TU WIEN Informatics

Mental Models of Cryptographic Protocols from Different Stakeholder Perspectives

DISSERTATION

zur Erlangung des akademischen Grades

Doktorin der Technischen Wissenschaften

eingereicht von

Dipl. Ing. Alexandra Mai, BSc
Matrikelnummer 01125691

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Univ.-Prof. Dipl.-Ing. Mag. Dr. techn. Edgar Weippl
Zweitbetreuung: Dipl. Ing. Dr. techn. Katharina Krombholz

Diese Dissertation haben begutachtet:

Corinna Schmitt	Isao Echizen

Wien, 14. November 2022

Alexandra Mai

Technische Universität Wien
A-1040 Wien ▪ Karlsplatz 13 ▪ Tel. +43-1-58801-0 ▪ www.tuwien.at

TU WIEN Informatics

Mental Models of Cryptographic Protocols from Different Stakeholder Perspectives

DISSERTATION

submitted in partial fulfillment of the requirements for the degree of

Doktorin der Technischen Wissenschaften

by

Dipl. Ing. Alexandra Mai, BSc
Registration Number 01125691

to the Faculty of Informatics

at the TU Wien

Advisor: Univ.-Prof. Dipl.-Ing. Mag. Dr. techn. Edgar Weippl
Second advisor: Dipl. Ing. Dr. techn. Katharina Krombholz

The dissertation has been reviewed by:

Corinna Schmitt	Isao Echizen

Vienna, 14th November, 2022

Alexandra Mai

Technische Universität Wien
A-1040 Wien ▪ Karlsplatz 13 ▪ Tel. +43-1-58801-0 ▪ www.tuwien.at

Erklärung zur Verfassung der Arbeit

Dipl. Ing. Alexandra Mai, BSc

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 14. November 2022

Alexandra Mai

Acknowledgements

The road to completion of this thesis was paved by the support and encouragement of my family and friends, to whom I am incredibly grateful. My parents enabled me to study the subject of my passion and together with my sister provided me with unfailing support and patience. Many thanks go to my friends for their moral support and the fun times and numerous teas (or cocktails/beers) we had. Most of all, I want to thank my boyfriend Alex, who I had many discussions with about my research and who cheered me up during tough times and celebrated with me milestones/achievements.

I am deeply grateful for the productive, welcoming, and relaxed atmosphere in the usable security group and the friendships which developed and were strengthened. Especially, I would like to thank Katharina Pfeffer for her enthusiasm and perseverance, the fun time we had while collaborating, the discussion of research ideas as well as our Schönbrunn meetings. I also want to thank Matthias Gusenbauer for challenging my points of view, the long discussions about scientific or philosophical topics, and the encouragement to join the usable security group. Furthermore, I want to express my deep gratitude to Katharina Krombholz and Adrian Dabrowski. Thank you Katharina for your enthusiasm, guidance, encouragement, and passion, which has inspired me to start and (finally) complete my PhD. And thank you Adrian for your support, both while collaborating and mentoring, especially during the last phase of my PhD.

I want to thank SBA Research for providing me with the environment and opportunity to pursue my research ideas and work with like-minded people on my thesis. I also would like to thank my colleagues at SBA Research, which have expanded and enriched my research and knowledge through insightful discussions and their feedback.

Kurzfassung

Heutzutage spielen Online-Interaktionen und vernetzte Informationstechnologien eine immer zentralere Rolle. Durch das Internet steht eine nie dagewesene Fülle an persönlichen Informationen online zur Verfügung, welche ohne Schutzmechanismen für jede(n) zugänglich wären. Um diese Informationen sowie andere Online-Interaktionen vor Privatsphäre und Sicherheitsangriffen zu schützen gibt es kryptographische Protokolle.

Eine Vielzahl von Personen kommt mit diesen kryptografischen Protokollen in Berührung - von den ExpertInnen, die diese Protokolle definieren und konzipieren, über die EntwicklerInnen, welche diese implementieren und verwalten, bis hin zu den EndnutzerInnen, die diese verwenden. All diese involvierten Personengruppen haben dabei unterschiedlich viel Erfahrung und technisches Vorwissen. Dadurch sind diese oft mit Anwendungen und komplexen Algorithmen konfrontiert, welche sie nicht (vollständig) verstehen, wodurch es zu falschen Annahmen und Sicherheitsproblemen bei der Benutzung bzw. Implementierung kommt. Daher ist es besonders wichtig, den menschlichen Faktor bei der Entwicklung von Anwendungen, die auf kryptographischen Protokollen basieren, mit einzubeziehen, um Verletzungen der Sicherheit und Privatsphäre zu verhindern, die auf mangelnde Usability zurückzuführen sind. Das Ziel dieser Dissertation ist es, zu untersuchen wie sich unterschiedliche Stakeholder kryptographische Systeme vorstellen und wie sich diese Vorstellungen auf die Sicherheit und Privatsphäre der Benutzer auswirken. Diese Erkenntnisse leisten einen Beitrag um diese komplexen Systeme benutzerfreundlicher zu gestalten und an die Bedürfnisse und Vorstellungen der TeilnehmerInnen anzupassen, um so einen sicheren und privatsphäre-wahrenden Umgang mit diesen Technologien zu ermöglichen.

Im Rahmen dieser Arbeit wurden die mentalen Modelle unterschiedlicher Personengruppen von vier Systemen untersucht, welche auf kryptographischen Protokollen basieren. Im ersten Teil haben werden die Vorstellungen von Kindern und deren Eltern über das Internet behandelt. Dabei wurde unterstrichen, dass trotz der großen Verbreitung des Internet, dieses immer noch nicht im (Schul-) Bildungssystem angekommen ist, so dass die jüngsten TeilnehmerInnen ein mangelndes Verständnis über die Möglichkeiten und vor allem die Gefahren des Internets aufweisen. Darüber hinaus wird das derzeitige Standardprotokoll für Kommunikation im Internet, HTTPS untersucht. In diesem Zusammenhang wurde eine quantitative Studie durchgeführt, um zu validieren wie sich Administratoren dieses Protokoll vorstellen und warum bzw. ob es immer noch zu Implementierungs- bzw. Wartungsschwierigkeiten kommt. Dabei konnte festgestellt werden, dass derzeitige Tools

die Implementierung und Wartung von HTTPS vereinfachen, es allerdings immer noch Herausforderungen gibt, welche zu Schwachstellen führen.

Im zweiten Teil dieser Arbeit wurden zwei Systeme untersucht, welche (teilweise) auf der Blockchain-Technologie aufbauen: i) Kryptowährungen und ii) SSI-Systeme. Beide Systeme wurden mit Hilfe von qualitative Studien untersucht um das implizite Wissen der Beteiligten über sie zu erforschen. Bei den Kryptowährungssystemen wurden EndnutzerInnen im Zuge einer qualitativen Studie befragt und herausgefunden, dass die sichere Schlüsselverwaltung eine große Herausforderung darstellt und die Benutzeroberfläche einen großen Einfluss auf die Wahrnehmung der Technologie hat. Um SSI-Systeme zu untersuchen, wurden Experten zu deren Verständnis und Erwartungen in Bezug auf diese Systeme befragt. Diese Studie legte den Grundstein für eine Definition von SSI und darauf aufbauenden Standards für diese Art von Systemen.

Die Erkenntnisse dieser Dissertation haben verdeutlicht wie wichtig es ist, zu verstehen, wie sich die unterschiedliche Personengruppen auf kryptographischen Protokollen basierende Systeme vorstellen, da falsche oder unvollständige mentale Modelle zu teils schwerwiegenden Sicherheits- und Privatsphäre-Risiken führen können. Basierend auf den Ergebnissen, wurden Richtlinien und Empfehlungen für die Gestaltung von benutzerfreundlichen kryptografischen Protokollen entwickelt.

Abstract

Nowadays, online interactions and connected information technologies play an increasingly central role. There is a plethora of personal information available via the Internet, which is accessible to anyone if no protective mechanisms are applied. In order to protect this information, as well as other online interactions, there are cryptographic protocols that prevent or reduce the risk of privacy and security intrusions (e.g., leakage of personal information or monetary losses).

A variety of people come in contact with cryptographic protocols ranging from the experts who define and conceptualize these protocols, to the developers who implement and manage them, to the end-users who use them. The stakeholders of cryptographic protocols have different levels of technical experience and knowledge. Thus, they are confronted with applications and complex algorithms that they do not (fully) understand, leading to wrong assumptions and insecure usage or implementations. Therefore, it is especially important to include the human factor in the development of cryptographic protocol-based applications in order to prevent security and privacy threats rooted in poor usability.

The goal of this thesis is to explore how different stakeholders perceive systems based on cryptographic protocols and how these perceptions affect the user's security and privacy. The findings of this thesis contribute to making these complex systems more user-friendly and adapting them to the needs and perceptions of the users in order to enable secure and privacy-preserving handling of these technologies.

This thesis investigates the mental models of different stakeholders of four systems based on cryptographic protocols. In the first part, the perceptions of children and their parents about the Internet are examined. The results underpin that despite the widespread use of the Internet, it has not yet reached the (school) education system, leading to a lack of understanding among the youngest participants about the possibilities and especially the dangers of the Internet. Furthermore, the current standard communication protocol of the Internet, Hypertext Transfer Protocol Secures (HTTPS) is examined. In this context, a quantitative study to validate how administrators envision this protocol and why or whether there are remaining implementation or maintenance difficulties was conducted. The findings of this study emphasize that current tools simplify the implementation and maintenance of HTTPS, but there are still challenges that lead to vulnerabilities.

In the second part of this thesis, the focus is laid on two systems that are (partially) built on blockchain technology: i) cryptocurrencies and ii) Self-Sovereign Identity (SSI) systems. For both systems, qualitative studies were conducted to explore the stakeholders' tacit knowledge of them. In order to investigate cryptocurrency systems, end-users were interviewed. The results emphasize that secure key management is a major challenge and that user interfaces greatly impact perceptions of the technology. The exploration of SSI systems was conducted by expert interviews, in order to examine their understanding and expectations regarding these systems. This study laid the foundation for a definition of SSI and standards for these types of systems.

The thesis findings highlight the importance of understanding how different stakeholders envision cryptographic protocol-based systems, as incorrect or incomplete mental models can lead to serious security and privacy risks. Based on these results, guidelines and recommendations have been developed for the design of usable cryptographic protocols.

Contents

CHAPTER 1

Introduction

In today's "always-online" society, all kinds of activities (e.g., banking and communication) are shifting from the analog to the digital space. Despite the enormous advantages of digitalization, security and privacy challenges need to be considered. Whistleblowers like Edward Snowden, who made mass surveillance public, as well as media reports on security flaws in messenger applications like Telegram or WhatsApp [172, 140] sparked the discussions on the value of online privacy and security.

The foundation for secure and privacy-preserving online interactions are cryptographic protocols that protect (sensitive) information from unauthorized modification, processing, destruction, and inspection. Due to the complexity of these protocols, the different stakeholders involved (e.g., administrators, developers, and end-users with different knowledge bases) are confronted with algorithms, tools, and applications that they often do not (fully) understand. Therefore, some of these stakeholders experience the usage of applications building on these complex protocols (e.g., encryption or key management) as a burden. This leads on the one hand to faulty implementations and on the other hand to insecure usage of cryptographic protocols. In order to prevent people from being overwhelmed with the complexity of those protocols and to foster safe usage, cryptographic systems should be designed to avoid security or privacy risks even when used by people with incorrect or incomplete perceptions of the systems.

The application and use of a cryptographic protocol happen at the hands of humans and due to incidents caused by human error, many security experts refer to humans as the weakest link in the security chain [192]. Therefore, the research field of usable security has become increasingly important in recent years. As Garfinkel [77] described: *"Today there is a wide consensus that systems that are not usable will inevitably suffer security failures when they are deployed into the real world. Only by simultaneously addressing both usability and security concerns will we be able to build systems that are truly secure."*

People's perceptions of a system have been shown to crucially influence their adoption of, and their interaction with it, which consequently affect their security and privacy. Therefore, the first step toward usable cryptographic protocols is to understand the requirements and mental models of the different stakeholders involved with the protocols and their corresponding systems. With this newfound understanding of the stakeholders' perceptions, the protocols and the systems can be designed to be inherently secure without the significance of a deeper understanding of the complex protocols.

In particular, I want to answer the following research questions with this thesis:

RQ1. What mental models of cryptographic protocol-based systems and their functional components do different stakeholders have?

RQ2. What are the key differences between the stakeholder's perception and the actual structure and functionality of the systems?

RQ3. Which mental models interfere with the secure and privacy-preserving usage or development of these systems?

RQ4. How could these systems be improved to be used securely despite possibly incorrect mental models?

1.1 Goal of this Thesis

The main goal of this thesis is to explore the relationship between people's perceptions of cryptographic protocols and the actual protocol's functionalities. Thereby, I want to investigate how incorrect perceptions influence the privacy-preserving and secure usage of these powerful security concepts. This thesis focuses on different stakeholder groups i.e., administrators, developers, and end-users to get novel insights into their mental representations of the cryptographic systems. With this thesis, I want to bridge the gap between the technological aspect of cryptographic protocols and the human factor.

Cryptographic protocols and systems are often perceived as unnecessarily challenging, a distraction from main tasks, or as an unknown black box for the stakeholders involved. Although the user doesn't need to understand all details, it is of utmost importance that their mental models do not interfere with privacy and security-preserving usage.

In order to cover different types of cryptographic protocols, I have chosen four systems as representatives for this thesis, which differ in their i) technological maturity, ii) the number of users, and iii) usability status (how much work has been conducted in this domain). In particular, I investigated the mental models of the following systems and protocols:

In the first part of this thesis, the mental models of the *Internet* and its current standard protocol for encrypted communication, *HTTPS* are explored. Both systems/protocols are comparably mature and provide high usability for end-users. The Internet is the

foundation for all online interactions and is now a common term that everyone knows, but a deeper understanding of its structure and potential dangers has not yet reached (most of) our education system. This is particularly problematic for today's generation of children, who come into contact with the Internet at an early age but are unable to properly assess its possibilities and, above all, its risks.
HTTPS is technologically a bit younger, and most importantly, many users do not know what this protocol does and why it is necessary for our daily (online) communication. Therefore, I decided to investigate the administrators' perceptions since they have to implement and maintain the protocol and consequently directly influence the security of the protocol. Although some usability research in this area was conducted, this thesis provides novel in-depth results discovered with the help of state-of-the-art quantitative methods, which additionally illuminate a different point of view.

In the second part of this thesis, two comparatively young systems in terms of their technical maturity and usability are investigated. Both systems are based on or fostered by a blockchain protocol. In particular, I explored *cryptocurrency systems* and *SSI*. Cryptocurrencies have become increasingly popular over the past decade, mainly due to rapid fluctuations in value and its broad media coverage. Animated by the Bitcoin hype, many people have acquired this new form of currency and lost money due to a lack of understanding of this cryptographic protocol. Therefore, I decided to take a closer look at the (lack of) understanding of end-users of this system.

SSI is the youngest of these four systems based on cryptographic protocols and due to its novelty, there is currently no application with a large user base. Thus, I decided to interview experts from the field to create a basis for standards and definitions and a common understanding of the principles.

1.2 Methodological Approaches

The research questions are inherently interdisciplinary and can only be answered by considering all disciplines involved. Therefore, in this thesis, a combination of state-of-the-art methods from security, human-computer interaction (HCI), and psychology research is used. Thereby, I followed a mixed-methods procedure for data acquisition to address the stated problems from different angles. In order to ensure scientific validity, I triangulated the findings.

For each system based on a cryptographic protocol I conducted the following four research steps:

1. *Comprehensive literature review:* I explored the current state-of-the-art of the technology and reviewed usable security research conducted in this area. Thereby, I examined both published and non-published literature as well as open-source available software. Based on this, I contextualized and systematized the findings to identify gaps in the currently existing literature.

2. *User studies:* In order to get deeper insights into how users perceive a system and how these perceptions influence the stakeholder's privacy and security, I conducted qualitative and quantitative user studies. Thereby, I used the following study types:

   - **Individual interviews** to gather qualitative insights into the mental models of the participants via semi-structured interviews. I used the interviews to build theories (Internet study) and explore the unknown problem spaces (cryptocurrency and SSI study). During the interviews, I used a combination of open- and closed-ended questions, drawing tasks, and card assignments. Due to the ongoing pandemic about half of the interviews I conducted were held online via Skype or Zoom and the other half in person.

   - **Focus Groups** to explore and discuss perceptions and hypotheses with a small group of experts. In contrast to interviews, focus groups offer the possibility of discussion amongst the participants, which encourages brainstorming and creative collaboration. In the focus group, I asked the experts open questions based on findings and assumptions from previously performed studies and let them discuss these in a moderated setting.

   - **Online Surveys** to gather quantitative data and test hypothesis. I used this quantitative method to validate the qualitative observations of HTTPS from Krombholz et al. [106] and to quantify specific challenges which administrators still have. In the online survey, I asked single and multiple-choice questions and Likert scales as well as follow-up open-ended questions in order to provide the participants with the possibility to express their thoughts and experiences in more detail.

3. *Evaluation:* To gain an understanding of the initially unknown problem spaces of cryptocurrency systems and SSI, I used a grounded theory approach. This approach allowed me to answer the research questions that are open and exploratory by nature, as it provides a set of suitable methodological steps. In comparison, I analyzed the Internet study using thematic analysis which emphasizes identifying, analyzing, and interpreting patterns of the meaning of already known problem spaces in order to get more detailed insights. The quantitative data I retrieved from the HTTPS online survey was analyzed with statistical tests and descriptive analysis, to investigate correlations, tendencies, and variations.

4. *Enhancement formulations:* Based on the findings from the previous three steps I formulated recommendations for future designs and emphasized the implications of current challenges with the perceptions of the cryptographic protocols. The recommendations lay the foundation for improving the usability of cryptographic protocols to prevent security and privacy risks for the stakeholders involved.

1.3 Scientific Contribution

This thesis contributes novel insights into mental models of i) currently highly used protocols and ii) cutting-edge cryptographic systems with the potential to shape the online realm in a more privacy-preserving direction. The insights I gathered from the studies I led (and executed/analyzed together with colleagues and students), were formulated into usable security guidelines that i) foster the formation of correct mental models and ii) enable safe usage and implementation despite wrong mental models of developers and end-users.

In order to gather insights into user ***mental models of the Internet*** I conducted semi-structured interviews with children (aged five to eight) and their parents to determine their perceptions including the Internet's perceived privacy and security landscape. Thereby, I found that children's mental models start to take shape beyond physically tangible components between the age of seven to eight years. Hence, I argue that it is important to educate children about the Internet and its security and privacy issues from an early age on. For younger children, I suggest using secure and privacy-preserving applications, as they are not yet able to grasp the bigger picture. Furthermore, I provide suggestions for visual interface cues, as children remembered them in great detail during the study.

The ***validation of challenges and mental models of HTTPS*** were examined with a large-scale study with administrators. Thereby, I investigated whether configuration problems explored in prior studies actually exist in the wild and whether the found mental models of HTTPS [106] are valid for administrators. The results of my study confirm that Let's Encrypt and ACME clients, such as Certbot, simplify configuration and maintenance for administrators, thus increasing the security of HTTPS configurations. However, there are still challenges remaining (e.g., the (deliberate) choice of weak ciphers or old versions due to compatibility reasons or the need for root privileges in the default setting for Certbot were stated as a barrier for usage), for which I provide improvement suggestions. Moreover, I extend the current body of work by examining the trust administrators put into Let's Encrypt and Certbot. I found that trust and usability issues are currently barriers to the widespread adoption of Certbot.

To understands end-users handling and understanding of cryptocurrencies, I investigated the ***mental models of cryptocurrency systems***. Thereby, I contributed a qualitative user study on user end-user perceptions of Bitcoin and Ethereum and their associated threat landscape. The results suggest that current cryptocurrency tools (e.g., wallets and exchanges) are not capable of counteracting threats caused by end-users misconceptions. Hence, users frequently fail to securely manage their private keys or assume to be anonymous when they are not. Based on the findings, I contribute actionable advice, grounded in the mental models of end-users, to improve the usability and secure usage of cryptocurrency systems.

The***mental models of Self-Sovereign Identity systems*** from an expert perspective were investigated to understand current commonalities and differences in SSI understanding. Therefore, I contributed the first qualitative expert study in this domain. The study

results highlighted 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 a broad adoption in the wild and improve privacy-preserving usage.

The research that is presented within this thesis has been published, presented, or is under submission at the following peer-reviewed conferences:

- A. Mai, L. Guelmino, K. Pfeffer, E. Weippl, K. Krombholz. **Mental Models of the Internet and its Online Risks: Children and their Parent(s)**. *In International Conference on Human-Computer Interaction. pp. 42-61. Springer (2022).* (Chapter 3)
- A. Mai, O. Schedler, E. Weippl, K. Krombholz. **Are HTTPS Configurations Still a Challenge?: Validating Theories of Administrators' Difficulties With TLS Configurations.** *In: International Conference on Human-Computer Interaction. pp. 173-193. Springer (2022)* (Chapter 4)
- A. Mai, K. Pfeffer, M. Gusenbauer, E. Weippl, K. Krombholz. **User Mental Models of Cryptocurrency Systems-A Grounded Theory Approach**. *In Sixteenth Symposium on Usable Privacy and Security (SOUPS). pp. 341-358 (2020).* (Chapter 5)
- A. Mai. **Expert Mental Models of SSI Systems and Implications for End-User Understanding - A Grounded Theory Approach** (Under Submission) (Chapter 6)

Other peer-reviewed publications of the author:

- K. Pfeffer, A. Mai, A. Dabrowski, M. Gusenbauer, P. Schindler, E. Weippl, K. Krombholz. **On the Usability of Authenticity Checks for Hardware Security Tokens**. *In 30th USENIX Security Symposium (USENIX Security 2021).*
- A. Dabrowski, K. Pfeffer, M. Reichel, A. Mai, E. Weippl, M. Franz. **Better Keep Cash in Your Boots-Hardware Wallets are the New Single Point of Failure**. *In Proceedings of the 2021 ACM CCS Workshop on Decentralized Finance and Security.*
- K. Pfeffer, A. Mai, E. Weippl, E. Rader, K. Krombholz. **Replication: Stories as Informal Lessons about Security**. *Eighteenth Symposium on Usable Privacy and Security (SOUPS) 2022.*
- M. Gusenbauer, A. Mai, K. Pfeffer, E. Weippl, K. Krombholz. **Qualitative Analysis of Barriers to the Adoption of Secure Multi-Party Computation**. (Under Submission)

1.4 Thesis Structure

The remainder of this thesis is structured as follows: In Chapter 2, a definition of mental models is provided. Furthermore, the concept and application of mental models in general and specifically in the domain of usable security are introduced. This information is required in order to understand the importance and methodological basis of the four studies (presented in Chapters 3-6) conducted in this thesis. Each user study is presented in an individual chapter allowing a more detailed discussion of the cryptographic protocol in focus, the methodology used, and the system-specific findings and recommendations in each case.

The first mental model study is introduced in Chapter 3. This study investigates how children (aged five to eight) and their parents perceive and deal with the Internet including their security and privacy awareness. In the beginning, a brief overview of the Internet and its fundamentals is given in order to provide the reader with background information to assess the found mental models. So far, research focused mainly on the threat models of "being online", while this study has a more holistic view, investigating general perceptions of the Internet in-depth. In contrast to prior studies, which were mainly conducted outside of Europe with highly-educated participants, I recruited $N = 26$ participants in Central Europe with diverse educational experiences. With a thematic analysis approach, I was able to contribute novel insights into Internet mental models and how they interfere with end-users privacy.

Following the Internet mental model study, the focus of Chapter 4 is on HTTPS, the current standard for securing online communication. In contrast to the first study, the target participants were not end-users but administrators, as they play an active part in configuring and maintaining the protocol. In the beginning, the reader is presented with a brief overview of the protocol and its configuration challenges. Despite the availability of tools to make the certificate renewal and configuration process easier (e.g., Let's Encrypt and Certbot), Internet scans show that still more than 50% of the most popular websites are poorly configured.Although a few recent studies looked at the remaining challenges for administrators in configuring HTTPS from a qualitative perspective, there is little work that produced quantitative results. Therefore, I conducted a quantitative survey with $N = 96$ experienced administrators (as opposed to a student sample, which was mostly used in prior studies) to investigate to which extent configuration problems revealed in prior studies actually exist in the wild and how administrators perceive HTTPS. Thereby, I validated previously found mental models of HTTPS [106] through thematic and statistical analysis and refined them into a mental model from an administrator perspective.

The last two mental model studies investigate cryptographic protocols which are built on or fostered by blockchain technology. Chapter 5 contains the mental model study of cryptocurrency systems. Almost ten years after the first Bitcoin transaction was performed, cryptocurrencies have gained popularity among different types of users, ranging from people who are simply curious to investors to gamblers. To provide the

reader with the necessary understanding of the blockchain and the cryptocurrencies Bitcoin and Ethereum, a brief overview is provided at the beginning of this chapter. While a detailed understanding of technical fundamentals is not essential to use a cryptocurrency system to perform basic transactions, I argue that incorrect mental models and knowledge gaps hinder users to operate such systems in the most secure and privacy-preserving manner and thereby make them susceptible to monetary loss and fraud. Therefore I contribute the first qualitative study ($N = 29$) on user mental models of cryptocurrency systems and the associated threat landscape. Using grounded theory, I revealed misconceptions affecting users' security and privacy and contributed actionable recommendations to improve the usability and secure usage of cryptocurrency systems.

In Chapter 6, the expert mental models of SSI are studied. 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. In the beginning, the current status of and background information on SSI is provided. Thereby, shortcomings such as missing definitions and standards as well as a lack of common understanding of SSI systems within the (IT) community are discussed. To investigate current commonalities and differences in SSI perceptions, I contribute the first qualitative user study ($N = 12$) on expert mental models of SSI and its associated threat landscape. The study results provide the basis for a general definition of SSI and further standards for such systems. Based on the grounded theory approach of conducting the expert interviews, a minimal knowledge map for (potential) SSI end-users was constructed.

Following the four studies, Chapter 7 discusses the overarching results of the mental model studies, both from a technical perspective (i.e., how to enable usable cryptographic protocols) and a methodological perspective (i.e. lessons learned for future mental model studies). Furthermore, the research questions mentioned at the beginning of this chapter, are answered to highlight the importance of the research impact of this thesis. Finally, Chapter 8 concludes this thesis, by briefly summarizing the conducted studies and highlighting the main findings of this thesis. Last but not least, an outlook is given for future research paths in order to improve (other) current and future systems based on cryptographic protocols.

CHAPTER 2

Fundamentals of Mental Models

This chapter describes background information about mental models and provides an overview of mental model studies in general and in the area of usable security.

> *"The mental image of the world around you which you carry in your head is a model. One does not have a city or a government or a country in his head. He/she has only selected concepts and relationships which he/she uses to represent the real system."*
> (Forrester [68])

The pioneer of mental models is considered to be the psychologist Kenneth Craik, who described them as a "small scale model" of the world and how it works, which is carried by each individual in their mind [45]. The idea of mental models was further developed and supported by studies and research from Johnson-Laird [94, 92] and Kearney et al. [100]. Within their studies, mental models have been shown to influence peoples' actions and behavior crucially. Johnson-Laird et. al expanded the scope of mental models to "psychological representations of real, hypothetical, or imaginary situations" [93]. Although there is not one general description of mental models, most researchers utilize mental models as simplified and often implicit internal representations of objects and/or processes, formed by humans in order to reason, form explanations, and anticipate events.

Besides mental models, some (usable security) researchers use the notion of folk models–a term used in anthropology–in order to refer to non-expert users (i.e., no experts, as the individuals were not trained formally in the researched area). In this thesis, I only use the term "mental models" as the conducted studies both included experts and non-experts. For each study, I explicitly state the level of expertise of the participants.

A mental model is constructed based on the environment surrounding an individual, their experiences, perceptions, and understandings of it. Each new piece of information

provided to an individual is evaluated, filtered, and classified. Based on these assessments appropriate actions (i.e decisions) are taken [94]. Appropriate actions might include, ignoring or dismissing the new information, adding a mental model, or adjusting an existing mental model. It is important to note, that mental models do not accurately or completely represent the real world, but rather abstractly describe it and are constantly evolving [95]. Mental models can be sparse or detailed, incorrect or incomplete, and can either represent only one individual understanding or can be shared among different people/groups/teams. The mental models can be either explicit or tacit, whereby the latter is harder to investigate.

Mental models help to understand how individuals think while performing different activities in order to reach their desired goals. With the help of this information, it is easier to understand the process and reasons behind specific behavior and actions. Among the main motivations for conducting mental model research in the area of natural science are:

- To understand problems (e.g., design or usability) and behavior of people
- To introduce new systems/processes
- To adapt and improve systems/processes
- To communicate risks and optimize the information process

The area of application of mental models ranges from concrete applications such as the use of thermostats [101] to the understanding of how people perceive and use music [32] or a language [152]. In the context of this thesis, the focus lies on studies in the realm of HCI and usable security, therefore, the following mental model types, study methodologies, and examples of user studies are addressing this area of research.

2.1 Types of Mental Models

There are various types of mental models described in the literature. In the following, three general types of mental models are described which are used in HCI research [33].

Surrogates: This type of mental model perfectly mimics a process's or system's input and output behavior. Therefore the assumptions of the mental model holders about the system's output are always in accordance with the system or process, however, the holder is not able to explain the system's internal functionality. Young [204] raised the question of whether such a model is possible in reality as it is very difficult to hold a surrogate without understanding its functionality even for simple systems such as a calculator.

Metaphors: This type of model sets a direct equivalence between a system or process familiar to the mental model holder and the target system or process. Thereby the notations used to describe the system or process are often influenced by the domain of

the analogical part. A typical metaphor often described in the literature is the typewriter analogy to the text editor systems [52].

Glass Box: This type of model is a combination of the aforementioned two model types. Thereby, the mental model holder uses a surrogate for the input and output of the system or process and argues about its functionality and internal structure with metaphors [53]. Those models are often used in a prescriptive context (i.e., mental model holders were introduced or taught to envision the systems or processes in a certain way–by design) and less often in a descriptive one.

2.2 Research Methods

In general, there are two approaches to investigating and identifying the mental models of users. The three methods are: direct and indirect elicitation [95, 114].

Direct elicitation: Thereby the participants are asked to actively define and form a mental model. The interviewer is allowed to assist with visualization tools, words (e.g., buzzwords or descriptions to help people remember or explain specific functionalities or parts of the system or process), or card assignment tasks. This explicit articulation and visualization of knowledge provide the possibility to be directly validated by the participant, however, for the participants, it can be difficult to visualize and articulate the process and system components.

Indirect elicitation: This type of elicitation includes a content analysis of the observation of tasks. In the former method, the researcher conducts a systematic analysis of either written or verbal statements. An advantage of this method is that it can be performed asynchronously with the participant, however on the downside, the researcher can not (directly) verify the mental model or ask further questions. The latter method is used for mental models of processes and tasks which are important to investigate without interruption.

In order to ensure higher mental model accuracy often the indirect elicitation methods are combined with direct methods and vice versa. Therefore, in this thesis, both elicitation methods are used in combination with all interview studies (Internet mental models - Chapter 3; Cryptocurrency systems mental models - Chapter 5; SSI mental models - Chapter 6), and the indirect elicitation was used for the quantitative HTTPS study in Chapter 4.

2.3 Mental Models Studies in Human-Computer Interaction

Donald Norman [141, 142, 143] was among the first researchers who pursued the idea of using mental models in the design processes in the domains of HCI. The mental model studies presented in this section represent only a small fraction of the complete body of research conducted in this area. In order to give an overview of some research, only

publications from ACM CHI Conference on Human Factors in Computing Systems are presented, as this conference has a wide influence and is one of the best-rated conferences in this field.

In the area of engineering and programming, Horvath et al. [87] investigated the mental models' users have of application programming interfaces (APIs) in order to improve their usability. The mental models of reverse engineering processes were investigated by Votipka et al. [184] in order to improve program analysis tools. Brackenbury et al. [23] investigated how participants interpret and handle bugs with the help of a Trigger-Action Programming (TAP) system.

To understand the perception and expectations of users from conversational agents like Siri or Alexa, Luger and Sellen [125] conducted a mental model study with current users. Schirra et al. [166] investigated the mental models of mobile applications through a sketch study. Thereby the sketches highlighted the most central elements for the user and the misunderstandings (e.g., non-existing features or misinterpreted design).

The Internet of Things (IoT) describes the communication between physical objects and systems over the Internet. Thereby it is important to resolve the problems of feature interaction in order to ensure the usability and functionality of the systems and objects. To understand current problems and misunderstandings of feature interaction in IoT, Yarosh and Zave [202] investigated the user's mental modes about them. In 2018 Cho [35] published his research investigating how people interact with the virtual assistance of Google Home. He found that people tend to speak with the system in a conversational manner and are confused by the inability of the system to answer when it did not understand the command or was not able to perform it. Another study investigating the mental models of IoT home assistants was conducted by William et al. [194].

Over the last couple of years, there were a great boom of Artificial Intelligence (AI) applications and therefore the research interest in this area also increased. In 2007 Tullio et al. [182] investigated how non-technical users interact with and perceive an intelligent system. Thereby, they conducted a six-week field study t see how and if mental models are changing over time. Kulesza et al. [109] conducted a study investigating mental models of a personalizable intelligent agent of a music recommender system. They found that it is important to help users understand the basic reasoning of a system in order to create the desired output through personalization. How language influences the perception of AI and its usage in a virtual teaching assistant, was investigated by Wang et al. [188]. Bansal et al. [17] examined the effects of AI explanations on the team performance and found mental models on when participants trust an AI.

In 2020 Gero et al. [79] conducted a mental model study on how people develop mental models of AI, with the help of a cooperative game setting. Another study investigating the mental models of AI, with a focus on AI players, was executed by Villareale et al. [183]. They used the drawing game *iNNk* for their study and based on their findings proposed a framework for player mental model development.

AI in the realm of health care was investigated by Okolo et al.[145] as a possibility to

increase the healthcare delivery in rural India. They investigated community health workers' perceptions of an AI application with the purpose of automatically diagnosing diseases.

2.4 Usable Security Mental Models Studies

Usable security is the connection between the research domains of the supposedly antagonistic areas of security and usability. It uses many methodologies from HCI in order to improve both the security and usability of the stakeholders involved and mental models are no exception. Therefore, in the following Section, a brief overview of mental model studies in the area of usable security is provided.

In order to provide the user with effective (security and privacy) systems they need to ensure their usability either with (usability) standards for the technology and the user interfaces or user-adjusted risk communication [191, 133, 120, 15]. Therefore it is important to understand the mental models of the stakeholder involved in the systems and processes in order to improve them. The presented studies in this chapter provide an overview of usable security studies which are important in the field, however, are not directly related to the cryptographic protocols investigated in this thesis. Related studies to the corresponding protocols are introduced in the respective study Chapters (3-6).

In 2010 Wash [189] identified in his user study eight different folk models of security threats to 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. Wash et al. [190] stated that instead of attempting to force users into more 'correct' mental models, technology should be shaped to work well with existing mental models. Furthermore, Wash et al. found that some wrong mental models can still lead to good security behaviors (e.g., wrong assumptions about attackers can lead to increased caution which is beneficial for security and privacy).

Bravo-Lille et al. [25] investigated in their mental model interview study, the perceptions and responses of users to security alerts, as users tend to ignore security warnings. They highlighted with their study effective ways to convey security information to users. Renaud et al. [162] 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 email architecture in general as well as related usability issues led to the refusal of end-to-end encrypted emails. In 2015, Kang et al. [96] conducted a user study that focused on the mental models of Internet users concerning their privacy and security. As a result, they proposed different systems and policies which do not presume technical knowledge to improve the safe handling of the Internet for users.

