Expert Mental Models of SSI Systems and Implications for End-User Understanding

Alexandra Mai
TU Wien, SBA Research
amai@sba-research.org

Abstract

Self-sovereign identity (SSI) systems have gained increasing attention over the last five years. In a variety of fields (e.g., education, IT security, law, government), developers and researchers are attempting to give end-users back their right to and control of their data. Although prototypes and theoretical concepts for SSI applications exist, the majority of them are still in their infancy. Due to missing definitions and standards, there is currently a lack of common understanding of SSI system within the (IT) community.

To investigate current commonalities and differences in SSI understanding, I contribute the first qualitative user study ($N = 13$) on expert mental models of SSI and its associated threat landscape. The study results highlight the need for a general definition of SSI and further standards for such systems, as experts’ perceptions of SSI requirements vary widely. Based on the expert interviews, I constructed a minimal knowledge map for (potential) SSI end-users and formulated design guidelines for SSI to facilitate broad adoption in the wild and improve privacy-preserving usage.

1 Introduction

The daily lives of most people in today’s society is characterized by a large number of online interactions with various services (e.g., email, social media, messaging, online shopping, etc.). In order to use those services, the users currently either need to duplicate their identity information for each service, or use central identity schemes, e.g. by Facebook or Google. The former reduces user experience due to a lack of usability and the latter leads to a loss of privacy and increases the risk of data compromise. Therefore, the next evolutionary and logical step to take the self-determination of analog data into the digital space seems to be Self-Sovereign Identity (SSI) systems.

SSI enables an entity (e.g., a person or organization) to have complete control over their digital identity and to decide with whom their data will be shared.
and who is allowed to use it. Therefore, users have the opportunity to regain their privacy, which is currently severely restricted in many web applications.

Over the last years, the technical aspects of SSI have been studied in great detail \cite{31, 29, 25, 32}, while human-centered research is still rare. This led to several hundreds of prototypes and proof-of-concepts (PoCs), which all provide various guarantees in terms of governance, privacy, security, and usability. Currently, one of the main challenges for SSI is the lack of a definition and corresponding standards (see Section 2), leading to various different assumptions of SSI requirements.

In order to emphasize the importance of jointly recognized standards and requirements in such a complex system, I conducted the first expert mental model study of SSI, to establish a common basis for such systems. Based on the interpretivist formative paradigm I explored and got insights into how experts (i.e., people working in the field of SSI, e.g. development, standardization, legal, etc.) perceive an SSI system and how (the lack of) current standards influences their mental models. Furthermore, I investigated which aspects of this complex system future end-users need to understand to enable reasonable and secure usage.

In particular, I sought to answer the following research questions:

\begin{enumerate}
\item[(RQ1)] What mental models of an SSI system do experts have?
\item[(RQ2)] What are the possible security and privacy risks of SSI systems?
\end{enumerate}

In order to answer those research questions, I:

\begin{enumerate}
\item[(i)] conducted the first qualitative expert study investigating expert mental models of SSI and its corresponding threat landscape,
\item[(ii)] constructed guidelines that need to be considered for future SSI designs in order to prevent security and privacy threats and,
\item[(iii)] extracted a minimal knowledge map for end-users, which need to be understood for safe usage and broad adoption of the technology.
\end{enumerate}

\section{Background and Current Status SSI}

Currently, there is no precise definition or standard for SSI. Therefore in this section, I reflect on the state of knowledge on which many publications of the last years are based. The first one to define requirements for SSI was Christopher Allen \cite{2} in 2016 and the Sovrin Foundation grouped them into three main categories as can be seen in Table 1. Those principles were refined by different aspects such as provability \cite{32} and secure transactions \cite{34}. However, none of them are used exclusively as "ground-truth". As a result, everyone makes an individual decision about which requirements to use for their "SSI" system.

Despite the different requirements and technologies used, there are certain consistent actors and components in an SSI system \cite{22}. A basic SSI architecture has three main actors:
Table 1: The 10 principles of SSI categorized by the Sovrin Foundation [33]

<table>
<thead>
<tr>
<th>Security</th>
<th>Controllability</th>
<th>Portability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Protection</strong></td>
<td><strong>Existence</strong></td>
<td><strong>Interoperability</strong></td>
</tr>
<tr>
<td>The rights of users must be protected.</td>
<td>Users must have an independent existence.</td>
<td>Identities should be as widely usable as possible.</td>
</tr>
<tr>
<td><strong>Persistence</strong></td>
<td><strong>Control</strong></td>
<td><strong>Transparency</strong></td>
</tr>
<tr>
<td>Identities must be long-lived.</td>
<td>Users must control their identities.</td>
<td>Systems and algorithms must be transparent.</td>
</tr>
<tr>
<td><strong>Minimization</strong></td>
<td><strong>Consent</strong></td>
<td><strong>Access</strong></td>
</tr>
<tr>
<td>Disclosure of claims must be minimized.</td>
<td>Users must agree to the use of their identity.</td>
<td>Users must have access to their own data.</td>
</tr>
<tr>
<td></td>
<td><strong>Portability</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Information and services about identity must be transportable.</td>
</tr>
</tbody>
</table>

- Issuer: an organization (e.g., bank, university, government agency, etc.) issuing certain verifiable credentials or claims for the holder/user (e.g., driving license, birth certificate, educational degree, etc.)

- Holder/User: an individual or company which controls and manages their credentials.

