the system under consideration.

Step 2: Finalize the threat list and weight the threats: Prune the list of threats to remove threats that are not relevant to the system under consideration. Determine the weight of the threats regarding their relevancy to the system under consideration.

Step 3: Sort the threats according to their weight.

Step 4: Determine mitigations for the threats.

Step 5: Determine how many threats each mitigation can cover.

Step 6: Aggregate the weight of the covered threats with the efficiency of the covered mitigations.

To determine the relevant threat groups in step 1 the motivation behind the known threat groups and their used techniques are analysed. Based on this, we can determine how well the victim would fit into the profile of known victims of those threat groups. As a result of this analysis, we get a list of threat groups that would be motivated to attack the system under consideration as well as have the ability and know-how to implement an attack. With the list of relevant threat groups comes a list of their used threats. These lists can be aggregated into a list of threats and weighted for step 2. In step 2 we prune the list of threats that could not be used to attack the system under consideration due to physical limitations. A lack of cloud infrastructure, for instance, would make cloud-based threats irrelevant. To weight the threats for step 2 their relevancy for the system under consideration is used. This can, for instance, be the number of threat groups that make use of that threat. After each threat has been weighted, they are sorted in step 3.

For step 4 and 5 we first determine the possible mitigation methods for the threats in the list from step 3 and aggregate them into one list of relevant mitigations. Following this, we can determine the efficiency of mitigations. This can be done by determining the number of threats covered by each mitigation. Step 6 completes the process by aggregating the efficiency of mitigations with the relevancy of the threats. This can be done by summing up the weights of the threats covered by a mitigation. For a more fine-

tuned output, an attack-defence tree as presented in [AN15] can be used. Alternatively, the method presented by [Sa13] could also be used for step 5 and 6. The resulting value shows the efficiency of a mitigation regarding relevant threats and threat groups.

3.1 Example

This methodology was used in the ONCE project to determine the efficiency of mitigations. ONCE is a wallet solution for digital identities. To determine the relevant threats and threat groups we used the Risk Analysis Platform RALF. RALF was configured with the technical components involved in ONCE and the results of a questionnaire concerning the motivation of different attack groups regarding ONCE that were answered by partners from different types of stakeholders in the ONCE consortium.

Threat Name	Threat ID (MITRE)	Threat Group Usage
Malicious File	T1204.002	17
Spearphishing Attachment	T1566.001	15
Windows Command Shell	T1059.003	11
Drive-by Compromise	T1189	11
Scheduled Task/Job	T1053	9

Tab. 1: Weight of Threats

As a result of this first step, we obtained a list of 22 attack groups that would be both motivated to attack and have been recorded exploiting threats that are relevant for ONCE. Based on the relevant threat groups, the relevant threats and the number of their uses by the threat groups could be determined, thus completing step 2. For step 3 the threats were sorted according to their number of uses. The top 5 threats can be seen in Tab. 1. For step 4 we determined the possible mitigation methods using MITRE ATT&CK and sorted them based on how many threats they would cover for step 5. Lastly, for step 6 we summed up the weights for the threats covered by each mitigation, to determine the efficiency value of each mitigation resulting in the list shown in Tab. 2.

Threat Name	Mitigation ID (MITRE)	Efficiency
Restrict Web-Based Content	M1021	54
Execution Prevention	M1038	49
User Training	M1017	46
Privilleged Account Management	M1026	31
User Account Management	M1018	30

Tab. 2: Efficiency of Mitigations

The methodology presented in this chapter uses the Risk Analysis Platform RALF which is based on the patent [Ku20] to determine the relevant threats and threat groups. The

patent covers an automated system for the evaluation of information security risks. Although RALF was used in ONCE and as the basis for step 1, the presented methodology is not dependent on RALF and can use other sources to determine the relevant threats.

4 Discussion

The approach presented in the previous chapters can prioritize mitigations for a system under consideration regarding the relevancy of the threats. This knowledge gives cybersecurity experts a better understanding over the mitigation landscape of the system they are supporting. Beyond that, this knowledge can be used to optimize the employed mitigations for a system. For instance, redundant mitigations that cover a threat that can also be covered by other mitigations that are already in place can be removed if they don't add any additional protection. This can reduce the available attack surface that an attacker can exploit. A prerequisite for this methodology to work properly is an automated and precise method for discovering threats and relevant threat groups.

While this methodology can be used as is and provide more insight into the mitigation landscape of a system under consideration, there is room for improvement with the methodology as it stands now. For instance, the weighing of the threats in step 2 can be done with the number of threat groups that make use of the threat alone. To provide more context to the relevancy of the threat, the likelihood of a threat occurring can be considered as well. Furthermore, techniques for evaluating the efficiency of mitigations in step 6 should be tested and optimised to determine the most optimal implementation.

5 Conclusion

This paper presents an approach for assessing the efficiency of mitigations for a system under consideration regarding the number of threats that a given mitigation can cover. The presented methodology is detailed and shown at the hand of an example, and it's advantages and room for growth are described. The next steps would be to optimize step 2 and step 6, test the approach with different systems under consideration and have the results be evaluated by experts in the cybersecurity field.

6 Bibliography

[AN15] ASLANYAN, ZARUHI ; NIELSON, FLEMMING: Pareto Efficient Solutions of Attack-Defence Trees. In: FOCARDI, R. ; MYERS, A. (Hrsg.): *Principles of Security and Trust, Lecture Notes in Computer Science*. <Springer><Berlin Heidelberg>95-114, 2015.