In 2016 Alaqra et al. [9] conducted a case study regarding the requirements and the general perception of privacy- and security-maintaining services of users. Their focus was

specifically on cloud services and their transparency. Naiakshina et al. [135] conducted a study on mental models of mobile messaging tools with a special focus on Short Message Service (SMS) and WhatsApp. The participants consisted of three groups: students with computer science degrees, students with non-technical degrees, and non-academic people. They investigated a general basic mental model for most users and found that most of them are skeptical towards the security of mobile messaging.

Gallagher et al. [73] published mental models about the onion routing system Tor from 17 participants including experts and non-experts. They found severe gaps in the participants' knowledge which could potentially lead to de-anonymization. Another study conducted by Zeng et al. [206] investigated smart home technologies and revealed fundamental mismatches between users' threat models and reality.

Yao et al. [201] 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 was conducted by Perry and Ferreira [155]. Oates et al. [144] proposed exploring mental models through illustrations and conducted a mental model study on users' privacy perception. Wu and Zappala [196] 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. Another study investigating encrypted communication was performed by Abu Salma et al. [3]. They quantified mental models and found that end-users often underestimate the security benefits of end-to-end encryption.

In 2019 Fulton et al. [72] published their findings on a mental model study investigating the effects of fictional television and film on the perceptions of computer security. Binkhorst et al. [20] examined the mental models of corporate VPNs. Thereby, they found that the participant had a high-level technical understanding, however, the deeper understanding especially about the security of VPNs diverged from the actual functioning.

CHAPTER 3

Mental Models of the Internet

Disclaimer: *The contents of this Chapter were published as part of the publication "Mental Models of the Internet and its Online Risks: Children and their Parent(s)." In: International Conference on Human-Computer Interaction 2022 [127]*

The global Internet population continues to grow daily and already has more than 4.9 billion active users [91]. Today's children grow up in a digital world. In contrast to their parents, they do not know a world without the Internet. From a very early age, many of them experience video streaming on their parents' smartphones/tablets/smart TVs, digitized toys, or computer/console games [85]. Despite the supposedly low level of direct interaction that children have with these devices, their privacy and security can be violated. For instance, incidents were reported where (default) settings allowed toys to unintentionally record children [132].

Using the Internet usually requires little technical knowledge about its components and how they interact with each other. However, in order to correctly assess one's own security and privacy on the Internet when using common applications (e.g., browsers, messaging applications) and behave accordingly, it is important to have at least a basic understanding of the system. This becomes especially crucial when security breaches or other problems occur (e.g., leaked passwords [70] or identity theft).

Multiple studies investigated Internet mental models of teenagers and adults, however, less research was done in the area of young children and their respective parents. Livingstone [123] found that children value their privacy both in offline and online communication with their peers, when seeking advice and when forming relationships. Therefore, it is important to respect and protect their desire for privacy. This study takes a step in this direction, by examining children's mental model of the Internet. Based on the findings, this Chapter discusses and provides advice on how to enhance i) privacy and security settings and ii) raise children's awareness.

The purpose of this study was to investigate the Internet mental models of children and their parents in Central Europe, in order to validate and extend prior studies by Kumar et al. [110] and Zhang-Kennedy et al. [208]. To deepen previous findings, this study used drawing tasks and scenarios to elicit mental models of the Internet, which are often based on tacit knowledge. This hard-to-express form of knowledge is generated by individual experiences and assumptions, which strongly influence behavior [100].

One of the main goals was to examine whether the education of the parents influences the children's view and awareness of the Internet and its risks. Thereby, the study particularly focused on aspects, that potentially interfere with the secure and privacy-preserving Internet usage of children. Another aim was to understand the European context, as prior work was conducted mostly in countries outside the European Union (e.g., Australia and the U.S.).

This study sought to answer RQ1-RQ4 in relation to the Internet and its stakeholders the end-users (both children and their parents). In addition to RQ1, this study investigated whether the mental models of children and their respective parent(s) are similar (RQ1a).

To answer the research questions, 26 semi-structured interviews were conducted with 13 families in the metropolitan and suburban areas of Austria. The methodology used in this study followed an iterative approach to conduct qualitative interviews and thematically analyze them until theoretical saturation [82] was reached.

The findings suggest that children's mental models of the Internet deepen between the ages of five and eight. In this process, they switch from simple, physically tangible descriptions (e.g., a TV or smartphone), to elaborations that go beyond them (e.g., describing that the Internet represents knowledge). The insights gathered during the study suggest that risk awareness also depends on the children's age. The reason for this is, on the one hand, that their parents shield and protect them from risks. On the other hand, the superficiality of their mental models is also a barrier to developing risk perception. Furthermore, the study underpinned that parents were more aware of security and privacy risks than children. However, most parents were not concerned about their own privacy, but more about their child's safety on the Internet. These concerns, however, restricted the privacy and freedom of children in some cases.

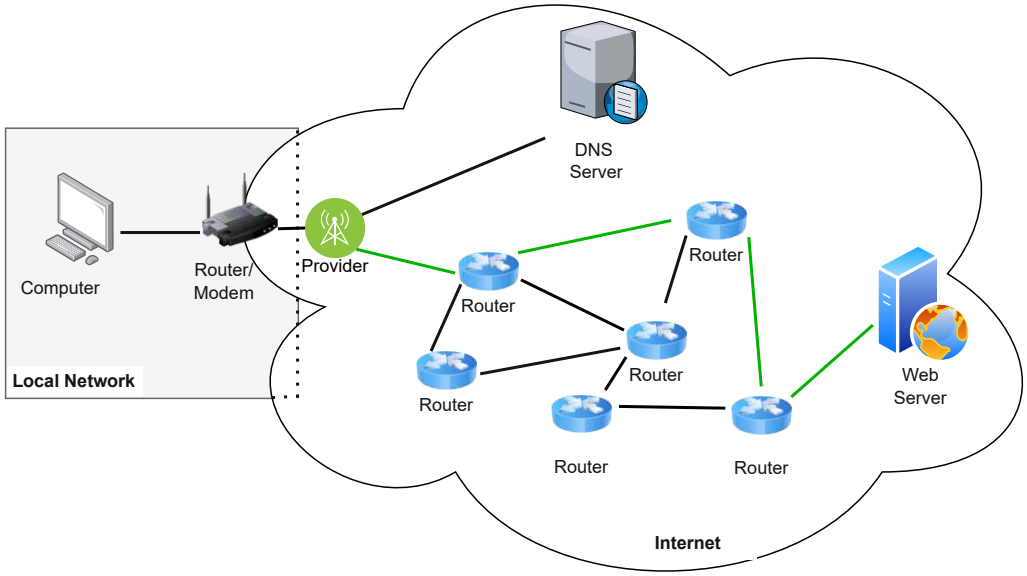


Figure 3.1: Simplified overview of contacting a website on the Internet

3.1 Background Information on the Internet

The Internet (**inter**connected **net**works) is a global network of billions of electronic devices (e.g., computers, mobile phones, and TVs). In order to *be online* an end-user simply must connect with an electronic device to the Internet. Due to various usability improvements, the processes of being online seem to be very simple to end-users, although a very complex system lies beneath the surface.

The fundamentals of the Internet were developed in the 1960s in the United States and the first websites were connected in the early 1970s. However, the broad usage and visibility of the Internet to the general public only started 20 years later [26]. The rapid spread of the "network of all networks" ensures that today more than 60% of the world's population already has access to the Internet [91]. The Internet consists of a network of physical cables and hardware, which are spread and connected all over the world. Such cables include telephone wires, fiber optic cables as well as TV cables, which are installed both on land and in water.

The following paragraph describes the Internet in a very simplified form which is the basis for the Internet mental model study. A visualization of the communication processes and connections of the Internet can be found in Figure 3.1. In order to visit a website, a computer needs to be connected to a modem/router either via LAN cable or WLAN. The modem/router assigns the computer an IP address. To access an Internet page, a request is sent via the modem which establishes a connection with the Internet Service Provider (ISP) e.g., Telekom or T-Mobile. The ISP looks up the IP address of the server on which the requested website is stored on the Domain Name Server (DNS) and sends a

request. The communication/data exchange between the various devices (e.g., servers and routers) works via standardized network protocols, i.e. these are independent of the hardware and the underlying operating systems. The most important protocols are: i) the Internet Protocol (IP), ii) the Transmission Control Protocol (TCP), iii) the User Datagram Protocol (UDP) and iv) the Internet Control Message Protocol (ICMP). In order to view the searched web page, its content is divided into individual data packets at the server and pieced together at the client with the help of these protocols. The data packets are then sent via various routers to reach their destination (i.e., the IP address of the requesting computer) as quickly as possible.

The Internet and its everyday usage emphasize the importance of this technology. It consists of many applications and therefore was investigated in various usable security studies, highlighting its potential and pitfalls [191, 192, 16]. In the following Section, the scientific context of this study is outlined in the realm of usable security, specifically of mental model studies.

3.2 Related Work on Internet User Studies

This Section examines related studies on Internet mental models as well as threat models of Internet-related actions/devices from a children's and an adolescent's perspective. An overview of related studies can be found in Figure 3.2, indicating the age and number of the participants, as well as the type of study (i.e. qualitative or quantitative study) and the year of publication.

3.2.1 Internet Mental Models

Thatcher and Greyling [178] conducted the first study about mental models of Internet users in 1998 with university students in South Africa. They created a questionnaire including a drawing task to explore their participants' understanding of the structure and functionality of the Internet. A methodologically similar study investigated the mental models of primary and high school students [151] in Greece.

Yan investigated in three large-scale studies [198, 199, 200] the effects of age and experience on the technical and social understanding children and adolescents have of the Internet. Hereby, he found that children have limited resources for their understanding of the Internet, which has implications on their Internet experience/usage and their perception of Internet-related devices (e.g., PC, telephones, TV). Diethelm et al.[51] conducted a study with secondary school students and found metaphors for Internet education at that age.

In 2017, Kumar et al. [110] conducted an Internet mental model study with highly educated U.S. families (children aged five to eleven). The interviews were semi-structured, with hypothetical scenarios to make it easier for the children to imagine potential risks. They found that children have some strategies to avoid certain security or privacy risks, however, they still heavily relied on parental support. The study with the smallest

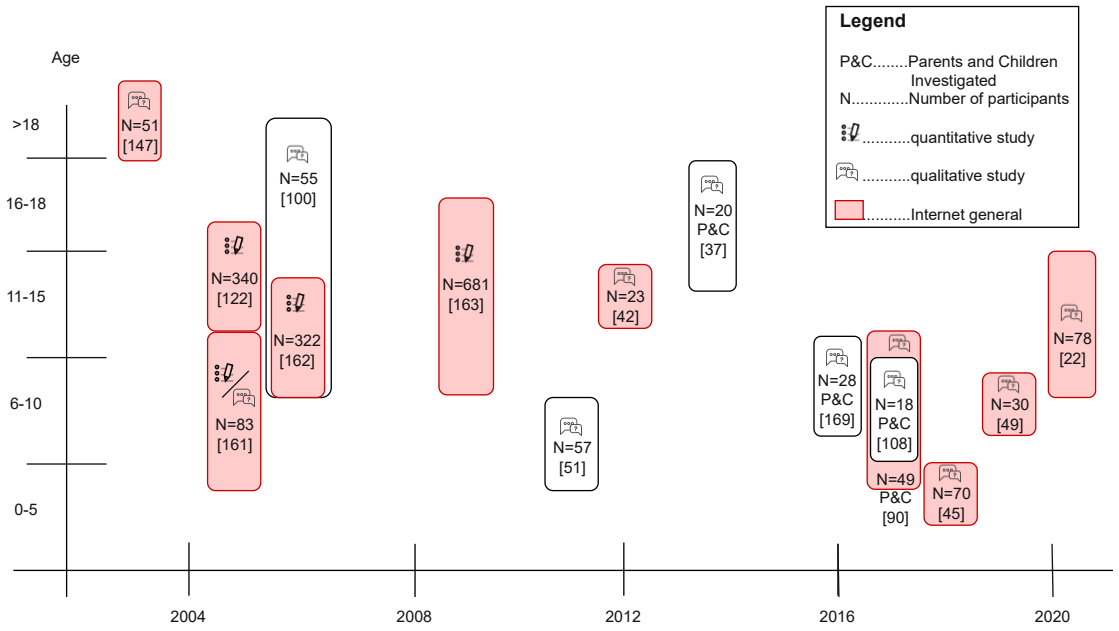


Figure 3.2: Overview of relevant children/adolescents mental model studies on the Internet and online threats/privacy perceptions

children (4–5) was conducted by Edwards et al. [56] to identify everyday concepts which can be used for Internet education. Eskelä-Haapanen and Kili [59] focused their study on the trustworthiness of information that can be obtained from the Internet. The most recent study was presented in 2021 by Brodsky et al. [28] which focused on Generation Z and the influential factors of age and usage. Thereby, they found no significant differences, however, determined that most participants described the Internet as ubiquity.

3.2.2 Privacy and Security Perceptions of Internet Related Actions/Devices

Besides mental model studies of the Internet also Internet-related actions and devices are relevant in usable security, to ensure safe and secure usage of the technology. Livingstone [123] gave insights into children's understanding and interpretation of privacy. Ey and Cupit [61] conducted group interviews with children to investigate their strategies when facing online threats. Their results showed that the children were able to handle dangerous situations appropriately if they received Internet-related education, otherwise they had difficulties identifying the potential dangers.

Cranor et al. [46] explored the boundaries of teens' privacy when using online services in comparison to their parents. They revealed that parents and their teenage children have different privacy concepts, as parents made incorrect analogies between the real and the digital world.

Zhang-Kennedy et al. [208] studied the perceived mobile threat models of parents and their children (aged seven to eleven) in Canada. They identified different threat models and found that parents perform different protection strategies to protect their children from exposure.

McReynolds et al. [132] uncovered both, children's and parents' mental models of online toys and their expectations. They found that children were not aware of their toys' connection to the Internet. Most of the parents had privacy concerns about the toys recording their child(ren), however, there were also some who approved of the possibility to oversee their children's activities.

The study presented in this Chapter differs in comparison to the aforementioned studies primarily in that it assessed the general perception of the Internet and its components and actors from both children and their parents. Based on those perceptions, further insights were gathered about the connection between mental models and experienced risks, and the prevention mechanisms used by participants. To further extend the current base of knowledge, this study also studied different levels of educational background and recruited participants within a different cultural background (Central Europe), which has so far been understudied.

3.3 Methodology of the Internet Mental Model Study

In the following, the methodology used for the Internet mental model study of children and their parents is described. First, the ethical considerations (Section 3.3.1 are explained, followed by the recruiting process of the participants (Section 3.3.2. Afterward, the study design and procedure including details about the pilot study and the interview procedure are presented. Last but not least, the data analysis is described in Section 3.3.4.

3.3.1 Ethical Consideration

The study followed the organizational guidelines of SBA Research, as the research center, unfortunately, has no separate institutional review board. The guidelines require preserving the participants' privacy by limiting the amount of sensitive data that is collected to a minimum. Before conducting the interviews, the purpose of the study was explained and the parents were asked to sign (for in-person interviews) or orally consent (for online interviews) to the study's data handling procedure. The data handling procedure strictly follows the EU's General Data Protection Regulation (GDPR) [43].

Some of the study participants were children. It was particularly important to treat them with respect and give them the feeling that they are in a protected environment. Therefore, the study design was adapted as much as possible to their needs. The children did not have to read anything as the study was designed with visual material and all the questions were asked verbally.

Table 3.1: Demographics of the study participants (26 interviews from 13 families)

Demographics		Parents	Children
Age	Min–Max	35–44	5–8
	Median	40	6.5
Gender	M	30%	46%
	W	70%	54%
Highest Education	Pre-School	-	40%
	Primary School	-	60%
	A-Level/Apprenticeship	50%	-
	Graduate education	50%	-

3.3.2 Child-Parent Pair Recruitment

It was the goal to recruit a diverse sample in regard to the participants' educational background and gender. Potential participants were approached via social media as well as local pre-schools, following prior studies [61, 46]. Furthermore, personal contacts and the referral principle was used to reach new participants. A short recruitment description of the study was distributed without disclosing the concrete purpose of the study. To describe the research project, the wording *digital media usage today* was used to not reveal too much information about the study beforehand and thus, prevent participants from reading up information that could distort the study results. The families were chosen based on the parent's educational background, in order to ensure a diverse sample.

The selection of participants was performed in three rounds. In the beginning, five families (10 interviews) were recruited focusing on parents without a university degree and children between 5-8 years. The decision to conduct the study with children in this age range was made, as other studies did not investigate this age group in depth and it is the time when children start to use the Internet (sometimes even on their own) and are able to explain their thoughts and perceptions. The interview data was explored through an initial open coding process. Based on those findings, the recruiting strategy was extended to higher educated participants in order to broaden the insights by including a more diverse sample (10 interviews). In the third round, six additional interviews were collected from three families, with one seven-year and two eight-year-old children, as an explanation shift of the components of the Internet was observable (from physical tangible components to more technical and intangible awareness) between seven- and eight-year-old children which needed further analyzing. The last two interviews did not bring any new insights, meaning that theoretical saturation was reached. Therefore, the interviews were stopped and an in-depth analysis was conducted.

Hence, the study's total sample was 26 participants from 13 families. The demographics of the final dataset can be found in Table 3.1, indicating the age, gender, and highest completed education distribution of the participants.

3.3.3 Study Design and Procedure

The study consisted of a semi-structured interview guideline and a short pre-study questionnaire (see Appendix A.1.1) containing closed-ended questions. The pre-study questionnaire covered the participants' demographics and their Internet usage. The interview guideline was used to understand the mental models of children and their parents about the Internet and its related threat landscape. In line with other mental model studies from usable security [206, 106, 128] and HCI [110, 132] a drawing exercise was used besides a verbal interview to elicit the mental models. The interview consisted of three drawing scenarios, which were guided by some questions to support the participants. The interview guidelines for parents and children differed only marginally, as the latter only used child-friendly language. The study materials used, such as pens and pictures, were also the same for both groups.

The drawing tasks covered three main themes: i) General Internet functionality, ii) specific types of activities the child performs on the Internet (in this study: watching a video), and iii) privacy and security concerns. The drawing tasks were used to help the participants visualize and organize their thoughts. To avoid misinterpretations of the drawings, the participants were asked to think aloud and explain their pictures while drawing. Based on the verbal explanation and the drawings, comprehensive insights into the mental models of the participants were gathered.

The opening question of the semi-structured interview was *"Do you know the (term) Internet?"*. In the *first scenario* the participants were asked to describe the Internet's functionality and its main components and actors. Depending on the participants' answers to this question, follow-up questions were asked concerning the components and actors which they mentioned, and about connections between them. To conclude the first scenario, the participants were asked about security and privacy risks they could think of.

The *second scenario* covered the setting of watching and streaming videos on YouTube. To facilitate immersion into the scenario, a picture of the YouTube Kids page was provided displaying two thumbnails of different videos (see Figure 3.3). After providing the picture, the participants were asked what they believe to happen when searching for and clicking on the displayed video. If participants encountered a mental block or did not mention where the video comes from and where it is stored, the interviewer asked specifically about this. To lead the participants to the security and privacy aspects of the Internet, the interviewer asked them about their opinion of who can view the video(s) and who knows if someone watched a specific video.

The *third scenario* investigated instant messaging. Similar to the second scenario, a picture was shown to accentuate the scenario. This time a screenshot from a message

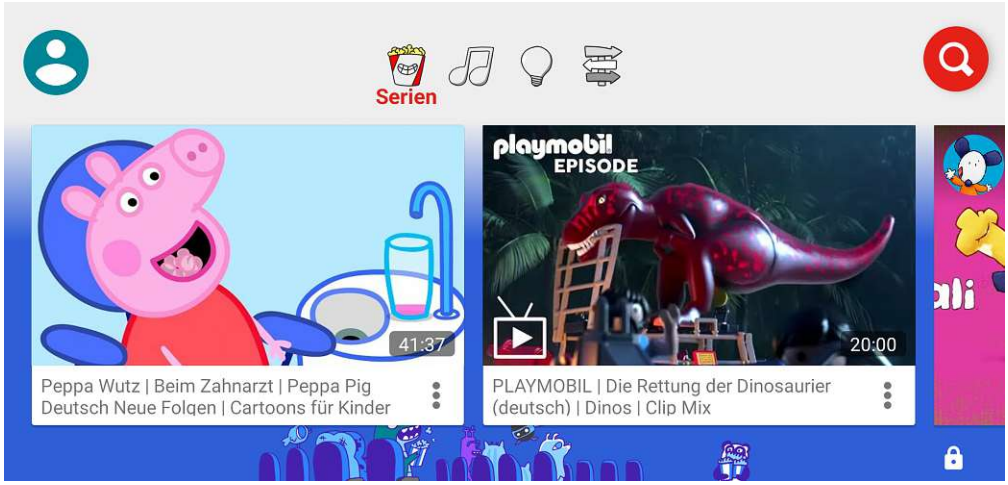


Figure 3.3: Picture of YouTube Kids for Scenario 2

exchange between the grandmother and the participant in a messaging app (see Figure 3.4) was shown. To not exclude children who were not yet able to read, the picture provided, both, for parents and children, only depicted emojis in the messages. The participant was asked to imagine sending a message to his/her relative. Based on this scenario, they were asked to draw the communication between them and their messaging partner. For assistance and to gather more in-depth information, the participants were asked about their thoughts on the connections, potential message loss, and the people who can read them.

During the study, the children were not directly asked about security or privacy threats, since it can be assumed based on prior studies [195, 208] that they are not familiar with such terms. Instead, they were given hints by asking questions about mean or scary people or inappropriate behavior they experienced. Thereby, the vocabulary used by the children in the pilot study was reused, to ensure familiarity and to avoid interviewer bias.

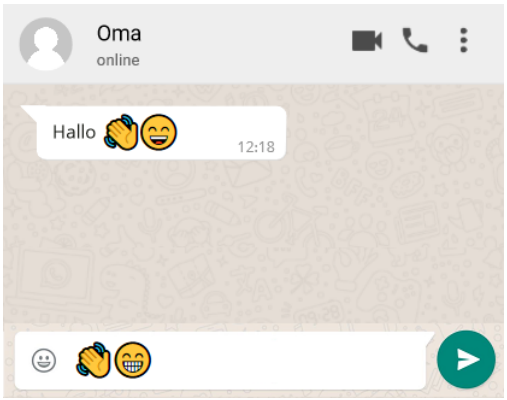


Figure 3.4: Picture of a messaging app for Scenario 3

Pilot study

Before conducting the study, the interview design was tested by interviewing two families (four interviews) as a pilot study. Thereby, two improvements of the initial study design were found. Each scenario of the initial design consisted of two to three questions, without any further prescribed guidelines and inspirational cues for the participants. Furthermore, only a single pen was provided for the drawing task, which seemed to hinder the creativity of several participants.

In the pilot study, the participants appeared to be overwhelmed by the vague questions and no further guidance. During the first interview when the interviewer asked more specific follow-up questions, the participants felt more comfortable and expressed their thoughts more easily. Therefore, the initial set of questions was extended with four concrete questions for each scenario. Although the study design had been changed, by adding questions to the study design, I decided to integrate two interviews into the final set of data, as the interviewer asked those questions during the pre-study.

After the pilot study, it became apparent that providing additional coloring material would be beneficial. This was especially true for the children, as they felt slightly irritated by the fact that there was only one choice of color. In order to encourage the participants' creativity, more colors were provided for the main study. To better distinguish between the general mental model and the threat model, the color red was only given when the participant started to elaborate on security and privacy risks.

Interview Procedure

This study consists of 26 participants from 13 different families in the metropolitan and suburban areas of Austria. 12 interviews were conducted in person and 14 via Skype, due to the ongoing Covid-19 pandemic at the time the study was conducted. The in-person interviews were performed at the participants' homes to ensure a comfortable setting for the children. All participants agreed to be recorded, enabling the interviewer to transcribe them afterward. Furthermore, pictures of the drawings were taken for the data analysis.

In the beginning, the parent-child pairs were briefed about the procedure, asked for their consent, and afterward, the interviews were conducted with parents and children separately. However, the parents were asked to be nearby when interviewing their children to make the children feel more at ease. This was targeted at enabling the children to speak more freely, but still have the possibility to engage with their parents, if they felt uncomfortable. First, the children were interviewed, and then their parents, to prevent children from picking up words or content from their parents during the interview, and simply repeating them. One study was an exception as it was conducted the other way around, however, the child did not stay in the same room to prevent any bias. The average interview lasted 25 minutes, whereby the interviews with children were shorter (an average of 15 minutes) than those with parents (an average of 30 minutes).

As compensation, the parent-child pairs received a €10 gift voucher at the end of the interview.

3.3.4 Data Analysis

All interviews were recorded and transcribed afterward. Based on these transcripts, two researchers followed a thematic analysis [39] approach and coded and searched for re-occurring themes in two different rounds. During this process, the researchers (author and student) discussed their findings and codes until a final agreement on the codebook was reached (inter-rater reliability Cohen's Kappa [40] $\kappa = 0.72$).

The codes are related to the Internet itself (including its actors, components, and their connections), the risks the participants faced (e.g., inappropriate content or unauthorized access to smartphones), and prevention strategies to avoid these (e.g., YouTube Kids or other filter mechanisms). The high-level themes of the codebook are shown in Figure 3.5 grouping the codes into the three categories: i) Internet, ii) Risks, and iii) Meta.

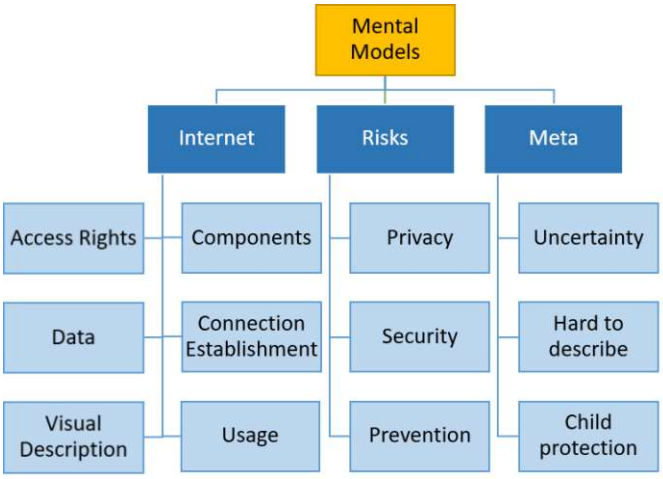


Figure 3.5: Overview of main themes of the codebook

3.3.5 Limitations of the Study

The study was conducted with a relatively small sample and limited geographic diversity, therefore the study provides some qualitative insights, which need to be further investigated in a larger study. The recruiting area was limited to one country as the first interviews were conducted in person. Furthermore, the sample is biased towards female parents (70% of the adult participants). Naturally, the methodology has its limitations, as the data is self-reported and, in comparison to quantitative studies, the results cannot be used for statistical analysis. Finally, the presence of some parents during their children's interviews, when conducted online, may have had an impact on their responses. However, I am confident that the children's comfort to answer questions honestly outweighs the potential parental bias, which is why I opted for this setting.

3.4 Findings of the Internet Mental Model Study

In the following, the findings related to the participants' perceptions of the Internet and security and privacy risks while performing online actions are described. Thereby, the qualitative insights and tendencies encountered during the study alongside with the participants' drawings and direct quotes are presented for illustration and deeper insights.

To better understand the participants' knowledge and perceptions, they were asked about their general Internet usage, their privacy concerns, and technical understanding/skill. An overview of their answers can be found in Table 3.2. Thereby, it was revealed that the majority of the parent participants used at least four Internet services regularly. In contrast, children used fewer online services. Only in the case of watching videos online, the children showed more activity than their parents. In Table 3.2, some fields are marked as "-" not applicable in the column of the children. In particular, no direct questions about privacy concerns were asked but instead a more children-friendly wording was used. For email and banking services, it was assumed that children do not use them on their own or don't possess one [50].

Both children and parents claimed to have an average knowledge of online devices and the Internet. However, only four participants learned specifically about the Internet during their education. Therefore, the knowledge was either self-taught or taught by reference persons (for children their caregivers), as indicated by the participants' statements about asking for help from others.

3.4.1 Internet - Mental Models

To explore the participants' understanding of the Internet and thereby answer RQ1 and RQ2, open-ended questions were formulated as the basic framework of the interview. Furthermore, the participants were asked to draw simultaneously to their explanations. Within the first question(s) it became apparent that four children (C4, C7, C8, C9) between five and six were not familiar with the term 'Internet', which made it difficult for them to directly talk about it. However, based on the two scenarios, the interviewer discovered that they did use some of the services on the Internet, but did not associate the word with them. For example, one five-year-old child stated, that they are still able to watch videos on YouTube without the Internet (C8).

In comparison, seven and eight-year-old children did link their explanation of the Internet to their individual activities, such as playing games and watching videos. One child stated without further direct explanation:

> *"That it [the Internet] is not good for children."* (C3)

Based on the remainder of the interview it became apparent, that C3 had a bad experience with a YouTube video that showed content that was not suitable for children. One eight-year-old child articulated a quite concrete perception of the Internet:

Table 3.2: Participants' Internet usage and knowledge based on Q6-Q8 from the pre-study questionnaire

		Parents		Children	
		Average	Median	Average	Median
Internet Services	Social Media	3.1	2.5	1.8	1
	Videos	3.8	4	4.2	4.5
	Online Shopping	3.5	4	1.9	1.5
	Instant messaging	4.9	5	1.7	1
	Video chatting	2.4	2	2.8	1
	Email	4.5	5	-	-
	Online Banking	3.6	3.5	-	-
	Games	-	-	2.4	1
Privacy Concerned[1]		3.5	3.5	-	-
Self-Assessed Skills	Knowledge smartphone/tablets	3.2	3	3.2	3
	Internet education	2.4	2	1.2	1
	Asking help from others	3.1	3	4.2	4.5
	Participant helps others	2.4	2	-	-

[1] Q7 from pre-study questionnaire; only parents were asked directly (see Appendix A.1.1)

> *"I imagine the Internet as a world, where I can do things, ask questions, and so on."* (C1)

Another eight-year-old child stated that:

> *"It [the Internet] is everywhere.... and everyone with a Laptop or a smartphone can use it."* (C12)

C1, C12 explained and drew the Internet as an earth where many (unknown) people are connected to each other and therefore, had a very similar mental model to three adults.

Another popular depiction of the Internet among the children participants was power poles or boxes which connect the user to the Internet. Three children (ages seven and eight) used those visual representations. One child explained while drawing their picture (see Figure 3.6):

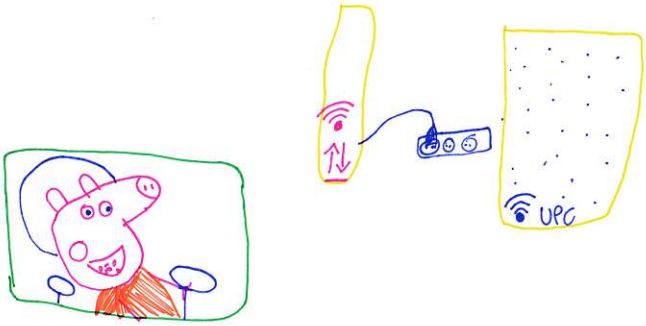


Figure 3.6: Internet depiction of participant C6 with a (Internet) box.

> *"[The Internet] is back there by the Internet box. It has to charge a bit, the Internet, so you have to turn up the power beforehand so that it works."* (C6)

The mental representations of the Internet verbalized and drawn by the participants were clustered in three categories based on how they are influenced by the knowledge of safety and privacy (see Section 3.4.2) of the participants:

- Activity/Interfaces: In this category participants first and foremost mentioned activities, which they perform while being online (e.g., finances, online shopping, social media, emails, games, videos). Thereby, they drew either user interfaces of the device they use or the activities themselves (three adults, nine children). Furthermore, especially the younger children explained their activities (watching a video or playing a game) in much detail about its content. The parents also mentioned activities, however, explained in more detail, what they were doing and knew that they were actively engaging with the Internet and its information/resources.

- Earth/Worldwide: Participants in this category explained the Internet as an earth, which has global connections. The communication is possible in all directions (four adults, two children).

- Network/Technology: The participants mentioned or drew (via cable/WIFI) connected components within a network of which they are a part. These components include smartphones, laptops, and servers (four adults, two children).

The categories of mental models defined are not mutually exclusive, as the mental model of one participant can contain parts of other categories. The assignment was based on the study transcripts and in the case of several possible assignments, the two researchers decided on the category based on the participant's drawing. Exemplary pictures painted by the participants of the three categories can be found in Figure 3.7.

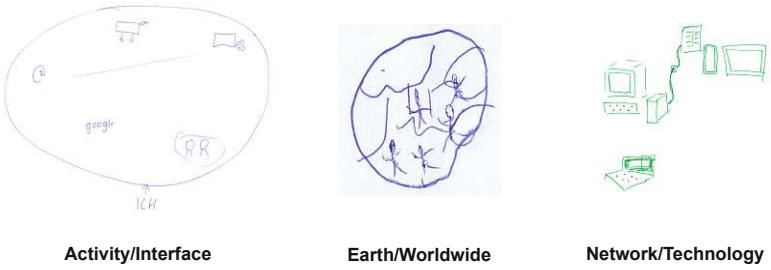


Figure 3.7: Drawing examples from one of the three categories: Activity/Interface, Earth/Worldwide, and Network/Technology

3.4.2 Online Risks - Mental Models

In this Section, the findings of perceived security and privacy risks from the children's and parent's perspectives are described. In order to answer RQ3, the risks are discussed in relation to the Internet mental model categories found in the Section above (see Section 3.4.1).

Children

During the study both, children and parents were asked about privacy and security risks. However, for children, the interviewer refrained from using these words as it can be assumed that children might not (fully) understand the terms. Instead, more tangible wording such as "bad guys on the Internet" or "people that want to steal your information" was used.

The participants were asked several times throughout the interview about threats that may jeopardize the privacy and security of them. Thereby, it was revealed that the children in the *Activity/Interface* category (age five to seven) showed little concern when asked about mean, unfriendly, or unpleasant people or incidents. They either were unsure (three children), did not answer (one child), said there are no bad people on the Internet (two children) or told about videos that were "not good/nice" (three children). Furthermore, two children (age seven) mentioned risks related to their health.

> *"If you look too long, you get a headache."* (C3)

It was interesting to observe that the youngest children assumed that only they and their mothers or families know when they do things on the Internet. In comparison, the older four children in the category explained that there are other people who see that you are watching a video, for example, by "a thumbs up" or the number of viewers.

The two children (age seven to eight) in the category *Earth/Worldwide* had a somewhat deeper understanding of the potential dangers of the Internet in comparison to the first

category. They also emphasized that besides their parents, friends or other people who uploaded the Videos know their actions. Furthermore, one child (age seven) noticed that:

> *"My dad said no (bad) person can change the message...Only if my grandma deletes it unintentionally [laughs]"* (C11)

In the category *Network/Technology* I experienced the most knowledgeable children. The two children (age eight) mentioned during their interview that a bad person could be in the system. One of them even called that person a "Hacker", however, they could not elaborate on what this person could do in more detail. Furthermore, both were aware that their online activity can be noticed by other people. One specified who knows about their online activity:

> *"Besides my parents...Magenta knows the videos I watch...and Steam also knows the games I play"* (C13)

Especially the younger children had a very sparse understanding of online risks. Their threat models about the Internet were always connected to personal experiences. Only two older children (age eight) had an awareness of hackers or other bad people. Furthermore, one child articulated that once their computer was broken due to a bad program which made their screen blue.

> *"We had a program on our computer which programmed everything blue and then it was broken."* (C11)

Parents

When asked about differences between security and privacy, eight parents gave a correct explanation, whereas five parents had a hard time distinguishing between these terms and/or used them as synonyms (three of the Activity/Interfaces category, two of the Earth/Worldwide category).

Adults articulated a lot of potential risks. Figure 3.8 highlights the threats and how many participants mentioned them. The most mentioned and discussed risk by all participants, parents, and children, was video-integrity exploits. Thereby, inappropriate content is disguised as being child-friendly. This phenomenon surfaced in 2017 and is known by the neologism *Elsagate* [24]. Ishikawa et al. [89] and Papadamou et al. [150] presented two studies to detect the disturbing content. However, until now no technical solution was found in order to protect the children. Therefore, the parents of the study developed their own mitigation strategies (e.g, pre-watching the video).