- Verifier: A company or individual that requests a specific credential (e.g. birthday to determine if a person is old enough to drink), and verifies the validity of the credential

Furthermore, an SSI architecture has a registry, which maintains the pairing of the identification and authentication by a technology-dependent authentication method (e.g., asymmetric cryptography). The registry can be either central or decentral (e.g., DLT or blockchain). The actual credential or verifiable claim is stored in a storage controlled by the holder/user. The storage of the data can be offline and local (e.g., on the user’s smartphone) and/or online (e.g., cloud).

Currently, there exist three standards for (possible) components within an SSI system that are agreed on by the community: i) verifiable credentials
(VCs\[^1\] ii) decentralized identifiers (DIDs\[^2\] and iii) decentral ledger technologies (DLTs\[^3\]) While these constitute a first step towards the standardization of SSI systems, there are still interoperability, communication, and usability standards missing among others.

3 Related Work

In this section, I describe related work in the field of SSI systems and mental model studies in the realm of usable security.

SSI systems Identity management is a research area that has existed for decades and continues to progress in new perspectives with the development of novel (security) technologies. One of the first proposals for a user-centric identity system was made more than one decade ago by Cameron et al.\[^4\]. They proposed an abstract design where the user is in control of their data, however without the security and portability characteristics of a decentralized system. The Sovrin Foundation was among the first to propose an SSI system in their whitepaper \[^33\] in 2016.

Following the whitepaper, a plethora of prototypes and proof of concepts were proposed, both from the research community and the industry. Thereby, the approaches reached from blockchain-based \[^32\] \[^25\] \[^9\] to local or distributed storage solutions \[^3\] \[^7\] \[^1\].

Mühl et al. \[^22\] gave an overview of SSI system components and discussed the state-of-the-art in 2018 based on Zooko’s Triangle \[^38\] (the three elements of the triangle: secure, decentral and human-readable). Until the development of distributed ledger technology (DLT), only two of these elements could be fulfilled at a time.

Stokkink and Pouwelse \[^32\] proposed an SSI solution for permissionless decentralized digitized passwords. Othman and Callahan \[^25\] developed a decentralized credential storage option via blockchains of biometric data for authentication. Freytsis et al. \[^9\] created an SSI-based prototype for a facility birth registration system in Kenya.

A non-blockchain-based SSI system is the IRMA (“I reveal my attributes) project which was introduced by Alpár et al. \[^3\] and is based on attribute-based credentials. Another non-blockchain-based solution is the Private Data System (PDS), where nodes manage local key-value databases and communicate through executable “choreographies”.

In 2020, Houtan et al.\[^13\] found that current state-of-the-art approaches lack i) standardization, ii) real-world usable solutions and iii) the consideration of the human factor (e.g., examination of the usability of the application and

\[^1\] VCs are tamper-evident credentials and their authorship can be cryptographically verified.

\[^2\] DIDs are a new type of identifier, which enables a verifiable, decentralized digital identity. The credential-checking method of a DID does not rely on a third party.

\[^3\] DLTs are cryptographic protocols and infrastructures, which enable decentral storage of information without a central authority.
Mental Model Studies Mental models are simplified and often implicit internal representations of objects and/or processes in the real world, formed by humans in order to make sense of their surroundings. Such mental models have been shown to crucially influence peoples’ actions and behavior [15]. In 2010 Wash [36] identified in his user study eight different folk models of security threats of home computer users. With these mental models, he explained why users ignored expert security advice and how botnets exploit knowledge gaps within those models. When examining mental models in computer science and security-related fields, imperfect mental models can be neglected as long as misconceptions do not lead to undesirable actions as clarified by Wash and Rader [37]. They proposed a way to shape mental models of non-experts to direct them to secure handling of the software.

Renaud et al. [27] emphasized the correlation between incomplete threat (mental) models and a lack of adoption of security-related applications. They showed that a poor understanding of the systems architecture in general, as well as related usability issues, lead to refusal of system usage. In 2015, Kang et al. [14] conducted a study that focused on mental models of Internet users and their privacy and security. Based on their results, they proposed different systems and policies which do not presume technical knowledge to improve the safe usage of the Internet.

Yao et al. [40] conducted a study on mental models of online behavioral advertising and how users perceive web trackers. A study comparing users’ behavior and understanding of analog and digital currency transaction systems were conducted by Perry and Ferreira [26]. Oates et al. [23] proposed the exploration of mental models through illustrations and conducted a mental model study on users’ privacy perception. Wu and Zappala [39] examined mental models of encryption and revealed that users frequently believe encryption would be an access control mechanism only, whereby some users were not even aware that encryption transforms the source data. Recent user studies on user mental models of cryptocurrency systems were conducted by Voskobojnikov et al. [35] and Mai et al. [21]. They found that users have problems understanding the underlying cryptography of blockchain-based cryptocurrency systems and that key management still poses a huge problem for many users.

As the main purpose of an SSI system is to give back the sovereignty of the user’s data, the burden of SSI key management is similar to cryptocurrencies. Although cryptocurrencies and SSI systems have similar basic principles, it remains unclear whether user mental models overlap in two such different application areas (i.e., monetary data vs. identity data).
4 Method

In the following, I describe the methodology used to address the research questions. I conducted a qualitative inquiry into how experts understand and think about an SSI system and its potential security and privacy threats. In order to get an in-depth exploration of the expert mental models, I used an iterative methodology of data collection and analysis (see Section 4.4) as it is common in qualitative research [24].

