Saarland University
Department of Computer Science

Challenges for Individual Digital Sovereignty in the Context of Security and Privacy

Dissertation
zur Erlangung des Grades
des Doktors der Ingenieurwissenschaften
der Fakultät für Mathematik und Informatik
der Universität des Saarlandes

von
Lea Theresa Gröber

Saarbrücken, 2025

Tag des Kolloquiums:	04.07.2025
Dekan:	Prof. Dr. Roland Speicher

Prüfungsausschuss:

Vorsitzender:	Prof. Dr. Antonio Krüger
Berichterstattende:	Dr. Katharina Krombholz
	Prof. Dr. Andreas Zeller
	Prof. Dr. Alena Naiakshina
Akademischer Mitarbeiter:	Dr. Matthias Fassl

Zusammenfassung

Digitale Souveränität ermöglicht es, selbstbestimmte Entscheidungen über digitale Technologien zu treffen, besonders hinsichtlich Sicherheit und Datenschutz. Diese Dissertation untersucht in vier Studien Sicherheits- und Datenschutzprobleme, die die digitale Souveränität einschränken.

Zunächst analysiere ich selbst gehostete Systeme, die maximale Souveränität bieten. Eine qualitative Studie zeigt, warum Menschen diese nutzen und welche Herausforderungen sie bewältigen müssen. Eine quantitative Studie untersucht die Verbreitung und Merkmale dieser Systeme und deren Betreibern. Beide Studien verdeutlichen das Potenzial, weisen aber auf Sicherheitsprobleme hin.

Weiterhin untersuche ich Mainstream-Technologien, die Nutzerautonomie begrenzen. Eine Studie über autonome Fahrzeuge beleuchtet, welche Informationen Nutzer in sicherheitskritischen Situationen benötigen, um informierte Entscheidungen zu treffen. Eine weitere Analyse zeigt, dass soziale Medien globaler Unternehmen Privatsphäre und Sicherheit nicht ausreichend für alle Nutzergruppen gewährleisten, wie etwa für pakistanische Content Creators. Die Ergebnisse zeigen, dass Automatisierung zwar unterstützt, aber Nutzer in kritischen Kontexten mitbestimmen wollen. Zudem unterstreichen sie die Notwendigkeit alternativer Ansätze, um Abhängigkeiten von Großkonzernen zu reduzieren und digitale Souveränität für alle Nutzergruppen zu stärken.

Abstract

Digital sovereignty empowers individuals to make self-determined decisions and actions regarding digital technologies, particularly concerning security and privacy. This dissertation explores the security and privacy challenges limiting users' digital sovereignty through four studies. First, I investigate self-hosted systems that offer maximum digital sovereignty. A qualitative study explores why people self-host and what challenges they face. A quantitative study examines the prevalence and characteristics of these systems and their operators. These studies highlight the potential for digital sovereignty but underscore significant barriers, especially in securing systems.

Next, I examine mainstream technologies that restrict user sovereignty. A study on autonomous vehicles, which require minimal user input but offer limited control, examines the information drivers need for security-critical situations, enabling digitally sovereign use. Lastly, I analyze global corporations' impact on non-Western populations through a study of Pakistani content creators on social media. The study reveals insufficient safeguards for vulnerable communities in a shifting threat landscape. These studies suggest that while automation can aid, users value informed decision-making in critical contexts. Additionally, reliance on large corporations fails to guarantee security and privacy for all users, emphasizing the need for alternative approaches to enhance digital sovereignty.

Background of this Dissertation

This dissertation is based on the following four papers. I contributed to all papers as the main author. One paper was based on my master's thesis and resulted from a significant extension. I collaborated with people from CISPA Helmholtz Center for Information Security, Nextcloud, Lahore University of Management Sciences, Max Planck Institute for Software Systems, and Georgetown University. One paper was published at the *CHI Conference on Human Factors in Computing Systems*, ranked A* in CORE2023, and three papers were published at the *Usenix Security Symposium*, ranked A* in CORE2023.

[P1] **Gröber, L.**, Mrowczynski, R., Vijay, N., Muller, D. A., Dabrowski, A., and Krombholz, K. To Cloud or not to Cloud: A Qualitative Study on Self-Hosters' Motivation, Operation, and Security Mindset. In: *32nd USENIX Security Symposium (USENIX Security 23)*. 2023, 2491–2508.

[P2] **Gröber, L.**, Lenau, S., Weil, R., Groben, E., Schilling, M., and Krombholz, K. Towards Privacy and Security in Private Clouds: A Representative Survey on the Prevalence of Private Hosting and Administrator Characteristics. In: *33rd USENIX Security Symposium (USENIX Security 24)*. 2024, 6057–6074.

[P3] **Gröber, L.**, Fassl, M., Gupta, A., and Krombholz, K. Investigating Car Drivers' Information Demand after Safety and Security Critical Incidents. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI. ACM, Yokohama, Japan, May 2021, 1–17.