[Be20] BENTLEY, MARK ; STEPHENSON, ALEC ; TOSCAS, PETER ; ZHU, ZILI: A Multivariate Model to Quantify and Mitigate Cybersecurity. Risk 8/2020, 61, 2020

[Ek23] EKSTEDT, MATHIAS ; AFZAL, ZEESHAN ; MUKHERJEE, PREETAM ; HACKS, SIMON ; LAGERSTRÖM, ROBERT: Yet another cybersecurity risk assessment framework. *International Journal of Information Security* 22/2023, 1713-1729, 2023

[GMP20] GOURISETTI, SRI NIKHIL GUPTA ; MYLREA, MICHAEL ; PATANGIA, HIRAK: Cybersecurity vulnerability mitigation framework through empirical paradigm Enhanced prioritized gap analysis. *Future Generation Computer Systems* 105/2020, 410-431, 2020

[Ku20] KUROWSKI, SEBASTIAN: Automatisches Abschätzen von Informationssicherheitsrisiken. DE: 10 2018 216 887.3, 2018/2020

[Le21] LEE, IN: Cybersecurity: Risk management framework and investment cost analysis.: *Business Horizons*, 64/2021, 659-671, 2021

[Mi19] MITRE: *MITRE ATT&CK*. https://attack.mitre.org/. 2019-10-14

[OTK08] OJAMAA, ANDRES ; TYUGU, ENN ; KIVIMAA, JYRI: Pareto-optimal situaton analysis for selection of security measures. In: *MILCOM 2008 <IEEE,>*< San Diego, CA, USA> 1–7, 2008

[Sa13] SAWIK, TADEUSZ: Selection of optimal countermeasure portfolio in IT security planning. *Decision Support Systems*, 55/2013, 156–164, 2013

[Wi00] *Today's Top Risk Management Frameworks*. https://www.splunk.com/en_us/blog/learn/risk-management-frameworks.html. 2024-02-13

[Za23] ZADEH, AMIR ; LAVINE, BRANDON ; ZOLBANIN, HAMED ; HOPKINS, DONALD: A cybersecurity risk quantification and classification framework for informed risk mitigation decisions. *Decision Analytics Journal*, 9/2023, 100328, 2023

Trustworthy QWACs – Fact or Fiction?

Kai Martius [1], Tina Hühnlein [2], Detlef Hühnlein [2], Tobias Wich [2]

Abstract: Qualified certificates for website authentication (QWACs) have been introduced in Art. 45 of the eIDAS-Regulation (EU) No. 910/2014 about ten years ago and an amendment of the regulation has been provided with (EU) 2024/1183. Within the scope of drafts of this text the applicable requirements for QWACs changed, as explained below, which resulted in an open letter, which has been signed by a substantial number of scientists and researchers around the world and many NGOs. The fear of the scientists, researchers, NGOs and browser vendors, such as Mozilla for example, was particularly the introduction of a legal backdoor to emit bogus root certificates into the trust store of browsers by malicious state actors in order to intercept the web traffic of citizen within Europe and beyond. Against this threatening background, the present contribution seeks to provide unbiased information with respect to this controversial topic in order to contribute to answering the central question of the paper raised in the title, whether QWACs are trustworthy (today and tomorrow) and how the standardisation bodies in charge might move on to further improve trust.

Keywords: eIDAS, QWACs, EV SSL, trust, conformity assessment, CAB, supervisory body, browser vendor

1 Introduction and problem statement

There are currently 52 Qualified Trust Service Providers (QTSPs) according to Art. 3 (20) (EU) No. 910/2014 , which issue Qualified certificates for website authentication (QWACs) according to Art. 45 and Annex IV (EU) No. 910/2014. There was an original proposal of the European Commission for the update of the eIDAS-regulation provided on 3rd of June 2021, a version prepared by the Council of the EU from 23rd June 2023, the agreed text from 10th of December 2023 and the final text (EU) 2024/1183 published in the Official Journal of the European Union, which in particular includes the highly controversial Art. 45 (1b), which introduces an *upper bound* on the security requirements of QWACs.

Looking at the details of the version prepared by the Council of the EU (Nr. 352a-352b, page 283) it seems that the highly controversial paragraph was inserted in a rush[3] *without* any conflicting position of the involved parties. This remarkable stipulation may or may not become a practical problem depending on the standards for the evaluation of compliance, which will be referenced in a forthcoming implementing act according to Art. 45 paragraph 2,

[1] secunet Security Networks AG, Ammonstrasse 74, 01067 Dresden, Germany, kai.martius@secunet.com
[2] ecsec GmbH, Sudetenstr. 16, 96247 Michelau, Germany, {tina.huehnlein, detlef.huehnlein, tobias.wich}@ecsec.de
[3] As there is no paragraph 2a within Art. 45, the displeasing paragraph should have in fact be named paragraph 2a. In the final text (EU) 2024/1183 it is paragraph 1b.

doi:10.18420/OID2024_18

which is due by 21 May 2025.

To better understand, whether there is a real problem or only a “storm in a teacup”, we will try to provide an unbiased discussion of different aspects related to trust and the trustworthiness of QWACs, where we start in Section 2 with a general discussion of basic aspects related to trust, before we compare in Section 3 the trust and compliance regime of QWACs issued by QTSPs according to eIDAS and based on ETSI standards with certificates shipped within the trust store of browser after a corresponding conformity assessment according to WebTrust, based on Baseline Requirements and Extended Validation Guidelines. This includes a short summary of the main differences, before we conclude the paper in Section 4 with suggesting potential future path for further harmonisation of the requirements and improvement with respect to trust for both QWACs and EV SSL certificates.

2 General Aspects of Trust

„Trust“ is a deeply social concept of humans (and assumingly many other creatures) to survive. In today's world, trust has taken on a broader meaning as we increasingly interact with non-human entities such as technology and user interfaces. These elements have become so integral to our daily lives that untrustworthy individuals can cause us harm, both mentally and physically, as well as financially. The rising number of cyber security incidents serves as evidence of this.

Efforts have been made to establish trust in our technological infrastructure, moving away from blind trust and towards security-by-design and default. Various methods, such as evaluation and certification processes, transparency of code, and open standards, serve as the foundation for building trustworthiness.

This paper focuses on one critical aspect of our technical infrastructure: the generation and distribution of cryptographic keys and digital certificates. Cryptography plays a central role in ensuring confidentiality, integrity, and authenticity in digital communication, and keys are essential for cryptography. While symmetric algorithms were previously the norm, the advent of asymmetric algorithms allowed for the creation of hybrid crypto systems, reducing the need for secure distribution of secret keys.