In general, the parents related to category *Activity/Interfaces* expressed the least risk awareness among the adult participants. Besides inappropriate video content, they

mentioned device theft (as it is not directly related to the Internet, it is categorized as "Other") and some sort of surveillance. The parents in this category have often stressed that they do not know or simply do not understand the Internet and its risks. Therefore, they could not answer more specific questions about where the videos are located or who sees/hears someone's activities.

In the category *Earth/Worldwide* the participants' risks perceptions were also centered around themselves and their close ones. In addition to the above-described risks, those parents expressed experience with phishing or similar social engineering emails/messages. Furthermore, one parent mentioned a hacker:

> *"My phone number is online available, so the [hacker] could also possibly hack/access my cell phone"* (P12)

The parents of category *Network/Technology* all identified some sort of omniscient actor or hackers that can interfere with all systems, as nothing is completely secure. In line with the findings of Krombholz et al. [106] also two parents (one in this category and one from the Earth/Worldwide) mentioned that messages via Signal or WhatsApp can also be interfered with/altered (although both messenger services are end-to-end encrypted). The parents of this category described risks from a more technical perspective, by mentioning viruses that infect systems or Monster-In-the-Middle (MITM) and Denial of Service (DOS) attacks. Technical details were provided only by parents which were professionally involved with technical aspects of the Internet.

The majority of all adult participants agreed on the privacy invasion of the Internet, however, felt powerless against this transparency of human beings. They stated that they can only try not to provide even more additional information.

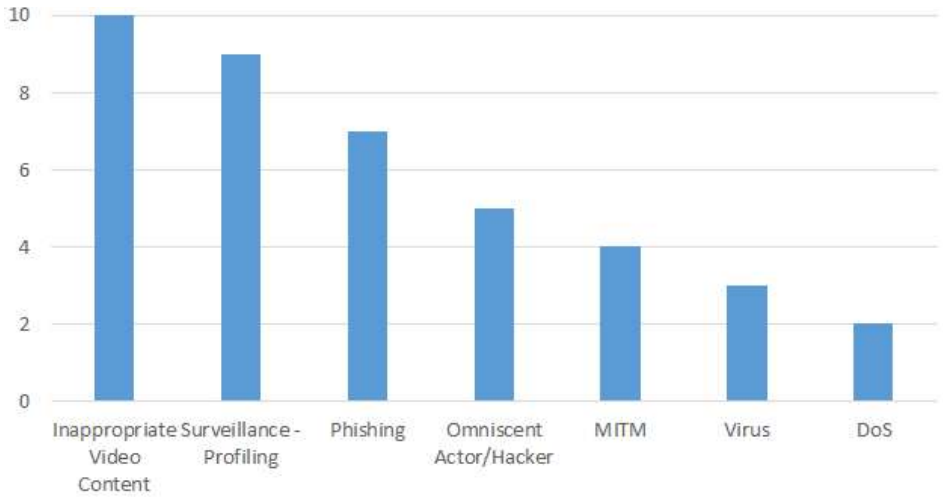


Figure 3.8: Perceived security and privacy threats of adult participants

Besides a high-risk awareness, eight participants stated that they don't think that neither their privacy nor their security is in direct threat.

> *"I can't control it [...], but I think, with a small fish like me they have nothing to get anyway—so they won't care."* (P5)

Although many participants expressed the opinion that most attacks are unlikely, five explicitly mentioned that they take precautions. Those were either related to the use of well-known and trusted services (e.g., password managers, antivirus software or firewalls), or resulted in them not putting too much private information on the Internet and storing passwords securely.

3.4.3 Child-Parent Pairs

In order to answer RQ1a (see the beginning of this Chapter) the mental models of all the child-parent pairs were compared. Thereby it became apparent that the mental models of the children are more sparse and differ greatly in their drawn and mentioned components. Only one child-parent pair was in the same category (CP11) namely *Network/Technology*. Interestingly, the other children in the categories *Network/Technology* and *Earth/Worldwide* had respectively the opposite category of their parents. This may be due to the fact that the other parent has taken over the part of the upbringing/parenting or because the children imagine things differently and deem other aspects more important. Four children mentioned that one or both parents explained to them incidents that occurred on the Internet and told them about some other online risks (e.g., bad/inappropriate content, IT viruses, hackers), which was also confirmed by the statements of the respective parents.

3.4.4 Meta findings

Uncertainty about the processes, protocols, and technologies of the Internet was a major factor during the interview, especially for children. However, with the help of the interview guidelines, they were able to articulate their thoughts.

In general, I observed that during the task where participants had to describe Internet aspects that are not physically tangible, children quickly reached the boundaries of imagination. Also parents had some problems describing in-depth the functionality and connections between many components. Some used buzzwords such as "encryption", "decentralized", or "rays/radiation waves", but their explanations of these terms were very vague when asked further. The participants mentioned different reasons for their uncertainty, including a lack of interest, transparency, training, and hands-on experience.

Due to the work with children, it was necessary to use different methodological approaches to gain valuable insight. It was important to get as much information as possible in a fairly short amount of time, because of the children's shorter attention span compared to

adults. In order to keep the children entertained, the drawing tasks and verbal questions were mixed, which turned out to be a fruitful approach. The aim of using pictures was to elicit smiles (particularly with the picture of Peppa Pig) and lighten the mood, especially of shy children.

Furthermore, despite the parents' willingness to participate in the study, the children's openness to being interviewed at the time of the appointment was still a risk for the study. Luckily, all children felt comfortable enough with the interviewees and the study setting.

3.5 Discussion of Internet Mental Models

This study emphasized that children's mental models shift from activity-based mental models of the Internet to more technical components and even intangible components between the age of seven to eight years, which is in line with children's cognitive development [47, 78, 185]. In line with the findings of Zhang-Kennedy et al [208] and Kumar et al. [110] this study highlighted that parents are concerned with the security and privacy of their children and therefore, actively or passively mediate their actions. The younger children indicated that they used the Internet only when adults are around. In comparison, children aged eight years felt more secure and shared during the interview, that they use the Internet mostly on their own. Five of them also possessed their own smartphone. Most parents confirmed that their children have to ask them for permission to use the services of the Internet.

Kumar et al. [110] found that parents saw security and privacy as a future concern, which is why they did not explain threats to their children, but rather blocked specific applications or protected information with passwords without further communication with their children. Opposed to Kumar et al., this study discovered that the reason why some parents refrained from discussing threats with their children was that they perceived security and privacy threats as unlikely and only as hypothetical possibilities. However, not all of the parent participants did coincide with this approach. They stated that they actively engaged with their child(ren) in order to explain security and privacy risks to them.

> *" I see it [something inappropriate] and then I can block it and we can talk about it."* (P8)

As the interviews with the children confirm, parental education and explanation of possible risks are important for the children, as they remember incidences vividly. Furthermore, some children stated explicitly that their parents did not educate them or just said they have to ask permission to use the Internet but the children were unsure what could happen and why.

Based on those findings, the following provides improvement suggestions, which answer RQ4. This study indicated that all children were able to articulate visual details

from videos and games or could describe details about the interfaces of the Internet tools/devices. Therefore, it can be argued that child-friendly designs and visualizations should be used to educate children about the Internet and its related security and privacy risks. Yan [200] revealed a lack of educational material for children. Therefore, suggestions are on the one hand to further enhance online activities for children in order to make them inherently secure and privacy-protecting and include visual cues designed especially for children (similar to the HTTPS security indicator) in the interfaces. On the other hand, it is important to encourage parents, legal guardians, and teachers to actively educate children. Concretely, the following improvements are recommended:

- Privacy and Security Enhancing Systems: The systems (mostly smartphones, tablets, and televisions) and the applications children use should preserve their privacy and security. The parental control of browsers is one example of such a mechanism. However, as some parents mentioned, their children manage to overcome/disable such systems, as they knew their passwords/had one account. Therefore, a solution could be that a fingerprint/face recognition of a child could enable a different user account with different settings than their parents. In the children's settings e.g., no purchases should be allowed (without parental authorization) and contents should be restricted based on certain (parental) defined rules.
- Visual Interface Cues: The children from this study remembered certain visual cues in great detail (e.g., the thumbs-up or thumbs-down symbol). Therefore, e.g., smileys or colorful borders should be added to videos that have been screened for child-friendliness or which indicate whether they are scary or otherwise inappropriate for children. With the help of such cues, children can easily decide for themselves whether they want to watch certain content or not.
- Education: In order to educate children about the Internet there should be two approaches. First, general education about the Internet and its risks should be performed with the help of visual supports [168, 71] as children are very influenced by them. Thereby books [149] or videos can help to provide the children with the necessary information, while still being fun to watch for children. It would be great if some child's favorite series would also create content with such information. Second, there should be incident-related education, as this study uncovered that children remember them well. It has been shown by Rader et al. [159] that anecdotal stories about security incidents help adults to learn secure Internet behavior. Thus, it can be assumed that also children's thinking and behavior about security risks on the Internet could be impacted by such stories when they are tailored to their knowledge levels. Therefore, parents can either relate to their own knowledge and experiences or retell stories they heard in the news or from friends. Possibly, books with visual representations or pictures in newspapers could be used for visualization. Both forms of education can either be performed by teachers, supervisors, or parents/legal guardians.

CHAPTER 4

Mental Models of HTTPS

Disclaimer: *The contents of this chapter were published as part of the publication "Are HTTPS Configurations Still a Challenge?: Validating Theories of Administrators' Difficulties With TLS Configurations." In: International Conference on Human-ComputerInteraction 2022 [129]*

Communication has increasingly shifted from the analogue to the digital realm, currently exacerbated by the pandemic as real-world contacts are limited. Cryptographic protocols like Transport Layer Security (TLS) form the basis of secure digital communication since they protect sensitive information against unauthorized modification, processing, destruction, or access. HTTPS is the current standard to guarantee secure communication between a client and a web server. Every day, TLS is used several million times by different applications such as email, websites, or messengers. However, to guarantee the security benefits of this protocol, it must be configured and maintained correctly by an administrator, which has shown to be prone to human error [108, 179, 19]. Free of charge tools such as Let's Encrypt and Certbot were introduced so that administrators are not anymore required to understand the full complexity of the cryptographic protocols. Let's Encrypt [2] is a certificate authority that issues cost-free certificates. Certbot is an Automatic Certificate Management Environment (ACME), which supports web administrators in generating and securely extending certificates and therewith, makes the process of configuring HTTPS more user-friendly. Several publications [65, 35] indicate an upward trend in HTTPS adoption and configuration security, but there is still potential for improvement, as many configurations remain insecure. Based on the Qualys SSL Lab Report of November 2021[1], about half of their scanned websites currently provide inadequate security.

[1]SSL Pulse —monitoring Alexa's top million websites for the quality of SSL and TLS support https://www.ssllabs.com/ssl-pulse/

Over the last years, several studies highlighted challenges of the HTTPS configuration process and the substantial impact of flawed configurations on security and privacy [38, 153, 130, 7]. Moreover, studies that shed light on the perspective of developers and administrators have increased in the last few years. Those user studies used different qualitative methods, such as qualitative interviews and lab studies, to explore the challenges of the HTTPS deployment process.

The theories developed from previously conducted user studies are used to formulate the research questions. With this study, the configuration problems revealed in prior studies are quantified through two administrator studies. To reach high external validity and get a realistic picture of the current situation, it was especially important to find participants who are actively configuring or maintaining web servers. This study extends the body of knowledge on administrator studies by (i) providing numbers that show the extent to which challenges discovered in prior studies (still) exist and (ii) investigating usability challenges with HTTPS configurations in a professional setting.

With this study the research questions RQ1-RQ4 presented in Chapter 1 are answered for the cryptographic protocol HTTPS and related configuration and maintenance system. The stakeholders for this study are administrators configuring and maintaining the protocol.

Therefore, the main contributions of this study are to:

- validate results of earlier qualitative studies with a quantitative survey,
- extend the body of administrator-studies and,
- investigate remaining usability issues with the usage of Certbot.

4.1 Background and Related Work of HTTPS

One of the most well-known and widely used protocol to secure online communication is the HTTPS protocol which uses the TLS. TLS was first defined in 1999 and was since revised with 4 new versions on the market bringing various security enhancements. Its purpose is to guarantee data privacy and integrity of a communication between a client and a server. Today, TLS is commonly used as the security layer in the HTTPS (**HTTP** over TL**S**) protocol.

An abstract depiction of HTTPS can be found in Figure 4.1. The certificate binds the public key to the identity of the server, which is verified by a challenge-response protocol (between step 2. and 3.). To establish a secure communication channel over TLS (1.3) between a client and a server, the client initializes a session by sending a "Client hello", the clients supported cipher suits and its key share. Then, the server answers with a return message which includes: i) a hello message, ii) the server key share, iii) the verified

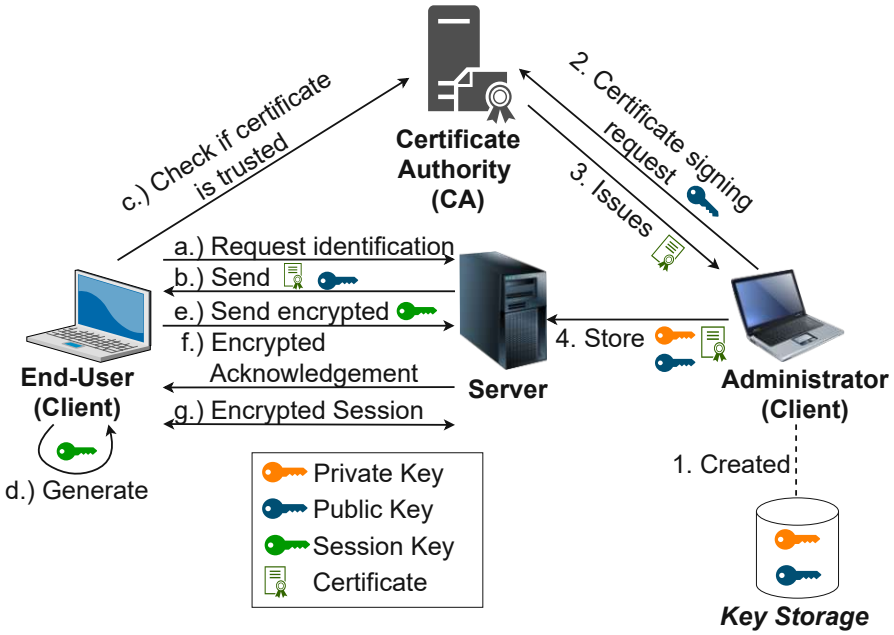


Figure 4.1: How to generate a certificate (1.-4.) and establish a encrypted communication session using TLS1.3 (a.-g.)

certificate and iv) the finished message[2]. Afterwards, the client generates the session key with the servers key share and checks the certificate. If the certificate is valid, the clients also sends a finished message (encrypted with session key) to the server and after that the communication takes place within an encrypted session.

The complexity of the HTTPS protocol lead especially in the beginning to faulty configurations and maintenance issues [54] in the wild. The introduction of Let's Encrypt (a certificate authority - an entity which provides digital certificates) and different frameworks, such as Certbot, Caddy and NGINX, have helped the administrators to provide secure configurations for the end-users. As these tools and the protocol itself have been upgraded and revised many times, an upward trend in securely configured communication channels on websites can be observed, however, with room for improvement.

Over the last years, several usable security studies [108, 63, 146] highlighted open challenges with HTTPS and the impact that insecure configurations have on security and privacy [55, 38]. In the following, the related work of usable security studies (including mental model studies) from the end user's and the administrators' perspectives are presented, as they are both affected by faulty configurations.

[2]The finished message is protected/encrypted with the just negotiated algorithms, and keys. It serves as a "checksum" for the other party before other (sensitive) information is shared via the session

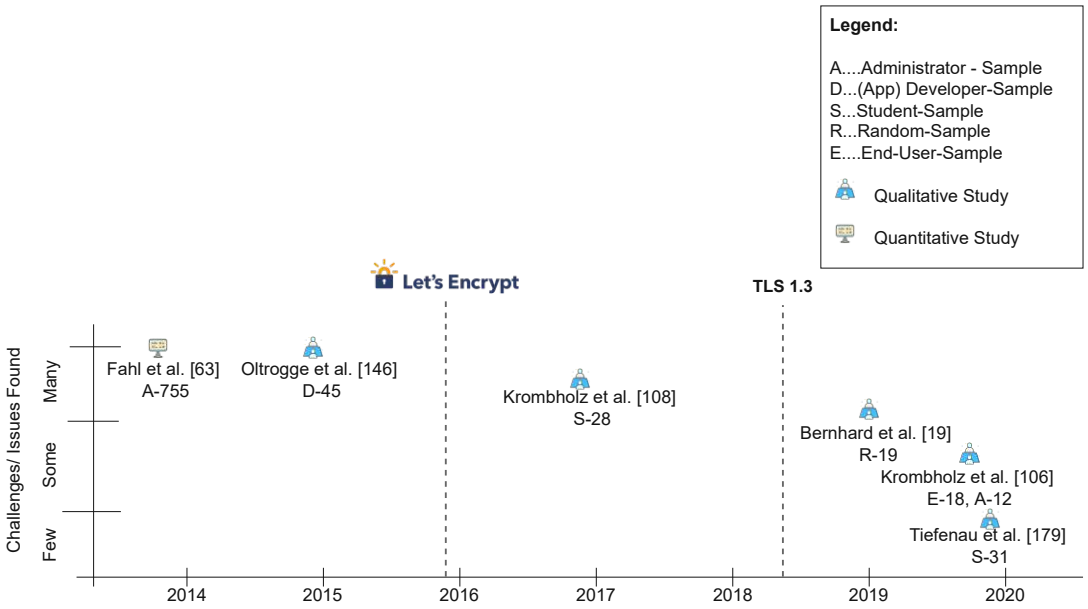


Figure 4.2: Overview of administrator studies on HTTPS over time and their abstracted amount of reported challenges.

4.1.1 End-User Studies

The focus of the end-user studies in the area of HTTPS lies in the secure use of the communication network. The task of end-users is to adequately react to security warnings and indicators in order to avoid insecure communication channels.

The first study examining the effectiveness of browser warnings was published in 2009 by Sunshine et al. [177]. They found that many participants have mishandled most of the warnings and therefore suggest an improvement and a minimization of them. In order to shed light on the challenges with those warnings, Harbach et al. [83] examined their linguistic properties and Akhawe et al [8] the correlation of the (in)effectiveness of warnings and the users' experiences. Felt et al. [67] conducted a study to highlight the differences between the SSL warnings of Google Chrome and Mozilla Firefox. Based on those results Felt et al. proposed new versions of SSL warnings [64] as well as a new set of security indicators [66] which were adopted and released by Google Chrome. The root causes for warnings and the reactions to them were studied by Reeder et al. [161] and Acer et al. [5]. Hereby they found that most of the warnings were caused due to client-side or network errors.

4.1.2 Administrator Studies

Within the last years, Secure Sockets Layer (SSL) and TLS user studies focused on the administrators in contrast to the end-users point of view. Figure 4.2 presents an overview of these user studies. The figure depicts the studies (including what kind of sample they

used and the size of it) over time and the number of challenges they found. The found challenges/issues are estimated based on the described results of the user studies (Few: $< 25\%$; Some: $\sim 25 - 50\%$; Many: $> 50\%$ of the participants faced challenges).

Fahl et al. [63] conducted a large-scale study with administrators to determine the reasons for non-validating security-critical X.509 certificates[3]. Their results showed that one third accidentally misconfigured the certificates and two thirds intentionally used non-validating certificates. Reasons for the use of non-validating certificates were that sites were either no longer in use or not designed to be accessible for users.

Oltrogge et al. [146] studied the applicability of certificate pinning for non-browser software. They found that pinning is considered very complex among developers, which results in poor adoption. Therefore, they implemented a web application that supports the developers by guiding them through a pinning-protected TLS implementation. However, users were still facing difficulties to deploy pinning correctly.

In 2017, Krombholz et al. [108] identified major pitfalls administrators face during the TLS deployment. They revealed that the configuration procedure is too complex and identified protocol components that are difficult to understand even for experts who managed to deploy valid configurations. In general, their results suggest that administrators heavily rely on online sources that, depending on their quality, often lead to faulty deployments.

The introduction of Let's Encrypt paved the way for the widespread use of encrypted communication [6]. However, still some challenges remained, that had to be overcome. Manousis et al. [130] found that only half of the domains, that were configured with Let's Encrypt certificates, had a valid LE certificate. Thus, they revealed that automation does not fully release administrators from the burden of dealing with the complexity of the protocol.

Bernhard et al. [19] extended the study of Krombholz et al. [108] with a more diverse sample and Let's Encrypt. Hereby, they confirmed the difficulty of the deployment process and confirmed that participants were not able to successfully deploy HTTPS. In order to understand the challenges faced by developers, Krombholz et al. [106] conducted a qualitative study with a focus on the general understanding of the HTTPS protocol. They investigated the mental models of 12 administrators and 18 end-users and found that many administrators only have a sparse understanding of the protocol, its components, and how they interact with each other. Tiefenau et al. [179] showed the positive effect of the improved usability through Certbot (and Let's Encrypt) on the security of HTTPS.

This study differs in comparison to the aforementioned publications primarily in that it quantitatively assessed the (remaining) usability issues and administrators' trust in Let's Encrypt and Certbot and validated the found mental models of Krombholz et al. [106]. Besides, this study provides high external validity by capturing an actual set of web server administrators (e.g., responsible for certificate managing, the set-up of TLS), in

[3] X.509 is a standard format of public key certificates used e.g., for TLS, which specifies an identity and a public key and is either self-signed or signed by a Certificate Authority (CA).

comparison to a student sample. Therewith, this study extends the current literature of developer studies with a real-world sample.

4.2 Methodology of the HTTPS Study

In the following, the methodology used to design, collect and analyze the data obtained from the HTTPS studies are described. First, the small-scale study conducted with 16 participants is described including its findings. Based on the findings from the small-scale study the quantitative online survey was designed which is described in Section 4.2.2. Thereby, the recruitment process, the sample validity, and the data analysis are presented. Finally, the ethical considerations of this study are discussed in Section 4.2.3).

4.2.1 Small-Scale Study

To flexibly explore the problem space before designing the online survey I decided to follow the approach by Jensen and Laurie [90] and conducted a small-scale study. Although some studies qualitatively investigated the administrative point of view [108, 106, 19, 179], this study replicated and challenged those results as the technology and the frameworks investigated in those studies have changed over time. In 2018 the latest version of TLS (1.3) became available and since 2021 versions older than TLS1.2 are deprecated. Furthermore, ACME clients such as Certbot have been revised several times and new ACME clients were released (e.g., Caddy2, lego) in the last five years. A questionnaire was developed to shed light on administrators' knowledge of the interplay between certificates, encryption, and keys. The questionnaire consisted of 31 questions (26 closed-ended and 5 open-ended questions), asking for demographic data, experiences with web server software, and TLS knowledge. At the end of the questionnaire, the participants were asked how they thought the TLS set-up process could be improved. The answers were coded by two researchers individually and discussed afterward until agreement on all coding conflicts was reached. The inter-rater reliability (Krippendorff's Alpha value [105] of $\alpha = 0.87$) indicates a high level of agreement.

For this study, $N = 16$ (the goal was to at least have 12 participants, to ensure study validity) administrators were recruited from one company in Central Europe, with more than 250 employees. The demographics of this study can be found in Table 4.1, displaying the gender, age, and highest completed education of the participants. The decision to recruit the small-scale study participants from one company was made, as it meant that the working environment and resulting biases and constraints of all participants were known. The participants were web administrators, either configuring the company's web servers (including certificate management) or working as consultants for clients. The study took place on two different days, in a laboratory environment provided by the company with a supervisor to ensure equal conditions. The questionnaire was distributed directly on-site by the supervisor. If there were any questions, there would have been pre-made aids, which the supervisor could have given, but they were not used.

Table 4.1: Descriptive quantitative analysis demographics (N=16)

Demographics		Participants (%)
Gender	Male	11 (75%)
	Female	2 (13%)
	Prefer not to say	2 (13%)
Age	18 – 32	5 (31%)
	33 – 42	6 (38%)
	23 – 52	3 (19%)
	>52	2 (13%)
Highest Education	High school	2 (13%)
	Bachelor degree	4 (25%)
	Master degree	8 (50%)
	Other	2 (13%)

Study Findings

In order to answer RQ1 and RQ2 the participants' knowledge about certificates and their connection to keys was asked with open questions, followed by closed-ended questions. When asked with an open question, 11 participants (69%) managed to correctly describe the purpose of a certificate (i.e., it is required for the proof of the server's identity during the authentication process and needed to establish a secure communication session). The rest of the participants either described only parts of the certificates' purpose (3; 19%) or mixed things up (2; 12%)). In line with findings by Krombholz et al. [106], two (12%) participants tended to confuse authentication and encryption. In the context of certificates and security indicators, four participants explicitly mentioned issues with (blind) trust which one have to put into the protocol. They explained that administrators have to trust the authenticity of the CA's certificates and end-users have to trust the security indicators indicating the trustworthiness of a website.

When asking them about keys and public-key cryptography, ten (63%) participants were able to explain this type of encryption correctly and three (19%) only explained parts of it. The remaining participants either stated that they did not know (1; 6%) or confused public and private keys (2; 12%), which could lead to severe problems (i.e., leaked information and insecure communication).

Besides the general functionality of the TLS protocol, questions were asked about the

configuration process and used technologies. 14 participants (88%) used a traditional CA (e.g., GoDaddy or Comodo), 12 participants (75%) obtained their certificate from a local self-maintained CA, and only one participant used Let's Encrypt. According to 13 participants (81%), the choice of certificate origin (i.e., which CA was used) was prescribed by the employer or client, the rest stated that they did not know (19%).

The study revealed that 6 participants (38%) still recommend outdated and insecure TLS/SSL protocols and are not sure which mechanisms are used (e.g., forward secrecy or HSTS) to ensure strong security. The test rating of SSL Labs [4] for the company's server configuration is in line with these findings as it only gets a B due to insecure settings (e.g., by supporting old protocols or ciphers).

To get a better understanding of the problems with the TLS configuration, the participants were asked directly about the difficulties they encountered during HTTPS configuration. The majority (14; 88%) stated, that they had no problems, and only two (12%) stated that they did. Occurring problems were on the one hand that the Microsoft-ISS registry settings did not have the appropriate setting levels and on the other hand, the participants missed deactivating fallbacks to older (TLS/SSL) versions. Their suggestions for improvement were staff training on state-of-the-art technologies, stricter (company) rules on HTTPS usage, and automatic configuration checks (ciphers, TLS versions).

4.2.2 Quantitative Study Design

The goal of the online survey was to get quantitative insights into the HTTPS configuration process from the administrators' point of view. The first study (see Section 4.2.1) informed the survey questions and answer possibilities. Therefore survey questions were rephrased, deleted, and added based on findings in the first study. For example, closed-ended knowledge questions were added followed up by open questions to provide the participants the possibility to express their thoughts in more detail. The answers of the first study where participants mentioned issues with trust, gave the impulse to investigate the administrators' trust in HTTPS, Let's Encrypt, and Certbot. English was used as the language for the questionnaire, as it is the working language in many IT-related jobs.

The survey consisted of a maximum of 47 questions, some of which were follow-up questions: (i) 28–34 closed-ended questions (multiple- and single-choice, 5 point Likert scale) and, (ii) 9–13 open questions

Most of the open questions were optional to answer in order to not discourage the participants. In this study, too, the answers were coded by two researchers and then discussed until agreement was reached (Krippendorff's Alpha value [105] of $\alpha = 0.83$). The survey was hosted on soscisurvey.de [81] and took on average 30 minutes to complete. The complete questionnaire can be found in Appendix A.2.

[4] https://www.ssllabs.com/ssltest/ - a service which analyses the configuration, the security and the supported TLS versions.

The online survey design was tested in two rounds of pilot studies with 19 participants (10 security researchers and 9 administrators) who gave feedback. Furthermore, the comprehension of the questions was checked as well as their order and unclear phrasings were removed as far as possible.

Recruitment

A common challenge when conducting an expert (administrator) study, is to acquire a satisfactory number of participants. In contrast to most previous studies in the area of HTTPS usability for administrators, the target participants of this study were administrators with HTTPS configuration experience and computer science students without experience were explicitly excluded. Although there are several studies [197, 4, 160] that indicate that students can be representative for administrators and developers, this has not yet been confirmed in the area of HTTPS configuration. Therefore, participants with professional knowledge of the HTTPS configuration and maintenance process were recruited. A prerequisite for the study was that the administrators had to have maintained at least one web server within 2020, in order to get an optimal view of current administrators' challenges and also, to avoid fading knowledge.

Following Pfeffer et al. [156] and Krombholz et al. [106], the survey was distributed via Twitter, LinkedIn, Facebook, Reddit, and sent it out via mailing lists. As it was particularly difficult to reach female administrators, I joined channels on social media and mailing lists that were primarily for women. To increase the response rate and as compensation for the participation, a lottery [113, 84] was announced consisting of four gift vouchers valued €50 each.

Sample Validity

To ensure statistical power, the effective sample size was calculated by following best practices for quantitative studies [118, 119]. Thereby, a significance level of 5%, a 95% confidence interval, and a power of 80% was chosen, which led to a minimum sample size of $N > 91$.

Altogether, 212 responses were received. From the total responses, unfinished questionnaires, data sets with incorrectly answered sanity questions, and inconsistently filled out check-up questions were filtered out. Furthermore, the average time spent per question was used in order to avoid automatic or random completion of the questionnaire. The final sample consisted of $N = 96$ participants, which exceeds the calculated minimum sample size. The demographics of the participant in the final data set can be found in Table 4.2, displaying the gender, age, current country of work, highest education, and profession of the participants as well as the number of web servers they configured or maintained within the last year.

Data Analysis

Four different approaches were used to analyze the data:

Table 4.2: Demographics of 96 study participants - HTTPS admin study

Demographics

Gender	Female (5%)	Male (76%)	Non-Binary (2%)
	Not disclosed (16%)	Self-described (1%)	
Age	18-24 (7%)	25-34 (42%)	35-44 (36%)
	45-54 (12%)	>55 (2%)	
Working Country	USA (28%)	Austria (24%)	Germany (20%)
	Rest Europe (17%)	Others (7%)	NA (4%)
Highest Education	No schooling (1%)	High school (18%)	College (8%)
	Technical training (3%)	Bachelor (31%)	Master (32%)
	Professional degree (1%)	PhD (5%)	
Profession	Programmer (30%)	System Admin (19%)	IT Consultant (11%)
	IT Architect (9%)	Manager (6%)	Web Admin (4%)
	Tester (2%)	Other (19%)	
# Web Server	1 (12%)	2–5 (36%)	6–10 (15%)
	>10 (37%)		

- *Descriptive Analysis:* to summarize all data according to frequencies, central tendencies, and dispersion or variation.
- *Exploratory Data Analysis:* to find relationships and patterns among the data from a bird-eye view (i.e., by visualizing the data).
- *Statistical Tests:* to investigate the correlation between different variables (pair-wise χ^2 or Fishers' exact tests F, with expected frequencies < 5, for nominally scaled single choice and multiple choice [5] questions, including an interpretation of the effect size Cramér's V [103]). Thereby, the null hypothesis of independence was rejected if $p<0.05$, within a 95% confidence interval.
- *Open-Coding:* to analyze the qualitative data (open questions), two researchers

[5] The Holm–Bonferroni correction [86] was applied to counteract the multiple comparisons problem for multiple-choice questions

coded these questions, by extracting recurring themes and statements. In the following some answers are provided for illustration and better insight (see Section 4.3).

4.2.3 Ethical Considerations

Following the series of guidelines of SBA Research, each data set was stored anonymously with an assigned ID. The email addresses collected for the lottery were stored separately from their corresponding data sets, without the possibility to link them, together. At the beginning of the questionnaire, the participants were informed which personal data will be collected and how (long) they will be stored, strictly following the EU's General Data Protection Regulation (GDPR[6]).

4.2.4 Limitations

The sampling strategy allowed me to recruit a diverse sample of web server administrators with experience in the deployment or maintenance of HTTPS. Despite the recruiting strategy, the sample still has limitations, as the participants work mostly in Central Europe or the US. Therefore, the results can't be generalized as they might be influenced by cultural factors, such as privacy and security awareness in different countries. Although the study sample contained some female participants (5%), this number is still slightly below the average in IT engineering jobs [121]. However, I found no dedicated studies on the gender distribution of system administrators, making it difficult to judge the representativeness of the sample. Naturally, the quantitative survey approach also has its limitations as the data is self-reported and could have been distorted by participants looking up information during answering the questionnaire. I tried to minimize this risk by asking knowledge questions and evaluating the average time taken by participants to answer them as described in Section 4.2.2.

4.3 Results of the HTTPS Study

In the following, the results of the quantitative online study are presented to answer RQ1-RQ3, by providing both quantitative and qualitative insights. The dataset is analyzed based on the hypothesis that Let's Encrypt positively influence administrators' experiences, as prior studies showed [179, 63, 19, 108]. Furthermore, I investigated whether there are correlations or differences in the data that can be explained by other (demographic) characteristics (besides different CA usage), but did not find any, except for the company size (small < 50; large > 50). Therefore, the results are presented with a focus on: i) company size of administrators' current employer (small < 50 vs. large > 50) and ii) usage of Let's Encrypt vs. others.

[6] https://gdpr-info.eu/

4.3.1 General TLS/HTTPS Knowledge

The administrators were asked how they acquired their knowledge, in order to understand different influencing factors. All participants stated, that they acquired it through online research. Almost 60% relied exclusively on online sources and one-fifth additionally on information gathered from colleagues. 18% received training of web server configurations for TLS through education or specific seminars. Only a minority of 6% obtained their knowledge from company specifications such as from configuration frameworks and their documentation. One administrator described the learning process as "failure, lots of failures". Thereby, no significant difference was found between large and small companies (χ^2: $p > .08$) and Let's Encrypt and other CA users (χ^2: $p > .2$).

Furthermore, the participants were explicitly asked whether public and private keys play a part in HTTPS. 90% of the participants answered correctly, others either stated the absence of knowledge (5%) or answered incorrectly (5%). Thereby, significant differences were found with large effect size V between Let's Encrypt and other CA users (F: $p < .03$), as Let's Encrypt users answered it more correctly than the other users. Between small and large companies, no significant differences were found (χ^2: $p > .05$). When asked with an open question how the keys are used, about one-third could explain neither the role of them in the encryption/decryption process nor their connection to certificates. Interestingly, the statements about the precise connection(s) between the keys and TLS were answered correctly by a much larger proportion (>75%) of the participants. Only the statement "TLS uses (PKI) certificates to authenticate parties communicating" led to some confusion. In fact, 48% marked it as wrong and 52% as correct. Some of the participants noted in the "Other" option that they were confused although they provided a correct explanation. Thus, these answers were rated as correct (in total 76% answered correctly).

4.3.2 Configuration and Maintenance

The services configured and maintained by the administrators are equally distributed between company internal, external, and private services. For work-related server configurations, most of the administrators used NGINX (60%), followed by Apache (48%). Microsoft ISS and Caddy2 were used by one-fifth each; HaProxy and Traefik were mentioned in the "Other" option. Internal web servers of the company were only used by 10% of the participants.

The participants were also asked which software they (would) use privately. Hereby, 60% answered NGINX. Less than a third (29%) would use Apache for private web server configurations. 28% of the participants selected Caddy2 and below 10% other software. Thereby, no significant differences was observed between private and professional server-software usage and the company size (χ^2: $p > .2$).