[P4] **Gröber, L.**, Arshad, W., Shanza, Goetzen, A., Redmiles, E. M., Mustafa, M., Krombholz, K., et al. "I chose to fight, be brave, and to deal with it": Threat Experiences and Security Practices of Pakistani Content Creators. In: *33rd USENIX Security Symposium (USENIX Security 24)*. 2024.

Further Contributions of the Author

I co-authored several other peer-reviewed papers with collaboration partners from CISPA Helmholtz Center for Information Security and the University of Vienna. For paper [S3], I share the main authorship with my colleague Matthias Fassl.

[S1] Anell, S., **Gröber, L.**, and Krombholz, K. End User and Expert Perceptions of Threats and Potential Countermeasures. In: *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW 20)*. IEEE. 2020, 230–239.

[S2] Fassl, M., **Gröber, L.**, and Krombholz, K. Exploring User-Centered Security Design for Usable Authentication Ceremonies. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI 21)*. 2021, 1–15.

[S3] Fassl, M., **Gröber, L.**, and Krombholz, K. Stop the Consent Theater. In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (alt.CHI 21)*. 2021, 1–7.

[S4] Hashmi, S., Sarfaraz, R., **Gröber, L.**, Javed, M., and Krombholz, K. Understanding the Security Advice Mechanisms of Low Socioeconomic Pakistanis. In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI 25)*. 2025.

[S5] Roth, S., **Gröber, L.**, Backes, M., Krombholz, K., and Stock, B. 12 Angry Developers–A Qualitative Study on Developers' Struggles with CSP. In: *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security (CCS 21)*. 2021, 3085–3103.

[S6] Roth, S., **Gröber, L.**, Baus, P., Krombholz, K., and Stock, B. Trust Me If You Can–How Usable is Trusted Types in Practice? In: *33rd USENIX Security Symposium (USENIX Security 24)*. 2024, 6003–6020.

Acknowledgments

I owe my deepest gratitude to my advisor, mentor, and friend, Katharina Krombholz. Your unwavering support and encouragement provided a foundation for my growth, both academically and personally. The freedom you granted me to pursue somewhat unconventional topics, the countless thought-provoking discussions that expanded my perspective, and the shared (academic) adventures have profoundly shaped my life.

Collaborating with interdisciplinary and international teams has been one of the most enriching aspects of my research. The diverse perspectives spanning psychology, sociology, computer science, industry, and cultural contexts from Pakistan, the US, and Europe were instrumental in shaping this work. I am deeply grateful to my awesome co-authors Rafael, Nimisha, Daphne, Adrian, Simon, Rebecca, Elena, Michael, Matthias, Abhilash, Waleed, Shanza, Angelica, Elissa, Maryam, Sebastian, Ben, Simon, Sumair, Rimsha, Mobin, and Katharina.

I sincerely thank CISPA for fostering an outstanding and supportive environment where I could thrive professionally. To my amazing colleagues and friends at CISPA—Matthias, Alexander, Carolyn, Divyanshu, Simon, Dañiel, Hossein, Jonas, Simeon, Abdullah, Jannis, Shubham, and Florian—thank you for adding joy and camaraderie to both work and life. Special thanks go to Adrian and Max, who kept encouraging me to finish this work.

My heartfelt thanks go to the people who guided and encouraged me in my first academic steps. Working with Sascha Fahl during my undergraduate years was a transformative experience, and he encouraged me to pursue a Ph.D. after my bachelor's thesis—a path I hadn't envisioned before. I am deeply grateful to Yasemin Acar for her patience, guidance, and mentorship and to Kristina Scherbaum for sparking my interest in research and offering opportunities early on. Mobin Javed, your mentorship and unwavering support have had a profound impact on my career, and I cannot thank you enough.

This journey would not have been possible without the boundless support of my family. To my parents, Simone and Georg, my brother Niels, and my sister Alejandra - thank you for always being there for me. Especially to my partner Hendrik: your love, support, and belief in me have been my greatest strength, even in moments of self-doubt. You are truly one of a kind, and I consider myself incredibly lucky to have you by my side.

Lastly, I am deeply grateful to my friends, whose moral support and encouragement have been a constant source of strength and joy throughout the years. Christoph, Jannick, Luisa, Sebastian, Kerstin, Jennifer, Nico, Felix Z., Felix W., Niklas, Anna, Jonathan, Nicolas, Nadine, and Marius—thank you for being there for me.

This thesis is the culmination of collective support, guidance, and inspiration from all these incredible individuals. I am truly privileged and forever thankful.

Contents

List of Figures

List of Tables

1 Introduction

Today's modern Internet is increasingly centralized and commercialized, which has led to many problems and prevalent privacy violations. These have triggered a discussion about digital sovereignty. My research focuses on cybersecurity as an enabling factor for digital sovereignty. I explore how gaps in human-centric security measures limit individuals' ability to perform their tasks independently and self-determinedly. To motivate my work, I will give insights into issues of the modern Internet and how those have triggered discussions about digital sovereignty. After defining the concept of individual digital sovereignty, I contextualize it with the security and privacy challenges Internet users face. Finally, I organize my contributions to illustrate security challenges that stand in the way of people's digital sovereignty.