Efforts to build scalable certificate infrastructures have been ongoing since the 1990s. The technical aspects involve binding keys and attributes such as owner, validity period, and issuer. However, verifying the trustworthiness of these bindings poses a challenge. To establish trust, policies and their validation are incorporated into the certificate generation and validation process. These policies address both security implementation and lifecycle management. Policy validation cannot be fully automated as long as humans are involved, making human audits a standard practice. Multiple "roots of trust" have emerged within these certificate infrastructures, requiring trust in their ability to carry out the necessary checks. Achieving a common agreement relies on standard bodies such as eIDAS and the CA Browser Forum, which set standards for certificate infrastructures. This paper examines the policy and security standards of these organizations, as well as the regulatory

environment.

Ultimately, the enforcement of trustworthiness in these infrastructures depends on whom the end user trusts. The paper explores how the CA/B Forum mandates audits for Certification Authorities, as well as the compliance infrastructure of EU Member States.

3 Comparing the audits of eIDAS/ETSI with that of CA/B Forum

3.1 The eIDAS Trust System

The "eIDAS Trust System" comprises various entities, including organizations, legal entities, and regulatory bodies, collaborating to maintain trust in the eIDAS ecosystem. It provides XML-based lists on service trustworthiness, certificate applicability, and cryptographic key validity. The technical format for the "Trusted Lists" according to (EU) No. 910/2014 Art. 22 is defined in (EU) 2015/1505, which refers to ETSI TS 119 612 (v2.1.1).

As depicted in Figure 1, the "eIDAS Trust System" comprises

- the European Commission (EC),
- the European co-operation for Accreditation (EA),
- the EU Member States (MS),
- the National Accreditation Bodies (NAB), which are appointed by the MS and EA according to Regulation (EC) No 765/2008,
- the Supervisory Bodies (SB) designated by the MS according to Art. 8 (1) (EU) 2022/2555,
- the European Union Agency for Cybersecurity (ENISA), which supports the SBs according to Art. 37 (EU) 2022/2555,
- the Conformity Assessment Bodies (CAB) according to (EU) No. 910/2014 Art. 3 (18), which are accredited by the NAB according to (EC) No 765/2008 for performing eIDAS-specific conformity assessments, and last but not least
- the (Qualified) Trust Service Providers ((Q)TSP) according to (EU) No. 910/2014 Art. 3 (19)-(20), which provide one or more eIDAS services.

According to (EU) No. 910/2014 Art. 22 (3), the MS notify the EC "on the body responsible for establishing, maintaining and publishing national trusted lists, and details of where such lists are published, the certificates used to sign or seal the trusted lists and any changes thereto." According to (EU) No. 910/2014 Art. 22 (4), the EC uses this information to publish the "List of the Lists" (LotL), which in turn refers to the various Member State Trust Lists (MS-TL).

In order to be listed in the MS-TL, a TSP needs to engage a CAB to audit the conformity of the TSP's Trust Service. The result is recorded in form of a Conformity Assessment

Report (CAR), which is submitted to the SB in charge together with a corresponding notice according to (EU) No. 910/2014 Art. 21. The SB verifies the CAR and includes the information related to the Trust Service under consideration into the corresponding MS-TL.

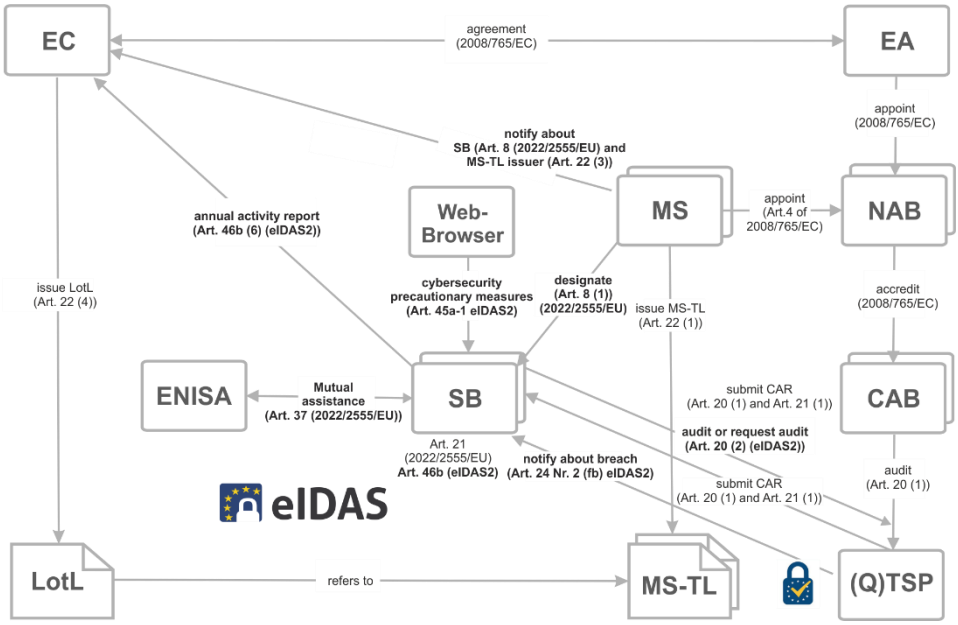


Figure 1: The eIDAS Trust System at a glance

While the Supervisory Bodies under (EU) No. 910/2014 have been specific authorities for the supervision of trust services, the amended regulation (EU) 2024/1183 foresees that the supervision of QTSPs is performed by the competent authorities for cybersecurity according to Art. 21 of the NIS2-Directive (EU) 2022/2555. Another important change has been introduced in Art. 20 (2) (EU) 2024/1183, which allows the Supervisory Bodies to perform audits of QTSPs or request that corresponding audits are performed by a CAB. For the specific case of QWACs the newly introduced Art. 45a (EU) 2024/1183 is worth to be mentioned here, as it stipulates that Web-browsers may, "only in the event of substantiated concerns related to security breaches or the loss of integrity of an identified certificate or set of certificates, providers of web-browsers may take precautionary measures in relation to that certificate or set of certificates." (see Art. 45a (2) (EU) 2024/1183).