In order to authenticate the HTTPS server, a certificate or another validation mechanism is used. Most participants (91%) use a free-of-cost alternative for obtaining certificates, such as Let's Encrypt (at least for some of their certificates). From these participants,

around one-third exclusively uses Let's Encrypt. The second most used certificate source is a local self-governing CA (41%). However, these CAs are used only in combination with another certificate source. Traditional CAs, external providers such as Cloudflare, or self-signed certificates, are used by less than a third. Altogether, 20% of the administrators were obliged by their company to choose a specific certificate origin. Significant differences were found concerning the obligation to use specific certificate origins depending on the CA usage (χ^2: $p < .03$, $V > .27$ [medium]). Let's Encrypt users are not obliged by their company to use them in comparison to other CA users. Furthermore, the results show a significant difference between the company size (F: $p < .01$) and the obligation of specific certificate origins, whereby large companies more often specify the place of origin.

The vast majority (84%) of the participants used an ACME client, to obtain certificates of Let's Encrypt. The usage of Let's Encrypt and Certbot is allowed by over two-thirds of the companies. A significant difference was found between large and small companies (F: $p < .02$), whereby small companies are more open to their usage. Altogether, 15% of the participants stated that it is prohibited by their company to use them, mostly due to mandatory software (company internal) or customer guidelines. About 82% of those using these tools reported that they changed their (working) routine. The other participants felt uncertain (11%) or were sure that they did not influence their workflow (7%).

The most important factors influencing the working process were on the one hand the automation and the associated security that comes certain ACME clients, and on the other hand, the free-of-cost certificates which are easy to handle (even when using multiple certificates). The participants perceived the higher update frequency both as positive, when used within an automation framework such as Certbot, and negative when used without one. Without an automation framework, the renewal process must be either self-automated or handled manually each time. In this context, one administrator stated:

> *"Although there is the automatic [certificate] renewal (cronjob), it happens that certificates expire, leading to broken services."* (P11)

The confidence in the security of most of the participants' configurations is very high (85%) with strong differences between Let's Encrypt and other CA users (χ^2: $p < .03$, $V > .28$ [medium]). In fact, Let's Encrypt users were not that confident about the security of their website. The main reason given by those participants was that they did not achieve an A grade in a TLS/SSL-test.

Certificates have different levels of validation, depending on the type of service they are used for. There are three different types:

- Domain Validation (DV): the CA verifies the owner of the certificate who has control over the domain (low-level validation).

- Organization Validated (OV): the CA verifies the organization and its rights to use the domain (medium level validation).
- Extended Validation (EV): the CA verifies the legal identity of the organization (high-level validation).

Administrators argued that the validation decision is often based on a cost-benefit factor. On the one hand, websites need to appear trustworthy in order for web users to trust their integrity. On the other hand, stronger validation is (more) expensive and often provides no noticeable benefit for the consumer. A significant difference was discovered between the participants' usage of CAs (F: $p < .04$) and their required certificate validation levels. I attribute this to the fact that Let's Encrypt does not provide EV and OV, as those validations require additional human interaction. Administrators using OV were required to do so due to company policies or customer requirements. The vast majority of administrators agreed that DV is the most important one (81%), as without it TLS would lose its meaning. The requirement of OV is stated by 16%, followed by 9% reporting a need for EV. 12% stated that they do not know which kind of validation they need and 3% were unwilling to disclose the requirements or answered that it depends on the use case.

TLS Deployment

During the set-up of TLS, several configurations can be made which influence the security and compatibility of the web server. Two-thirds of the participants have actively switched off HTTP as part of the TLS configuration in order to guarantee a connection via HTTPS. Most web server frameworks provide the administrator with certain software defaults, however, nearly two-thirds (64%) stated that they changed them. The answers to the question of why they have changed the default settings were mostly, that the default settings often do not meet the security standards of the administrators (73%) or the companies (17%).

> *"No sane defaults, those defaults prevent desirable security properties."* (P3)

The answers significantly differ between the different company sizes (χ^2: $p < .03$, $V > .28$ [medium]), as administrators of large companies, changed the default settings more often. Many changed the allowed ciphers and TLS versions to either stricter or looser settings depending on their requirements. Hereby, some explicitly stated that they enabled older versions for increased compatibility, thereby accepting the greater security risk. Furthermore, some enabled forward secrecy and HSTS headers to harden security and OCSP stapling to reduce costs for validation. Many changes were motivated by recommendations of Mozilla or various SSL/TLS scanners and guidelines such as bettercrypto.org.

In a professional environment, security audits help to reduce the risks of vulnerabilities. Half of the participants stated that their company performed audits when an exploit was found, 26% performed regular security audits, and 21% perform audits before live deployment.

Challenges

In order to investigate the problems administrators currently face, the participants were asked to elaborate on their experiences. Slightly more than half (57%) of the participants reported, that they had no problems during the configuration. The remaining participants (43%) stated, that they had experienced at least some problems.

> *"No [problems] when setting up standard web servers, but lots when setting up custom stacks."* (P71)

These challenges were mostly experienced by administrators, who did not use Certbot (60%)(χ^2: $p < .03$, $V > .27$ [medium]). The main challenges were caused by compatibility issues (23%), errors with the certificate chain (12%), the certificates order as well as trust issues with other (self-signed) certificates (9%). Those issues often appeared due to operating systems that were too old or clients which did not support new TLS versions. Also, errors from former administrators (5%) were mentioned, as some forgot to send the certificate chain when configuring the TLS server, which can lead to issues finding the correct root. Another problem for some administrators was poorly designed error messages, which were not perceived as helpful, as well as error reports, which did not show up in the console. In order to generate a (valid) certificate, the administrator needs to send a certificate signing request. However, if one puts a wrong value into the request or forgets information, the issued certificate is faulty.

4.3.3 Trust

The confidence a person places in a technology (i.e., trust) is an important factor for its adoption [117]. Due to the complexity of trust establishment with multiple factors involved, this study only focus on self-reported trust users put in HTTPS, the security indicator, Let's Encrypt, and Certbot. Therewith, I do not intend to simplify the concept of trust, but rather provide the first impetus to examine its influence on the use and configuration of HTTPS.

In order to enable HTTPS, some sort of authentication (via a CA) is necessary. The participants were asked whether they trust in the security of the HTTPS communication channel, the security indicator of web browsers (lock symbol), and the CA Let's Encrypt as well as the ACME client Certbot. The results of these questions are summarized in Table 4.3.

In general, the trust of the participants in **HTTPS** was very high. Reasons for that were the strong cryptography used in state-of-the-art TLS versions, its open-source nature,

Table 4.3: Results for Trust Questions

Trust in	Yes	No	I don't know
HTTPS	80%	13%	7%
Security indicator	69%	23%	8%
Let's Encrypt	90%	4%	6%
Certbot	57%	6%	37%

and the strong community behind it. The main argument against trust in HTTPS was that there are still too many poorly configured web servers in the wild. The study revealed that administrators of larger working environments have higher trust in the security of HTTPS than those of smaller ones (χ^2: $p < .02$, $V > .28$ [medium]). I explain this by their increased use of (higher) security standards and certificate validations (see Section 4.3.2), which strengthens the security of their HTTPS servers. One participant stated:

> *"It's the best we have. It's probably not 100% secure, but it's secure enough for my purposes.* (P55)

The participants' trust in **Let's Encrypt** was even higher than in HTTPS. Thereby Let's Encrypt users have more trust in it than users of other CAs. This can be explained by the fact that the use of technology can strongly influence trust [141, 117].

Some participants said, that although Let's Encrypt might not be technically different from other CAs, its non-technical features surpass others. These include its open-source nature, the expert community supporting it, the non-profit motive, the documentation, and the transparency when incidents happen. However, some administrators reported that a problem with the free CA is the (ab)use of its certificates for phishing websites, as confirmed by media reports [37].

Certbot's tasks range from certificate ordering to domain validation and automatic certificate renewal. The open source ACME client Certbot is trusted 36% less than Let's Encrypt. The trust in Certbot depends on whether administrators use Let's Encrypt or other CAs as significant differences were observed between those two characteristics (χ^2: $p < .01$, $V > .33$ [medium]). The reasons to trust this ACME client are very similar to those of Let's Encrypt. For Certbot's default usage, root privileges are required, which is not appreciated by some of the participants. Furthermore, one participant stated

> *"I don't trust it to not completely trash my server configuration. (Which has happened.)"* (P8)

The **security indicators** in the browsers are not trusted by 23% of the administrators. Again, problems with phishing websites were mentioned. Moreover, participants complaint about pre-installed (rogue) root certificates, which are distributed in operating systems, although a green lock symbol still appears.

> "The lock symbol is one of the things you should check, but not the only one. [It's] not good enough as a single factor of trust"

4.4 Discussion of the HTTPS Study

In this Section, the results are discussed and compared to the results of previously conducted studies. Thereby, I will go into detail about current web server usage, the administrators' knowledge of TLS, their confidence in the technology (and their skills/knowledge), and their challenges with HTTPS. With this study, prior findings were quantified and it was shown that administrators choose outdated TLS-versions or weak ciphers due to various reasons and that although Certbot provides high usability there are some adoption/usage barriers. Furthermore in this Section recommendations for improvements for the configuration and maintenance process of the HTTPS protocol are provided and thereby answer RQ4.

4.4.1 Web Server usage

Current websites mostly use web servers of Apache, NGINX, or Microsoft ISS [186]. In 2020, NGINX managed to surpass the use of Apache. The sample shows a similar distribution for web servers used in professional environments as shown on the left-hand side in Figure 4.3. Furthermore, the study revealed that more than half of the participants would not choose Apache or Microsoft ISS for their private web server (right-hand side in Figure 4.3), which indicates dissatisfaction with these web servers. In comparison, I hypothesize that the users of NGINX and Caddy2 are more satisfied with the software, as they would also use it for their private web server. Participants explained their usage of Apache with long-time experience, its dynamically generated content, and the PHP support. I hypothesize, based on individual responses from participants, that another possible explanation for the rise of NGINX is its flexibility and usability (due to easily understandable configurations) as well as the combined usage with Apache.

> *"[I use] Apache just for PHP support, NGINX for everything else."* (P45)

This is also reflected by the fact that Apache has increasingly lost market share over the past few years [187]. Despite the current low adoption of Caddy2, the participants have shown great interest in its ease of use and automation. *Given this trend, further studies need to be done to investigate the usability of NGINX and Caddy2.*

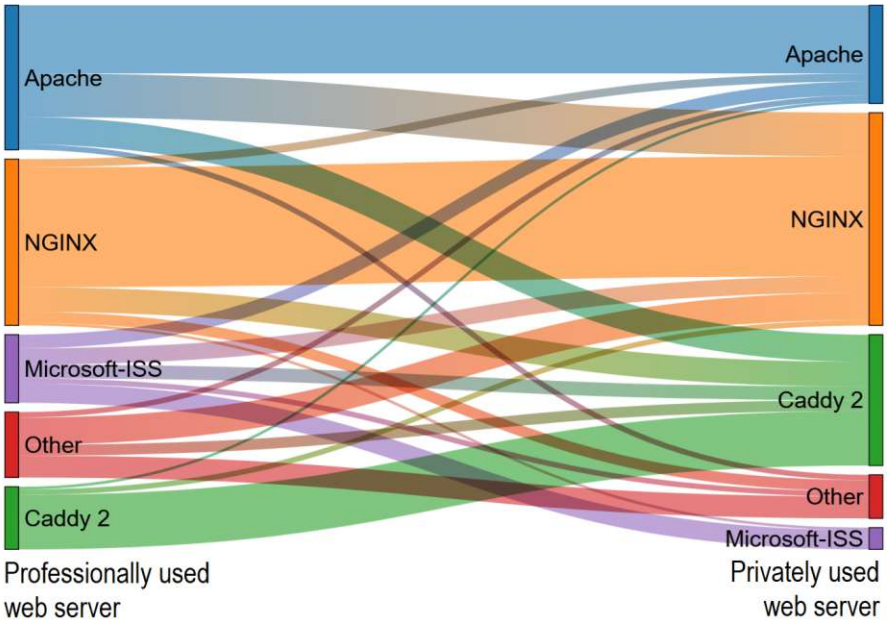


Figure 4.3: Web server usage for business vs. private usage

4.4.2 Knowledge of Certificates and Keys

Certificates have an essential role in the configuration of TLS, therefore administrators must understand their basic concept and functionality. The majority of the participants were able to correctly answer the questions (both open- and closed-ended) related to keys and how they interact with certificates.

In contrast to this study, Krombholz et al.[106] found in their study that administrators did not mention and thereby not acknowledge the existence of keys, while explaining HTTPS. Furthermore, they found that although the administrators mentioned some protocol components and commands, less than half of them talked about server authentication and its link to end-to-end encryption. I theorize that these "knowledge gaps" arose because they did not explicitly ask for explanations of fixed keywords, such as *certificates*, *public* and *private keys*. In comparison, the participant in this study were specifically asked about these components and thus, it was possible to query the (possibly) passive knowledge of the administrators. In line with Krombholz et al. [106], knowledge gaps of the administrators were found related to how the server and the CA interact, confusions of encryption and authentication, and a negative perception of HTTPS and the security indicator. Based on the findings, I created a new mental model of HTTPS from an administrators perspective. Thereby, the correct mental model from Krombholz et al. was used and extended by overlaying it with the found misconceptions of administrators, which are highlighted in blue (see Figure 4.4).

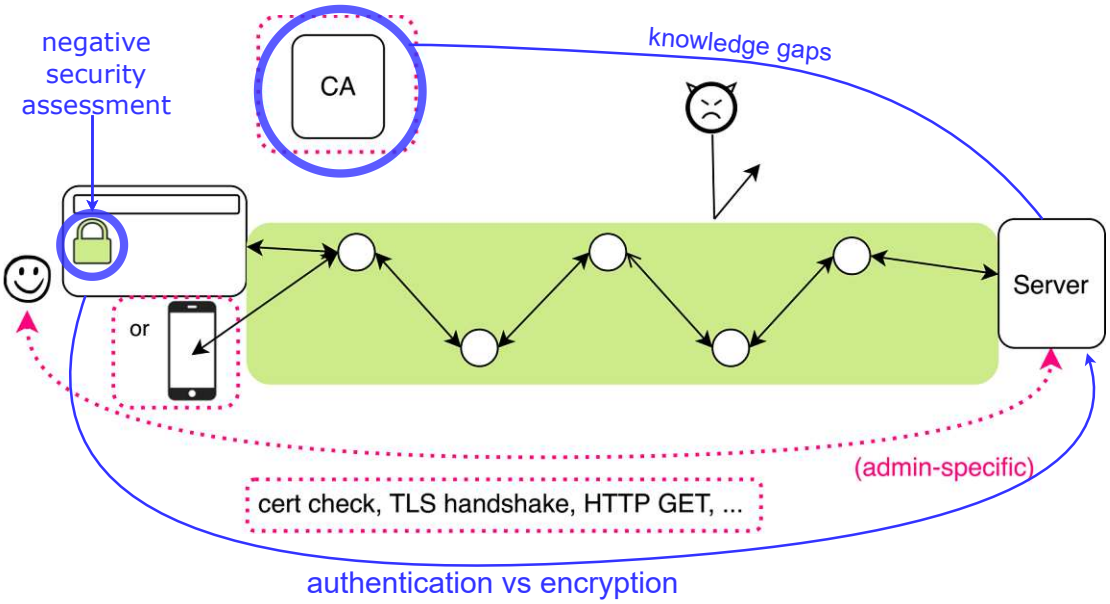


Figure 4.4: Correct mental model of HTTPS from Krombholz et al. [106] overlaid with the found misconceptions (highlighted in blue) which do either not correctly describe the actual protocol or negatively influence the users perception of HTTPS security.

4.4.3 Trust

As mentioned in Section 4.3.3, trust is a very complex concept and until now, to the best of my knowledge, no study did investigate the confidence of administrators in HTTPS and its most widely used configuration software. I closed this gap with this study and will discuss the drawn insights in the following. However, I do not claim exhaustiveness and emphasize the need for future in-depth investigations on this topic. The results suggest that administrators have high trust in the compound protocol HTTPS. Furthermore, it can be concluded that administrators are very satisfied with the use of Let's Encrypt, due to its automation and transparency in case of security breaches. Moreover, the study emphasized that the participants put high trust in Let's Encrypt's free CA. In contrast, Certbot is currently not very trusted among administrators. Reasons might include that (i) using Certbot is not necessary to obtain a free Let's Encrypt certificate and therefore, it is not that well known, (ii) users had negative experiences with its auto-configuration functionality (see Section 4.3.3, quote of participant P8) and (iii) it requires root privileges in its standard-setting putting users off.

Recommendation 1: *To overcome entry barriers, a slimmed-down version of Certbot could be provided, which among others does not require root-privileges in its standard settings.*

4.4.4 Let's Encrypt and Certificate Validation

The introduction of Let's Encrypt made it possible to secure connections to web servers using HTTPS without any financial burden. Due to its open-source code, the strong community of experts behind it and its popularity, more than 150 M Let's Encrypt certificates have been issued so far. The study showed that the use of Let's Encrypt is generally supported by the majority of companies, but is somewhat more common among smaller companies. The free creation of certificates means that even smaller websites, regardless of whether they are created for business or private use, can offer HTTPS connections. To provide more security by limiting the damage of key compromise and misissuance, the lifetime of Let's Encrypt certificates is set to 90 days. However, the short(er) lifetime also poses some challenges for the administrators. The results show that although there is the possibility to automatically update certificates, some participants do not do so since they consider it as an additional burden. Users are encouraged by Let's Encrypt to automate the certificate renewal process [1], however, for some web servers, this needs to be done manually or by changing the default settings. Some administrators (5%) forgot to automate their certificate renewal process, which could result in invalid certificates and broken services.
Recommendation 2: *Therefore, a suggestion for improvement would be to automatically enable the renewal process for all ACME clients and web servers and visually alert the administrator when this option is not enabled.*

The validation levels of certificates and their security guarantees were known by most of the participants. However, still, 12% did not know which validation their certificate provides. A possible reason for this knowledge gap could be that those administrators used Let's Encrypt, which only provides DV, hence no in-depth knowledge of certificate validation is needed by its users. When using Certbot, manual DV or an appropriate DNS plugin is the only way to obtain wildcard certificates[7]. Unfamiliarity of the validation guarantees of a certificate, which was discovered by 12% of the participants, could lead to wrong decisions for web servers with higher certificate requirements.
Recommendation 3: *In order to overcome this lack of knowledge, the validation provided within the process of certificate generation in the ACME client, should be visually or textually emphasized.*

4.4.5 HTTPS Challenges

Although many participants are well aware of the security risks of older TLS versions, there are still many websites allowing older (i.e., insecure) versions. Fahl et al. [63] found that two-thirds of their participants deliberately used non-validating certificates due to compatibility issues or because their web server was used exclusively for testing purposes. Furthermore, some web servers were not intended to be publicly accessible and their

[7] A certificate that contains a wildcard character (*) which can be used for multiple sub-domains of a domain.

administrators were not even aware of the fact, as they could only be found with the help of a web crawler.

In this study, one-fifth of the administrators reported using certificates with poor security settings since they were forced to do so to support old servers and operating systems, or company or client requirements. I argue that the discrepancy between the higher number of insecure servers found in Fahl et al.'s study and this might have two main reasons: i) I did not crawl the internet and therefore, the data solely relies on self-reporting, and ii) the introduction of Let's Encrypt and Certbot (or other ACME clients) helped administrators to freely and correctly configure their certificates. Fahl et al. revealed as part of their findings a wish list of administrators in order to make the configuration process of X.509 certificates for HTTPS web servers easier. Let's Encrypt has realized this wish list.

Challenges regarding the TLS deployment process with an Apache web server were first investigated by Krombholz et al. [108] whose work was later extended by Bernhard et al. [19]. Both found that the process is very complex and only a part of their participants was able to deploy a secure configuration. Bernhard et al.'s comparison between Let's Encrypt and manual deployment methods suggests more effective HTTPS deployments when using Let's Encrypt, which takes less time and achieves slightly better SSL Lab grades. The study revealed that administrators using Let's Encrypt felt less confident that their website is configured in a secure way in comparison to other CA users. In line with Krombholz et al.s' results, one-third of the administrators did not strictly enforce HTTPS, 7% had problems with properly securing their communication (usage of weak ciphers), and 5% perceived error messages as too generic and unclear.

Tiefenau et al. [179] concluded that Certbot is an asset in the field of fast, easy, and secure TLS configuration. One of the remaining weaknesses they mentioned is its transparency. This study cannot substantiate or quantify this weakness as no administrator explicitly mentioned it. However, some participants mentioned that it is sometimes difficult to change (standard) settings as it is unclear what the current settings include.
Recommendation 4: *Therefore, providing an "advanced" option, which allows the user to change settings in more detail and get a clear overview of current settings is suggested.*

Some participants indicated that they reviewed their website with a rating tool, such as SSL Labs, to determine if they have configured it securely.
Recommendation 5: *Since there are still some websites with insufficient TLS security, ACME clients like Certbot should automatically perform a rating during the configuration process. This way administrators can actively decide if they are satisfied with the rating and the associated settings.*

4.4.6 Administrator vs. Student Sample

Computer science students are often used as a convenience sample for expert users [197, 4, 136, 108]. Since related work used different methodologies than used in this study (quantitative vs. qualitative) and had different technical assumptions/conditions, no

direct comparisons can be made. However, in the following, interesting observations are discussed obtained from triangulating the findings.

In order to get the most accurate picture of the problems faced by administrators, it is necessary to understand their prerequisites. Students have a different demographic distribution compared to administrators. On average, the administrators in this sample are ~ 15 years older. Furthermore, not all student participants of prior studies (30-40%) had experience with the work of a system administrator or the configuration process of TLS.

Although this study has yielded similar results compared with previous user studies which used student samples, there is no conclusive evidence that administrators can be replaced by students. In fact, this study uncovered additional results including the need for "advanced" settings in ACME clients, the barrier of requested root privileges of Certbot (in its standard settings), and the confidence in Let's Encrypt and Certbot. Therefore, I conclude that in order to get an accurate picture of the specific challenges faced by professional administrators, a convenient sample of computer science students should not be used as a substitution.

However, I argue that student samples can be helpful for the detection of difficulties with HTTPS in a broader sense, since Let's Encrypt and ACME clients enable anyone to configure HTTPS due to reduced complexity and no financial cost.

CHAPTER 5

Mental Models of Cryptocurrency Systems

Disclaimer: *The contents of this Chapter were published as part of the publication "User Mental Models of Cryptocurrency Systems-A Grounded Theory Approach". Sixteenth Symposium on Usable Privacy and Security (SOUPS) 2020 [128]*

More than ten years after the first Bitcoin transaction was performed [60], cryptocurrencies have gained popularity among different types of users, ranging from technology enthusiasts to investors, gamblers, and people who are simply curious. Media reports often contain anecdotes of negative user experiences with security and privacy in cryptocurrency systems. Cryptocurrencies obviate the need for central control by maintaining a decentralized public ledger. While technical aspects of cryptocurrencies have been heavily studied (e.g., [22], [75], [154], [14]), human-centered studies are still rare. Gao et al. [74] used semi-structured interviews to explore specific aspects of the Bitcoin system out of context. Krombholz et al. [107] quantitatively examined user perceptions of Bitcoin security mainly based on closed-ended questions. However, so far no research investigated mental models which are based on users' *tacit knowledge*. Such knowledge consists of implicit and subjective assumptions that cannot easily be verbalized but are heavily influencing human behavior [100].

This study extends the state of the art by providing the first qualitative insights ($N = 29$) of people's **mental models of cryptocurrencies**[1] **and associated security and privacy threats**. Therefore, drawing and card assignment tasks were used to investigate the mental models. The study methodology follows an inductive approach based on Grounded Theory (GT) [116, 80, 175].

[1]The focus of this study was on Bitcoin and Ethereum since they were the most prevalent cryptocurrency systems in terms of market capitalization [41] at the time of this study.

This study answered the presented research questions RQ1-RQ4 regarding the cryptographic protocol-bases systems Bitcoin and Ethereum which are both based on a Proof-of-Work (PoW) protocol. The stakeholders of this study are similar to the first study (see Chapter 3 end-users. However, as these monetary systems are not suitable and intended for children, the target participants for this study are adult users.

The goal of this study is to explain (some of) the reasons for user-caused security incidents. The findings are necessary in order to re-design tools and create effective strategies for behavior change. I argue that cryptocurrency tools (e.g., wallets, online exchanges) should be designed in a way to avoid security or privacy risks even when used by people with incorrect or incomplete mental models. This is in line with Wash et al. [190] claiming that instead of attempting to force users into more 'correct' mental models, technology should be shaped to work well with existing mental models.

Through this study, it was possible to identify the gaps between the actual protocol functionality and users' mental representations. Although not all the incorrect or incomplete mental models found imply security pitfalls, some partly explain why users of current cryptocurrency tools fail to securely manage their digital assets and have wrong assumptions about privacy and anonymity. Mental models with negative consequences include an erroneous understanding of cryptocurrency systems with regards to (i) cryptographic keys, (ii) anonymity, and (iii) fees.

5.1 Background Information on Blockchain and Cryptocurrencies

The blockchain is an example of a Distributed Ledger Technology (DLT) which stores the records in a decentral and publicly verifiable manner. Thereby, the chain continuously grows by storing records in interlinked blocks as can be seen in Figure 5.1. Each block contains the hashed header of its previous block and the Merkle root, which is the (pairwise) hash of all transactions included in this block.

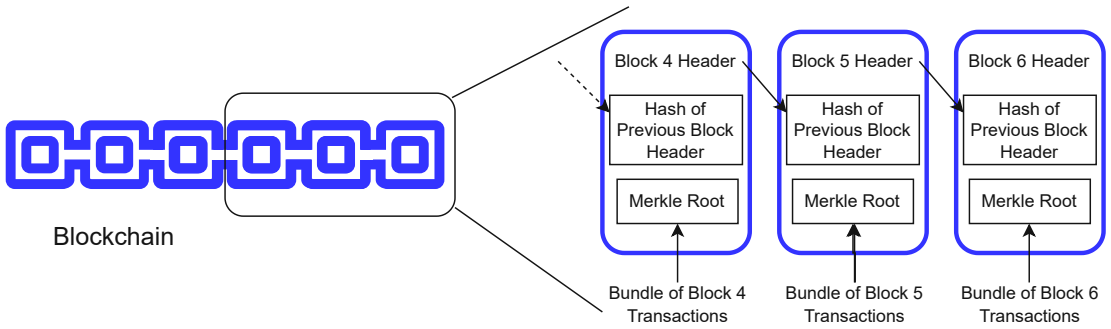


Figure 5.1: Basic structure of blocks in the blockchain

Currently, the most widely used application of blockchain technology is cryptocurrencies. Cryptocurrencies require, in contrast to conventional monetary systems, no central control, and offer mostly pseudonymity (there are some cryptocurrencies and additional measurements offering anonymity, however, they are not within this thesis research scope). Bitcoin and Ethereum are the leaders among existing cryptocurrencies in terms of market capitalization [41].

Bitcoin was introduced in a white paper by Satoshi Nakamoto [138] and launched at the beginning of 2009. Although Bitcoin was not the first attempt to introduce a virtual currency, it was and still is the most successful one, and thus became a pioneer and predecessor for all other cryptocurrencies. Ethereum was introduced by Vitalik Buterin [29] and launched in 2015. It is an open-source distributed software platform, which enables amongst others the implementation of smart contracts and decentralized applications.

Both Bitcoin and Ethereum currently use the Proof of Work (PoW) consensus protocol, which is also known as the mining process. It works by proving that a certain amount of computation was necessary to calculate the next block. Therewith, it ensures that a block can only be created with a certain amount of effort, preventing manipulation attacks.

5.2 Related Work on Cryptocurrency User Studies

Cryptocurrency systems differ from other public key systems (e.g., secure messaging or time-stamping services), as keys are used to sign transactions that are transparently published in the blockchain, instead of ensuring confidentiality through encryption. The threat model is entirely different as well: losing a private key leads to severe problems in the context of cryptocurrency systems as monetary assets are involved. A plethora of research has been carried out to study the security and privacy aspects of the Bitcoin system [203]. Nevertheless, several user-centric challenges remain, providing a breeding ground for security and privacy threats.

Only a few studies have examined the usability and user perceptions of cryptocurrency systems, mainly focusing on Bitcoin. Baur et al. [18] found that users attribute a high potential to cryptocurrencies, but perceive the usefulness of current cryptocurrencies as low. According to Khairuddin et al. [102], the major motivation for users to buy Bitcoin is its potential for financial revolution, increased user empowerment, and its use as an investment. Sas and Khairuddin [165], as well as Lustig and Nardi [126], explored the trust issues of Bitcoin users. Elsden et al. [57] proposed a typology of emerging blockchain applications making it easier for users to understand them. Gao et al. [74] conducted a qualitative study where they found that the major entry barrier for non-users is a perceived necessity for profound technical knowledge.

The first large-scale quantitative user study was presented by Krombholz et al. [107], revealing that many users neither understand nor use the security capabilities of coin

management tools correctly. Although the authors also conducted a small number of qualitative interviews, those were only used to contextualize their quantitative findings, not to construct an inductive theory. Kazerani et al. [99] investigated the influence of (poor) usability of cryptocurrency management tools on the adoption of Bitcoin by lay people. Eskandari et al. [58] compared the usability of different cryptographic key management approaches.

However, these earlier studies either opted for a quantitative study design [107] or asked questions that the interviewees deemed too complex to answer given their background as non-users [74]. At the time of this study, it was the first mental model study on cryptocurrencies that aimed at discovering the tacit knowledge of the participants. Therewith, answering open questions on why users commonly fail to manage private keys safely in the context of cryptocurrencies and which parts of current cryptocurrency tool interfaces put users with incorrect mental models at security or privacy risk. Based on the findings, suggestions for future designs of cryptocurrency tools on how to ensure that user behavior does not compromise the users' security and privacy are provided.

5.3 Methodology of the Cryptocurrency Study

The overall goal was to understand user perceptions and misconceptions of functional principles, and whether they prevent users from using cryptocurrencies in the most secure and privacy-preserving manner. In this study Bitcoin and Ethereum are used as examples of prevalent cryptocurrencies. The Ethereum's smart contract functionality is excluded to only focus on its native currency ether. Furthermore, payment channels are out of the scope of this study. Therefore, general assumptions can be made about user perceptions with regard to the majority of cryptocurrencies that build on the same functional principles as Bitcoin and Ethereum (i.e., in relation to key generation and usage, transaction generation and confirmation, blockchain application, and mining operations). For the remainder of this Chapter, I will thus use the term *cryptocurrency* to refer to bitcoin, ether, and similar cryptocurrencies.

5.3.1 Grounded Theory

The study follows an inductive approach and uses GT [116, 80, 175] to explore user perceptions based on qualitative data. GT is a set of systematic inductive methods to develop theories that are grounded in qualitative research data. A key characteristic is that it merges data collection and analysis in an iterative approach until (theoretical) saturation is reached [175]. Therefore, different phases of recruitment and coding are necessary (see below). By following a process that emphasized direct and immediate analysis of the collected data, it was possible to generate descriptive theories that are as close to reality as possible. GT is traditionally used in social sciences and has gained popularity in human-computer interaction and usable security research [106, 73, 96].

5.3.2 Recruitment

The goal was to recruit a diverse sample of current and potential future cryptocurrency users. Potential interviewees were approached through Bitcoin mailing lists and social media as well as personal contacts, in order to get in touch with organizations that work with blockchain technology.

In the beginning, a short description of the study was distributed including a questionnaire (Appendix A.3.1) for preselection. To prevent potential participants from reading up on the technical intricacies of blockchain technology, the concrete purpose of the study was not disclosed, only that it deals with cryptocurrencies. From the people who completed the questionnaire a subset of participants fitting the recruitment criteria were selected.

The participants were chosen according to their self-reported level of knowledge about cryptocurrencies and information technology (ranging from lay users to experts) as well as their usage of cryptocurrencies. Thereby the emphasis was to recruit participants with diverse exposure to and interaction with cryptocurrencies. The recruited participants consisted of 7 people who were not actively using cryptocurrencies but who were working with cryptocurrencies in their professional life (e.g., organizing cryptocurrency meetups, conferences or projects with wallet/exchange operators). Further 10 participants considered cryptocurrencies mainly as an investment, 5 used them mainly for trading, and 7 actively used cryptocurrencies as a payment method.

While the self-reported data might not fully reflect the actual knowledge level of participants, I am confident that these measures are sufficiently accurate to reflect the inherently diverse target population and that a diverse sample was obtained.

5.3.3 Sampling

GT [175] requires going back and forth between data collection and analysis in order to construct a theory that is derived from data and not chosen a priori (as is the case in quantitative studies). Following GT, the selection of participants was conducted in two rounds (two weeks apart). First, an initial sample of 18 cryptocurrency users (experts and non-experts) was collected, and then explored the obtained data through open coding.

Based on the concepts derived from the analysis, the initial sample was extended with people who are not actively using cryptocurrencies themselves, but work in institutions that use or deal with cryptocurrencies or blockchain technology (see Section 5.3.2). Since these people were confronted with cryptocurrency tools, at least at a superficial level, they have certain mental models but are not influenced by cryptocurrency tool interfaces. These mental models are particularly interesting as they represent perceptions of (potential future) first-time cryptocurrency users for whom cryptocurrency tools should be designed as well. By comparing non-users to users, it was possible to explore how cryptocurrency tool interfaces might influence mental models (cryptocurrency tool bias) and also investigate biases of non-users (bank bias). For the second round of recruitment,

Table 5.1: Participant demographics. Total N=29

Demographics		Participants (%)
Gender	Male	19 (65.5%)
	Female	10 (34.5%)
Age	18 – 22	1 (3.4%)
	23 – 27	12 (41.4%)
	28 – 32	10 (34.5%)
	33 – 37	4 (13.8%)
	38 – 42	2 (6.9%)
Highest Education	High school	5 (17.2%)
	Bachelor degree	10 (10.4%)
	Master degree	14 (72.4%)

additional data was collected by recruiting a sample of 11 participants based on emerging theories.

The final set consisted of 29 participants. The study participants' age, gender, and education distribution are summarized in Table 5.1. In order to protect the privacy of the participants, their age was queried, beginning with 18 years, in intervals of five years.

5.3.4 Design and Procedure

As shown by Kearney and Kaplan [100], people commonly construct implicit knowledge maps to understand complex systems when the systems' functionality goes beyond their technical knowledge. They argue that such tacit knowledge influences people's decision-making and behavior in critical situations, although they are often not aware of it. I opted for a study design that encourages participants to expose their tacit knowledge and functional understanding by engaging them in drawing and card assignment exercises. During these exercises the participants had to assign cards with a function (e.g., "sign transaction" or "generate public key") to the entities in their drawings.

Based on related work on human factors of Bitcoin [155, 58, 102, 57] and recent mental model studies in usable security [96, 74, 73, 206, 190, 162], an interview script was constructed for semi-structured interviews including a short pre-assessment questionnaire covering demographics and the participants' general cryptocurrency usage patterns, two drawing tasks, and a card assignment task. The complete study material can be found in the Appendix A.3.

The final dataset is based on 29 interviews which were conducted in person in two Austrian cities, namely Vienna and Graz. The interviews took place either in a lab room at SBA Research, the participants' workplace, or their home. The majority of the interviews were conducted by two researchers (one interviewer and one assistant taking notes). Two interviews were performed by only one interviewer due to scheduling conflicts. With the informed consent of the participants, all interviews were recorded, and fully transcribed, and all drawings and card assignments were photographed. The pictures (along with the transcribed verbal explanations) served as the baseline for coding.

Pilot Studies

Altogether three pilot studies were conducted to test the expressiveness and practicability of the interview script. At the end of each iteration, feedback was requested from the respective participant. Thereby the participants were asked to explain their understanding of selected functional concepts and well-known buzzwords – e.g., blockchain, [de-]centralized system, miner – if the participants had not drawn or mentioned them during the interview. After completing the other two pilot studies (with a slightly changed interview guideline incorporating the insights from the first iteration) the study design phase was conculed as the interview guideline yielded good insights into user mental models. As the interview guideline was not modified anymore after the first pilot study the last two pilot studies were included in the final sample (29 studies in total), a.