4.1 Recruitment

For this study, I sought to have a diverse sample of (SSI) experts in order to get meaningful insights. As SSI experts I defined, on the one hand, individuals from the industry who develop or work on SSI technologies (e.g. developers, SSI consultants, legal persons) and on the other hand scientists from the fields of SSI, DLTs (Distributed Ledger Technologies) and DIDs (Decentralized Identifiers).

I distributed an Email to potential participants without disclosing the concrete purpose of the study. After the interview I asked the participants whether they have further contacts to other SSI experts, thus using a snowballing sample technique [18] to recruit further people following approaches from peer-reviewed papers in the area of usable security [34, 17, 21].

I conducted the study in two rounds with a period of the preliminary analysis in between. In the initial round, I recruited six participants which were personal contacts (two from the industry and four from the research sector). The second round consisted of further seven participants (six from the industry and one researcher) chosen from the referred contacts provided by my initial sample, leading to a total of 13 interviews (demographics are summarized in Table 2). The focus of the second round was on experts from the industry and people having experience with working groups for the preparation of standards (e.g., W3C, ESSIF/EBSI, DIF, etc.).

My recruiting method might not generate a representative sample of all SSI experts in terms of residence and gender. However, I believe that the selected participants due to their various backgrounds allow insights into expert mental models of SSI and shade light on the basic understanding an end-user needs to have. In addition, I did not observe any new insights from the last two interviews. This suggests theoretical saturation [12] which is why I did not conduct any additional interviews.

4.2 Procedure

I developed a semi-structured interview protocol that helps participants to expose their mental models, based on drawing exercises used in recent mental model studies [14, 17, 21]. The interviews lasted on average 45 minutes and were held online, due to the ongoing pandemic. The interview was either held in English or in German as all participants are currently working in Central Europe.
### Table 2: Participant demographics ($N = 13$)

<table>
<thead>
<tr>
<th>Demographics</th>
<th># Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Education</strong></td>
<td></td>
</tr>
<tr>
<td>Computer Science</td>
<td>6</td>
</tr>
<tr>
<td>Law</td>
<td>2</td>
</tr>
<tr>
<td>Economics</td>
<td>2</td>
</tr>
<tr>
<td>Engineering</td>
<td>2</td>
</tr>
<tr>
<td>Mathematics</td>
<td>1</td>
</tr>
<tr>
<td><strong>Current Profession</strong></td>
<td></td>
</tr>
<tr>
<td>Researcher SSI/DLTs/DIDs</td>
<td>5</td>
</tr>
<tr>
<td>Consultant</td>
<td>3</td>
</tr>
<tr>
<td>Developer/Data Scientist</td>
<td>3</td>
</tr>
<tr>
<td>CEO/CTO</td>
<td>2</td>
</tr>
<tr>
<td><strong>Worked with standardization working group</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>8</td>
</tr>
<tr>
<td>No</td>
<td>5</td>
</tr>
</tbody>
</table>

For a smooth start, I first asked the participants about their educational background as well as their current job. Then I asked non-technical questions about their experiences and impressions of SSI. Afterward, I used three tasks to specifically probe their understanding of an SSI system, its components, actors, and their connections:

- In the first task I asked the participants to use a piece of paper and draw their idea of an SSI system in a rough sketch. They were asked to explain their drawing, either simultaneously or afterward.

- Based on their drawing, the second task was to think of possible security and privacy risks that can (theoretically) occur in such a system. Again, they had to mark or draw the mentioned risks in their sketch, in order to help them to visualize.

- The last task was to imagine, based on their drawn system, what a non-tech savvy end-user (e.g., a friend or parents) should understand about this system in order to use it in a secure and privacy-preserving manner.

During each of the tasks, the majority of the time was spent on follow-up questions in order to get a deeper understanding of the participant’s perception. The follow-up questions depended on the participant’s explanations and were therefore adapted individually without a predetermined guideline. This method provided me with the opportunity to dig deeper into specific details of the participant’s mental model.
In the first round of interviews, my focus was to gain general insights about SSI and explore the experts’ mental models. I probed to discover security and privacy aspects that concern the end-user, as well as other actors involved in the system. Furthermore, I asked what is necessary to understand SSI from an end-user perspective. After I finished the first six interviews, I made a preliminary open coding in order to determine whether specific topics needed further investigation or clarification. Therefore, in the second round of interviews, I used a slightly modified interview protocol to focus on standards that are currently lacking but under development in different working groups. The new interview guideline included two additional questions about existing standards and standards that are under development, including potential issues and challenges.

4.3 Prestudy

In order to test the comprehensibility and practicability of the interview guideline, I conducted two prestudies. I requested feedback from both participants after the interview by asking whether questions were unclear or made them feel uncomfortable. I explicitly asked for suggestions for improvement. Both participants noted they felt comfortable with the interview questions and liked the idea of the sketching part for visualization, although they are usually rather reluctant to draw or sketch. Only for the third task one felt overwhelmed to draw something and expressed that only a verbal description should be enough. Therefore, I decided, that it should also be valid to just verbally describe the answers during the tasks (especially the third one) if the participants do not want to draw or sketch. Otherwise, no further feedback was provided and therefore the interview guideline stayed the same, besides the encouragement to draw.

4.4 Data Analysis

In line with other qualitative studies in usable security [17, 21, 10, 42], I followed a grounded theory-based approach as proposed by Corbin and Strauss [6] to analyze the interview data. It is an iterative method with the goal of systematically analyzing and interpreting qualitative data to form theories that are grounded in data.