3.2 CA / Browser Forum compliance regime

On the other hand, the „trust and compliance framework" of the CA / Browser Forum is

much simpler, as there is only the WebTrust Auditor, which is trusted by the Web-Browser vendor, and audits a Certification Authority (CA) according to the Baseline Requirements and Extended Validation Guidelines developed by the CA / Browser Forum.

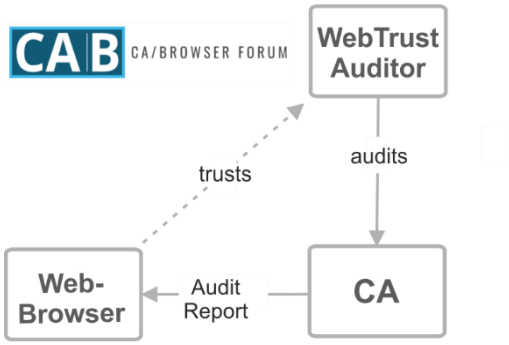


Figure 2: CA/Browser Forum Trust System at a glance

3.3 ETSI standards for QWACs and CAB-Forum requirements

In a similar manner the set of standards for QWACs developed within ETSI ESI is somewhat more complex and general than the Baseline Requirements and Guidelines developed within the CA / Browser Forum.

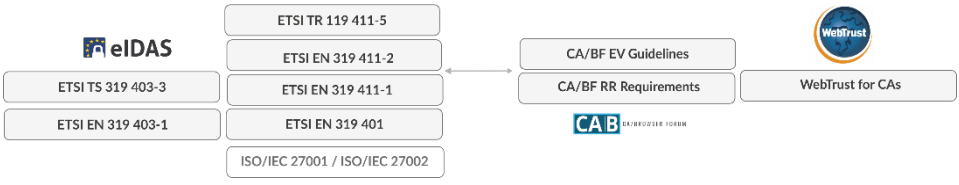


Figure 3: ETSI standards for QWACs and CA/Browser requirements

The set of requirements for auditing QWACs is composed of the basic requirements, applicable for any QTSP in EN 319 401, in turn often referring to general requirements defined in ISO/IEC 27001, ISO/IEC 27002, and CA-specific requirements defined in EN 319 411-1, EN 319 411-2. Furthermore, there is a technical report[4] [TR119411-5], providing guidelines for the coexistence of web browser and EU trust controls.

For eIDAS / ETSI there are also standards for general requirements for conformity assessment bodies assessing Trust Service Providers ETSI EN 319 403-1 and

[4] Note, that there is currently ongoing work within ETSI ESI, which aims at producing a technical specification ETSI TS 119411-5 for this purpose, which is planned to build upon and improve the existing technical report. The current draft of this technical specification includes the novel "2-QWAC Approach", in which a regular domain validated certificate, which is issued based on CA/B Forum requirements is combined with a JAdES-based TLS Certificate Binding according to ETSI TS 119 182-1.

corresponding requirements for auditing public CAs based on ETSI TS 119403-2.

For auditing CAs according to the requirements of the CA / Browser Forum there are only the Baseline Requirements, the Guidelines for Extended Validation Certificates and the WebTrust for CA audit criteria [WebTrust].

3.4 The main difference of eIDAS/ETSI compared to CAB-Forum

While the detailed requirements for auditing QTSPs, which issue QWACs according to eIDAS/ETSI are more or less comparable to the requirements defined within the CAB-Forum, there is one important difference, which should be highlighted here. The Baseline Requirements (Section 7.1.2.9) allow that a CA may construct and sign a "Precertificate" according to RFC 6962 for the purpose of submitting it to a Certificate Transparency (CT) log. The operator of the CT log arranges the logs in a Merkle hash tree, which in the end allows to check whether there has been a certificate issued for a particular domain. If there would be an authorised certificate for a particular domain, the domain owner would be informed accordingly and could act with appropriate measures, such as revoking a bogus certificate.

4 How to further improve trust in QWACs and Browsers

Therefore, we propose some basic principles and measures to further improve trust in QWACs and browsers:

- A user-friendly interface for end systems to understand the "built-in" trust infrastructure of Web Browsers and other related applications should be provided, as well as an easy opt-out mechanism to personalize these trust relationships, according to user preferences.
- There should be transparency and user-friendly representation of certificate policies applying to a specific certificate chain, which is especially important for QWACs.
- Applying certificate transparency according to RFC 6962[5] will substantially minimise the risk that a bogus CA would be able to issue a fraudulent certificate for a particular domain without being detected.
- Finally, the publication of human and machine-readable audit reports to enable independent expert validation and automated analysis of these reports to foster trust and transparency in the certification infrastructures and the implemented audit processes for QWACs and EV SSL certificates.

It is planned to discuss these recommendations for potential improvements with respect to trust within ETSI ESI and the EU Web Authentication Task Force.

[5] See also https://certificate.transparency.dev/howctworks/ for an accessible explanation of the basic principle of certificate transparency and RFC 9162 for an updated version of the certificate transparency standard.

Qualified Ledgers – Breakthrough for proven security and legal trust in DLT through eIDAS2 Regulation?

Ignacio Alamillo [1], Steffen Schwalm[2] , Carsten Stoecker[3] , Ricky Thiermann[4]

Abstract: eIDAS 2.0 as a legal and technical framework for trustworthy, decentralized identities in conjunction with the EU digital wallet and various trust services could lead to a rise in distributed ledger technologies (DLT) and European Blockchain Services and Infrastructure (EBSI). A variety of possible uses of distributed ledger technologies in conjunction with the EU digital wallet under the regulatory requirements of eIDAS 2.0 are conceivable and could also lead to broader use of EBSI with the qualified trust service for electronic ledgers.