Interview Procedure

Before the actual interview started, participants were briefed, signed a consent form, and received their compensation (20 Euro Amazon voucher). Each interview lasted roughly 30 minutes and consisted of semi-structured questions, two drawing tasks, and one card assignment task. These tasks were based on a concrete scenario, namely transferring a certain amount of bitcoin or ether to a fictional friend called Alice.

In the course of the first drawing task, the participants were asked to visualize and verbally express all components and actors involved in the transaction process, as well as their connections. The participants were encouraged to think aloud while drawing to gather additional insights into their reasoning. Afterwards, the participants were provided with 15 cards with short descriptions of selected functions (e.g., "generate private key", "generate transaction", "validate transaction"; see Appendix A.3.2). Depending on which cryptocurrency the participants were more familiar with (self-assessment), the cards reflected the terminologies from either bitcoin or ether. The task was to assign the cards to the components and actors in their drawings (see Figure 5.2). There was no full definitions but asked the participants to verbalize their own understanding of these technical terms and the associated context.

The card assignment task was added to assist the participants in refining and contextualizing their tacit knowledge. In order to eliminate the possibility of misunderstood terms and random guessing during the interview, the participants were encouraged

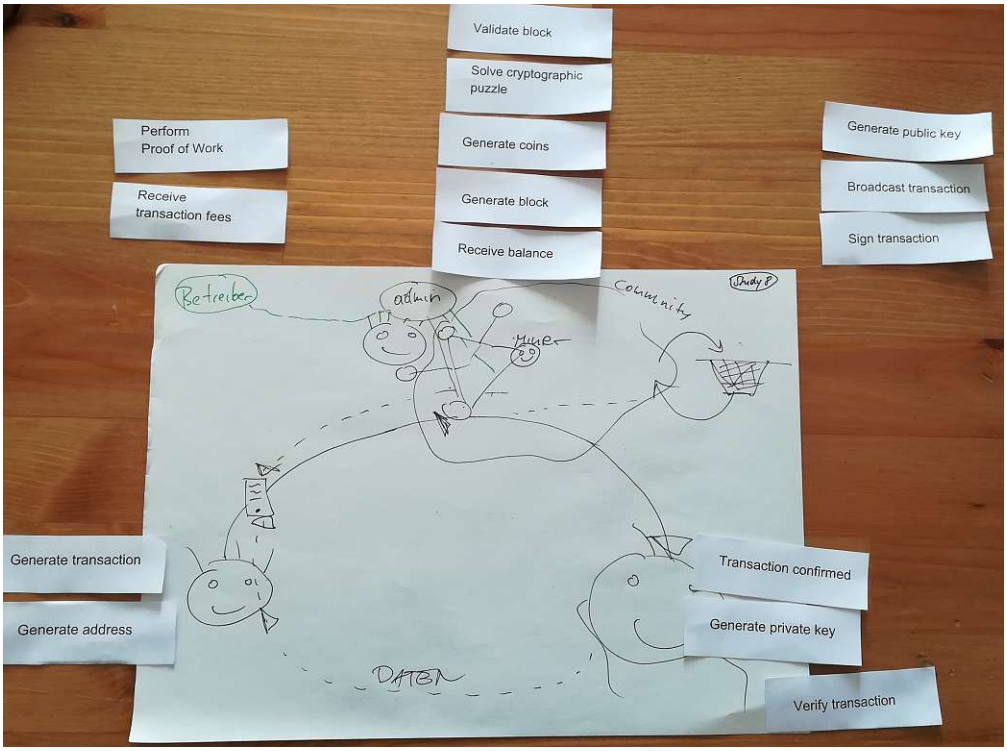


Figure 5.2: Card assignments to the components and actors of the drawn transaction process (S8).

to provide detailed definitions as far as possible and the interviewers asked follow-up questions if further clarification was needed. I am confident that this procedure was beneficial to evaluate the participants' mental representations of cryptocurrencies and their understanding of associated terms.

The second drawing task was used to elicit understandings of attackers and threats, and how specific security and privacy risks are contextualized in transaction processes. The interview protocol can be found in the Appendix A.3. In order to ensure a correct understanding of the wordings used by the participants, the interviewer paraphrased the statements and asked detailed questions if necessary for clarification.

5.3.5 Ethical Considerations

One of the fundamental requirements of the institutional ethics guidelines is to preserve the participants' privacy and limit the collection of sensitive data as much as possible. Therefore, IDs were assigned to study participants which were used to refer to participants throughout the study and analysis process. Before the interview, all participants were asked to sign consent forms in which the goals of the study and data handling procedures were described. Those consent forms were stored securely and do not contain any links to the IDs assigned to each participant. The study furthermore strictly followed the EU's

General Data Protection Regulation (GDPR), and the information gathered throughout the interviews (transcripts, notes, and drawings) was only stored until publication, but at longest for one year after the study.

5.3.6 Coding of the cryptocurrency mental model study

In the following the coding procedure of the study data (transcripts, pictures, and drawings) is presented. Thereby, a GT-based approach was followed to interpret the data. The coding steps of: i) open and axial coding, ii) selective coding, and iii) final coding are described in detail. Finally, the step of theory building and mental model construction is described.

Open and Axial Coding

After the first 18 interviews, two researchers independently coded the data (initial *open coding*) with the aim to group recurring statements and assertions relating to the same phenomena (preliminary categories). The created codes are based on the drawings and think-aloud protocols. I refrained from assigning codes based on single denotations or terms (e.g., verify, confirm, encrypt). Instead, entire statements were coded and hence included the context in which a term was used.

In line with the full GT approach and while discussing the results, it was decided to extend the participant pool by including people who do not actively use Bitcoin but work in a field related to cryptocurrencies. I am confident that the perceptions and opinions of those people add a new perspective to the study outcome since they have no practical experience in using Bitcoin or Ethereum – and, as such, are not influenced by interfaces of exchange platforms or wallets – but possess some (theoretical) knowledge about blockchain technology. Moreover, this sample's knowledge structures are particularly relevant when thinking about the design of future technology for managing cryptocurrencies, since those people are either potential new users or decision-makers in corporate environments.

Following the GT methodology, the second round of open coding was performed to refine the results of the first round. Three researchers independently coded the entire data set in three rounds (Round 1: 10 interviews, Round 2: 10 interviews, Round 3: 9 interviews). In order to systematize the process, *affinity mapping* was applied, whereby the interview transcripts were cut into snippets and sticky notes were used (see the affinity mapping process in Figure 5.3) to label newly found categories that emerged from recurring or related statements. The interviews were coded based on contextualized statements (instead of single terms) and based on those a codebook was created denoting the participants' mental representations and reasoning. After each round of individual coding, the relations between the newly found categories were discussed until an agreement upon a set of higher-level categories wa reached (*axial coding*). For instance, a decision was to group the categories "public key generation external", "key and address independent", and "one public key for all" to the meta-category "address/key

generation". The categories of the third round of coding served as a baseline for *selective coding*.

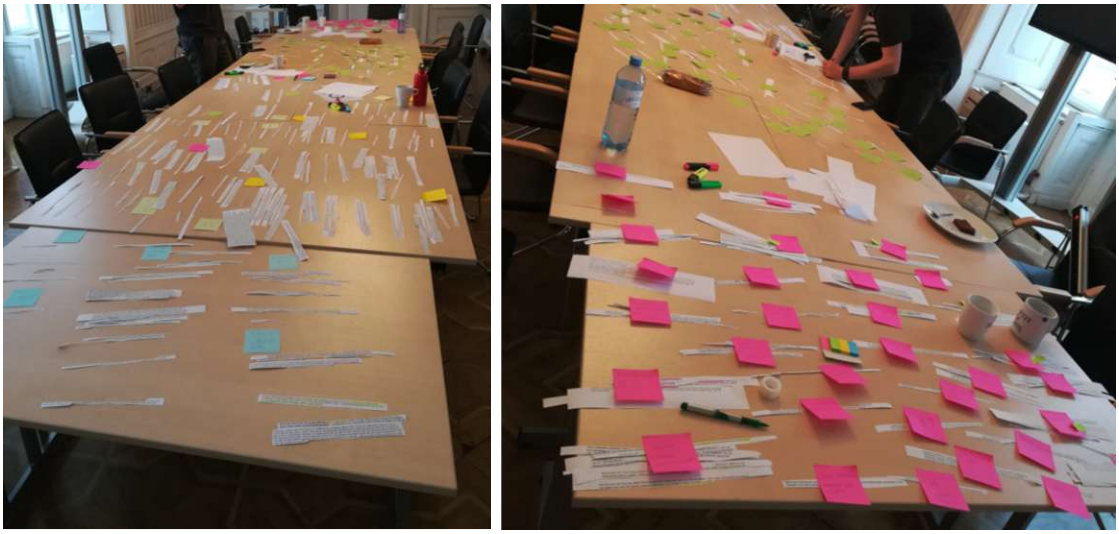


Figure 5.3: The process of affinity mapping with sticky notes.

The final sample size was determined by reaching saturation [82], i.e., no new insights could be gathered from the interviews. As saturation was achieved in the newly emerging categories, the interviewing process was stopped after 29 participants.

Selective Coding

It is the process to figure out the core categories which include all data. During this process, three independent researchers agreed upon a set of final core categories centered around the identified misconceptions which might compromise the users' security and/or privacy. The misconceptions are grouped into the following four different top-level categories which consist of multiple subcategories (final codebook see Figure 5.4[2]):

- ***Meta*** – This category includes statements that were meaningful for building theories, but are not directly related to cryptocurrency systems and their functionality. It comprises general opinion changes during the interviews, prerequisites, statements about the control or power of the system, biases that influenced participants' descriptions, and misconceptions related to encryption and hashing.
- ***System*** – The system category includes statements describing the blockchain (*Blockchain Description*) as well as where and how it is stored (*Location*). Additionally, this category is split into *Structure*, *Behavior*, and *Function*. *Structure* includes descriptions about the connection between users and miners. The category *Behavior* refers to the behavior of the system (e.g., who receives fees or who

[2]For category B all correct answers are marked with an asterisk *

A Meta	B.4.4 Verification	**C Privacy**
A.1 Bias	B.4.4.1 all peers*	C.1 Attacks
A.1.1 economy	B.4.4.2 blockchain	C.1.1 anonymous
A.1.2 wallet	B.4.4.3 peers: majority/n	C.1.2 identity disclosure @ 3rd party
A.1.3 bank	B.4.4.4 central	C.1.3 address mapping
A.1.4 exchange	B.4.4.5 user	C.1.4 doxxing
A.2 Preassumptions	B.4.5 Transactions: Generation	C.1.5 endpoints
A.3 Opinion Change	B.4.5.1 directly written in blockchain	C.1.6 attacker: state
A.4 Control/Power	B.4.5.2 by user/wallet*	C.1.7 attacker: system participant/s
A.5 Encryption Misconception	B.4.5.3 by blockchain	C.1.8 attacker: external
B System	B.4.5.4 by miners	C.2 Prevention
B.1 Blockchain Description	B.4.6 Transactions: Propagation	C.2.1 self-initiated: mining
B.1.1 system/software	B.4.6.1 client sends to all	C.2.2 self-initiated: info inquiry (3rd party)
B.1.2 algorithmus/actor	B.4.6.2 client sends to part*	C.2.3 self-initiated: address generation
B.1.3 deletable	B.4.6.3 central	C.2.4 self-initiated: identity hiding
B.1.4 datastructure: all transactions*	B.4.6.4 storage pool	C.2.5 self-initiated: anonymized shopping
B.1.5 datastructure: parts	B.4.6.5 direct	C.2.6 user does not care
B.2 Location	B.4.7 Transactions: Confirmation	**D Security**
B.2.1 storage: chunked	B.4.7.1 n blocks*	D.1 Attacks
B.2.2 storage: copied*	B.4.7.2 blockchain	D.1.1 no/secure
B.2.3 central	B.4.7.3 Alice	D.1.2 human failure
B.2.4 internet	B.4.7.4 central	D.1.3 man-in-the-middle
B.2.5 distributed: all*	B.4.7.5 n miners verify	D.1.4 hacking: endpoints
B.2.6 distributed: nodes with shares	B.4.7.6 through fees	D.1.5 hacking: central entity
B.3 Structure	B.4.8 Coin generation	D.1.6 hacking: exchange
B.3.1 User-system connection	B.4.8.1 miner*	D.1.7 DoS
B.3.1.1 automatic *	B.4.8.2 coins equal fees	D.1.8 mining majority
B.3.1.2 cloud	B.4.8.3 exchange	D.1.9 future technology/ theoretical
B.3.1.3 central	B.4.8.4 central	D.1.10 price: volatility
B.3.2 Miner: Connection internal	B.4.8.5 all/system/blockchain	D.1.11 price: manipulation
B.3.2.1 fully connected	B.4.9 Address/Key Generation	D.1.12 attacker: state
B.3.2.2 connected graph*	B.4.9.1 pub key generation: miner	D.1.13 attacker: external
B.3.2.3 pools*	B.4.9.2 pub key generationl: system	D.1.14 attacker: miner
B.3.2.4 not connected	B.4.9.3 one public key for all	D.1.15 technical failure
B.3.2.5 master/server	B.4.9.4 key/address: independent	D.1.16 social engineering
B.3.3 Miner: Connection external	B.4.9.5 send priv key: cloud	D.2 Prevention
B.3.3.1 miner equals blockchain	B.4.9.6 send priv key: between users	D.2.1 system initiated
B.3.3.2 separate system	B.4.9.7 client/wallet generates keys*	D.2.2 self initiated: software
B.3.3.3 user cannot mine	B.4.9.8 key/address: dependent*	D.2.3 self initiated: behavior
B.4 Behavior	B.4.10 PoW/Crypto Puzzle	D.2.4 self initiated: hardware
B.4.1 Fees: recipient	B.4.10.1 encrypted code solving	D.2.5 helplessness
B.4.1.1 cryptocurrency operator	B.4.10.2 find pre-computed value	D.2.5 helplessness
B.4.1.2 exchange	B.4.10.3 compute blockhash*	
B.4.1.3 user	B.5 Function	
B.4.1.4 all	B.5.1 Key	
B.4.1.5 miner*	B.5.1.1 sign: me*	
B.4.1.6 broker and miner	B.5.1.2 sign: Alice	
B.4.1.7 rich participants	B.5.1.3 sign: Alice and me	
B.4.2 Fees: amount	B.5.1.4 sign: Miner	
B.4.2.1 user selected*	B.5.1.5 sign: other system participants	
B.4.2.2 admin selected	B.5.1.6 signing equals transaction verification	
B.4.2.3 miner selected	B.5.1.7 approval	
B.4.2.4 fixed	B.5.1.8 access blockchain	
B.4.3 Block Generation	B.5.1.9 private key equals address	
B.4.3.1 user	B.5.1.10 private key is wallet/account password	
B.4.3.2 blockchain	B.5.2 Address	
B.4.3.3 central	B.5.2.1 nickname	
B.4.3.4 miner*	B.5.2.2 payment destination*	

Figure 5.4: Cryptocurrency mental models - Final Codebook

generates a block). *Function* on the other hand categorizes the tasks of the keys and the addresses.

- ***Privacy*** – This category codes all mentioned attacks and possible prevention mechanisms on users' privacy.

- ***Security*** – This category includes all attacks and possible prevention mechanisms specific to the users' security.

Final Coding

With a final set of codes grouped into categories, two researchers independently went through all 29 interviews and assigned one or multiple codes, thus generating a comprehensive codebook. Thereby, the transcripts, drawings, and outcome of the card assignment task served as a baseline. The inter-rater reliability was reported with a Krippendorff's Alpha value [105] of $\alpha = 0.89$, indicating a high level of agreement among the coders. This relatively high number is fostered by the technical classification and the granularity of the codebook. Conflicts mostly appeared due to slightly different interpretations of the drawings, which sometimes conflicted with the think-aloud protocols. When a conflict was detected, the drawings and transcripts were consulted, and the results were discussed again. In these cases, the researchers agreed that the verbal explanations should weigh more than the card assignments since the latter were sometimes less expressive than the participants' verbal descriptions. All conflicts among the coders were resolved.

Theory and Mental Model Construction

The last step of the GT approach was to form theories including the overarching mental models which describe how the participants perceive cryptocurrency systems. First, two independent researchers generated two draft mental models: one incomplete model and one inaccurate model for the structure, function, and behavior of components in cryptocurrency systems. The models were constructed based on the results, centered around those categories which resulted from the selective coding (codebook). Then, the two coders met in person to reach an agreement. The constructed mental models were validated through negative case analysis [27] by going through all interviews to check whether the participants' statements can be assigned to one of the draft mental models. If not, I sought to understand how they diverged from the draft and adopted it accordingly. In doing so, I iteratively refined the draft mental models until all statements could be assigned. A participant's mental model can contain aspects of the incorrect and incomplete mental model. In order to (i) construct a theory, (ii) examine whether the mental models interfere with secure and privacy-preserving usage of cryptocurrencies, and (iii) understand how resulting issues can be solved, a focus group (Section 5.4.7) was conducted with four experts in the field of cryptocurrencies and blockchain technology. The two final models are presented in Section 5.4.2 and Section 5.4.3.

5.3.7 Limitations of the Cryptocurrency Study

Participant recruitment via mailing lists, social media, and personal contacts ensured a diverse sample regarding age and profession for the study. However, the sample still has its limitations as it is biased towards a higher educated social stratum; also, non-users without any connection to cryptocurrencies were excluded. Furthermore, the recruiting area was limited to two cities in Austria. Therefore cultural differences to other countries/continents cannot be made. Furthermore, the European legal landscape with regard to security and privacy (GDPR) also most likely influenced the participants.

The interviews were conducted in German, which is why language-specific expressions in direct participant quotes may have been lost in translation. However, all direct translations were double-checked by a translator, which is why I am confident that such issues have been kept to a minimum.

This study followed an inductive approach to gather insights into user perceptions of cryptocurrency systems. However, the methodology also has its limitations as the data is self-reported and, in comparison to quantitative studies, the sample size is fairly small. Still, I feel confident that the sample is sufficiently large to observe general tendencies, as theoretical saturation was reached.

5.4 Findings of the Cryptocurrency Study

In this Section, first, a simplified description is presented of the Bitcoin and Ethereum system to provide the appraisal factors for the assessment of the data. Then, the participants' veritable mental models are presented in order to answer RQ1 and RQ2 (see Chapter 1). These models represent incomplete and inaccurate descriptions from the participants in correspondence to the structure, functionality, and behavior of cryptocurrency systems. Direct participant quotes (translated into English) are provided for illustration. Since quantitative results (numbers) in qualitative research cannot be used to generalize findings, all statements are discussed without providing numbers.

5.4.1 Appraisal Factors

Before conducting the study, a ground truth model was constructed together with two cryptocurrency experts. These experts were also part of the focus group. I do not claim exhaustiveness of the expert mental models which can be incomplete and diverse as well. To increase the validity, two experts were interviewed and one mental model was constructed incorporating both statements. Similar to the user interviews, both experts were asked to draw all components and actors involved in a transaction process and verbalize their thoughts. Afterward, a simplified representation of their mental models was constructed (see Figure 5.5). This model serves as a basis for the evaluation of user mental models and only focuses on important parts of user transactions. The study revealed that the participants' mental models were consistently sparser than the expert

mental model. The comparison of users' and expert mental models is purely illustrative and non-judgmental. The assessment basis was defined as follows:

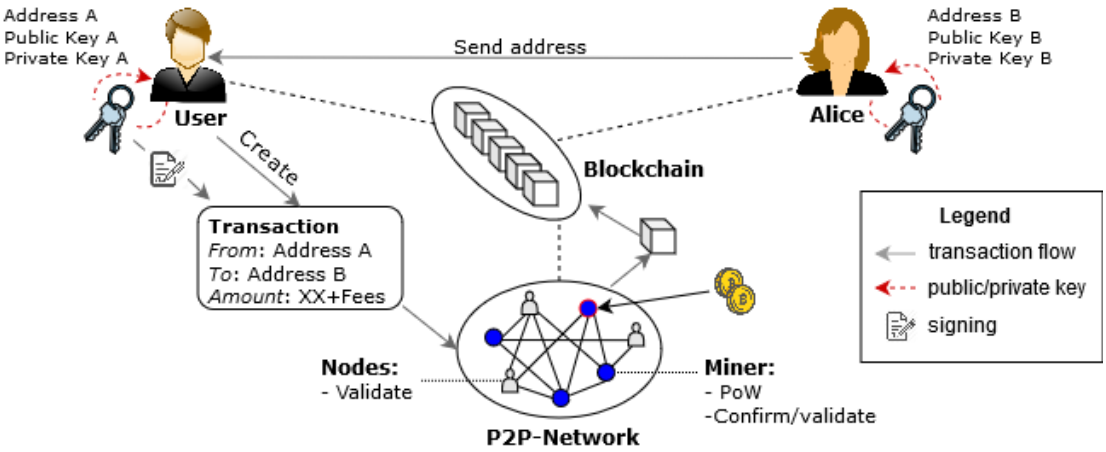


Figure 5.5: Ground truth of the transaction process of cryptocurrency systems based on blockchains.

Bitcoin and Ethereum are blockchain-based, Peer-to-Peer (P2P) networks which enable users to perform transactions with virtual (crypto-)currencies. The system consists of multiple participants (peers) that are grouped in four different roles: (i) sender, (ii) receiver, (iii) miner, and (iv) other users. Each participant can hold multiple roles.

The basic requirement to perform a transaction within a cryptocurrency system is that sender and receiver must have some kind of wallet or are enrolled with an online exchange service. A wallet consists of public keys, private keys, and addresses. The private key is randomly generated and permits the user to spend cryptocurrency units. In the case of offline wallets, it is the user's responsibility to securely store and back up the private key. The address is a hashed version of the public key and acts as a public identifier of the asymmetric key pair.

Prior to performing a transaction, the receiving party communicates its address to the sending party. The sender creates the transaction which comprises the sending and receiving address as well as the transferred amount, including fees. The amount of the fees can be selected by the user and determines the processing speed of the transaction (transactions with higher fees are more likely to be included within the next block). Afterwards, the sender signs the transaction with the private key and broadcasts it to the P2P network. The verification – for both the transactions and the blocks – is performed by peers in the network. Thereby, not all peers necessarily perform full validation (e.g., Simplified Payment Verification (SPV) wallets or thin clients do not check whether transactions are valid, but they rather evaluate whether full nodes have validated them correctly).

A specific transaction t is considered to be confirmed when (i) a miner successfully constructed the PoW for a block b containing t, (ii) b ends up in the heaviest chain

(i.e., the chain with the most cumulative PoWs), and (iii) a certain amount of blocks is succeeding b (as the blockchain gets longer, the confirmation can be considered to be more secure). The miner (or mining pool, i.e., a cluster of miners that work together) who solves the PoW first gets rewarded with newly created currency (a specific amount depending on the implementation of the system) and the transaction fees. As soon as the transaction is confirmed, the amount is credited from the sender's to the receiver's wallet.

5.4.2 Incomplete Mental Model

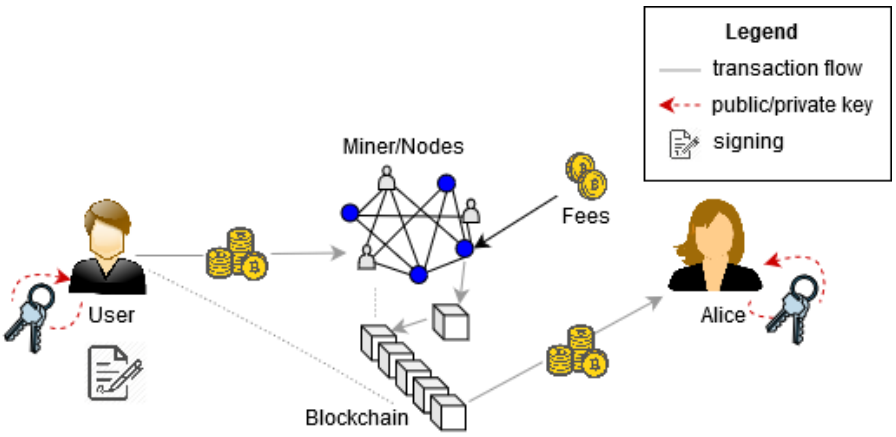


Figure 5.6: Incomplete cryptocurrency mental model

Figure 5.6 depicts the best-case mental model grounded in the qualitative analysis. It includes technically correct yet sparse perceptions compared to the ground truth (see Figure 5.5). I did not encounter poor decision-making as a result of incomplete mental models, hence the missing details are not crucial for the secure usage of the cryptocurrency system.

Several users correctly stated that cryptocurrency systems are decentralized with annotations reflecting an outline of a P2P system, and a transaction flow matching the ground truth (illustrated through lighter grey continuous lines in Figure 5.6). The majority correctly stated that the user's wallet software generates the public/private key pair (illustrated by a dotted red line in Figure 5.6). Some of them also knew that in order to send coins to another party, the sender has to sign the transaction with the generated keys. They correctly mentioned that an address is the payment destination in the proposed scenario. Many participants correctly understood that miners receive transaction fees. However, only a few participants knew how fees are actually calculated and could give a correct explanation of the mining process.

5.4.3 Inaccurate Mental Model

The mental model presented in Figure 5.7 incorporates the participants' misconceptions of cryptocurrency systems. However, not all illustrated components are reflected in

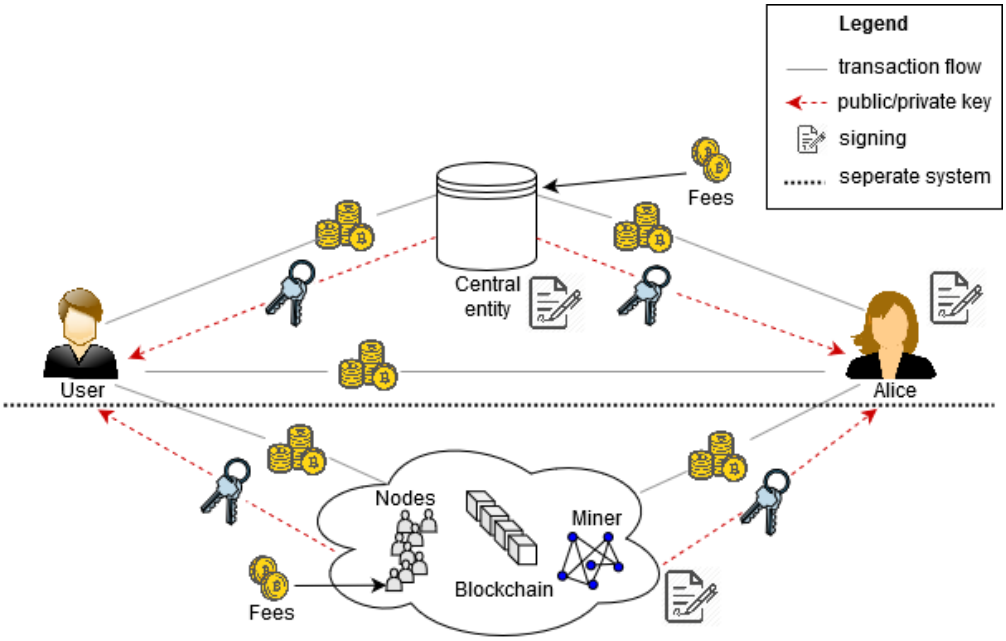


Figure 5.7: Inaccurate cryptocurrency mental model

all mental models of the participants. Misconceptions related to the transaction flow are illustrated by a grey, continuous line, and those related to the key generation are shown through dashed, red lines in Figure 5.7. I found that many misconceptions do not jeopardize users' security or privacy. In the following crucial and non-crucial misconceptions are discussed.

Some participants assumed a central management entity as part of a cryptocurrency system, such as a server or broker. Others thought that a direct end-to-end connection existed between sender and receiver, via which transactions are performed.

> *It is a decentralized system because there is no pivotal element. Only the two accounts interact with each other directly without a third person interfering.* (S6)

One participant hypothesized that in addition to an end-to-end user connection, a further connection to a cloud exists through which users can get initial approval for transactions in order to afterward send confirmed transactions directly to the receiver. Participants with incorrect mental models often described the blockchain, other nodes, and the miners either only through keywords without being able to explain them, or as a separate system or cloud. Therefore in the graphic, they are depicted as a cloud and (partly) separated system (see the bottom half in Figure 5.7). In the following Sections, these misconceptions and their impact on security and privacy are discussed in detail.

Cryptographic Keys

Many misconceptions related to the keys used in cryptocurrency systems were identified during this study. Although users' problems with cryptographic keys (and their management) have already been investigated for other application areas – for example secure messaging and Pretty Good Privacy (PGP) – the effects of mistakes from the users' perspective are different for cryptocurrency systems (i.e., direct monetary impact). In particular, it became apparent that participants do not understand who generates the keys. Some participants claimed that the miners carry out key generation or expected the whole cryptocurrency system to generate keys.

> *Hmmm well, I don't generate my private keys myself, they are saved in my smartphone app. It is... generally the blockchain who generates it [the key] for me, or the network, the blockchain. It is floating in the air somehow. I don't know. It comes from the internet.* (S19)

One participant thought that all parties in the Bitcoin system share one common key. This would break cryptocurrency systems because everybody would have access to everybody else's funds. Other participants presumed that in order to send coins between two parties, the users' private keys have to be sent to "the cloud".

> *I generate my private key and send it to the cloud. Then I get back [from the cloud] a public key... I must be able to rely on the channel to be secure, e.g., encrypted, when I send my private key to the cloud.* (S22)

Through contextual information from this interview, it became clear that S22 was not referring to storing a key in the cloud (which would be a correct mental model), but sending it to the cloud in order for the blockchain to get decrypted.

One participant assumed that they have to send the private key directly to the recipient. This would crucially harm the user's security as it would enable the receiver to have access to the sender's account. As most of the participants were not aware of the fact that the private key is generated on their side, they also did not understand that the private key should never be exposed to external entities (such as miners, central entities, or other system participants).

The participants also lacked an understanding of the signing process. Some stated that the receiver has to sign the transaction. Others thought that both, the sender and the receiver have to sign. A few participants inaccurately stated that the miners have to sign transactions. Several participants assumed that other end-users in the system are signing transactions in order to validate them. One participant stated that other end-users as well as the miners have to sign a transaction. A participant claimed that a user's keys were necessary in order to access the blockchain. As a result, users frequently did not

understand why and how they should keep their private key safe, given that they did not understand what a private key can be used for.

During the study, many incorrect card assignments and descriptions not matching the ground truth model in relation to cryptocurrency addresses could be observed. It was unclear to many participants what a cryptocurrency address actually is. One participant thought that the private key is a user's Bitcoin address. This misconception is especially severe as it might encourage a participant to share the private key with other participants. Many participants assumed that the generated keys and cryptocurrency address are entirely independent. Some participants assumed that the address is a form of nickname, similar to a pseudonym that you choose on a message board.

Such misinterpretations of key generation and usage can have a major security impact if cryptocurrency tools delegate the responsibility of key generation or management to the users without providing guidelines. If due to misconceptions, users make their keys accessible to others, they become susceptible to theft.

Fees

The participants expressed many incorrect assumptions about how fees are calculated and what their purpose is. A few participants explicitly stated a lack of knowledge in this regard. One participant thought that fees are defined by an administrator, two said that the miners select the amount. Others stated that the amount of fees is fixed.

> *Miners ask for transaction fees, I don't know if I can choose the amount...If I want to send money and I am in a hurry, for example in the case of smart contracts, then it is possible that the miner knows that I am in a hurry and the miner adds an exorbitant amount of transaction fees [to my transaction].* (S20)

As a result of such misconceptions, users might pay transaction fees that are too high in comparison to the amount that would have been needed to fulfill their requirements, if no guidelines are provided by cryptocurrency tools.

Anonymity Misconceptions

During the coding process, further themes related to anonymity in cryptocurrency systems emerged from the collected data which are not directly related to the generated mental models. A few participants assumed that transactions stored in the blockchain are deleted after some time.

> *After 8 blocks one blockchain is ready and it becomes one instance...Then, the old one is deleted.* (S8)

This entails a wrong assessment of privacy features offered by the blockchain. Participant S8 perceived the blockchain as oblivious and drew a garbage can where old transactions are disposed/recycled (this is only correct within the Lightning network [157], which S8 was not referring to). S8 stated that it is not possible to store a too big amount of data in the blockchain. Many participants incorrectly assumed that the cryptocurrency system applies some form of encryption by default. The participants imagined that either the blockchain, the transaction, or the transaction channel between the end points is encrypted.

> *The transaction itself must be encrypted to ensure a secure connection between server and client.* (S29)

One participant argued that the cryptographic puzzle or hashing is an en-/decryption operation necessary to get access to the money which was sent.

> *Alice receives the cryptographic puzzle, but I don't know what happens if she can't solve it...Because, I mean the bottom line is, I encrypt it [some kind of transaction code] and she receives it.* (S17)

Furthermore, some participants thought about encryption as a major factor used for security purposes. However, those participants commonly also stated that it is necessary to have some kind of additional knowledge in order to pursue this kind of cryptographic task.

> *I guess you can encrypt them [the transactions], however, I do not know how.* (S20)

These misconceptions violate participants' privacy as they incorrectly assume that information in the blockchain is unreadable by the public. Moreover, in line with the findings by Gao et al. [74], it might discourage people from using cryptocurrencies when they are under the assumption that only participants with cryptographic knowledge are able to correctly apply privacy or security measures (i.e., encrypting the blockchain).

5.4.4 Mental Models of Security Threats and Prevention

Most of the participants were able to explain a broad spectrum of (potential) security risks. The majority of the participants mentioned threats related to compromised end points (e.g., mobile phones or computers), which are indeed present as shown in a Kaspersky [98] report. However, this threat is not limited to cryptocurrency applications. Furthermore, the participants named mining majority attacks (i.e., an attacker controlling more than 50% of the mining power in the network). No mining majority attack has yet been performed on Bitcoin or Ethereum, although Bitcoin Gold experienced a 51% attack

in May 2018, and a theoretical approach of an Eclipse attack on Ethereum has been described by Yuval et al. [131]. Therefore, there is a possibility that such an attack could happen in a larger cryptocurrency system, especially when ownership and mining are increasingly concentrated on a small group of people [158].

Many participants referred to attacks related to human failure, such as people losing their private keys or failing to store keys in a secure way. This is in line with results from Krombholz et al. [107] and newspaper articles [97] providing evidence that key loss is often caused by the users themselves. Some mentioned the threat of online exchanges being hacked, which has indeed been reported frequently [115]. Others mentioned price fluctuations or intentional price manipulation (e.g., through fake news or other social media posts) as risk factors. Furthermore, some participants correctly stated that Denial-of-Service (DoS) attacks on cryptocurrency systems [31] pose a potential security risk.

In contrast, some participants revealed an incorrect understanding of the threat landscape in cryptocurrency systems and described attacks which are not feasible in a decentralized system. A few participants stated that hacking of central entities, such as the miners, full nodes, or parts of the P2P network is feasible (e.g., with a master computer, see Figure 5.8). Some described Monster-in-the-Middle (MITM) attacks as a possibility, where an attacker interferes or manipulates the transaction process and possibly alters information (e.g., the recipient's address). Related to that, one participant stated that attacking one single peer would critically harm the whole system.