Following a grounded theory approach for data analysis, the three steps to form those theories are:

- **Open Coding** is the process of finding descriptive codes for all statements within the transcripts of each interview. Its goal is to discover recurring themes and properties within the data.

- **Axial Coding** is the process of compiling the open codes into more abstract categories and corresponding subcategories. Its goal is to uncover relationships between individual codes.

- **Selective Coding** is the process of refining the codes into a final codebook. The goal is to formulate theories grounded in data.
All interviews were i) recorded and transcribed afterwards and/or ii) notes and pictures of the drawing were taken during the study.

After finishing the data collection and transcriptions, I coded the prestudies in order to get a better understanding of emerging themes that correspond to expert mental models of SSI. Following the approach of Wash [36], I assembled a list of main themes, which I expected to see in the upcoming interviews. Those themes included information about the registry, issuer, and verifier, as well as credentials and identifiers, and opinions about SSI principals. Once I had the main themes, I coded the first round of interviews. After the first round of open coding, I found that standards/standardization and metaphors needed further investigation in the second round of interviews (see Section 4.2). Following the second round of interviews, I conducted another round of open coding and afterward combined all codes. I abstracted the found codes (axial coding) and summarized them using affinity mapping into a final codebook (selective coding). With the codebook, I coded all interviews again and created frequency tables in order to identify patterns and formulate theories (e.g., minimal knowledge map, recommendations). I used the drawings for visualization purposes which additionally informed the codebook, however, as some drawings needed extra explanation I used the transcripts when depictions were unclear (e.g., no textual explanation of a depicted "person", therefore I asked the participant during the interview to explain their drawing in more detail) or did not match the verbal description (e.g., a participant explained more details or components than what was drawn).
4.5 Ethical Considerations

The research center I am working for, which is located in central Europe, has, unfortunately, no institutional review board. However, they have a series of guidelines that should be followed when conducting user studies. One of the main requirements is to preserve the security and privacy of participants. Therefore I assigned IDs to each study participant and limited the collection of sensitive information as far as possible (only the educational background and current profession were asked), strictly following the EU’s General Data Protection Regulation (GDPR). Also, the interviews were only recorded when explicitly permitted by the participant.

5 Findings

In the following section, I describe the identified mental models which I found in the data from the expert interviews (transcripts, notes, and drawings). The purpose of qualitative research is to explore phenomena and perceptions in-depth, rather than to generalize and quantify. Therefore, I do not report numbers for the different mental models, but instead, describe in detail the range of expert perceptions.

Note that by describing and categorizing the mental models, I do not intend to imply that the models are incorrect or bad. Due to the lack of a definition of SSI (i.e., there is no "perfect" or "correct" model to compare against) and the diverse backgrounds and foci of the participants, all models are incomplete to a varying degree.

5.1 Expert Mental Models of SSI

Although the participants did not have a specific application in mind when describing an SSI system, they mentioned different application areas such as educational certificates, online shopping, or driving licenses. All emphasized that one main goal of SSI is that it should be usable for all online systems that require some sort of identification. This goal is described as "interoperability" in the 10 principles that were established by Christopher Allen [2] in 2016 and they seem to be often used as a foundation in the SSI community. When asked directly about Allen’s principles, all stated that they think that they provide a good basis. Participant #5 said, "I did not question the 10 principles, he [Christopher Allen] is the father of SSI. It [the principles] is also accepted by my colleagues.". However, not all agreed that the principles are sufficient to describe the complete requirements in enough depth. Participant #9 mentioned "unlinkability should be explicitly stated in the principles as it would severely affect the user’s privacy if it is not guaranteed... and SSI would be obsolete, as current [central] systems allow the gathering of personal information". Others argued that certain trade-offs need to be made for real-world applications. Participant #3 stated that "You have to compromise on security and privacy due to the legal basis and because of the usability".
Based on the participants’ drawings and descriptions, I constructed an expert mental model of SSI (see Figure 2). It shows all the components and connections which were mentioned by the participants. Thereby, all blue components show the parts that were explained by all participants and the dashed lines represent parts that were mentioned only by some participants. In the following, I will describe the components and connections of the mental models in more detail.

**The Actors and Communication** The mental models of all participants consisted of the three main actors as described in Section 2. Furthermore, verifiable credentials (VCs) were mentioned by most of the participants. VCs represent a piece of personal information that is cryptographically trustworthy. Interestingly, the VCs were not depicted as a separate component by the participants. Therefore, I refrained from applying a VC component to the expert
mental model but instead wrote the VC on the communication path.

How the connection and/or communication between the actors and components were explained and presented was different among the participants. Everyone emphasized that direct communication between the issuer and the verifier is not allowed. However, whether there was direct communication between issuer and holder and holder and verifier was not uniformly described. The communication depended on whether the participants mentioned: i) a communication layer, ii) DIDs as a separate component, or iii) no extra components. With the extra communication layer, there was no direct communication between the actors, but communication was exclusively done via "some kind of central communication layer" as participant #1 explained. If DIDs were an active part of the system, there was communication between the holder and the verifier (e.g., see Figure 1), and the other communication was via the created DID of the user. For example, participant #4 explained that "the users can create the DIDs and they can register them in a system, that can be something decentralized, like a blockchain [...] or also something else". Those participants mentioning DIDs within the system also described some kind of "resolver" which is necessary to get the information on how to interact with the entity represented by the identifier. This information includes for example the cryptographic keys and the technical specification for the communication.