Keywords: eIDAS2, EU Digital Wallet, Architecture Reference Framework, DLT, EBSI, Electronic Ledger

1 Introduction and status of DLT in Europe

Until 2019 the Distributed-Ledger-Technology (DLT) and its most famous representative blockchain generated a real hype particularly the well-known use case Bitcoin [Ko21]. After the bitcoin crash and especially the security concerns of German National Cybersecurity Authority [TO19] as well as the issues around the German ID Wallet [ID21] first doubts about the real capacity, security and trust of DLT occurred. In this context standardization on DLT increased and industry as well as public sector used the chance to enable the technology for high-regulated industries with corresponding requirements on records management and trust [Le17] [AS22]. Within framework of the European Blockchain Partnership (EBP), the European Commission established a European DLT-infrastructure provided by the Member States [EBSI]. This means that the DLT nodes are under the responsibility of the Member States and so ensure a government trust anchor. Since [EBSI] contains its own governance and technical specifications together with conformance tests for wallets it could solve the trustworthiness issues in DLT but, as it lacks security standards and independent audit processes, the growth of [EBSI] was limited. Beside [EBSI] also other national or private DLT networks have appeared e.g. [Alastria] in Spain, [ID Union] in Germany, [Findynet] in Finland or [Comercio] in Italy. In most cases DLT was used as a form of decentralized PKI for the execution of the new SSI paradigm [AS22] based in wallets as well as in the issuance and verification of verifiable credentials, such as digital

[1] Universidad de Murcia. Facultad de Derecho, Campus La Merced. 30001 Murcia.
[2] msg group, Robert-Bürkle-Str. 1, 85737 Ismaning, Germany
[3] Spherity GmbH, Emil-Figge-Straße 80, 44227 Dortmund, Germany
[4] Spherity GmbH, Emil-Figge-Straße 80, 44227 Dortmund, Germany

doi:10.18420/OID2024_19

diploma, mobile driver license or power of attorney. Other use cases such as cryptocurrencies, supply chain or notarization can be mentioned.
To use DLT for trustworthy digital transactions, it is necessary to make transactions and their records evident against third parties, to fulfil burden of proof and documentation needs [AS22] [Ko20]. Due to the lack of appropriate measures to fulfil such requirements of state-of-the-art record management it was not possible to use DLT for trustworthy digital transactions in general and decentralized identities in particular as needed in regulated environments [Ko20], [AS22]. Those shortages and the lack of proven security of DLT networks and their providers lead to the de facto ban of DLT for regulated industries in some EU member states like e.g. Germany [AC24], [TO19], [EC21].

The [eIDAS2] establishes, as an amendment of [eIDAS1], a legal and technical framework for trustworthy decentralized identities with the EU Digital Wallet (EUDIW) and related (qualified) trust services, using or not DLTs, on one hand but with a new dedicated QTSP for Electronic Ledger on the other hand. Although the term Electronic Ledger in [eIDAS2] does not necessarily mean only DLT – even less, blockchain – this regulation seems like a step forward to close the gaps and to enable DLT to be used in regulated environments with typically comprehensive requirements on proven security and legal trust [AS22], [Ko20]. But what's the role of DLT within [eIDAS2]? How to differentiate the different possibilities in using DLT for EUDIW and QTSP but especially the new QTSP for Electronic Ledger? As [EBSI] already exists the question on its integration in [eIDAS2] occurs too.

The paper describes based on introduction on electronic ledger in general and DLT in particular (Section 2), the main changes of [eIDAS2] and the role of DLT in the new ecosystem, especially the new QTSP for Ledger. This description will lead to the role of [EBSI] as a European DLT network within the [eIDAS2] and its transformation according to the trust model of the regulation (Section 3). The paper closes (Section 4) with considerations on the future of DLT in high-regulated industries based on [eIDAS2], including possible use cases and an outlook on necessary standardization and research in the development of the [eIDAS2] ecosystem.

2 Distributed Ledger Technology

Basically, DLT is a decentralized distributed peer-to-peer network of technical nodes for data exchange and transaction execution. According to [IS20] a distributed ledger is in this case shared across a set of DLT nodes and synchronized between the DLT nodes using a consensus mechanism, which ensures that all transactions are valid and unaltered. Once written to the ledger the transactions are immutable, mainly based on hash protection of data stored on the chain. Any transaction can reliably be tracked on the chain. In case the DLT is organized in blocks it's called blockchain, so basically a blockchain is a special kind of DLT [AS22], [Ko21]. Blockchain is not a simple

algorithm, but a technological construct and enabling protocol that facilitates the decentralized intermediation of data between participants [HKH20]. In comparison to the original ideas of blockchain, DLT does not mandatorily require the elimination of an operator or consortium providing the distributed network, this depends on the kind of DLT which can be distinguished regarding the access rights and transparency of the transactions.

If DLT is to be used for trustworthy digital transactions, it is mandatory to fulfil requirements on records management including long-term preservation of the evidence of authoritative records also against 3rd parties, until the end of the retention periods in force and to keep them provable, as it is required for any business IT-system. This means a valid records management ensuring integrity, authenticity, reliability, confidentiality, and transferability of so authoritative records by trusted 3rd parties incl. evidence preservation for the whole retention period. Additionally proven security of a DLT network done by independent 3rd party based on international standards is an additional core requirement to use DLT in regulated environments with the need to fulfil burden of proof. Without additional measures like given in [DI21] DLT is currently not able to fulfil those as comprehensively described in [Ko20], [Ko21], [AS22] [IS23].

3 A QTSP for Electronic Ledger. The EBSI portfolio challenge

Section 11 [eIDAS2] introduces (qualified) trust services for Electronic Ledger (Art. 45k and following). [eIDAS2] mandates that qualified ledger "are created and managed by one or more qualified trust service provider or providers, establish the origin of data records in the ledger, ensure the unique sequential chronological ordering of data records in the ledger and record data in such a way that any subsequent change to the data is immediately detectable, ensuring their integrity over time". Although [eIDAS2] is technology neutral the description in Art. 45l is in line with the definition of DLT in international standards [IS20] and contains core properties of DLT. As [eIDAS2] contains the requirement of mandatory implementing acts referring to European standards it ensures coherent technical framework for DLT. Since the requirements on QTSP also apply for QTSP for Ledger these standards will also be the basement for certification by independent conformity assessment body and so ensure proven security and trust in DLT. It must be stated that Section 11 focus on all use cases not covered by the European digital identity wallet (EUDIW) nor any other (qualified) trust services so e.g. (qualified) signatures, seals, timestamps, attestations electronic delivery etc. This means that DLT can be used as infrastructure for any EUDIW as well as any other QTSP too, as the security will be proven within the conformity assessment of the CAB, but there's no need to use QTSP for Ledger as precondition to provide another (qualified) trust service nor an EUDIW [Sc24] [eIDAS2]. This differentiation is important as it leads to the core use cases for QTSP for Electronic Ledger as e.g. tokenization or digital assets, cryptocurrencies or traceability in supply chains and digital product pass. [Al24] shows the possible use case scenarios for electronic ledger (DLT) within [eIDAS2]

ecosystem.