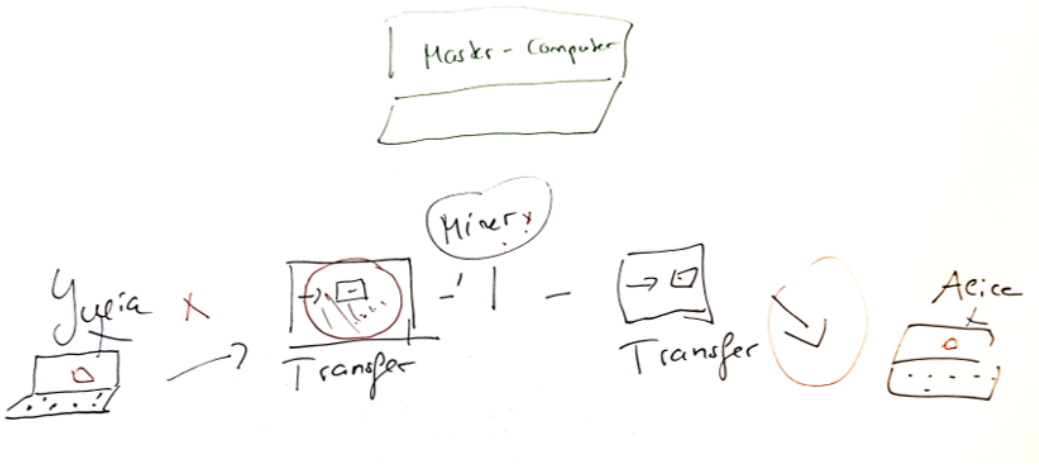


Figure 5.8: Risk assessment of the transaction process (S17)

> *It doesn't matter which point is attacked, as soon as one point is attacked a chain reaction is started and everything starts to collapse like dominoes.* (S2)

Other participants reported not to be aware of any security risks and to consider cryptocurrency systems to be secure by design. Some of the participants assumed

theoretical threats such as broken or weak cryptography that might expose users to a security risk.

Related to *prevention mechanisms against security threats*, more than half of the participants mentioned self-initiated behavior (such as storing private keys securely). Moreover, many referred to the usage of specific hardware (e.g., hardware wallets) and mentioned software (e.g., secure wallets) as a remedy against security breaches. In relation to that, participants described possible prevention mechanisms initiated by the cryptocurrency system, thinking that users cannot influence their execution. Many participants described feeling helpless as they do not think that (technically non-adept) users can actively apply any measures to circumvent such threats.

> *Maybe I can keep a low profile and I shouldn't sit in the tram with the app because of shoulder-surfing... As a non-professional I cannot really do more.* (S22)

5.4.5 Mental Models of Privacy Threats and Prevention

Some participants assumed that they are anonymous when using cryptocurrencies. However, the majority mentioned address mapping as a possible privacy threat, which is indeed possible [13, 104]. Thereby they either explained that social media and internet activity can be used to correlate the location and activity of a user to a cryptocurrency payment, or that as soon as an address is connected to a specific person, all other addresses of this person can be correlated as well.

The second biggest privacy threat people mentioned was identity disclosure through third parties, since it is often mandatory to provide identification when purchasing or exchanging cryptocurrencies. Doxxing (writing private data into the blockchain) and a privacy-threatening attack of the end points (e.g., hacking of the private smartphone or computer) were also mentioned. Notably, potential future attacks with the help of quantum computers or artificial intelligence were referred to by several participants. Some thought that the state might be a possible attacker or named external persons with bad intentions as relevant attackers. In contrast, others thought that the system participants themselves might carry out attacks on their privacy.

With respect to *prevention mechanisms against arising privacy threats*, participants referred to the possibility to mine themselves in order to prevent identity disclosure when buying cryptocurrencies. A few participants explained that it is possible to buy cryptocurrencies from a specific third party that does not require identity disclosure. One participant assumed that the usage of two-factor authentication would ensure privacy:

> *To secure myself against the threat that IP addresses can be mapped [to bitcoin addresses], I use two-factor authentication.* (S7)

Some explicitly stated not to care about the prevention of privacy threats as they do not consider them important or do not assume that privacy issues exist in decentralized systems.

5.4.6 Tool Bias

Many cryptocurrency users focused their explanations and drawings of the transaction process on the graphical user interface which they are exposed to when performing transactions, either via a mobile wallet, a PC wallet, or an online exchange. The study revealed that wallet interfaces shaped the way participants perceived the blockchain location (centralized vs. decentralized), its functionality (persistent, transparent), and the users' role within the cryptocurrency system.

Figure 5.9 shows a drawing (example 1) that is influenced by the interface displayed to users when carrying out transactions via mobile phone. In particular, this study highlighted that the participants were frequently influenced by a feature of the interfaces currently used by many online exchanges and wallets (Figure 5.9 example 2). Thereby, the current number of confirmations is displayed to show how many blocks are already successfully mined and incorporated in the heaviest chain of the blockchain. After a specific number of succeeding blocks the current transaction is marked as "accepted". However, from the study, it can be deduced that users commonly misinterpret these confirmations as a specific number of miners or peers who signed, approved, or validated their transactions. Even among experts, a specific fixed number of confirmations is assumed, although the security actually depends on the weight of the longest chain (see Sompolinsky and Zohar [171]).
In contrast to the cryptocurrency tool bias for cryptocurrency users, a **bank bias** was discovered for non-users. They often stated that the blockchain is centrally managed or that transactions are conducted directly between users.

5.4.7 Expert Focus Group

In order to construct a theory and answer RQ3 (see Chapter 1), the security and privacy impact of the participants' mental models were discussed in an expert focus group which consisted of four members from a different research group at the institution who are primarily researching blockchain technology. One researcher led the discussion and two researchers took notes and asked follow-up questions. First, the incorrect model was presented to all participants including printouts of the visualization. The experts were asked whether they could think about any security or privacy risks related to this model. Then, the incorrect model was discussed in three rounds based on the categories resulting from the selective coding (keys, fees, anonymity misconceptions). In each round, first, the identified misconceptions were presented, and then the participants were asked whether they think that these categories interfere with security and privacy. If the answer was yes, the experts' opinions were asked on how these security problems could be prevented. The discussion and improvement suggestions for cryptocurrency tools are based on the outcome of this focus group.

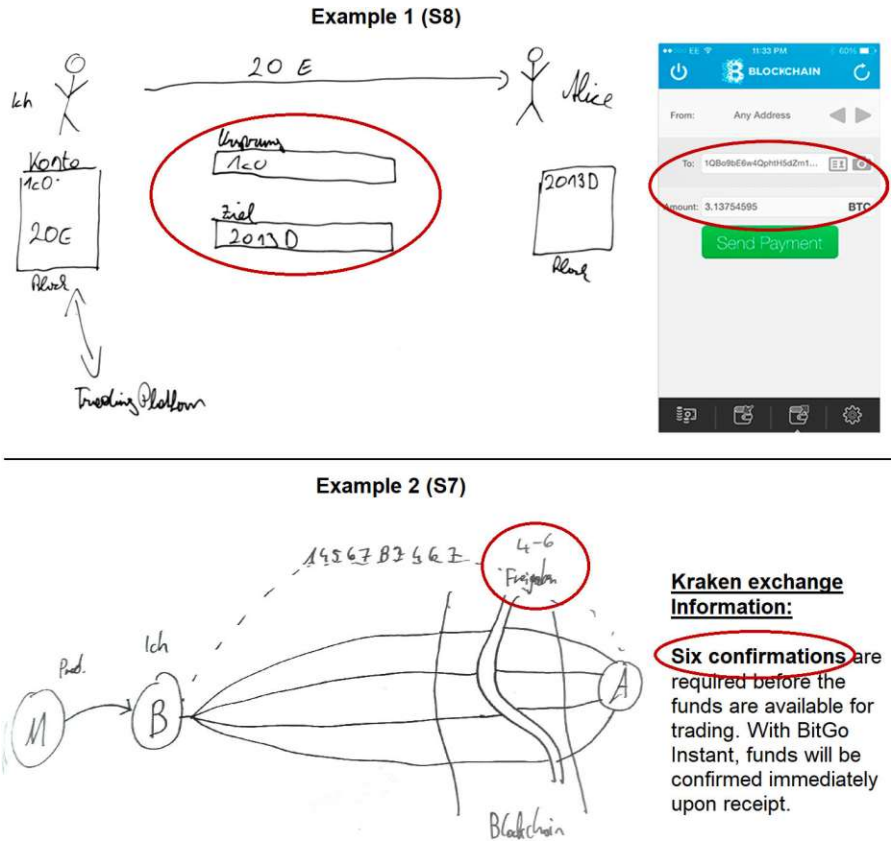


Figure 5.9: Illustrating cryptocurrency tool bias

The researchers decided on a final set of categories that are important for the theory generation since they have a direct impact on users' security and privacy. These categories are (i) **keys**, (ii) misconceptions regarding **anonymity**, and (iii) **fees**. The resulting mental models are centered around these aspects of the participants' mental representations. The anonymity misconceptions only emerged from the participants' descriptions of the transaction process and were not reflected in their drawings.

Regarding the questions of how cryptocurrency tools could prevent security problems caused by incorrect mental models, the focus group brought up the challenge of designing tools that are adapted to diverging mental models. There is a thin line between an easy-to-use system and a system that gives (expert) users the feeling of being too simple to be secure, and also provides too little information to evaluate the system. Therefore, the focus group proposed that the user interfaces of cryptocurrency tools should have options to switch between different levels of complexity, providing the user with the chance to interact with the system and obtain detailed information about it only if desired. This approach has been (partly) implemented by Coinomi [42] (see Appendix A.3.3) and should be a standard feature for all (future) wallets.

It is currently hard for users to decide how secure a cryptocurrency is since its security characteristics are not evident. The study's results suggest that most users base their decision of which cryptocurrency to buy on the amount of fees or their popularity in the media. Some effort has been made by Zhang and Preneel[207] to evaluate the attack persistence of Proof of Work consensus protocols. However, at the time of this study, there was no publicly available platform or ranking that could serve as a guideline for users to easily compare the security of different cryptocurrencies.

5.5 Discussion of the Cryptocurrency Study

The results explain the roots of several misconceptions with impact on security and privacy found in related work [74, 58, 107] and can be directly linked to concrete improvement suggestions for cryptocurrency tools (e.g., wallets or exchanges).

I claim that modifications of the interface of cryptocurrency management tools can prevent security and privacy threats caused by incorrect mental models. This claim is based on the observation that there is a **cryptocurrency tool bias** of cryptocurrency users (see Section 5.4.6). The results indicate that users' mental models are influenced by the interfaces of tools and technologies they use, which will be subject to further research.

While I believe that it should not be expected from (lay) users to understand complicated details of cryptocurrency systems, I argue that as much knowledge is necessary as crucially needed to circumvent security and privacy flaws. Results are classified as important, which actively affect a user's behavior in the cryptocurrency system. This study suggests that (i) the function of the private key, (ii) the transparency of the blockchain, and (iii) the function of fees have to be understood correctly by users to act in a secure and privacy-preserving way when using cryptocurrencies.

5.5.1 Challenges and Improvement Suggestions

This study revealed a wide range of mental models, from very detailed to sparse and from correct to incorrect. In order to answer RQ4 this Section describes improvement suggestions and remaining challenges. In general, improvement suggestions include – in line with the outcome from the focus group – designing cryptocurrency tools adapted to diverging mental models and different user groups (e.g., experts and non-experts). Therefore, cryptocurrency tool providers ought to offer **different levels of complexity**.

In the following, based on the results from the study and the expert focus group, a discussion is provided on how current cryptocurrency tools should be adapted to allow people to use them in a secure and privacy-preserving manner, irrespective of their (incorrect) mental models.

Anonymity

About a quarter of the participants used the term "encryption" when describing a transaction process in cryptocurrency systems. Many participants stated that the blockchain is encrypted. I hypothesize that these users mixed up authentication/signing (which indeed takes place during a transaction process) and encryption. Most of these participants assumed that encryption is a safety measure against security- or privacy breaches. Moreover, many participants presumed that transactions cannot be tracked due to the encryption of the blockchain. Such misconceptions can potentially jeopardize people's privacy as some participants were incorrectly assuming that their information is hidden from the public or that all information is deleted after some time. The results suggest that people with these misconceptions refrain from taking measures to safeguard their privacy while believing that they are anonymous.

Furthermore, this study revealed misconceptions about the persistence of the blockchain. From discussions with industrial partner institutions, it can be inferred that blockchain technologies are commonly applied in areas where it does not make sense, such as for ephemeral data. The mental models found in the course of this study explain such a contradiction.

Recommendation: *Interfaces of cryptocurrency tools should illustrate the openness, persistence, and transparency of the blockchain. For example, a block explorer could be integrated, visualizing in which block a transaction is integrated and how many succeeding confirmed blocks currently exist. Some wallets (see Appendix A.3.3) provide access to textual block explorers as an additional feature; however, there is no graphical visualization integrated into the wallets. Furthermore, a pop-up could be shown before pursuing a transaction, stating that this transaction will be broadcasted in clear text to the cryptocurrency network and no information can be altered later on.*

Cryptographic Keys

Previous research on public key cryptography for e-mail encryption has shown that users have difficulties managing and understanding asymmetric keys [192, 163]. The study supports this finding as less than half of the participants were able to correctly describe how keys are generated and used. Until the time of writing, no holistic solution has been proposed to solve these issues. Bitcoin and Ethereum use keys differently than for example PGP (i.e., it is only used to sign data instead of also encrypting it) and come with an unexplored and diverse user group. Nevertheless, no research has been conducted so far to examine how people understand the function of keys in the context of cryptocurrency systems.

The results show that less than half of the participants were able to correctly describe how keys are generated and used. These misconceptions directly influence the way users manage their keys, thus putting them at risk for monetary loss and fraud.It could be observed that many users did not draw a connection between their private key and the ability to carry out transactions from their accounts. They either (i) did not know

that a private key existed (e.g., when using online or offline wallets), (ii) misunderstood the functionality of the private key, or (iii) only had an abstract idea of the private keys without reflecting on its purpose in this system. Moreover, this study discovered misconceptions in relation to the key generation. Those ranged from assuming the public key has to be sent to the cloud to obtain a private key, to the supposition that the miners or the blockchain generate keys. I suppose that these incorrect perceptions interfere with secure key management if users are not aware of the fact that private keys give access to their funds and should be known only to their owner, hence being kept safe locally.

In line with research on usable key management in other domains [164, 163], it can be suggested to automate tools as far as possible so that users do not have to deal with key generation or key back-ups, while still providing as much transparency and information as needed to not expose users to security or privacy risks (for a feature overview of key storage and back-up systems from popular wallets, see Appendix A.3.3).

For cryptocurrency systems this means that users must at least understand that their seed phrase (or private key) (i) should not be shared with anybody else, and (ii) can currently not be recovered in case of loss, leading to the loss of all funds. These facts should be emphasized to the user during wallet initialization and whenever the wallet is used, as discussed in the above Sections.

Recommendation: *In order to avoid that users lose their seed phrases, wallets should enforce seed phrase back-ups by asking the users to input a certain number of words from their phrase after making a copy (e.g., writing it down on paper, taking a picture, copying it on a USB device). Furthermore, wallets should ask users to enter their seed phrases in specific time intervals to ensure that they maintain access. Many wallets did not implement these features at the time of this study (see Appendix A.3.3). Alternatively, using automatic key recovery is suggested (e.g., similar to trusted friends [62]).*

Fees

Currently, many cryptocurrency tools only offer one fixed amount for fees. The results show that due to this practice, the majority of participants do not know that users can actively select how much they want to pay as mining fees during the creation of a transaction. Therefore, users are not aware that it is in their power to select how quickly their transaction will be included in the blockchain. As a result, users might pay transaction fees that are too high in comparison to the actual amount needed for their requirements. Furthermore, when a specific wallet always uses the same amount of fees, this also comes with a privacy problem, since it is then possible to track back the used wallet.

Recommendation: *User interfaces of cryptocurrency tools should remind users that by choosing the amount of transaction fees they can influence how quickly their transaction will be included in the next block. I recommend providing the user with two options i) precomputed fees based on heuristics (leading to different amounts for each user and transaction) which are labeled with understandable terms (e.g., "slow—low fees", "default"*

and "fast—high fees") and ii) the option to set the fees him/herself. A comparable approach is provided by the Blockchain [21] and Coinomi[42] wallet (see Appendix A.3.3).

Security and Privacy Threats and Prevention

This study revealed that while the participants showed a basic understanding of the threat landscape in cryptocurrency systems, their knowledge about possible prevention mechanisms was poor and led to a feeling of helplessness among half of them. These participants either believed that users cannot take any measures, but need to rely on the system, or they assumed that prevention mechanisms (e.g., wallet encryption, hardware wallets) can only be pursued by technologically knowledgeable users. This coincides with the results found by Krombholz et al. [107] which showed that many users do not apply security measures offered by state-of-the-art cryptocurrency tools.

Recommendation: *Cryptocurrency tools should perform encryption by default and inform the users about this safety measure (see Appendix A.3.3 for the status of popular wallets). Most people use cryptocurrencies only rarely, especially when the main purpose of the use is investment or curiosity. This makes it difficult for users to build up trust in the system and develop a deeper understanding of it. Moreover, they should add cues and visualizations to explain to the users which security measures (e.g., encryption, backups) are implemented so that users can make informed trust decisions.*

CHAPTER 6

Mental Models of Self-Sovereign Identity Systems

Disclaimer: *The contents of this Chapter are part of the paper "Expert Mental Models of SSI Systems and Implications for End-User Understanding" which is currently under submission*

The daily lives of most people in today's society are characterized by a large number of online interactions with various services (e.g., email, social media, messaging, and online shopping). 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 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 [173, 169, 148, 174], while human-centered research is still rare. This led to several hundreds of prototypes and Proofs-of-Concept (PoC), 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 6.1), 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, an expert mental model study of SSI was conducted, 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) 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.

With this study I sought to answer RQ1, RQ3 and RQ4 for the cryptographic protocol-based system SSI. Similar to the HTTPS study presented in Chapter 4 the stakeholders investigated are experts, however, it was not a pre-requisite to actively implement the protocol. In contrast to the other studies, RQ2 can't be directly answered, as currently no standards and widely used systems exist in the wild of SSI, in order to make a general comparison between the actual structure and functionality and the found mental models.

In order to answer those research questions, I:

- conducted a qualitative expert study investigating expert mental models of SSI and its corresponding threat landscape,
- constructed guidelines that need to be considered for future SSI designs in order to prevent security and privacy threats and,
- extracted a minimal knowledge map for end-users, which need to be understood for safe usage and broad adoption of the technology.

6.1 Background and Current Status Self-Sovereign Identity

SSI empowers an entity (e.g., individual person or organization) to have complete control over the usage of their digital identity without requiring the permission of an intermediary or a central party. Through this control, users have the opportunity to regain their privacy, which is currently being severely invaded by e.g., Facebook, Instagram, or Google [176, 167]. In 2016 the Sovrin Foundation was among the first to propose an SSI system in their whitepaper [180].

Currently, most identity management architectures in the wild are central or federated solutions and are designed for a single purpose without possible application extensions. The functionalities of these solutions range from authentication to single sign-on and digital identity management. One example of such an authentication application is the widely used "A-Trust" smartphone signature in Austria. Other applications in the DACH region are for example NetID, and Verimi (single sign-on solutions). In contrast to the currently used centralized and federated solutions of identity management, in a decentralized system, the risk, trust, and privacy factors shift (see Figure 6.1) from a third party to the user(s) themselves.

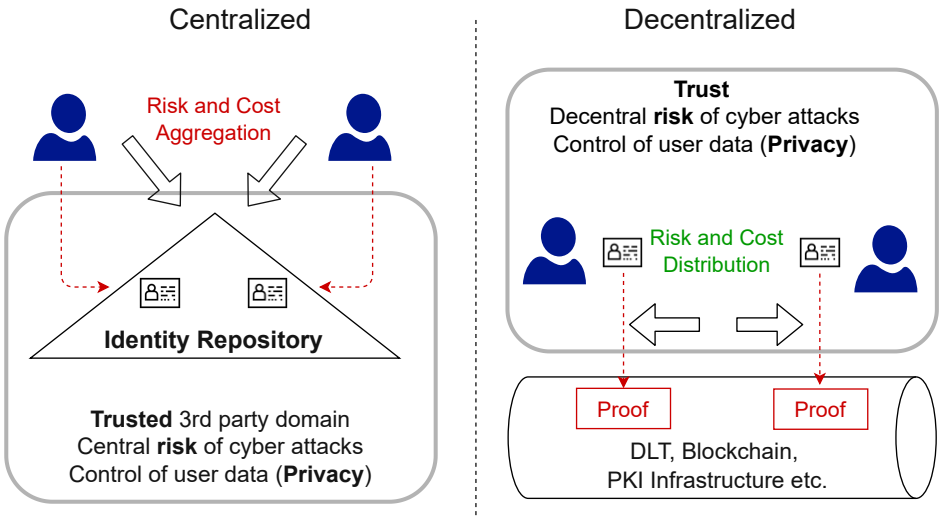


Figure 6.1: Digital Identity - Centralized vs. Decentralized Systems [1]

At the time of this study, no comprehensive and generally accepted definition of SSI existed. Therefore this Section reflects 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 [11] in 2016 and the Sovrin Foundation grouped them into three main categories as can be seen in Table 6.1. Those principles were refined by different aspects such as provability [174] and secure transactions [181]. 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 [134]. A basic SSI architecture has three main actors:

- Issuer: an organization (e.g., bank, university, government agency) issuing certain verifiable credentials or claims for the holder/user (e.g., driving license, birth certificate, educational degree)
- 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 or the driving license allowance during a vehicle stop/inspection), and verifies the validity of the credential

[1] Graphic adapted from Source: `https://www.citi.com/ventures/perspectives/opinion/digital-identity.html`

Table 6.1: The 10 principles of SSI categorized by the Sovrin Foundation [180]

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

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 storage controlled by the holder/user. The storage of the data can be offline and local (e.g., on the user's smartphone or a hardware token) and/or online (e.g., Google cloud, OneDrive).

Currently, there exist three standards for (possible) components within an SSI system that are agreed on by the community: i) Verifiable Credential (VC) [2], ii) Decentralized Identifier (DID) [3] and iii) decentral ledger technologies (DLTs) [4]. While these constitute a first step towards the standardization of SSI systems, there are still interoperability, communication, and usability standards missing among others.

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

[3]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.

[4]DLTs are cryptographic protocols and infrastructures, which enable decentral storage of information without a central authority.

6.2 Related Work on Self-Sovereign Identity 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.[30]. 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 [180] in 2016.

Following the whitepaper, a plethora of prototypes and PoC were proposed, both from the research community and the industry. Thereby, the approaches reached from blockchain-based [174, 148, 69, 137] to local or distributed storage solutions [12, 48, 10].

Mühle et al. [134] gave an overview of SSI system components and discussed the state-of-the-art in 2018 based on Zooko's Triangle [193] (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 [174] proposed an SSI solution for permissionless decentralized digitized passwords. Othman and Callahan [148] developed a decentralized credential storage option via blockchains of biometric data for authentication. Freytsis et al. [69] 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. [12] 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.[88] found that current state-of-the-art approaches lack i) standardization, ii) real-world usable solutions and iii) the consideration of the human factor (i.e., examination of the usability of the application and the understanding of the benefits and downsides of a decentral system from the user's perspective). In general, the concept of SSI has received very little attention in the realm of usable security, and no mental model study has been conuducted at the time of writing. One of the very few studies in usable security was conducted by Laatikainen et al. [112] investigating the adoption barriers in a holistic approach by using field research methods to investigate practitioners and researchers. Thereby they found that one of the adoption barriers is diverse interpretations and understandings of SSI. At the time of this study no mental model study investigating the systems functionalities was published, to the best of my knowledge. The study presented in this Chapter builds on the findings of Laatikainend et al. [112] by investigating in depth the understanding and perception of SSI from an expert point of view, to pave the path for standards and a general definition of SSI, where current and future systems can be built upon.

6.3 Methodology of the Self-Sovereign Identity Study

In the following, the methodology used for the study of mental models of SSI is described. Thereby the goal of the study was to qualitatively inquire 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 experts' mental models, an iterative approach of data collection and analysis (see Section 6.3.4) was used, as it is common in qualitative research [147].

6.3.1 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, DLT and DID.

The recruiting was performed by distributing an Email to potential participants without disclosing the concrete purpose of the study. After the interview, the participants were asked whether they have further contacts with other SSI experts, thus using a snowballing sample technique [111] to recruit further people following approaches from peer-reviewed papers in the area of usable security [189, 106, 128].

The study was conducted in two rounds with a period of the preliminary analysis in between. In the initial round, six participants were recruited which were personal contacts (two from the industry and four from the research sector). The second round consisted of a further seven participants (six from the industry and one researcher) chosen from the referred contacts provided by the initial sample, leading to a total of 13 interviews. The participants demographics are summarized in Table 6.2, displaying their education, current profession, and whether the participant worked with(in) a standardization working group. 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).

My recruiting method might not generate a representative sample of all SSI experts in terms of residence, gender, and culture. However, I am confident 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, no new insights from the last two interviews were observed, suggesting theoretical saturation [80] which is why no additional interviews were conducted.

6.3.2 Procedure

The study design consisted of a semi-structured interview protocol that helps participants to expose their mental models, based on drawing exercises that proved to be very helpful in the previous studies (see Chapter 3 and Chapter 5). The interviews lasted on average

Table 6.2: Participant demographics ($N = 13$)

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

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.

The complete study design can be found in the Appendix A.4.1. For a smooth start, I first asked the participants about their educational background as well as their current job. Then non-technical questions were asked about the participant's experiences with and impressions of SSI. Afterward, three tasks were used to specifically probe their understanding of an SSI system, its components, actors, and their connections:

- In the first task the participants were asked 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 participants' 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 the interviews, the 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 the first six interviews were finished, preliminary open coding was performed in order to determine whether specific topics needed further investigation or clarification. Therefore, in the second round of interviews, a slightly modified interview protocol was used 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.

6.3.3 Prestudy

In order to test the comprehensibility and practicability of the interview guideline, two prestudies were conducted. Feedback was requested from both participants after the interview by asking whether questions were unclear or made them feel uncomfortable and thereby 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.

6.3.4 Data Analysis

In line with other qualitative studies in usable security [106, 128, 73, 206], a grounded theory-based approach was followed as proposed by Corbin and Strauss [44] to analyze the interview data. It is an iterative method with the goal to systematically analyze and interpret 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 the 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, the prestudies were coded in order to get a better understanding of emerging themes that correspond to expert mental models of SSI. Following the approach of Wash [189], I assembled a list of main themes, which I expected to see in the upcoming interviews. Those themes included information about the registry, issuer, 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 6.3.2). Following the second round of interviews, another round of open-coding was conducted and afterward all codes were combined. The found codes (axial coding) were abstracted and afterward, they were summarized by using affinity mapping into a final codebook (selective coding). With the codebook, all interviews were coded again and frequency tables were created in order to identify patterns and formulate theories (e.g., minimal knowledge map, recommendations). The drawings were used for visualization purposes which additionally informed the codebook, however, as some drawings needed extra explanation the transcripts were used when depictions were unclear (e.g., no textual explanation of a depicted "person", therefore the interviewer 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).

6.3.5 Ethical Considerations

One of the main requirements, in order to follow the ethical guideline of SBA Research, is to preserve the security and privacy of participants. Therefore throughout the study IDs were assigned to each study participant and the collection of sensitive information was limited as far as possible (only age, gender, the current area of residence and work, as well as the educational background and current profession were asked) which are not stored together with the IDs, strictly following the EU's General Data Protection Regulation (GDPR). Also, the interviews were only recorded when explicitly permitted by the participant. Furthermore, the recordings were deleted after transcription and the transcripts were kept at most 6 months after the study was held.

6.3.6 Limitations

By choosing a qualitative user study and an GT approach over other possibilities, I accepted certain trade-offs. Due to the qualitative nature of this study, statistically valid statements can not be made and therefore the results can't be generalized for every SSI expert and every (future) SSI system.

Furthermore, the sample was not particularly large. 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.

Besides the methodological limitations, 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.

6.4 Results of the Self-Sovereign Identity Study

In the following Section, I describe the identified mental models which were found in the data of the expert interviews (transcripts, notes, and drawings) in order to answer RQ1 and RQ3. The purpose of qualitative research is to explore phenomena and perceptions in-depth, rather than to generalize and quantify. Therefore, in this Section, no numbers are reported for the different mental models, but instead, the range of expert perceptions is described in detail.

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.

6.4.1 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 [11] 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, an expert mental model of SSI was constructed (see Figure 6.2). 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 describe the components and connections of the mental models in more detail.

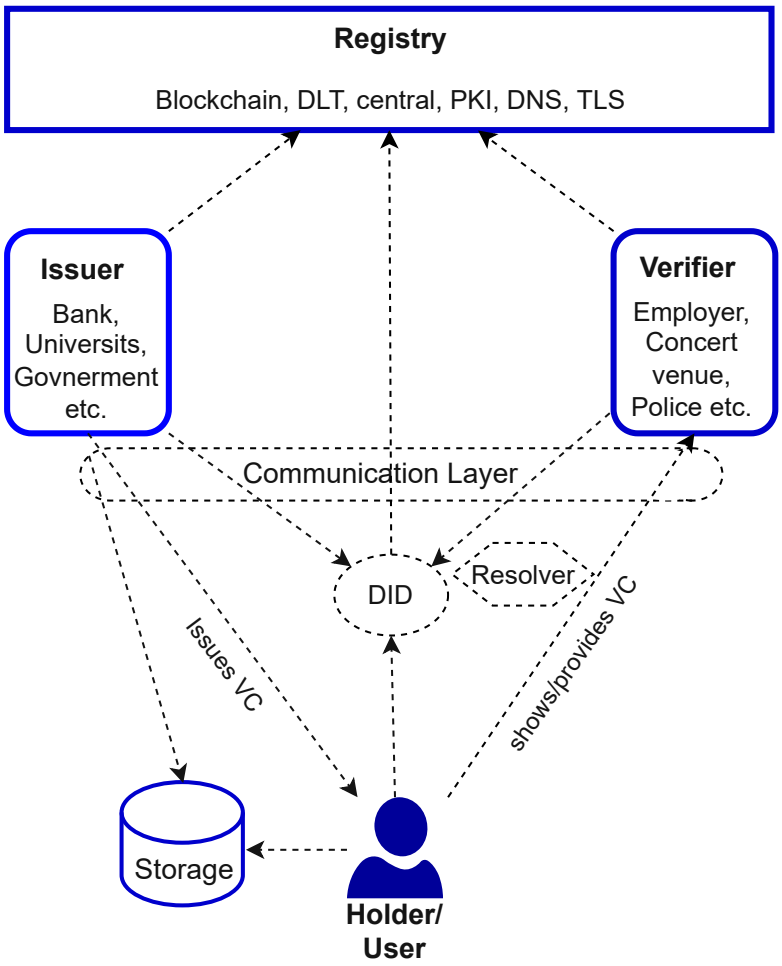


Figure 6.2: Expert Mental Model of SSI. Blue components were mentioned by all participants, the dotted components and communication paths only by some.

The Actors and Communication The mental models of all participants consisted of the three main actors as described in Section 6.1. Furthermore, VC were mentioned by most of the participants. VC represent a piece of personal information that is cryptographically trustworthy. Interestingly, the VC 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 was

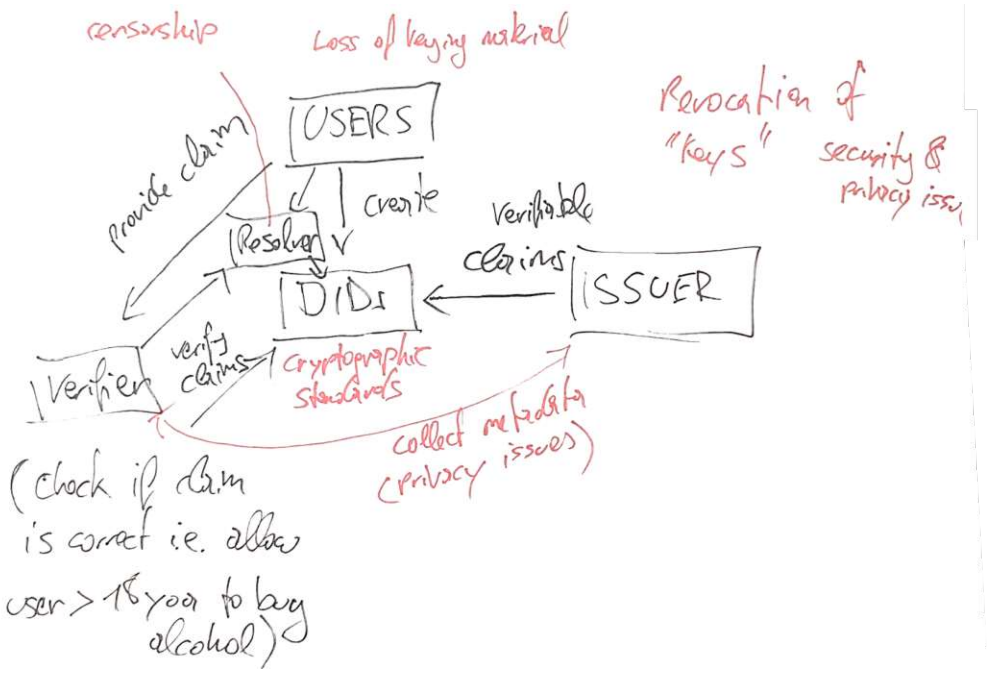


Figure 6.3: Drawing of participant #4

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) DID 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 DID were an active part of the system, there was communication between the holder and the verifier (see Figure 6.3), and the other communication was via the created DID of the user. For example, participant #4 explained that *"the users can create the DID and they can register them in a system, that can be something decentralized, like a blockchain [...] or also something else"*. Those participants mentioning DID 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, various descriptions were discovered of the registry itself, its location, and which technology should 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 DID) 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 DLT to DNS and TLS. The latter were 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. DLT 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"*.

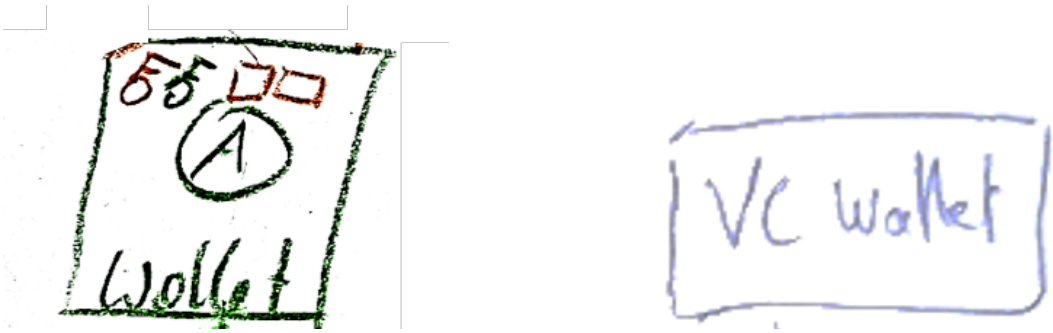
QR Code and Deep Links QR codes and deep links were mentioned when the participants talked about a communication layer or when the conversation turned to the user interface. In both cases, these two things served as link establishments between two actors (issuer and holder or holder and verifier). A 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.

6.4.2 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 the coding process, I grouped these into five pictorial categories.

The (digital) wallet: In this category the participants described SSI like a digital version of a person's wallet. Some participants even depicted a wallet in their drawing as can be seen in Figure 6.4. 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 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.



(a) Participant #1 (b) Participant #7

Figure 6.4: Depiction of wallet in participants drawings

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, all PET systems mentioned by the participants were combined. 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 how to develop new technology in order for it to be adopted.

6.4.3 Security and Privacy Risks of SSI

In the data, a variety of different threat models concerning SSI systems were found. The models were divided 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 (i.e., 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 of 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 were 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 of 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). 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 for which 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 full-filled. If DIDs or the resolver use 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 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.

6.4.4 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"*.

6.4.5 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 the 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 questions, the participants began to give examples of their professional activities. For example one researcher began to describe findings from a current project 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) background 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 options. 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 especially detailed from research 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"*.

6.5 Discussion of the Self-Sovereign Identity Study

This study was inspired by the rise of SSI research and its potential to shape users' online identities while safeguarding their privacy. Understanding the mental 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 6.1) 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 (VC vs. DID) 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, it became apparent that experts' explanations and their points of view reflected whether they were currently active in a (specific) standardization working 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.

6.5.1 Controversial Points of View of Study Participants

In this study, a big controversy was found 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 discovered in this study was 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 [76, 205]) 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) that have taken this path, but as Liu

et al. [122] have shown in their study, there are still some difficulties to overcome with current (beta version or prototype) systems.