Registry All participants mentioned some sort of registry during the course of their interviews. Thereby, I found various descriptions of the registry itself, its location, and which technology should be or must be used. The descriptions of the registry were mostly very general and referred to a place where, for example, the public keys are stored. However, some participants went into more technical details about the registry and explained how verifiable data registries (used for DIDs) work. Participant #3 also presented a possible future "meta registry", which is a kind of "a pool [...] a bit like KYC [know your customer] companies and that [pool] just has every one of the ecosystems and has access to these [individual] registries". Interestingly, most participants started to describe the registry as "a decentral" system or component. However, later on, they specified that the registry can also be central. Participant #5 explained, "the important thing is, that the registry meets certain criteria, such as a high level of availability".

The technologies mentioned for the registry ranged from blockchain and DLTs to DNS and TLS. The latter was only referenced generally as a possibility for the registry system without further explanation, as they are "already known and working infrastructures which can be used", according to participant #6. DLTs and blockchain were also either described superficially or in more detail when the participant was very familiar with the technological stack of a certain blockchain.

Storage All participants mentioned that holders possess some sort of storage where their credentials are stored. Many described or associated the storage
with a wallet (application). The participants with this perception, also described a direct connection between the wallet storage as the issuer sends the credential to the wallet of the holder. Some participants were not sure, whether the storage must be local or if there are other (secure) options. Participant #2 stated, "As a Holder, I never need to store or file anything remotely. I store the credentials locally.". Some gave evasive answers to more specific questions about where the data is stored, such as participant #8 "Actually it doesn’t matter where you store the data [credentials] ... as long as it is secure and you have access to it."

**QR Code and Deep Links** QR codes and deep links were mentioned when the participants talked about a communication layer or when the conversation turned to the user interface. In both cases, these two things served as link establishments between two actors (issuer and holder or holder and verifier). A QR code or Deep Link can be received either online or offline by directly photographing the QR code. This has an impact on the physical distance that can exist between the actors. However, the distance (personal vs. online contact) has nothing to do with the technology itself but rather with processes in relation to authentication as some participants highlighted. For instance, they mentioned that there are different guidelines for issuing certain types of documents. In most countries, an individual has to go in person to a government office to get an analog passport. This likely would not change with an online passport unless the policy and therefore the process is changed. Therefore, it will depend on the legal basis for SSI whether a personal visit to the authorities is necessary for certain identifications or not (e.g., passport).

5.2 Pictorial understanding and metaphors

Throughout the interviews, the participants repeatedly used different metaphors and pictorial language to consciously or unconsciously describe SSI. The examples used to describe this comparatively new and complex system were taken from both the analog and the digital world. During my coding process, I grouped these into five pictorial categories.

**The (digital) wallet:** In this category the participants described SSI as a digital version of a person’s wallet. Some participants even depicted a wallet in their drawing as can be seen in Figure 3. Similar to a physical wallet the users have all their documents with them to identify themselves, as long as they have their device (most likely a smartphone) with them. Participant #7 explained, "if you want to go to a club, you have to show your driver’s license or ID card at the entrance. You do the same thing with an SSI system, only it’s on your smartphone". The participants from this category used their explanation to show the "simplicity and easy use" of the system and its opportunities.

**The operating system:** In this category participants described the operating system of a computer or smartphone as the standardized base, which is the goal of interoperability in SSI. SSI developers or other individuals can "implement their own application based on the operating system of the smartphone and use it for whatever they want [different use cases]" as participant #10 explained.
Physical documents: The physical documents mentioned in this category were birth certificates and passports. Participant #6, for example, explained that "the holder has the same responsibility as with physical documents; if I lose my passport then it’s gone, I have to get a new one". Within this category, the participants emphasized the risks as well as its opportunities which go hand in hand with SSI systems and with what the user has to pay attention to. Another participant (#2) mentioned "In the real world, you would never consider it okay for an official to stand behind you with a document folder and show your certificate, for example. And if you want to go in somewhere, the official arranges it with the guard; you would be incapacitated.”

Privacy Enhancing Technologies: In this category, I combined all PET systems mentioned by the participants. Among others, TLS and PGP were mentioned, because “you have to define a complete protocol stack and the community has to agree on it” as participant #8 stated. The systems in this category were mainly used as positive or negative examples of how to develop new technology in order for it to be adopted.

5.3 Security and Privacy Risks of SSI

In the data, I found a variety of different threat models concerning SSI systems. I divided the models into four broad categories based on who caused the risk. Every threat model was shared by multiple participants of this study.

Bad Actor In this category, participants mentioned that some actors with the intention to exploit the system cause privacy risks. These actors can either be an active part (i.e., issuer or verifier), an indirect part (e.g., a developer of the software), or an external part of the system (i.e., a bad third party). The active actors of the system can, when working together, collect (meta) information about a user. The indirect privacy threat to a user’s information appears when a provider or developer intentionally implements issuer and/or verifier software that collects and combines data. Another attack in this category was Monster-In-the-Middle (MITM) attacks were external bad actors who either eavesdrop on the communication between active actors of the system or intervene by playing
the role of a certain person in the system. Although respondents with this model were concerned that the user’s privacy could theoretically be in danger, they often proposed technical possibilities to circumvent them. Participant #5 said, “to prevent a MITM from intervening, you can use TLS for example”. However, for a rogue actor, the participants agreed that a certain risk cannot be avoided.