There's no trust by default in the European Union. Trust only occurs based on European law, supervised by European and national supervisory bodies, accreditation of conformity assessment bodies under European standards, certification of trust services by CAB under supervision of national supervisory bodies and verifiable via European wide trusted lists [AS22], [Sc23]. As DLT can be used as infrastructure for all QTSP but also EUDIW and especially the new QTSP for Electronic Ledger the role of [EBSI] as European Blockchain Service Infrastructure needs to be analysed further.

The European Blockchain Services and Infrastructure ([EBSI]) is a project initiated by the European Commission and a group of 29 European countries. The project, which was set up in 2018, aims to lay the foundation for future EU public services. The [EBSI] project is currently run by nodes operated by member states. Each country is expected to operate at least one node of [EBSI] at full scale. This approach aligns with the decentralized nature of blockchain technology and is suitable for multi-party cooperation. [EBSI] on one and it ensures a governmental trust anchor and so clear responsibility on the other hand this approach leads to the question on how such a network might be provided (QTSP for Electronic Ledger) or use (by EUDI Wallet Issuer or QTSP using DLT) by a certain provider. With the introduction of [eIDAS2] and the concept of qualified electronic ledgers, the [EBSI] could potentially not only evolve from an 'electronic ledger' into a 'qualified electronic ledger' enhancing security and reliability of the network, and providing legal certainty for use cases that build on the EDIC's electronic ledger. [EBSI] could also act as decentralized, pan-European Infrastructure for other (qualified) trust services such as issuance of (qualified) certificates, eDelivery (as e.g. planned in [TRACE4EU] project) or Archiving as well as the EUDI Wallets but also for infrastructure components like a trust issuer registry as possibly more scalable replacement of the trust list [ET21].

Much more complex in case of DLT is the portfolio of a QTSP for Electronic Ledger. This applies especially on [EBSI] where the main nodes remain in responsibility of member states and so the possible QTSP must deal with already existing authorities taking one main task in the DLT network – running the main nodes – per default. As it's currently not planned to change this governmental trust anchor in the EDIC it limits the portfolio of the future QTSP for Electronic Ledger in case of [EBSI]. One possibility could be that the QTSP provides only the validating nodes and so controls the execution of transactions in the network, similar approach would be the provision of the consensus mechanism and/or the responsibility for the whole security and trust in the network. As [EBSI] is designed as pan-European network it's also thinkable that 1-n QTSP may provide certain parts like validating nodes or sub-nodes or e.g. the implementation and operation of special applications like smart contracts.

4 Considerations for the future of DLT in eIDAS ecosystem

[eIDAS2] defines the legal and through mandatory implementing acts for de facto all components also the technical framework for trustworthy decentralized ecosystem in Europe. As the regulation is technology neutral it also allows the utilization of DLT for each component from EUDI Wallet and all QTSP. With the QTSP for Electronic Ledger [eIDAS2] establishes a dedicated (qualified) trust service for DLT. Due to the integration of DLT in the eIDAS trust framework all requirements on EUDI Wallet and QTSP like liability (EUDIW = member state), conformity assessment by independent CAB apply which ensures the proven security, legal trust and so solves the main gaps mentioned in Section 1 which limited a broad utilization of DLT in Europe. QTSP for ledger can be a game-changer for Industry 4.0, particularly in sectors such as energy, supply chain, and manufacturing. For instance, qualified DLT can be used for secure authentication, authorisation, service discovery, and data sharing in the energy system. It can support use cases such as flexibility aggregation, load shifting, EV charging forecasting and settlement, Guarantee of Origin, smart dispatch, smart city, and customer switching processes. In the manufacturing sector, qualified DLT can also play a significant role. Industrial use cases such as Third-Party Risk Management, smart manufacturing, digital product passports (DPPs), and supply chain optimisation can benefit from a qualified DLT infrastructure. In the industry section qualified electronic ledgers may enable execution of European Supply Chain Regulation for proof of origin of products but also product-/data and document traceability. Trusted Digital Product Passports can be enabled using qualified ledger. Document traceability might be also exciting for public sector and e-commerce for audit trails on any online service - a combination with eDelivery could also ensure evident confirmation of receipt in decentralized ecosystems. Public sector applications will benefit as well from the QTSP for Ledger service [SA21] [RCG19].

In summary [eIDAS2] creates the basement for possible breakthrough of DLT to be used in high-regulated industries while ensuring burden of proof and documentation requirements as the main requirements must be fulfilled by QTSP for Ledger or Providers of EUDIW resp. QTSP using Ledger. [EBSI] can act as common European infrastructure as its governmental trust anchor ensures an additional advantage in comparison to complete private networks. As regarding regulation, the implementing acts have to be published not later than May 20th, 2025, the research shall focus on definition of concrete security and technical requirements for certification of EUDIW/QTSP using DLT as well as QTSP for Ledger. Especially the portfolio definition of QTSP for Ledger and in this context the adjustment of [EBSI] regarding [eIDAS2] seem to be most important issues to be solved.

Bibliography

[AS22] Alamillo, Dr. I., Schwalm, S.: Self-Sovereign-Identity & eIDAS: a Contradiction?

Challenges and Chances of [eIDAS2]. European Review of Digital Administration & Law - Erdal2021, Volume 2, Issue 2, pp. 89-108

[Al24] Alamillo, Dr. I., Schwalm S., Stoecker, C., Thiermann, R.: Qualified Ledgers: Bridging the Gap between Blockchain Technology and Legal Compliance. 2024.