6.5.2 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). Many experts expressed physical wallet and document metaphors. The former corresponds to the terms used to describe current cryptocurrency and SSI applications (e.g., "hardware wallet" and "software 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 the study, related work and online presentations of SSI systems were examined. Thereby I found that most papers and presentation materials used similar metaphors, and in line with the findings, especially *the (digital) wallet and the (physical) documents* metaphors were popular. Other metaphors found during this research, were, for example, the dot metaphor [34] used by Sovrin to describe the online identity of a user. Another example is the ring metaphor [124], 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 the 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 6.4.2).

6.5.3 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, a minimal knowledge map for end-users was designed. 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, specifically the answers to 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, decentral). 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 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., deep links) the user needs to understand how to manually use them.

6.5.4 Implications for SSI Designs

In order to determine the implications of the findings and thereby answer RQ4, a triangulation of the findings was performed. Therefore, the minimal knowledge map, the security and privacy risks, and the general findings (described in Section 6.4.2 and Section 6.4.5) 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 complex 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).

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 could be maintained. However, in order to avoid confusion between cryptocurrencies and SSI another metaphor could be used e.g. a document map or folder. To emphasize the significance of the stored information, their importance and properties could be highlighted by visual markers such as small document images or certificate stamps.

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 the study indicated. Therefore, more convenient methods, like individual selections of the seed phrase from a pool of words [205] or a form of Shamir's secret sharing by sharding the recovery key between multiple people or items [170] should rather be used.

CHAPTER 7

Discussion of User Studies

This chapter recapitulates the findings presented in this thesis and discusses their impact on a meta-level (thereby answering RQ1-RQ3) in order to incorporate these learnings into future research. Furthermore, a discussion of the recommended improvements for usable cryptographic protocols is provided in order to answer RQ4 in general, in contrast to the detailed discussions in the individual study Chapters 3-6. Thereby, directions and recommendations for further research in this area are outlined.

This thesis presented four user studies to investigate the mental models of different stakeholders involved in cryptographic protocol-based systems. Two studies were conducted with experts, i.e. administrators, developers, and researchers, and the other two with end-users from different age groups. When comparing the studies, it is important to note that the first expert study was conducted online, with a large focus on quantitative data, as opposed to the second study, which was also conducted online, but with qualitative interviews. The end-user studies were both qualitative in nature. The cryptocurrency study was conducted exclusively in person, whereas the Internet study was partially conducted online due to the pandemic. All studies used an iterative approach to gather and analyze the data, whereby the analytical approach differed, depending on whether the problem space of the protocol was known (thematic analysis - Internet and HTTPS) or unknown at the time of the study (GT - Cryptocurrencies and SSI). The characteristics of the studies in this thesis are summarized in Table 7.1.

In the following paragraphs, I will answer the research questions introduced in Chapter 1.

RQ1: *What mental models of cryptographic protocol-based systems and their functional components do different stakeholders have?*.

Throughout the four studies, it became apparent that most expert stakeholders hold a *glass box* mental model (see Section 2.1) of the system or protocol. On the other hand, non-expert users had a metaphor mental model (see Section 2.1) to describe the rather complex cryptographic protocols and systems based on them. As in the nature of mental

Table 7.1: Overview of study characteristics

Properties \ Study	**Internet**	**HTTPS**	**Cryptocurrencies**	**SSI**
# Participants	26	96	29	13
Study Type	semi-structured interview	quantitative questionnaire	semi-structured interview	semi-structured interview
Stakeholders	end-users	administrators	end-users	developers & researcher
Study setting	in person / online	online	in person	online
Methodology	thematic analysis	thematic analysis	GT	GT

models, all of them were to some degree either incomplete or incorrect. The level of detail of the mental models found depended on the age of the stakeholders as well as on their involvement in the protocols and their education. The protocol-specific mental models can be found in Section: i) 3.4 for the Internet, ii) 4.4 for HTTPS, iii) 5.4 for cryptocurrency systems and iv) 6.4 for SSI.

The next two research questions will be answered together, in order to highlight which differences between the mental model and reality concretely interfere with secure and privacy-preserving usage.

RQ2: *What are the key differences between the stakeholder's perception and the actual structure and functionality of the systems?*

RQ3: *Which mental models interfere with the secure and privacy-preserving usage or development of these systems?.*

In all four studies, deviations to varying degrees between the mental models and the actual technical functioning of the protocols and their systems were found as highlighted in the answer above to RQ1. In the case of expert participants who actively work with the systems, i.e. implement, maintain, or research them, I rarely found inaccurate mental models of the protocol/system and its functionality, which also led to more accurate awareness of the risk landscape. However, not all expert mental models were correct, as some administrators in the HTTPS study (see Section 4.3) for example had knowledge gaps on how the certificates are communicating with the server and how the renewal of them works. These gaps might lead (especially when no ACME clients are used who perform those steps (semi-)automatically) to severe security threats to the online communication of the end-users. The SSI study (see Section 6.4) highlighted that the lack of standards and different requirement priorities potentially lead to privacy and security risks for the end-users. In contrast to the expert mental models, there were

major differences between the actual functionality of the protocols and the end-users perceptions thereof.

In both of the studies (Internet study see Section 3 and cryptocurrency study see Section 5) with end-users, the participants each had problems with the correct understanding of encryption and/or key management. These comprehension challenges have already been highlighted in several studies with an explicit focus on End-to-End Encryption (E2EE) [196, 49, 135] or key management [58, 164, 163]. The studies conducted within this thesis confirmed that E2EE and key management challenges still exist and cause problems for the security and privacy of the user despite the other use cases and cryptographic protocols, which I looked at.

In line with other studies [36, 190, 100], I have found that despite mental models reflecting false impressions of the actual protocols, not all do not necessarily have negative consequences. For example in the cryptocurrency study (see Chapter 5.4) some participants misinterpreted the confirmation count as a representation of the number of miners or peers who validated the transaction. This number actually represents the number of blocks that were created after the transaction on the blockchain, which makes the transaction more secure. Despite this misinterpretation, the participants knew that a larger number would provide more security and therefore would not pose any risk to them.

RQ4: *How could these systems be improved to be used securely despite possibly incorrect mental models?.*

The findings of my studies suggest several conclusions and recommendations for usable cryptographic protocols, which I grouped into three categories:

- **Visual cues**: In the studies, I found that the visual interfaces and cues provided dramatically influence the formation of mental models. Therefore, I recommend minimal visual cues that clearly communicate to the users of the system what they are supposed to do or what effects their actions will have on them. However, this represents a major challenge, as users are quickly overwhelmed or oversaturated by too much visual information. Therefore, it is an important task of usable security to find visual cues that foster correct mental models and safe usage. In this thesis, I proposed several concrete enhancements to the systems and tools with the purpose of protecting users with misconceptions from security and privacy threats. For example, in the Internet study, I proposed to mark or highlight child-friendly content with smileys or colorful borders. In the HTTPS study I recommended visual cues about the security status and an inherent rating of the security. As visual advice for cryptocurrency wallets or exchanges a depiction of the transaction within the blockchain to emphasize transparency was proposed. In order to test the recommendations, the next important step is to conduct A/B tests [139] to determine their applicability and usability.

- **Level of abstraction**: In the studies, I found that most (lay) participants felt overwhelmed with security and privacy measures. Therefore, I suggest abstracting and automating certain security-critical measurements (e.g., backups or key recovery). However, these should not be completely hidden but displayed to users through visual cues (see above) as they sometimes crucially influence the mental models of end-users. However, since there are cases in which the participant wants/needs to make changes, there should be the possibility of "advanced" options to change specific parameters and settings. In general, one of the biggest challenges of usable security is to balance automation and security on the one hand, and on the other hand to leave enough freedom to not inhibit the possibilities of the system and its usability. As future research, it is important to investigate the necessary level of abstractions (necessary to inform mental model and foster adoption vs. security-critical interactions) which is necessary in order to foster security- and privacy-preserving usage.
- **Encryption**: In the studies, I found that the concept of encryption and where it is applied often does not match the actual functioning of the cryptographic protocols. Therefore, it is particularly important to pay attention to these (potential) misunderstandings and to foster correct communication of fundamental properties of the cryptographic system to the stakeholders. Thereby, both education and visual cues play an important role as users will come into contact with cryptographic fundamentals on a daily basis. Without a basic understanding, as in the case of the cryptocurrency study participants, users incorrectly assumed that cryptocurrency systems are encrypted and that they are therefore by default anonymous. Therefore, in the area of encryption further research is needed on why encryption is assumed in some systems even though it is not present.

During this thesis, I also encountered some methodological implications, which I will discuss in the following.

In the course of the studies, I noticed that participants without a technical or security background tended to use buzzwords to superficially describe technologies. They may have heard these words in connection with the systems in question or in other security or IT contexts without being sure exactly what they mean. By asking more detailed questions during the interviews, I was able to determine that most of the lay users did not have a clear idea or explanation of buzzwords such as "encrypted", "blockchain", or "cloud". This demonstrates the importance of qualitative studies in order to investigate difficulties with technologies in more detail and to filter out what is actually understood and what is just a buzzword to give a name to an unknown component.

Qualitative findings are also essential for expert studies, besides quantitative findings. In quantitative studies, the participants tended to use fewer buzzwords that they did not understand, but their misunderstandings often lay mostly in the technical details. Therefore, it was also important to me to offer the participants opportunities to provide explanations (open questions) during the quantitative study.

In the qualitative study designs, I used drawing tasks or card assignment tasks to help the participants express their (tacit) knowledge and perceptions of the system in question. I found that the majority of the participants felt comfortable (after sometimes initial skepticism or reluctance) with the drawing tasks, and it seemed that the visual depiction often helped them, especially when formulating threats. Furthermore, about one-third of the participants have added actors, connections, or components to their drawings although I already moved on to other questions. Those additions came to mind while talking about other things and after they have checked whether they have already included them in their system, they adjusted it accordingly. Based on the results of the studies, I am confident that additional visual components (drawings, screenshots, assignment cards) make the interview more interesting and engaging for the participants. Furthermore, they can be good support particularly when a participant in the study is stuck and the interviewer can guide them with these components. Following some interviews, the participants had questions about the systems and the correctness of their answers. In these cases, I took their drawings to discuss the system's functionality and explain possible shortcomings or misunderstandings of them.

Based on the encountered findings during the different studies, I formulated some recommendations for future (mental model) user studies in the area of usable security.

- Researchers should be aware of buzzwords from participants and should question them critically.
- Researchers should keep in mind that for both experts and laypeople, qualitative insights contain many important findings.
- Researchers should use additional materials or methods in interviews besides the normal conversation, for the purpose of loosening up and supporting the participant (e.g., drawing, visual use cases, card assignments)
- Researchers should select the research method appropriate to previous knowledge in the field, i.e. when there is less knowledge, qualitative methods should be used, and quantitative methods for validating hypotheses and general aspects of systems/cryptographic protocols.
- Researchers should investigate a technology/system from different stakeholder perspectives to get novel insights and compare e.g. tech-savvy and non-tech-savvy perspectives.
- Researchers should test their interview guidelines thoroughly, as participants might not understand them or their answers might not answer the research questions.
- Researchers should be aware of personal and priming biases by asking suggestive questions or using terms (e.g., encryption or authentication) not mentioned by the participants and thereby influencing their answers.

- Researchers should consider which group is targeted by their research to ensure ecological validity (as student sample cannot represent administrators in all cases see Section 4.4).

CHAPTER 8

Conclusion

This thesis contributes to existing knowledge on usable security research of peoples' perceptions of cryptographic protocols. Thereby, the findings shed light on different stakeholders' mental models of those protocols and systems based on them. When used properly, these protocols protect the security and privacy of users. However, due to their complexity and their usability challenges, they are often implemented or used incorrectly, which reduces or, in the worst case, eliminates the positive effect of the protocols. Therefore, it is essential to understand on the one hand the experts (developers, administrators, researchers) perspective to overcome security and privacy mistakes that can be threatening to a multitude of people. On the other hand, the end-users perception is important as it affects the security of one's own data and also the adoption of the respective system and protocol.

In the first part of this thesis, two widely used systems with well-established cryptographic protocols were investigated: i) the Internet in general from the perspective of both very young and adult users, and ii) its encrypted communication protocol HTTPS from the point of view of administrators, who are responsible for its implementation and maintenance.

I conducted a qualitative user study both with children and their respective parents to examine their mental models of the Internet and its security and privacy threat landscape. Thereby, it was revealed that the age of the children and their education (through parents and school) have a relevant influence on their perception of the Internet and its possibilities and dangers. The results suggest that the comprehension and knowledge of this intangible technology deepen between the ages of five and eight. Based on these results, recommendations are provided that can inform the design of future (child-friendly) Internet services and the way this topic is addressed in education.

In the context of the encrypted communication protocol HTTPS I closed the gap between prior qualitative studies and real-world administrators' challenges by conducting a

descriptive quantitative analysis and an online survey. Therefore, administrators were recruited who configured or maintained at least one web server at the time of the study to ensure that the experience with the servers and frameworks used is up to date. The results quantify the challenges administrators faced during configuration and enabled me to validate and refine the administrators' mental model of HTTPS discovered in a prior study by Krombholz et al. [106].

In the second part of this thesis, two cryptographic protocols were examined based on or fostered by blockchain technology: i) cryptocurrencies and ii) SSI.

To assess perceptions and misconceptions of cryptocurrency users, I conducted a semi-structured interview study enriched with drawing and card assignment tasks.The results of this study showed that flaws and inconsistencies in user mental models of cryptocurrency systems expose users to security and privacy risks when using current cryptocurrency tools. The findings explain why cryptocurrency users fail to manage their private keys securely and, as a result, frequently fall victim to money loss and fraud. Based on those findings several concrete enhancements to state-of-the-art cryptocurrency tools (e.g., wallets or exchanges) were provided with the purpose of protecting users with misconceptions from security and privacy threats.

To determine experts' mental models of SSI systems, I conducted a qualitative expert study. The study results highlighted 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, a minimal knowledge map for (potential) SSI end-users was constructed and design guidelines for SSI were formulated to facilitate a broad adoption in the wild and improve privacy-preserving usage.

While this thesis offers valuable insights into the mental models of different stakeholders of four specific cryptographic protocol-based systems, more research is needed on whether the recommendations lead to improved usability of those systems. In addition, further studies are needed to investigate whether this thesis's findings can be used to inform other designs of cryptographic protocols. This is especially interesting since it would provide the possibility of a general guideline for such systems, as end-users and developers/administrators often seem to have similar difficulties, such as misconceptions or misinterpretations of encryption and key management.

APPENDIX A

Additional Study Material

A.1 Internet Mental Model Study

A.1.1 Pre-Study Questionnaire

1. Age
2. Gender
3. Highest completed education
4. Do after-school care/grandparents or others look after your child regularly? If yes, how often?
5. Are you professionally involved with technical aspects of the internet?
Yes; No
6. How applicable are the following statements?
Never(1), Rarely(2), Sometimes(3), Often(4), Daily(5)
I use the following Internet services: [*only asked parents]

- *Social Media (Facebook, Xing, Instagram, etc.)*
- *VOD (YouTube, Netflix, Amazon Prime Video, etc.)*
- *Online Shopping (Amazon, ebay, willhaben, etc.)*
- *Instant messaging (WhatsApp, Facebook Messanger, Viber, etc.)*
- *Online video chat services (Skype, Facetime, etc.)*
- **Email*
- **Online banking*
- *Games (just asked during the children's interview)*

7. I am concerned about my privacy and the security of my data when using any of the services listed above.
Not at all(1)—I am very concerned (5)
8.How much do the following statements apply to you?
Does not apply(1)—Applies very much (5)

- *I am experienced with technical devices such as computers, smartphones and tablets.*
- *I learned about the internet in the course of my education.*
- *I often ask other people for help when I have problems with my computer/smartphone/tablet.*
- *I am often asked by other people for help when they have problems with their computer/smartphone/tablet.*

A.2 Quantitative Administrator Study - HTTPS

A.2.1 Survey Questionnaire

Multiple choice questions are marked with a ^{M}. In the case of questions with correct and incorrect choices, the options that we have assessed as correct are marked with * and options which only apply in certain exceptional cases/versions marked with **.

- **1. What is your profession?**
  Manager ;Programmer; Web Administrator; IT Consultant; Tester; IT Architect; System Administrator; Other, please specify
- **2. Please give an estimate about the size of the company you work for**
  < 10 employees; 11–50 employees; 51–250 employees; >250 employees
- **3. How many web servers did you administrate/configure HTTPS within the last year?**
  None; 1; 2–5; 6–10; >10
- **4^{M}. What are the services provided by the web servers under your responsibility?**
  Company internal services, Company external services; private services; IoT services; Others, please specify
- **5^{M}. Which web server software do you use at your current job for maintaining or configuring HTTPS?**
  Apache; NGINX; Microsoft ISS; Caddy; Caddy 2; Company internal software; Other, please specify; I do not currently administrate or maintain web server software at work; I don't know

- **6[M]. Which web server software would/do you use for your own/private web server for maintaining or configuring HTTPS?**
  Apache; NGINX; Microsoft ISS; Caddy; Caddy 2; Other, please specify; I don't know
- **7. Why would/do you use this web server software for HTTPS configuration/administration?**
  Open answer
- **8[M]. Where do you obtain certificates when setting up TLS for a web server?**
  Traditional certificate authority, e.g., GoDaddy, or Comodo; Free-of-cost alternative, e.g., Let's Encrypt; External provider, e.g., Cloudflare, or Amazon CloudFront; Local self-governing certificate authority; I use self-signed certificates; Other, please specify; I don't know
- **9[M]. Where did you learn how to create a server configuration for TLS?**
  Online; Company Guidelines; Colleagues; School/University; Seminar; Other, please specify
- **10. Do you use Certbot an Automatic Certificate Management Environment (ACME) client to obtain certificates?**
  Yes, No, I don't know
- **11. Did Certbot and/or Let's Encrypt change your configuration/maintenance routine?**
  Yes, only Certbot did; Yes, only Let's Encrypt did; Yes, both of the above did; Not sure; No
- **12. You said Certbot and/or Let's Encrypt changed your workflow. Please describe in which way it changed your workflow.**
  Open answer
- **13. Are you confident that the most recent website you configured encrypts the communication securely?**
  Yes; Partially, I tried to follow guidelines; No; I don't know; Others, please specify
- **14. What are the security guarantees your last website's HTTPS configuration provides?**
  Open answer
- **15. Do you trust that communication with a server over HTTPS provides a secure communication channel?**
  Yes; No; I don't know
- **16./17. Why do/don't you trust that communication with a server over HTTPS provides a secure communication channel?**
  Open answer

- **18. Do you trust the security indicator lock symbol found in the address bar of web browsers?**
  Yes; No; I don't know
- **19. Why do you not trust the security indicator?**
  Open answer
- **20. Which color does the HTTPS Everywhere symbol have? This is a data sanity check: Please make sure to select purple as the right answer.**
  Yellow; Orange; Purple; Red; Green; Blue; Black; White
- **21. Do you trust the security of Let's Encrypt?**
  Yes; No; I don't know
- **22./23. Why do/don't you trust the security of Let's Encrypt?**
  Open answer
- **24. Do you trust the security of Certbot?**
  Yes; No; I don't know
- **25./26. Why do/don't you trust the security of Certbot?**
  Open answer
- **27[M]. Which kind of validation do you (or your company) require for the certificates you use?**
  Domain validation; Extended validation; Organization validation; Other, please specify; I don't know
- **28. Why do you or your company want this validation/these validations?**
  Open answer
- **29. Do public or private keys play a part in TLS?**
  Yes, both; Yes, public keys; Yes, private keys; No; I don't know*
- **30. How are keys used in HTTPS?**
  Open answer
- **31. Are certificates and PKI (Public-Key-Infrastructure) relevant for setting up TLS?**
  Yes; No; I don't know*
- **32[M]. How do certificates and Public-Key-Infrastructure (PKI) interact with each other?**
  The PKI is used to sign the certificate(s); The certificate contains an identity (e.g., server name) and a public key; The certificate authority (CA) signs the certificate*; TLS uses PKI certificates to authenticate parties communicating*; The certificate is used to sign the PKI; They do not interplay with each other; Other, please specify*

- **33. Please indicate your level of agreement with the following statements about HTTPS (HTTP over TLS).** ***likert scale with 5 options: strongly agree–strongly disagree; I don't know***
  PKI stands for public key infrastructure [data sanity check: please select "strongly disagree" to confirm that you read the whole message]; Under TLS, messages can be manipulated and read while being transmitted on the Internet; TLS ensures message confidentiality; During transmission, communication is securely encrypted using properly configured TLS*; Messages are encrypted with asymmetrical encryption algorithms**[1]; Public/Private keys are generated uniquely for each connection between client and server when using HTTPS**[2]; Symmetric cryptography is used for the TLS key exchange; TLS does not work without certificates*; Authentication is necessary for encryption*; Encryption is required to guarantee message authenticity*
- **34. Did you disable the HTTP access to your website when setting up TLS?**
  Yes; No; I don't know
- **35. When you last deployed TLS, did you change any of your software defaults?**
  Yes; No; I can't remember
- **36. Which TLS default settings did you change?**
  Open answer
- **37[M]. Why did you change the TLS default settings?**
  Prescribed by the company; Security reasons; Support of old protocols; Other, please specify
- **38. Does your employer oblige you to choose a specific certificate origin?**
  Yes; No; I don't administer/configure web servers at my company; I don't know
- **39. Does your employer allow the use of Let's Encrypt?**
  Yes; No; I don't know
- **40. Does your employer allow the use of Certbot?**
  Yes; No; I don't know
- **41. Does your company have security policies for securing web servers?**
  Yes, they're mandatory; Yes, they're available; No; Other, please specify
- **42[M]. When do security audits happen at your company?**
  After the first setup; When setup changes are made; Regularly scheduled. Please indicate how often; When required, e.g., for newly discovered exploits; Other, please specify; Never; I don't know

[1] only if used with old RSA key exchanges
[2] only when using ephemeral key exchange methods

- **43[M]. Have you had any problems setting up TLS? Please tick all that apply and shortly describe what the problem was.**
  Unsupported platform(s); Server communication failure; Reports didn't show in the console; Configuration manager problems; Server authentication failed; Client authentication failed; Specification not understandable; Certificate generation; Others, please specify; I had no problems

- **44. What is the highest level of education you finished?**
  No schooling completed; High school; College; Technical training; Bachelor's degree; Master's degree; Professional degree; Doctorate

- **45. What is your gender?**
  Woman; Man; Non-binary; Prefer not to disclose; Prefer to self-describe

- **46. What is your age?**
  18–24; 25–34; 35–44; 45–54; 55–64; Older than 64

- **47. Where are you currently working?**
  Drop-down selection

A.3 Bitcoin Mental Model Study

A.3.1 Demographics gathered via a pre-study questionnaire

- Age/ Gender
- Profession/ Highest completed level of education/ Recent professional status
- I have a good understanding of Computers and the Internet: Likert Scale from 5 (agree) - 1 (disagree)
- I often ask other people for help when I am having problems with my computer: Likert Scale from 5 (agree) - 1 (disagree)
- I am often asked for help when other people have problems with their computer. Likert Scale from 5 (agree) - 1 (disagree)
- Which cryptocurrencies have you heard of?
- Was the subject of cryptography and/or cryptocurrencies part of your education or your profession?
- If yes, briefly outline the topics you heard of.
- Do you use Bitcoin/Ethereum?
- For which matters do you mainly use Bitcoin/Ethereum?

A.3.2 Interview Protocol

General

- Which kind of education do you have and what is your current profession?
- When and how did you become aware of cryptocurrencies?
- How have you been dealing with cryptocurrencies so far?
- Why do you use Bitcoin/Ethereum? (just asked if the participant owns a cryptocurrency)
- What is in your opinion the cryptographic part of cryptocurrencies?

Mental Models

- [**Drawing Task 1**] Please draw a picture of how you think the transaction process works between you and a second person called Alice. Imagine you transfer BTC/ETH 20 to Alice. Remember to include all relevant persons and components into your drawing.

- [**Card Assignment Task**] We prepared some cards which describe various functionalities of a cryptocurrency system. Please assign these cards to the components you drew in Phase 1. If you feel you missed a component before, please draw them with green colour. The cards we provided during this task:
  - Generate address
  - Generate public key
  - Generate private key
  - Transaction confirmed
  - Generate transaction
  - Sign transaction
  - Broadcast transaction
  - Verify transaction
  - Generate block
  - Validate block
  - Perform Proof of Work
  - Solve cryptographic puzzle
  - Receive transaction fees
  - Generate coins
  - Only Bitcoin: Receive unspent transaction output (UTXO)
  - Only Ethereum: Receive balance

Attacker Models

- There are two words which are lately frequently used in the media in relation to cryptocurrencies, namely "security" and "privacy". What do these two words mean to you and what are the differences between them?

- [**Drawing Task 2**] Please have a look on the model you created during Phase 2. Take a red marker for drawing security risks and a blue marker for drawing privacy risks. While drawing, keep the following two questions in mind:

  – Where do you think the potential threats occur?

  – Who is causing those threats?

  After the participant has finished the drawing, ask: "What countermeasures do you know to prevent those risks?"

A.3.3 Wallet Feature Overview

Table A.1: Feature overview of 4 popular software wallets at the time of our study

	Blockchain.com	**Coinbase.com**	**Coinomi**	**Exodus**
Founded	2011	2012	2013	2015
Supported Cryptocurrencies	BTC, ETH, BCH, XLM, USD-D	BTC, ETH, BCH, ETC, LTC, ERC-20 tokens	BTC, ETH, BCH, ETC, LTC, etc.	BTC, ETH, BCH, ETC, LTC, etc.
Type	wallet/exchange	wallet/exchange	wallet/exchange	wallet/exchange
Private key storage	local	local	local	local
Back-ups	user initiated (seed phrase)	user initiated (seed phrase, gdrive with PIN)	user initiated (seed phrase)	user initiated (seed phrase)
Force seed phrase back-up	yes	no	no	no
Transaction Fees	options (pre-calculated/custom)	no options	options (low/normal/high priority)	no options
Wallet encryption	password (forced)	fingerprint/ PIN (forced)	password (standard)/ biometric/ none	none (standard)/ PIN / fingerprint
Periodic seed phrase querying	no	no	no	no
Block explorer included	yes (textual)	yes (textual)	yes (textual)	no
Different complexity levels	no	no	yes (creation: "fast", "advanced")	no

A.4 SSI Mental Model Study

A.4.1 Interview Guideline

Demographics

- Age/ Gender
- Current area of residence and work
- Education
- Current profession

General

- When was the first time you heard of SSI or decentral identities? Where did you hear it?
- Have you ever used a SSI system/ application? *If yes, what kind of application? If not, where/how have you been dealing with SSI systems so far?*
- How would you describe a self-sovereign identity system?
- What are principles/basics that should be fulfilled by a SSI system? (Hints if nothing was mentioned: 10 principles based on C. Allen existence, control, access, transparency, persistence, portability, interoperability, consent, minimization, and protection)

Mental Models

- [**Drawing Task 1**] Please draw a picture of an SSI system including all relevant persons and components (Buzzwords: Verifier, Service Provider, Issuer, Decentral)
- [**Drawing Taks 2**] Please have a look on the model you created. Take a marker in another color and please mark or draw possible security and privacy risks within the system.
  - Who causes the threats/ Who is the attacker?
  - What is the actual risk and who is involved?
- [**Task 3**] Please mark or describe what you deem as absolutely necessary for and end users to understand about this system.

Added questions for second interview round:

Which standards for SSI do currently exist?

Which standards are your currently discussed in your working group? (Why? Are there any issues/conflicts with other (proposed) standards?)

List of Figures

List of Tables

Acronyms

ACME Automatic Certificate Management Environment. 35, 40, 47, 49, 50, 54–56, 108

AI Artificial Intelligence. 12, 13

APIs application programming interfaces. 12

CA Certificate Authority. 39, 41, 42, 45–50, 52, 53, 55

DID Decentralized Identifier. 88, 90, 96, 97, 100–102, 105

DLT Distributed Ledger Technology. 58, 88–90, 97, 101, 105

DNS Domain Name Server. 17, 97

DoS Denial-of-Service. 76

E2EE End-to-End Encryption. 109

GDPR General Data Protection Regulation. 65, 69

GT Grounded Theory. 57, 60, 61, 65, 68, 93, 107, 108

HCI human-computer interaction. 3, 10, 11, 13

HTTPS Hypertext Transfer Protocol Secures. xi, 2, 4, 5, 7, 35–39, 42, 43, 45, 46, 48–56, 108, 109, 113, 114, 125

ICMP Internet Control Message Protocol. 18

IoT Internet of Things. 12

IP Internet Protocol. 18

ISP Internet Service Provider. 17

MITM Monster-in-the-Middle. 76, 99

P2P Peer-to-Peer. 70, 71, 76

PGP Pretty Good Privacy. 73, 81, 98

PII personally identifiable information. 100, 102

PoC Proofs-of-Concept. 85, 89

PoW Proof-of-Work. 58, 70, 71

SMS Short Message Service. 14

SPV Simplified Payment Verification. 70

SSI Self-Sovereign Identity. xii, 3–6, 8, 11, 85–94, 97–105, 107, 108, 114

SSL Secure Sockets Layer. 38, 42, 47, 48

TAP Trigger-Action Programming. 12

TCP Transmission Control Protocol. 18

TLS Transport Layer Security. 35, 36, 38–42, 46–49, 51, 52, 54–56, 97–99, 105

UDP User Datagram Protocol. 18

VC Verifiable Credential. 88, 95, 101, 102, 105

Bibliography

[1] Josh Aas. Why ninety-day lifetimes for certificates? `https://letsencrypt.org/2015/11/09/why-90-days.html`, Accessed: 2021-01-08.

[2] Josh Aas, Richard Barnes, Benton Case, Zakir Durumeric, Peter Eckersley, Alan Flores-López, J Alex Halderman, Jacob Hoffman-Andrews, James Kasten, Eric Rescorla, et al. Let's encrypt: An automated certificate authority to encrypt the entire web. In *ACM SIGSAC Conference on Computer and Communications Security (CCS'19)*, pages 2473–2487, 2019.

[3] Ruba Abu-Salma, Elissa M Redmiles, Blase Ur, and Miranda Wei. Exploring User Mental Models of End-to-End Encrypted Communication Tools. In *8th USENIX Workshop on Free and Open Communications on the Internet (FOCI 18)*, 2018.

[4] Yasemin Acar, Michael Backes, Sascha Fahl, Doowon Kim, Michelle L Mazurek, and Christian Stransky. You get where you're looking for: The impact of information sources on code security. In *IEEE Symposium on Security and Privacy (S&P'16)*, pages 289–305. IEEE, 2016.

[5] Mustafa Emre Acer, Emily Stark, Adrienne Porter Felt, Sascha Fahl, Radhika Bhargava, Bhanu Dev, Matt Braithwaite, Ryan Sleevi, and Parisa Tabriz. Where the Wild Warnings Are: Root Causes of Chrome HTTPS Certificate Errors. In *ACM SIGSAC Conference on Computer and Communications Security (CCS'17)*, 2017.

[6] Maarten Aertsen, Maciej Korczyński, Giovane CM Moura, Samaneh Tajalizadehkhoob, and Jan van den Berg. No domain left behind: is let's encrypt democratizing encryption? In *Applied Networking Research Workshop*, pages 48–54, 2017.

[7] Devdatta Akhawe, Bernhard Amann, Matthias Vallentin, and Robin Sommer. Here's my cert, so trust me, maybe? understanding tls errors on the web. In *22nd International Conference on World Wide Web*, pages 59–70, 2013.

[8] Devdatta Akhawe and Adrienne Porter Felt. Alice in warningland: A large-scale field study of browser security warning effectiveness. In *22nd USENIX Security Symposium (USENIX'13)*, pages 257–272, 2013.

[9] Alaa Alaqra, Simone Fischer-Hübner, Thomas Groß, Thomas Lorünser, and Daniel Slamanig. Signatures for privacy, trust and accountability in the cloud: Applications and requirements. In *Privacy and Identity Management. Time for a Revolution?*, pages 79–96. Springer, 2016.

[10] Sinică Alboaie and Doina Cosovan. Private data system enabling self-sovereign storage managed by executable choreographies. In *IFIP International Conference on Distributed Applications and Interoperable Systems*, pages 83–98. Springer, 2017.

[11] Christopher Allen. The Path to Self-Sovereign Identity. `http://www.lifewithalacrity.com/2016/04/the-path-to-self-soverereign-identity.html`, 2016. Accessed: 2022-01-19.

[12] Gergely Alpár, Fabian van den Broek, Brinda Hampiholi, Bart Jacobs, Wouter Lueks, and Sietse Ringers. Irma: practical, decentralized and privacy-friendly identity management using smartphones. *HotPETs 2017*, 2017.

[13] Elli Androulaki, Ghassan O. Karame, Marc Roeschlin, Tobias Scherer, and Srdjan Capkun. Evaluating user privacy in bitcoin. *International Conference on Financial Cryptography and Data Security (FC'13)*, 2013.

[14] Maria Apostolaki, Aviv Zohar, and Laurent Vanbever. Hijacking bitcoin: Routing attacks on cryptocurrencies. *IEEE Symposium on Security and Privacy (S&P'17)*, 2017.

[15] Farzaneh Asgharpour, Debin Liu, and L Jean Camp. Mental models of computer security risks. In *WEIS*, 2007.

[16] Dirk Balfanz, Glenn Durfee, Diana K Smetters, and Rebecca E Grinter. In search of usable security: Five lessons from the field. *IEEE Security & Privacy*, 2(5):19–24, 2004.

[17] Gagan Bansal, Tongshuang Wu, Joyce Zhou, Raymond Fok, Besmira Nushi, Ece Kamar, Marco Tulio Ribeiro, and Daniel Weld. Does the whole exceed its parts? the effect of ai explanations on complementary team performance. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–16, 2021.

[18] Aaron W. Baur, Julian Bühler, Markus Bick, and Charlotte S. Bonorden. Cryptocurrencies as a Disruption? Empirical Findings on User Adoption and Future Potential of Bitcoin and Co. In Marijn Janssen, Matti Mäntymäki, Jan Hidders, Bram Klievink, Winfried Lamersdorf, Bastiaan van Loenen, and Anneke Zuiderwijk, editors, *Open and Big Data Management and Innovation*, pages 63–80, Cham, 2015. Springer International Publishing.

[19] Matthew Bernhard, Jonathan Sharman, Claudia Ziegler Acemyan, Philip Kortum, Dan S Wallach, and J Alex Halderman. On the usability of https deployment. In *2019 CHI Conference on Human Factors in Computing Systems*, pages 1–10, 2019.

[20] Veroniek Binkhorst, Tobias Fiebig, Katharina Krombholz, Wolter Pieters, et al. Security at the end of the tunnel: The anatomy of vpn mental models among experts and non-experts in a corporate context. In *USENIX Security Symposium*, number 31, 2022.

[21] Blockchain. Blockchain.com. `https://www.blockchain.com/wallet`. Accessed: 2020-05-26.

[22] Joseph Bonneau, Andrew Miller, Jeremy Clark, Arvind Narayanan, Joshua A Kroll, and Edward W Felten. Sok: Research perspectives and challenges for bitcoin and cryptocurrencies. *IEEE Symposium on Security and Privacy (S&P'15)*, 2015.

[23] Will Brackenbury, Abhimanyu Deora, Jillian Ritchey, Jason Vallee, Weijia He, Guan Wang, Michael L Littman, and Blase Ur. How users interpret bugs in trigger-action programming. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pages 1–12, 2019.

[24] Russell Brandom. Inside elsagate, the conspiracy-fueled war on creepy youtube kids videos. `https://www.theverge.com/2017/12/8/16751206/elsagate-youtube-kids-creepy-conspiracy-theory`. Accessed: 2020-11-17.