Usability Risks The threat models of this category both affect the privacy and security of the system actors. These risks are induced by the decisions of the user and the design of the software. Therefore, in order to prevent them, on the one hand, the user has to actively perform actions and make decisions and on the other hand, the system needs to be designed in a way to encourage or even enforce a secure/privacy-preserving usage.

The most frequently mentioned risk is the loss of verifiable identity material (e.g., verifiable claims, the keys, etc.). As the user has complete control, in most cases “when the data is gone, it’s gone!”. However, to prevent data loss there are some options that the user has to decide whether to take them. On the one hand, there is the possibility to make back-ups of the verifiable identity material. Thereby, the data must be encrypted and safe at a secure (non-public) place. Some wallets provide a backup feature, however, “the standards used and [the] location where the backup is stored should be secure”, as participant #10 stated. Another thread was the compromise of the keys, due to self-inflicted unintentional data disclosure through the use of insecure online systems while signing. A possible security measurement would be to use “secure hardware modules, similar to hardware wallets from cryptocurrencies” as participant #5 explained.

Technical Risks Technical risks form the third category of threat models, which can cause major security and privacy risks. These technical vulnerabilities are the starting point for the attack vectors from the first two risk categories. Due to the rapidly evolving nature of technology, serious problems can arise both in the implementation as well as in the agreement of standards. Participant #5 stated, “as there exists no technically perfect (concerning security, privacy, and usability) system, general flaws in the network and the storage can affect the integrity of the data.” Therefore, (unintentionally) faulty implementations of SSI applications have the same risks as other software projects. An example was provided by participant #1, “ID Wallet; it was just not good engineering what they did. It was a disaster [as mentioned even in the news]. Not enough testing and pen-testing.”. Another security risk evolves around minimal security standards which need to be fulfilled. If DIDs or the resolver uses weak protocols the information is not safe. Those minimum standards need to be defined and decided on, as well as enforced. In this context, “zero-knowledge proofs” were mentioned by some participants as they provide the possibility for minimal information disclosure.

SSI-specific technical risks were mentioned in the context of key revocation
and the storage of personally identifiable information (PII) on blockchains, as both could result in privacy issues. Some prototypes use lists or bit-arrays in order to enable the revocation of keys. Participant #4 explained, “the verifier can then always check by themselves if I still have a driver’s license [although it’s revoked],” resulting in a severe invasion of privacy. Some participants mentioned that there exist approaches that (want to) store DIDs or PII with timestamps on the blockchain. As the data stored on a blockchain is either publicly available or at least visible to a specific amount of people, the extraction of information comes with major privacy risks.

Legal and Government Actors In this fourth category, legal and governmental restrictions pose a risk to the security and privacy of the participants involved in SSI. The processes for authentication as well as the decision of who is allowed to issue specific credentials pose an organizational challenge that might influence the security and privacy of users. When changes to a specification happen, this might pose security or privacy risks when certain parts of legal or governmental restrictions are circumvented or ignored. When asked for more details on the effects, the circumstances, or examples of these risks, the participants did not want to or could not provide more details.

5.4 Custodial vs. Non-Custodial Wallets

In the interviews, I was able to examine two fundamentally different attitudes regarding custodial (web) wallets. A custodial wallet is a digital wallet where the users decide to give their private keys to a trusted service provider (e.g., for fiat currencies to PayPal or in the realm of cryptocurrencies to an exchange like Kraken). On the one hand, some thought that these were good and important because they increase usability and thus promote adoption. Moreover, they explained that by having a choice of different wallet providers, one has the freedom to decide which service(s) to trust and change accordingly. When asked for a more detailed explanation of this opinion, participant #3 explained, ”You don’t have the problem [managing seed phrase or keys], the provider does that for you. [...] Unlike Facebook for example, with Facebook I can’t just go away to, I don’t know, Foodbook. [...] The moment the wallet provider does something wrong, then they [the users] just go somewhere else, that’s the important thing.”

On the other hand, some had concerns about custodial wallets, because in their opinion they lead to a central third party in the system. They explained that this would contradict the self-sovereignty part of SSI and therefore would be a major intrusion into the privacy of the user. In this context, participant #7 said ”custodial (web) wallets are in fact no longer SSI but centralized”.

5.5 Meta findings

Interestingly, when asked to present an SSI system, a large proportion of participants showed existing images or slides. This means that many have already worked with different graphics in their professional or private environment in
order to be able to discuss/present the system to other people. It was my goal to
dig deeper into their knowledge instead of gathering pre-prepared slides, which
were possibly designed by others and would bias them. Therefore, I asked them
to draw freely and tell me what comes to their mind. I noticed that after they
had mentioned and/or drawn the three actors, they sometimes faltered a bit
and thought about whether they had forgotten something. With further ques-
tions, the participants began to give examples of their professional activities.
For example, one researcher began to describe findings from a current project
and whether DIDs or VCs are preferable over one another, depending on the
technology. Another participant (industry; worked with standardization working
group) described how they want to establish a standardized SDK for SSI
and the difficulties of acceptance, as all would have to use it in order to be
interoperable. It was noticeable that participants which are either developing
or actively engaging with the technology, tend to explain the system in detail,
thereby especially DIDs methods and registry systems were described in depth.
Furthermore, people with DLT (working) backgrounds tend to emphasize the
importance of DLT and blockchain technology more in comparison to others
who often mentioned other PKI systems which are already in use as viable op-
tions. In general experts with educational backgrounds in a technical direction
gave deeper insights into the SSI architecture and specific protocol details. The
elaboration of security risks based on technology was specially detailed by re-
search participants. In comparison, participants with no technical background
or current positions without deeper technical insights, tend to explain the SSI
system more superficially. They expressed their lack of knowledge or discomfort
with some follow-up questions by stating "I can not express that well [...] this
is all really very technical" or "I am not quite sure how this is in detail".