[AC24] eIDAS 2.0 Architecture Concept – Public, https://gitlab.opencode.de/bmi/eudi-wallet/eidas-2.0-architekturkonzept-v1, accessed: 30/01/2024

[To19] Federal Office for Information Security (BSI): Towards Secure Blockchains. Concepts, Requirements, Assessments. 2019

[Ec21] Eckpunktepapier für Self-sovereign Identities (SSI) unter besonderer Berücksichtigung der Distributed-Ledger-Technologie (DLT). Bundesamt für Sicherheit in der Informationstechnik. Bonn 2021

[DI21] DIN TS 31648:2021. Criteria for trusted transaction. Records Management and Evidence Preservation in Distributed Ledger Technologies and Blockchain.

[ET21] ETSI TS 119 612 Electronic Signatures and Infrastructures (ESI); Trusted Lists.

[HKH20] Hellwig, D., Karlic, G., & Huchzermeier, A. Build Your Own Blockchain. Springer International Publishing. 2020

[ID24] ID Wallet des Bundeskanzleramts, ein Projekt der Bundesregierung: Datenschutzrechtliche Aspekte, https://fragdenstaat.de/anfrage/id-wallet-des-bundeskanzleramts-ein-projekt-der-bundesregierung-datenschutzrechtliche-aspekte/#nachricht-643462, accessed: 30/01/2024

[IS20] ISO 22739:2020: Blockchain and distributed ledger technologies - Terminology, 2020

[IS23] ISO DTR 24332. Information and documentation — Blockchain and DLT in relation to authoritative records, records systems, and records management

[Ko20] Korte, U. et. al.: Criteria for trustworthy digital transactions – Blockchain/ DLT between eIDAS, GDPR, Data and Evidence Preservation. OpenIdentity Summit 2020. Lecture Notes in Informatics (LNI). Proceedings. Bonn 2020 S. 49-60

[Ko21] Korte, U. et. Al.: Records Management and Long-Term Preservation of Evidence in DLT. In: Roßnagel, H., Schunck, C. H. & Mödersheim, S. (Hrsg.), Open Identity Summit 2021. Bonn: Gesellschaft für Informatik e.V.. (131-142)

[RCG19] Reddick, C. G., Cid, G. P., & Ganapati, S. Determinants of blockchain adoption in the public sector: An empirical examination. Information Polity, 24(4), 379–396.

[Sc23] Schwalm S.: Trusted transaction in Electronic Ledger?. Overview on international standardization in DLT. Seeblock Webinar DLT Standardization. 10.11.2023.

[SA21] Sobolewski, M., & Allessie, D. Blockchain Applications in the Public Sector: Investigating Seven Real-Life Blockchain Deployments and Their Benefits. In M. P. and S. H. J. Reddick Christopher G. and Rodríguez-Bolívar (Ed.), Blockchain and the Public Sector: Theories, Reforms, and Case Studies (pp. 97–126). Springer International Publishing.

GI-Edition Lecture Notes in Informatics

P-316 Andrea Kienle, Andreas Harrer, Jörg M. Haake, Andreas Lingnau (Hrsg.)
DELFI 2021
Die 19. Fachtagung Bildungstechnologien der Gesellschaft für Informatik e.V.
13.–15. September 2021
Online 8.–10. September 2021

P-317 M. Gandorfer, C. Hoffmann, N. El Benni, M. Cockburn, T. Anken, H. Floto (Hrsg.)
Informatik in der Land-, Forst- und Ernährungswirtschaft
Fokus: Künstliche Intelligenz in der Agrar- und Ernährungswirtschaft
Referate der 42. GIL-Jahrestagung
21.–22. Februar 2022 Agroscope, Tänikon, Ettenhausen, Schweiz

P-318 Andreas Helferich, Robert Henzel, Georg Herzwurm, Martin Mikusz (Hrsg.)
FACHTAGUNG SOFTWARE MANAGEMENT 2021
Fachtagung des GI-Fachausschusses Management der Anwendungsentwicklung und -wartung im Fachbereich Wirtschafts-informatik (WI-MAW), Stuttgart, 2021

P-319 Zeynep Tuncer, Rüdiger Breitschwerdt, Helge Nuhn, Michael Fuchs, Vera Meister, Martin Wolf, Doris Weßels, Birte Malzahn (Hrsg.)
3. Wissenschaftsforum:
Digitale Transformation (WiFo21)
5. November 2021 Darmstadt, Germany

P-320 Lars Grunske, Janet Siegmund, Andreas Vogelsang (Hrsg.))
Software Engineering 2022
21.–25. Februar 2022, Berlin/Virtuell

P-321 Veronika Thurner, Barne Kleinen, Juliane Siegeris, Debora Weber-Wulff (Hrsg.)
Software Engineering im Unterricht der Hochschulen SEUH 2022
24.–25. Februar 2022, Berlin

P-322 Peter A. Henning, Michael Striewe, Matthias Wölfel (Hrsg.))
DELFI 2022 Die 20. Fachtagung Bildungstechnologien der Gesellschaft für Informatik e.V.
12.–14. September 2022, Karlsruhe

P-323 Christian Wressnegger, Delphine Reinhardt, Thomas Barber, Bernhard C. Witt, Daniel Arp, Zoltan Mann (Hrsg.)
Sicherheit 2022
Sicherheit, Schutz und Zuverlässigkeit
Beiträge der 11. Jahrestagung des Fachbereichs Sicherheit der Gesellschaft für Informatik e.V. (GI)
5.–8. April 2022, Karlsruhe

P-324 Matthias Riebisch, Marina Tropmann-Frick (Hrsg.)
Modellierung 2022
Fachtagung vom 27. Juni - 01. July 2022, Hamburg

P-325 Heiko Roßnagel, Christian H. Schunck, Sebastian Mödersheim (Hrsg.)
Open Identity Summit 2022
Fachtagung vom 07. - 08. July 2022, Copenhagen

P-326 Daniel Demmler, Daniel Krupka, Hannes Federrath (Hrsg.)
INFORMATIK 2022
26.–30. September 2022
Hamburg

P-327 Masud Fazal-Baqaie, Oliver Linssen, Alexander Volland, Enes Yigitbas, Martin Engstler, Martin Bertram, Axel Kalenborn (Hrsg.)
Projektmanagement und Vorgehensmodelle 2022
Trier 2022