[25] Cristian Bravo-Lillo, Lorrie Faith Cranor, Julie Downs, and Saranga Komanduri. Bridging the Gap in Computer Security Warnings: A Mental Model Approach. *IEEE Symposium on Security and Privacy (S&P'11)*, 2011.

[26] Britannicy. Internet computer network. `https://www.britannica.com/technology/Internet`, Accessed: 2022-03-04.

[27] Anne E Brodsky. Negative case analysis. *The SAGE encyclopedia of qualitative research methods*, page 553, 2008.

[28] Jessica E Brodsky, Arshia K Lodhi, Kasey L Powers, Fran C Blumberg, and Patricia J Brooks. "it's just everywhere now": Middle-school and college students' mental models of the internet. *Human Behavior and Emerging Technologies*, 3(4):495–511, 2021.

[29] Vitalik Buterin et al. A next-generation smart contract and decentralized application platform. *white paper*, 3(37):2–1, 2014.

[30] Kim Cameron. A user-centric identity metasystem. *Microsoft Corp*, 2008.

[31] David Canellis. Crippling DoS vulnerability put the entire Bitcoin market at risk. `https://thenextweb.com/hardfork/2018/09/20/bitcoin-core-vulnerability-blockchain-ddos/`, 2018. Accessed: 2020-01-31.

[32] Clément Canonne and Jean-Julien Aucouturier. Play together, think alike: Shared mental models in expert music improvisers. *Psychology of Music*, 44(3):544–558, 2016.

[33] John M Carroll and Judith Reitman Olson. Mental models in human-computer interaction. *Handbook of human-computer interaction*, pages 45–65, 1988.

[34] Full Moon Cartoon. Sovrin. https://www.youtube.com/watch?v=Hg7psADNcVU, 2016. Accessed: 2022-01-25.

[35] Chia-ling Chan, Romain Fontugne, Kenjiro Cho, and Shigeki Goto. Monitoring tls adoption using backbone and edge traffic. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 208–213. IEEE, 2018.

[36] Jing Chen. Risk communication in cyberspace: A brief review of the information-processing and mental models approaches. *Current opinion in psychology*, 36:135–140, 2020.

[37] Catalin Cimpanu. 14,766 Let's Encrypt SSL Certificates Issued to PayPal Phishing Sites. https://www.bleepingcomputer.com/news/security/14-766-lets-encrypt-ssl-certificates-issued-to-paypal-phishing-sites/, Accessed: 2020-12-11.

[38] Jeremy Clark and Paul C Van Oorschot. Sok: Ssl and https: Revisiting past challenges and evaluating certificate trust model enhancements. In *IEEE Symposium on Security and Privacy (S&P'13)*, pages 511–525. IEEE, 2013.

[39] Victoria Clarke and Virginia Braun. Thematic analysis. In *Encyclopedia of critical psychology*, pages 1947–1952. Springer, 2014.

[40] Jacob Cohen. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46, 1960.

[41] Coinmarketcap. Coinmarket. Cryptocurrency Market Capitalizations. https://coinmarketcap.com/coins/, Accessed: 2019-05-06.

[42] Coinomi. Coinomi. https://www.coinomi.com/en/. Accessed: 2020-05-26.

[43] Intersoft Consulting. General Data Protection Regulation GDPR . https://gdpr-info.eu/, Accessed: 2021-01-14.

[44] Juliet M Corbin and Anselm Strauss. Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology*, 13(1):3–21, 1990.

[45] Kenneth James Williams Craik. *The nature of explanation*, volume 445. CUP Archive, 1967.

[46] Lorrie Faith Cranor, Adam L Durity, Abigail Marsh, and Blase Ur. Parents' and teens' perspectives on privacy in a technology-filled world. In *10th Symposium On Usable Privacy and Security (SOUPS' 14)*, pages 19–35, 2014.

[47] William Damon, Richard M Lerner, Deanna Kuhn, and Robert S Siegler. *Handbook of child psychology, cognition, perception, and language*. John Wiley & Sons, 2006.

[48] Yves-Alexandre De Montjoye, Erez Shmueli, Samuel S Wang, and Alex Sandy Pentland. openpds: Protecting the privacy of metadata through safeanswers. *PloS one*, 9(7):e98790, 2014.

[49] Sergej Dechand, Alena Naiakshina, Anastasia Danilova, and Matthew Smith. In encryption we don't trust: The effect of end-to-end encryption to the masses on user perception. In *2019 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 401–415. IEEE, 2019.

[50] Kj Dell'Antonia. When Should a Child Get an E-Mail Account? `https://parenting.blogs.nytimes.com/2013/01/11/when-should-a-child-get-an-e-mail-account/`, Accessed: 2021-02-08.

[51] Ira Diethelm, Henning Wilken, and Stefan Zumbrägel. An investigation of secondary school students' conceptions on how the internet works. In *Proceedings of the 12th Koli Calling International Conference on Computing Education Research*, pages 67–73, 2012.

[52] Sarah A Douglas and Thomas P Moran. Learning text editor semantics by analogy. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 207–211, 1983.

[53] B Du Boulay, T O'Shea, and J Monk. The glass box inside the black box: Presenting computing concepts to novices. studying the novice programmer. e. soloway and jc spohrer, 1989.

[54] Zakir Durumeric, James Kasten, Michael Bailey, and J Alex Halderman. Analysis of the https certificate ecosystem. In *2013 conference on Internet measurement conference*, pages 291–304, 2013.

[55] Zakir Durumeric, Zane Ma, Drew Springall, Richard Barnes, Nick Sullivan, Elie Bursztein, Michael Bailey, J Alex Halderman, and Vern Paxson. The security impact of https interception. In *Network and Distributed System Security Symposium (NDSS'17)*, 2017.

[56] Susan Edwards, Andrea Nolan, Michael Henderson, Ana Mantilla, Lydia Plowman, and Helen Skouteris. Young children's everyday concepts of the internet: A platform for cyber-safety education in the early years. *British journal of educational technology*, 49(1):45–55, 2018.

[57] Chris Elsden, Arthi Manohar, Jo Briggs, Mike Harding, Chris Speed, and John Vines. Making Sense of Blockchain Applications: A Typology for HCI. *SIGCHI Conference on Human Factors in Computing Systems (CHI'18)*, 2018.

[58] Shayan Eskandari, Jeremy Clark, David Barrera, and Elizabeth Stobert. A first look at the usability of bitcoin key management. *arXiv preprint arXiv:1802.04351*, 2018.

[59] Sirpa Eskelä-Haapanen and Carita Kiili. 'it goes around the world'–children's understanding of the internet. *Nordic Journal of Digital Literacy*, 14(3-04):175–187, 2019.

[60] Blockchain Explorer. Block #0. `https://www.blockchain.com/btc/block-index/14849`, 2019. Accessed: 2019-09-17.

[61] Lesley-Anne Ey and C Glenn Cupit. Exploring young children's understanding of risks associated with internet usage and their concepts of management strategies. *Journal of Early Childhood Research*, 9(1):53–65, 2011.

[62] Facebook. How can I choose friends to help me log in if I ever get locked out of my account? `https://www.facebook.com/help/119897751441086/`, 2019. Accessed: 2019-08-27.

[63] Sascha Fahl, Yasemin Acar, Henning Perl, and Matthew Smith. Why eve and mallory (also) love webmasters: A study on the root causes of ssl misconfigurations. In *ACM Symposium on Information, Computer and Communications Security*, pages 507–512, 2014.

[64] Adrienne Porter Felt, Alex Ainslie, Robert W Reeder, Sunny Consolvo, Somas Thyagaraja, Alan Bettes, Helen Harris, and Jeff Grimes. Improving ssl warnings: Comprehension and adherence. In *33rd annual ACM conference on human factors in computing systems*, pages 2893–2902, 2015.

[65] Adrienne Porter Felt, Richard Barnes, April King, Chris Palmer, Chris Bentzel, and Parisa Tabriz. Measuring https adoption on the web. In *26th USENIX Security Symposium (USENIX'17)*, pages 1323–1338, 2017.

[66] Adrienne Porter Felt, Robert W Reeder, Alex Ainslie, Helen Harris, Max Walker, Christopher Thompson, Mustafa Embre Acer, Elisabeth Morant, and Sunny Consolvo. Rethinking connection security indicators. In *Twelfth Symposium on Usable Privacy and Security (SOUPS'16)*, pages 1–14, 2016.

[67] Adrienne Porter Felt, Robert W Reeder, Hazim Almuhimedi, and Sunny Consolvo. Experimenting at scale with google chrome's ssl warning. In *SIGCHI conference on human factors in computing systems*, pages 2667–2670, 2014.

[68] Jay W Forrester. Counterintuitive behavior of social systems. *Theory and decision*, 2(2):109–140, 1971.

[69] Maria Freytsis, Iain Barclay, Swapna Krishnakumar Radha, Adam Czajka, Geoffery H Siwo, Ian Taylor, and Sherri Bucher. Development of a mobile, self-sovereign

identity approach for facility birth registration in kenya. *Frontiers in Blockchain*, 4:2, 2021.

[70] Charlie Fripp. Check this list: 3.2 billion leaked usernames and passwords. `https://www.komando.com/security-privacy/3-billion-leaked-passwords/777661/`, Accessed: 2021-02-10.

[71] Jana Fruth, Carsten Schulze, Marleen Rohde, and Jana Dittmann. E-learning of it security threats: A game prototype for children. In *IFIP international conference on communications and multimedia security*, pages 162–172. Springer, 2013.

[72] Kelsey R Fulton, Rebecca Gelles, Alexandra McKay, Yasmin Abdi, Richard Roberts, and Michelle L Mazurek. The effect of entertainment media on mental models of computer security. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, pages 79–95, 2019.

[73] Kevin Gallagher, Sameer Patil, and Nasir Memon. New me: Understanding expert and non-expert perceptions and usage of the tor anonymity network. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS' 17)*, pages 385–398, 2017.

[74] Xianyi Gao, Gradeigh D. Clark, and Janne Lindqvist. Of Two Minds, Multiple Addresses, and One Ledger: Characterizing Opinions, Knowledge, and Perceptions of Bitcoin Across Users and Non-Users. *SIGCHI Conference on Human Factors in Computing Systems (CHI'16)*, 2016.

[75] Juan Garay, Aggelos Kiayias, and Nikos Leonardos. The bitcoin backbone protocol: Analysis and applications. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 281–310. Springer, 2015.

[76] Paco Garcia. Biometrics on the blockchain. *Biometric Technology Today*, 2018(5):5–7, 2018.

[77] Simson Garfinkel and Heather Richter Lipford. Usable security: History, themes, and challenges. *Synthesis Lectures on Information Security, Privacy, and Trust*, 5(2):1–124, 2014.

[78] Susan A Gelman, Megan Martinez, Natalie S Davidson, and Nicholaus S Noles. Developing digital privacy: Children's moral judgments concerning mobile gps devices. *Child development*, 89(1):17–26, 2018.

[79] Katy Ilonka Gero, Zahra Ashktorab, Casey Dugan, Qian Pan, James Johnson, Werner Geyer, Maria Ruiz, Sarah Miller, David R Millen, Murray Campbell, et al. Mental models of ai agents in a cooperative game setting. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2020.

[80] Barney G Glaser and Anselm L Strauss. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Transaction publishers, 1967.

[81] SoSci Survey GmbH. SoSci Survey – the Solution for Professional Online Questionnaires. https://www.soscisurvey.de/, Accessed: 2021-01-08.

[82] Greg Guest, Arwen Bunce, and Laura Johnson. How many Interviews are enough? An Experiment with Data Saturation and Variability. *Field methods*, 18(1):59–82, 2006.

[83] Marian Harbach, Sascha Fahl, Polina Yakovleva, and Matthew Smith. Sorry, i don't get it: An analysis of warning message texts. In *International Conference on Financial Cryptography and Data Security*, pages 94–111. Springer, 2013.

[84] Ian A Harris, Oliver K Khoo, Jane M Young, Michael J Solomon, and Hamish Rae. Lottery incentives did not improve response rate to a mailed survey: a randomized controlled trial. *Journal of Clinical Epidemiology*, 61(6):609–610, 2008.

[85] Donell Holloway, Lelia Green, and Sonia Livingstone. Zero to eight: Young children and their internet use. 2013.

[86] Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, 1979.

[87] Amber Horvath, Mariann Nagy, Finn Voichick, Mary Beth Kery, and Brad A Myers. Methods for investigating mental models for learners of apis. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–6, 2019.

[88] Bahar Houtan, Abdelhakim Senhaji Hafid, and Dimitrios Makrakis. A survey on blockchain-based self-sovereign patient identity in healthcare. *IEEE Access*, 8:90478–90494, 2020.

[89] Akari Ishikawa, Edson Bollis, and Sandra Avila. Combating the elsagate phenomenon: Deep learning architectures for disturbing cartoons. *arXiv preprint arXiv:1904.08910*, 2019.

[90] Eric Jensen and Charles Laurie. *Doing real research: A practical guide to social research*. Sage, 2016. ISBN 978-1446273883.

[91] Joseph Johnson. Internet usage worldwide - statistics & facts. https://www.statista.com/topics/1145/internet-usage-worldwide/#dossierContents__outerWrapper, Accessed: 2022-01-28.

[92] Philip N Johnson-Laird. Mental models and thought. *The Cambridge handbook of thinking and reasoning*, pages 185–208, 2005.

[93] Philip N Johnson-Laird, Vittorio Girotto, and Paolo Legrenzi. Mental models: a gentle guide for outsiders. *Sistemi Intelligenti*, 9(68):33, 1998.

[94] Philip Nicholas Johnson-Laird. *Mental models: Towards a cognitive science of language, inference, and consciousness*. Number 6. Harvard University Press, 1983.

[95] Natalie A Jones, Helen Ross, Timothy Lynam, Pascal Perez, and Anne Leitch. Mental models: an interdisciplinary synthesis of theory and methods. *Ecology and Society*, 16(1), 2011.

[96] Ruogu Kang, Laura Dabbish, Nathaniel Fruchter, and Sara Kiesler. "my data just goes everywhere:" user mental models of the internet and implications for privacy and security. In *Eleventh Symposium On Usable Privacy and Security (SOUPS' 15)*, pages 39–52, 2015.

[97] Michael Kaplan. I accidentally threw away $60M worth of Bitcoin . https://nypost.com/2018/05/26/i-accidentally-threw-away-60m-worth-of-bitcoin/, 2019. Accessed: 2019-05-02.

[98] Kaspersky. The number of mobile malware attacks doubles in 2018. https://www.kaspersky.com/about/press-releases/2019_the-number-of-mobile-malware-attacks-doubles-in-2018-as-cybercriminals-sharpen-their-distribution-strategies, Accessed: 2019-05-06.

[99] Ali Kazerani, Domenic Rosati, and Brian Lesser. Determining the Usability of Bitcoin for Beginners Using Change Tip and Coinbase. pages 1–5. ACM Press, 2017.

[100] Anne R. Kearney and Stephen Kaplan. Toward a Methodology for the Measurement of Knowledge Structures of Ordinary People: The Conceptual Content Cognitive Map (3CM). *Environment and Behavior*, 29(5):579–617, 1997.

[101] Willett Kempton. Two theories of home heat control. *Cognitive science*, 10(1):75–90, 1986.

[102] Irni Eliana Khairuddin, Corina Sas, Sarah Clinch, and Nigel Davies. Exploring Motivations for Bitcoin Technology Usage. *SIGCHI Conference on Human Factors in Computing Systems (CHI'16)*, 2016.

[103] Hae-Young Kim. Statistical notes for clinical researchers: chi-squared test and Fisher's exact test. *Restorative dentistry & endodontics*, 42(2), 2017.

[104] Philip Koshy, Diana Koshy, and Patrick McDaniel. An analysis of anonymity in bitcoin using p2p network traffic. *International Conference on Financial Cryptography and Data Security (FC'14)*, 2014.

[105] Klaus Krippendorff. Content Analysis: An Introduction to It's Methodology. pages 241–243. SAGE Publications, 2004.

[106] Katharina Krombholz, Karoline Busse, Katharina Pfeffer, Matthew Smith, and Emanuel von Zezschwitz. " if https were secure, i wouldn't need 2fa"-end user and administrator mental models of https. In *2019 IEEE Symposium on Security and Privacy (S&P'19)*, pages 246–263. IEEE, 2019.

[107] Katharina Krombholz, Aljosha Judmayer, Matthias Gusenbauer, and Edgar Weippl. The Other Side of the Coin: User Experiences with Bitcoin Security and Privacy. *International Conference on Financial Cryptography and Data Security (FC'16)*, 2016.

[108] Katharina Krombholz, Wilfried Mayer, Martin Schmiedecker, and Edgar Weippl. "I Have No Idea What I'm Doing"-On the Usability of Deploying HTTPS. pages 1339–1356, 2017.

[109] Todd Kulesza, Simone Stumpf, Margaret Burnett, and Irwin Kwan. Tell me more? the effects of mental model soundness on personalizing an intelligent agent. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1–10, 2012.

[110] Priya Kumar, Shalmali Milind Naik, Utkarsha Ramesh Devkar, Marshini Chetty, Tamara L Clegg, and Jessica Vitak. 'no telling passcodes out because they're private' understanding children's mental models of privacy and security online. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW):1–21, 2017.

[111] Anton J Kuzel. Sampling in qualitative inquiry. 1992.

[112] Gabriella Laatikainen, Taija Kolehmainen, Mengcheng Li, Markus Hautala, Antti Kettunen, and Pekka Abrahamsson. Towards a trustful digital world: Exploring self-sovereign identity ecosystems. In *Pacific Asia Conference on Information Systems*. Association for Information Systems, 2021.

[113] Jerold S Laguilles, Elizabeth A Williams, and Daniel B Saunders. Can lottery incentives boost web survey response rates? findings from four experiments. *Research in Higher Education*, 52(5):537–553, 2011.

[114] Kelsey LaMere, Samu Mäntyniemi, Jarno Vanhatalo, and Päivi Haapasaari. Making the most of mental models: Advancing the methodology for mental model elicitation and documentation with expert stakeholders. *Environmental Modelling & Software*, 124:104589, 2020.

[115] Eric Larcheveque. 2018: A record-breaking year for crypto exchange hacks. `https://www.coindesk.com/2018-a-record-breaking-year-for-crypto-exchange-hacks`, Accessed: 2019-05-06.

[116] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. *Research Methods in Human-Computer Interaction*. Morgan Kaufmann, 2017.

[117] Ji-Hwan Lee and Chi-Hoon Song. Effects of trust and perceived risk on user acceptance of a new technology service. *Social Behavior and Personality: an international journal*, 41(4):587–597, 2013.

[118] Russell V Lenth. Some practical guidelines for effective sample size determination. *The American Statistician*, 55(3), 2001.

[119] Mark W Lipsey. *Design sensitivity: Statistical power for experimental research*. 1989. ISBN 978-0803930636.

[120] Debin Liu, Farzaneh Asgharpour, and L Jean Camp. Risk communication in security using mental models. *Usable Security*, 7:1–12, 2008.

[121] Shanhong Liu. Software developer gender distribution worldwide as of early 2020 . `https://www.statista.com/statistics/1126823/worldwide-developer-gender/`, Accessed: 2021-01-10.

[122] Yang Liu, Debiao He, Mohammad S Obaidat, Neeraj Kumar, Muhammad Khurram Khan, and Kim-Kwang Raymond Choo. Blockchain-based identity management systems: A review. *Journal of network and computer applications*, 166:102731, 2020.

[123] Sonia Livingstone. Children's privacy online: experimenting with boundaries within and beyond the family. 2006.

[124] Mick Lockwood. An accessible interface layer for self-sovereign identity. *Frontiers in Blockchain*, page 63, 2021.

[125] Ewa Luger and Abigail Sellen. " like having a really bad pa" the gulf between user expectation and experience of conversational agents. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 5286–5297, 2016.

[126] C. Lustig and B. Nardi. Algorithmic Authority: The Case of Bitcoin. In *2015 48th Hawaii International Conference on System Sciences*, pages 743–752, Jan 2015.

[127] Alexandra Mai, Leonard Guelmino, Katharina Pfeffer, Edgar Weippl, and Katharina Krombholz. Mental models of the internet and its online risks: Children and their parent (s). In *International Conference on Human-Computer Interaction*, pages 42–61. Springer, 2022.

[128] Alexandra Mai, Katharina Pfeffer, Matthias Gusenbauer, Edgar Weippl, and Katharina Krombholz. User mental models of cryptocurrency systems-a grounded theory approach. In *Sixteenth Symposium on Usable Privacy and Security (SOUPS' 20)*, pages 341–358, 2020.

[129] Alexandra Mai, Oliver Schedler, Edgar Weippl, and Katharina Krombholz. Are https configurations still a challenge?: Validating theories of administrators' difficulties with tls configurations. In *International Conference on Human-Computer Interaction*, pages 173–193. Springer, 2022.

[130] Antonis Manousis, Roy Ragsdale, Ben Draffin, Adwiteeya Agrawal, and Vyas Sekar. Shedding light on the adoption of let's encrypt. *arXiv preprint arXiv:1611.00469*, 2016.

[131] Yuval Marcus, Ethan Heilman, and Sharon Goldberg. Low-resource eclipse attacks on ethereum's peer-to-peer network. *IACR Cryptology ePrint Archive*, 2018(236), 2018.

[132] Emily McReynolds, Sarah Hubbard, Timothy Lau, Aditya Saraf, Maya Cakmak, and Franziska Roesner. Toys that listen: A study of parents, children, and internet-connected toys. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 5197–5207, 2017.

[133] M Granger Morgan, Baruch Fischhoff, Ann Bostrom, Cynthia J Atman, et al. *Risk communication: A mental models approach*. Cambridge University Press, 2002.

[134] Alexander Mühle, Andreas Grüner, Tatiana Gayvoronskaya, and Christoph Meinel. A survey on essential components of a self-sovereign identity. *Computer Science Review*, 30:80–86, 2018.

[135] Alena Naiakshina, Anastasia Danilova, Sergej Dechand, Kat Krol, M Angela Sasse, and Matthew Smith. Poster: Mental models–user understanding of messaging and encryption. In *Proceedings of European Symposium on Security and Privacy. http://www. ieee-security. org/TC/EuroSP2016/posters/number18. pdf*, 2016.

[136] Alena Naiakshina, Anastasia Danilova, Christian Tiefenau, and Matthew Smith. Deception task design in developer password studies: Exploring a student sample. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS'18)*, pages 297–313, 2018.

[137] Nitin Naik and Paul Jenkins. uport open-source identity management system: An assessment of self-sovereign identity and user-centric data platform built on blockchain. In *2020 IEEE International Symposium on Systems Engineering (ISSE)*, pages 1–7. IEEE, 2020.

[138] Satoshi Nakamoto. Bitcoin whitepaper. *URL: https://bitcoin. org/bitcoin. pdf-(: 17.07. 2019)*, 2008.

[139] Peter M Nardi. *Doing survey research: A guide to quantitative methods*. Routledge, 2018.

[140] Alfred Ng. WhatsApp, Telegram had security flaws that let hackers change what you see. `https://www.cnet.com/tech/mobile/whatsapp-telegram-had-security-flaws-that-let-hackers-change-what-you-see/`, 2019. Accessed: 2022-03-14.

[141] Khalil Md Nor and J Michael Pearson. The influence of trust on internet banking acceptance. *The Journal of Internet Banking and Commerce*, 12(2):1–10, 1970.

[142] Donald A Norman. Twelve issues for cognitive science. *Cognitive science*, 4(1):1–32, 1980.

[143] Donald A Norman. Design rules based on analyses of human error. *Communications of the ACM*, 26(4):254–258, 1983.

[144] Maggie Oates, Yama Ahmadullah, Abigail Marsh, Chelse Swoopes, Shikun Zhang, Rebecca Balebako, and Lorrie Cranor. Turtles, Locks, and Bathrooms: Understanding Mental Models of Privacy Through Illustration. Number 4. De Gruyter Open, 2018.

[145] Chinasa T Okolo, Srujana Kamath, Nicola Dell, and Aditya Vashistha. "it cannot do all of my work": community health worker perceptions of ai-enabled mobile health applications in rural india. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–20, 2021.

[146] Marten Oltrogge, Yasemin Acar, Sergej Dechand, Matthew Smith, and Sascha Fahl. To pin or not to pin—helping app developers bullet proof their tls connections. In *24th USENIX Security Symposium (USENIX'15)*, pages 239–254, 2015.

[147] Anthony J Onwuegbuzie and Nancy L Leech. Validity and qualitative research: An oxymoron? *Quality & quantity*, 41(2):233–249, 2007.

[148] Asem Othman and John Callahan. The horcrux protocol: a method for decentralized biometric-based self-sovereign identity. In *2018 international joint conference on neural networks (IJCNN)*, pages 1–7. IEEE, 2018.

[149] Hello Ruby Oy. Hello Ruby Books Series. `http://www.helloruby.com/books`, Accessed: 2021-12-12.

[150] Kostantinos Papadamou, Antonis Papasavva, Savvas Zannettou, Jeremy Blackburn, Nicolas Kourtellis, Ilias Leontiadis, Gianluca Stringhini, and Michael Sirivianos. Disturbed youtube for kids: Characterizing and detecting inappropriate videos targeting young children. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14, pages 522–533, 2020.

[151] Marina Papastergiou. Students' mental models of the internet and their didactical exploitation in informatics education. *Education and Information Technologies*, 10(4):341–360, 2005.

[152] Fabio Parente. *Moving through language: a behavioural and linguistic analysis of spatial mental model construction.* PhD thesis, University of Nottingham, 2016.

[153] Arnis Parsovs. Practical issues with tls client certificate authentication. In *Network and Distributed System Security Symposium (NDSS'14)*, volume 14, pages 23–26, 2014.

[154] Rafael Pass, Lior Seeman, and Abhi Shelat. Analysis of the blockchain protocol in asynchronous networks. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 643–673. Springer, 2017.

[155] Mark Perry and Jennifer Ferreira. Moneywork: Practices of Use and Social Interaction around Digital and Analog Money. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 24(6):41, 2018.

[156] Katharina Pfeffer, Alexandra Mai, Adrian Dabrowski, Matthias Gusenbauer, Philipp Schindler, Edgar Weippl, Michael Franz, and Katharina Krombholz. On the usability of authenticity checks for hardware security tokens. In *30th USENIX Security Symposium (USENIX'21)*, 2021.

[157] Joseph Poon and Thaddeus Dryja. The bitcoin lightning network: Scalable off-chain instant payments, 2016.

[158] Burgess Powell. Not Only Is A 51% Attack On Blockchain Possible, But It's Coming. `https://blocklr.com/news/51-attack-blockchain-more-likely-than-you-think/`, 2018. Accessed: 2019-04-24.

[159] Emilee Rader, Rick Wash, and Brandon Brooks. Stories as informal lessons about security. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*, pages 1–17, 2012.

[160] Elissa M Redmiles, Amelia R Malone, and Michelle L Mazurek. I think they're trying to tell me something: Advice sources and selection for digital security. In *IEEE Symposium on Security and Privacy (S&P'16)*, pages 272–288. IEEE, 2016.

[161] Robert W Reeder, Adrienne Porter Felt, Sunny Consolvo, Nathan Malkin, Christopher Thompson, and Serge Egelman. An experience sampling study of user reactions to browser warnings in the field. In *2018 CHI conference on human factors in computing systems*, pages 1–13, 2018.

[162] Karen Renaud, Melanie Volkamer, and Arne Renkema-Padmos. Why doesn't Jane Protect her Privacy? In *International Symposium on Privacy Enhancing Technologies Symposium*, pages 244–262. Springer, 2014.

[163] Scott Ruoti, Jeff Andersen, Scott Heidbrink, Mark O'Neill, Elham Vaziripour, Justin Wu, Daniel Zappala, and Kent Seamons. "we're on the same page": A usability study of secure email using pairs of novice users. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 4298–4308, New York, NY, USA, 2016. Association for Computing Machinery.

[164] Scott Ruoti, Jeff Andersen, Tyler Monson, Daniel Zappala, and Kent Seamons. A comparative usability study of key management in secure email. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS'18)*, pages 375–394, 2018.

[165] Corina Sas and Irni Eliana Khairuddin. Design for Trust: An Exploration of the Challenges and Opportunities of Bitcoin Users. 2017.

[166] Steven Schirra, Shraddhaa Narasimha, Sasha Volkov, and Justin Owens. Understanding user mental models through app sketches from memory. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pages 1–6, 2022.

[167] Shwadhin Sharma and Babita Gupta. Information privacy on online social networks: illusion-in-progress in the age of big data? In *Analytics and data science*, pages 179–196. Springer, 2018.

[168] Diana LM Sharp, John D Bransford, Susan R Goldman, Victoria J Risko, Charles K Kinzer, and Nancy J Vye. Dynamic visual support for story comprehension and mental model building by young, at-risk children. *Educational Technology Research and Development*, 43(4):25–42, 1995.

[169] Reza Soltani, Uyen Trang Nguyen, and Aijun An. A new approach to client onboarding using self-sovereign identity and distributed ledger. 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 1129–1136. IEEE, 2018.

[170] Reza Soltani, Uyen Trang Nguyen, and Aijun An. Practical key recovery model for self-sovereign identity based digital wallets. In *2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*, pages 320–325. IEEE, 2019.

[171] Yonatan Sompolinsky and Aviv Zohar. Bitcoin's Security Model Revisited. *arXiv preprint arXiv:1605.09193*, 2016.

[172] Anthony Spadafora. Researchers discover security flaws in Telegram encryption protocol. `https://www.techradar.com/uk/news/researchers-discover-security-flaws-in-telegram-encryption-protocol`, 2021. Accessed: 2022-03-14.

[173] Quinten Stokkink, Georgy Ishmaev, Dick Epema, and Johan Pouwelse. A truly self-sovereign identity system. In *2021 IEEE 46th Conference on Local Computer Networks (LCN)*, pages 1–8. IEEE, 2021.

[174] 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.

[175] Anselm Strauss, Juliet Corbin, et al. *Basics of Qualitative Research*, volume 15. Newbury Park, CA: Sage, 1990.

[176] Stuart Sumner. *You: For sale: Protecting your personal data and privacy online*. Syngress, 2015.

[177] Joshua Sunshine, Serge Egelman, Hazim Almuhimedi, Neha Atri, and Lorrie Faith Cranor. Crying wolf: An empirical study of ssl warning effectiveness. In *USENIX security symposium (USENIX'09)*, pages 399–416. Montreal, Canada, 2009.

[178] Andrew Thatcher and Mike Greyling. Mental models of the internet. *International journal of industrial ergonomics*, 22(4-5):299–305, 1998.

[179] Christian Tiefenau, Emanuel von Zezschwitz, Maximilian Häring, Katharina Krombholz, and Matthew Smith. A usability evaluation of let's encrypt and certbot: Usable security done right. In *ACM SIGSAC Conference on Computer and Communications Security (CCS'19)*, pages 1971–1988, 2019.

[180] Andrew Tobin and Drummond Reed. The inevitable rise of self-sovereign identity. *The Sovrin Foundation*, 29(2016), 2016.

[181] Kalman C Toth and Alan Anderson-Priddy. Self-sovereign digital identity: A paradigm shift for identity. *IEEE Security & Privacy*, 17(3):17–27, 2019.

[182] Joe Tullio, Anind K Dey, Jason Chalecki, and James Fogarty. How it works: a field study of non-technical users interacting with an intelligent system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 31–40, 2007.

[183] Jennifer Villareale, Casper Harteveld, and Jichen Zhu. " i want to see how smart this ai really is": Player mental model development of an adversarial ai player. *Proceedings of the ACM on Human-Computer Interaction*, 6(CHI PLAY):1–26, 2022.

[184] Daniel Votipka, Seth Rabin, Kristopher Micinski, Jeffrey S Foster, and Michelle L Mazurek. An observational investigation of reverse engineers' process and mental models. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–6, 2019.

[185] Lev Semenovich Vygotsky. *The collected works of LS Vygotsky: Problems of the theory and history of psychology*, volume 3. Springer Science & Business Media, 1997.

[186] W3Techs. Usage of web servers broken down by ranking. `https://w3techs.com/technologies/cross/web_server/ranking`, Accessed: 2021-02-01.

[187] W3Techs. Web server usage. `https://kinsta.com/wp-content/uploads/2019/06/web-server-usage.png`, Accessed: 2021-02-01.

[188] Qiaosi Wang, Koustuv Saha, Eric Gregori, David Joyner, and Ashok Goel. Towards mutual theory of mind in human-ai interaction: How language reflects what students perceive about a virtual teaching assistant. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2021.

[189] Rick Wash. Folk models of home computer security. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, pages 1–16, 2010.

[190] Rick Wash and Emilee Rader. Influencing Mental Models of security: A Research Agenda. In *Proceedings of the 2011 workshop on New security paradigms workshop*, pages 57–66. ACM, 2011.

[191] Alma Whitten and J Doug Tygar. Usability of security: A case study. Technical report, CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF COMPUTER SCIENCE, 1998.

[192] Alma Whitten and J Doug Tygar. Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0. In *USENIX Security Symposium*, volume 348, 1999.

[193] Zooko Wilcox-O'Hearn. Names: Decentralized, secure, human-meaningful: Choose two. *online] https://web. archive. org/web/20011020191610/http://zooko. com/distnames. html[retrieved 2018-04-21]*, 2003.

[194] Kristin Williams, Rajitha Pulivarthy, Scott E Hudson, and Jessica Hammer. The upcycled home: Removing barriers to lightweight modification of the home's everyday objects. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

[195] Maxine Wolfe. Childhood and privacy. In *Children and the environment*, pages 175–222. Springer, 1978.

[196] Justin Wu and Daniel Zappala. When is a Tree Really a Truck? Exploring Mental Models of Encryption. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS'18)*. USENIX Association, 2018.

[197] Khaled Yakdan, Sergej Dechand, Elmar Gerhards-Padilla, and Matthew Smith. Helping johnny to analyze malware: A usability-optimized decompiler and malware analysis user study. In *IEEE Symposium on Security and Privacy (S&P'16)*, pages 158–177. IEEE, 2016.

[198] Zheng Yan. Age differences in children's understanding of the complexity of the internet. *Journal of Applied Developmental Psychology*, 26(4):385–396, 2005.

[199] Zheng Yan. What influences children's and adolescents' understanding of the complexity of the internet? *Developmental psychology*, 42(3):418, 2006.

[200] Zheng Yan. Limited knowledge and limited resources: Children's and adolescents' understanding of the internet. *Journal of Applied Developmental Psychology*, 30(2):103–115, 2009.

[201] Yaxing Yao, Davide Lo Re, and Yang Wang. Folk models of online behavioral advertising. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 1957–1969, 2017.

[202] Svetlana Yarosh and Pamela Zave. Locked or not? mental models of iot feature interaction. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 2993–2997, 2017.

[203] Jesse Yli-Huumo, Deokyoon Ko, Sujin Choi, Sooyong Park, and Kari Smolander. Where Is Current Research on Blockchain Technology? - A Systematic Review. *PLOS ONE*, 11(10), October 2016.

[204] Richard M Young. Surrogates and mappings: Two kinds of conceptual models for interactive devices. In *Mental models*, pages 43–60. Psychology Press, 2014.

[205] Razieh Nokhbeh Zaeem, Manah M Khalil, Michael R Lamison, Siddhartha Pandey, and K Suzanne Barber. On the usability of self sovereign identity solutions. 2021.

[206] Eric Zeng, Shrirang Mare, and Franziska Roesner. End User Security & Privacy Concerns with Smart Homes. In *Symposium on Usable Privacy and Security (SOUPS'17)*, 2017.

[207] Ren Zhang and Bart Preneel. Lay down the common metrics: Evaluating proof-of-work consensus protocols' security. *IEEE Symposium on Security and Privacy (S&P'19)*, 2019.

[208] Leah Zhang-Kennedy, Christine Mekhail, Yomna Abdelaziz, and Sonia Chiasson. From nosy little brothers to stranger-danger: Children and parents' perception of mobile threats. In *Proceedings of the The 15th International Conference on Interaction Design and Children*, pages 388–399, 2016.