6 Discussion

This study was inspired by the rise of SSI research and its potential to shape
users’ online identities while safeguarding their privacy. Understanding the men-
tal models that experts have of SSI and its corresponding threat landscape sheds
light on i) requirement and standardization shortcomings as well as differences
of opinion in this process and ii) how SSI systems need to be understood to
favor adoption and secure and privacy-preserving usage.

The mental models of the experts varied widely in the depth and accuracy
of their representation of an SSI system. The basic building blocks of SSI (see
Section 2) were present in all participants. However, the communication flows
and details of individual components such as storage and the registry showed
considerable differences in explanation. On one hand, the differences are due
to different technologies (VCs vs. DIDs) and, on the other hand, due to the
lack of definitions and standards, such as requirements for a storage/wallet or
a registry.

While coding, I noticed that experts’ explanations and their points of view
reflected whether they were currently active in a (specific) standardization work-
ing group. Participants working with standards described and argued with them almost twice as often as participants without this background property. In order not to make any denunciations, I will not disclose any committee names and will try to write as non-judgmental and objective as possible in the following parts.

6.1 Controversial Points of View of Study Participants

In my study, I found a big controversy about whether an SSI system can exist with a custodial wallet or not. In the case of the pro custodial wallet opinion, the principle of control and access is interpreted as giving the user control to move their data at any time if something is wrong, and access to the data is theoretically (assuming an honest provider) always available. Furthermore, usability is seen as more important than the other two principles, as otherwise (in their opinion) there will be no broad adoption. In comparison, the contra custodial wallet experts consider it important that the control and access of the user’s data must not be handed over to a third party under any circumstances. They argue that this would otherwise resemble a current centralized system.

The second controversial issue I found in my study is the storage of PII data on the blockchain. One example mentioned was the idea of storing biometric data in the blockchain, to make key recovery more user-friendly. This would make it possible to access data again after the loss of a device. Another example was the centralization of a validation server. Both could have advantages in terms of usability, but some experts expressed concerns (which are in line with related published findings [11, 41]) about SSI criteria and privacy.

Therefore, I argue that missing definitions and requirements are the main issues that the SSI community will have to address in the near future in order to move forward. The aforementioned applications and technologies have different advantages and disadvantages and fulfill certain (SSI) requirements. Therefore, a path needs to be found that meets the security and privacy needs of SSI while still maintaining the usability of the system. There are some applications (e.g., Sovrin, uPort, etc) that have taken this path, but as Liu et al. [19] have shown in their study, there are still some difficulties to overcome with current (beta version or prototype) systems.

6.2 Metaphors

This study revealed that all participants used some kind of metaphorical explanation while describing SSI. Therefore, I hypothesize that those metaphors can be used to help end-users to imagine the complex system more easily and understand at the same time key features and risks (e.g., self-responsibility, loss of data means you have to issue a new one, etc). Many experts expressed physical wallet and documents metaphors. The former correspond to the terms used to describe current cryptocurrency and SSI applications (e.g., ”wallet”). The latter could be used to emphasize the uniqueness and importance of a certain credential to the user.
In order to contextualize the found metaphors used to describe SSI from my study, I examined related work and online presentations of SSI systems. Thereby I found that most papers and presentation materials used similar metaphors, and in line with my findings especially the (digital) wallet and the (physical) documents metaphors were popular. Other metaphors I found during my research, were for example, the dot metaphor used by Sovrin to describe the online identity of a user. Another example is the ring metaphor, where the user is surrounded by a ring of sovereignty which protects his/her data from the rest of the internet. I hypothesize that the participants from my study did not use those metaphors as they did not explain the risks or benefits as accurately and visually catching as the metaphors used (see Section 5.2).

6.3 Minimal Knowledge Map for End-Users

In order to prevent that SSI systems are misunderstood by their users which might lead to security and privacy weaknesses in their mental models, I designed a minimal knowledge map for end-users. This knowledge map is kept as generic as possible so that it covers as many use cases as possible from the current point of view of an SSI system. The basic assumption for the knowledge map is that the user wants to use an SSI system, but the motivation for doing so is irrelevant. The knowledge map is based on the findings of the expert study, especially the answers of the third task. When asked what users need to understand to use an SSI system the first reaction mostly resembled the statement of participant #1 “The end-user does not need to understand anything”. However, after targeted inquiries, it became apparent that a certain understanding is required to be able to use SSI.

- **Ownership of data:** The control over the data is with the user. Therefore, the users are responsible for their digital identities and need to make some decisions concerning their security and privacy.

  1. **Storage of data:** In order to ensure the control of the data, the user needs to store them somewhere accessible. There are some (beta) wallet solutions on the market, which provide different security features and storage options (e.g., local, cloud, decentralized). The minimum requirements should be: the wallet/storage should be secured by at least a strong password (as the user only has to remember one), the user’s bio-metrics, or even multi-factor authentication.