P-328 Volker Wohlgemuth, Stefan Naumann, Hans-Knud Arndt, Grit Behrens, Maximilian Höb (Editors)
Environmental Informatics 2022
26.–28. September 2022,
Hamburg, Germany

P-329 Arslan Brömme, Naser Damer, Marta Gomez-Barrero, Kiran Raja, Christian Rathgeb, Ana F. Sequeira, Massimiliano Todisco, Andreas Uhl (Eds.)
BIOSIG 2022
14. - 16. September 2022,
International Conference

P-330 Informatik in der Land-, Forst- und Ernährungswirtschaft
Fokus: Resiliente Agri-Food-Systeme
Referate der 43. GIL-Jahrestagung
13.–14. Februar 2023Osnabrück

P-331 Birgitta König-Ries, Stefanie Scherzinger, Wolfgang Lehner, Gottfried Vossen (Hrsg.)
Datenbanksysteme für Business, Technologie und Web (BTW 2023)
06.–10. März 2023, Dresden

P-332 Gregor Engels, Regina Hebig, Matthias Tichy (Hrsg.)
Software Engineering 2023
20.–24. Februar 2023, Paderborn

P-333 Steffen Becker & Christian Gerth (Hrsg.)
SEUH 2023
23.–24. Februar 2023, Paderborn

P-334 Andreas Helferich, Dimitri Petrik, Gero Strobel, Katharina Peine (Eds.)
1st International Conference on Software Product Management
Organized by „GI Fachgruppe Software Produktmanagement im Fachbereich Wirtschaftsinformatik (WI PrdM)“, Frankfurt, 2023

P-335 Heiko Roßnagel, Christian H. Schunck, Jochen Günther (Hrsg.)
Open Identity Summit 2023
15.–16. June 2023, Heilbronn

P-336 Lutz Hellmig, Martin Hennecke (Hrsg.)
Informatikunterricht zwischen Aktualität und Zeitlosigkeit
20.-22. September 2023, Würzburg

P-337 Maike Klein, Daniel Krupka, Cornelia Winter, Volker Wohlgemuth (Hrsg.)
INFORMATIK 2023
Designing Futures: Zukünfte gestalten
26. – 29. September 2023, Berlin

P-338 René Röpke und Ulrik Schroeder (Hrsg.)
21. Fachtagung Bildungstechnologien (DELFI)
11.-13. September 2023, Aachen

P-339 Naser Damer, Marta Gomez-Barrero, Kiran Raja, Christian Rathgeb, Ana F. Sequeira, Massimiliano Todisco, Andreas Uhl (Eds.)
BIOSIG 2023
20.-22. September 2023, Darmstadt

P-340 Axel Kalenborn, Masud Fazal-Baqaie, Oliver Linssen, Alexander Volland, Enes Yigitbas, Martin Engstler, Martin Bertram (Hrsg.)
Projektmanagement und Vorgehensmodelle 2023
16. und 17. November 2023, Hagen

P-341 Gunnar Auth und Tim Pidun (Hrsg.)
6. Fachtagung Rechts- und Verwaltungsinformatik (RVI 2023)
26.–27. Oktober 2023, Dresden

P-342 Volker Wohlgemuth, Dieter Kranzlmüller, Maximilian Höb (Eds)
EnviroInfo 2023
11.–13. October 2023, Garching, Germany

P-343 Rick Rabiser, Manuel Wimmer, Iris Groher, Andreas Wortmann, Bianca Wiesmayr (Eds.)
Software Engineering 2024
26. Februar – 1. März 2024, Linz

P-344 C. Hoffmann, A. Steinm E. Gallmann, J. Dörr, C. Krupitzer, H. Floto (Hrsg.)
Informatik in der Land-, Forst- und Ernährungswirtschaft
27.–28. Februar 2024, Stuttgart-Hohenheim

P-345 Steffen Wendzel, Christian Wressnegger, Laura Hartmann, Felix Freiling,Frederik Armknecht, Sicherheit 2024
Sicherheit, Schutz und Zuverlässigkeit
09.–11. April 2024, Worms

P-346 Axel Schmolitzky, Stefan Klikovits (Hrsg.)
SEUH 2024
29. Februar – 01. März 2024
Linz, Österreich

P-348 Mathias Weske, Judith Michael (Hrsg.)
Modellierung 2024
12.–15. März 2024, Potsdam

P-349 Ralf Laue, Stephan Fahrenkrog-Petersen (Hrsg.)
Enterprise Modeling and Information Systems Architecture(EMISA 2024)
13–14 March 2024, Potsdam

P-350 Heiko Roßnagel, Christian H. Schunck, Filipe Sousa (Hrsg.))
Open Identity Summit 2024
20.–21. June 2024, Porto

All volumes of Lecture Notes in Informatics can be found at https://dl.gi.de/handle/20.500.12116/21.

The titles can be purchased at:

Köllen Druck + Verlag GmbH
Ernst-Robert-Curtius-Str. 14 · D-53117 Bonn
Fax: +49 (0)228/9898222
E-Mail: druckverlag@koellen.de

Gesellschaft für Informatik e.V. (GI)

publishes this series in order to make available to a broad public recent findings in informatics (i.e. computer science and informa-tion systems), to document conferences that are organized in co-operation with GI and to publish the annual GI Award dissertation.

Broken down into
• seminars
• proceedings
• dissertations
• thematics
current topics are dealt with from the vantage point of research and development, teaching and further training in theory and practice. The Editorial Committee uses an intensive review process in order to ensure high quality contributions.

The volumes are published in German or English.

Information: http://www.gi.de/service/publikationen/lni/

ISSN 1617-5468
ISBN 978-3-88579-744-9

Open standards and interfaces as well as open source technologies play a central role in the current identity management landscape as well as in emerging future scenarios in the area of electronic identification and trust services for electronic transactions according to the eIDAS regulation (2014/910/EU), innovative payment services according to the second payment services directive (PSD2) (2015/2366/EU), trustworthy and privacy enhancing solutions according to the general data protection regulation (2016/679/EU) and other innovative applications in the area of e-health, e-government, cloud computing and the internet of things for example.