  2. **Backup:** In case the storage wallet or the access to it gets lost, a backup of the data is recommended, however not necessary. The scenario is similar to a lost passport, where the owner needs to have it newly issued when it is lost (which is associated with more or less time and possibly financial expenses). Some wallets back up the data automatically, while other wallets require this process to be performed manually or even do not provide a backup option. Therefore, it is important for the end-user to make an informed decision about whether a backup is important and based on that, decide which wallet to use.
• **Trust:** The protocols and implementations are (theoretically) open-source, therefore you do not need to blindly trust one specific company. Through cryptographic protocols, you only show minimal information (which you control) to the verifiers and therefore they never get your (complete) information.

• **Usage:** Current SSI solutions mostly use QR codes to connect the actors (issuer and holder or holder and verifier). Therefore, it is necessary to know how to scan them and that they are for connection purposes only. However, due to the pandemic situation and the related explosion of QR code usage, I feel confident that many potential end-users are familiar with their usage. With other connection mechanisms (e.g., DeepLinks) the user needs to understand how to manually use them.

### 6.4 Implications for SSI Designs

In order to determine the implications of the results, a triangulation of the findings was performed. Therefore, the minimal knowledge map, the security and privacy risks, and the general findings (described in Section 5.2 and Section 5.5) were taken into consideration. For each part of the minimal knowledge map two questions were asked: 1) Do any of the described risks pose a threat to this part of the minimal knowledge map? 2) If yes, how can it be prevented by design? If not, how could the minimal knowledge map requirements be incorporated into a (future) SSI design? The second question was answered with insight from the general findings. Based on the answers to these questions recommendations were extracted and grouped into four actionable recommendations.

1. **Hide complexity:** Based on the expert agreement, that users do not need to understand the system itself to use it, and should not be bothered with intricate details, it is necessary to hide the complexity of the system as far as possible. This, on the one hand, increases the usability of the system and on the other hand drives adoption, as current (central) systems are very convenient to use (e.g., single-sign-on, log-in with Google or Facebook, etc.).

   For example, a user does not need to understand how a DID, VC, or DLT works. Furthermore, the communication paths and the cryptographic authentication protocols should be hidden. However, it is important that the user gets enough information to trust the system, for example, security indicators can be used to indicate if certain data is encrypted (similar to the lock symbol used in encrypted TLS communication) or whether data was sent. One possibility to avoid too frequent requests for user consent would be to introduce a white list to define certain credentials that may always be shown.

2. **Metaphor usage:** The metaphor of an online wallet has become relatively widespread due to cryptocurrency wallets and should be maintained. To
emphasize the importance of the stored information, their importance and properties could be highlighted by visual markers of small document images.

3. Implicit backups: Several experts highlighted that backups are important for usability and adoption. Therefore, an implicit backup function should be used which securely encrypts the data e.g., every month, and provides the possibility to automatically store it at a place of user preference.

4. Key recovery: Currently most wallets use seed phrases in order to provide the possibility of key recovery. Users are often overwhelmed with seed phrases as some experts in my study indicated. Therefore, more convenient methods, like individual selections of the seed phrase from a pool of words [41] or a form of Shamir’s secret sharing by sharding of the recovery key between multiple people or items [30] should rather be used.

7 Limitations and moving forward

In order to answer the research questions, I chose to conduct a qualitative user study. As I chose a mental model study that is analyzed by GT over other possibilities, I accepted certain trade-offs. With this study, I aimed to explore the (until now) unknown problem space of expert perceptions of SSI systems, to build a basis for a definition and standards of SSI. Due to the qualitative nature of this study, I can not make statistically valid statements and therefore can’t generalize the findings for every SSI expert and every (future) SSI system.

Furthermore, I did not have a particularly large sample. However, I conducted the study in an iterative manner including two prestudies and two rounds of interviews with a total of 13 participants. Thereby, the results from the prestudy and the actual interviews were consistent. In fact, the last two interviews did not add any new codes, which suggests theoretical saturation.

The findings of this study might be influenced by the participants’ cultural background and the legal landscape of Central Europe, as I did not use a generalizable sampling method. Therefore, it remains future work to conduct further studies with different cultural backgrounds in order to validate or refine the mental models.

The findings of this study can be used as a foundation to define SSI and formulate (further) standards. With the knowledge map for end-users and the guidelines, I created a basis that can be used for future end-user studies in this realm. As a next step, it is necessary to validate the mental models of experts through quantitative studies. In order to test the design implications (e.g. metaphor usage) for SSI systems, a PoC is needed, which can be tested in an end-user study. Furthermore, as soon as there are productive systems of SSI on the market, it is important to explore the mental models of non-experts. Thereby a comparison between the risk and knowledge assessments of end-users and the minimal knowledge will shed light on its applicability and whether the map needs refinement.
Other publications point out, that some prototype or beta version implementations lack usability [19, 8, 28, 41]. Therefore, a definition and framework with usability guidelines for SSI systems would be important for future work. This will also increase the likelihood that usage problems will be minimized and that there will be few or no end-user mental models that are vulnerable to security or privacy risks. This is in line with the advice of Kempton [16], who suggests designing technologies that encourage the users despite incomplete or incorrect mental models to be more secure. An SSI application should, on the one hand, protect the user from attacks and, on the other hand, highlight potential risks to motivate the user to use it safely.

References


[32] Quinten Stokkink and Johan Pouwelse. Deployment of a blockchain-based self-sovereign identity. In 2018 IEEE international conference on Internet of Things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData), pages 1336–1342. IEEE, 2018.