1.1 The Commercialization of the Internet

The Internet is one of the most outstanding technical and cooperative achievements of modern times. As of October 2024, 5.52 billion people - 67.5% of the world's population - are active users of the Internet [282]. The Internet revolutionized communication with its decentralized architecture where the "intelligence" lies in the clients, and the network is used solely for transmitting data [312]. This high flexibility and expandability of the infrastructure and applications led to an enormous surge in innovation. As a popular example, the World Wide Web (WWW) [46] is the most important application that triggered the commercialization of the Internet.

In the beginning, Internet technology was highly complex, had poor usability, and required extensive domain knowledge. Thus, it was initially only possible for large institutions, such as universities, to host services. The Internet was only made accessible to the general public through entry-level applications such as browsers. In the course of commercialization, Web applications like Facebook and YouTube have enabled people on a large scale to transition from mere content consumers to content creators. While there was initially fierce competition between companies providing such applications, today, only a few multinational enterprises control the majority of user data. This centralization of the Internet runs counter to the core idea of a decentralized network and leads to serious problems. First, most monopolistic tech companies are based in the US, but their services reach international audiences. Their economic dominance has been criticized for reproducing inequalities in the Global South through dominating the digital ecosystem in these countries [33, 206]. Second, the monetization of the Internet largely runs on advertisements and the ability of companies to predict users purchasing behaviors. Thus, companies are incentivized to harvest as much user and metadata as possible, enabling them to profile their users effectively. Shoshana Zuboff coined this phenomenon "surveillance capitalism" and criticizes it for its disregard for user privacy and as a threat to democratic societies [386].

1.2 Digital Sovereignty

The concept of *digital sovereignty* seeks to counteract the prevalent privacy violations on the Internet by giving people meaningful controls over the technology they use and the data they produce. The term is ambivalent and spans a multitude of contexts and

actors [208]. Among others, it has been discussed in the context of consumer protection and as a means to strengthen people in their role as democratic citizens [208, 287]. While there is no universally agreed-upon description of the concept, the German Competence Center for Public IT offers the following general definition:

> *"Digital sovereignty is the sum of all* ***abilities*** *and* ***options*** *of individuals and institutions to be able to exercise their role(s) in the digital world in an independent, self-determined and secure manner."* [128, 61]

Following the definition, I explain what the terms *abilities* and *options* mean for individuals and contextualize them with security and privacy, as these are central aspects of digital sovereignty according to the definition above.

Ability for secure use An individual's *ability* significantly influences the extent to which they can achieve digital sovereignty. *Ability* refers to the expertise and skills needed for an actor to successfully carry out a task on the Internet, such as creating an account for social media or hosting a personal Web site. Notably, everyday tasks often require security-related knowledge and actions. For example, for account creation, users usually need to create and manage passwords. That is why if one wants to enable people to be technologically self-determined, one also needs to enable them to use said technology securely. In particular, this includes assessing risks, making security-critical decisions, and implementing the corresponding security measures. Even today, these remain hard tasks. In the past 30 years, research on security and privacy increasingly focused on human factors. Previously, humans have only been regarded as the weakest link and root cause of insecurities within technical systems. With three seminal papers from the late 90s [8, 387, 377], the perspective shifted towards user-centered security and a human-centric approach to design security technology. The premise was that security technology was so unusable that it let users down and thus facilitated insecure behavior. Since then, huge efforts have been invested in studying the root causes of security weaknesses by viewing security practices as a socio-technical problem space. A key issue is that security is usually a secondary task to users' primary interaction goal [84]. This can lead to the feeling that security mechanisms stand in the way, tempting users to bypass them. Moreover, security is highly complex, and even experts do not agree on the selection of appropriate defensive mechanisms [301]. In research, there is no consensus on how much a user should or must be able to decide and how much of the security-critical decisions and processes can be automated away [110, 83, 100]. Thus, even today, the secure use of Internet technologies requires domain knowledge, which stands in the way of users' digital sovereignty.

Options for Technologies that Protect or Violate Privacy Complementary, *options* are the technological choices offered to the actor, which may or may not facilitate a digitally sovereign use. For example, a service may grant users control over how data is collected, processed, and stored. The lack of such options, however, will hinder the users' possibility to exert control, even if they have the ability. While there are plenty of protection mechanisms for users in the area of security—even if they require

Table 1.1: Overview of the studies and their contributions which lay the foundation for this thesis.

Privacy Option	Security Abilities	Study	Contribution
✓	●●●	[P1]	Qualitative exploration of security challenges of self-hosting on the case of an open-source project.
✓	●●●	[P2]	Quantifying self-hosting on a representative US sample and statistical description of the self-hosting population.
✗	●○○	[P3]	Identifying need for and information to communicate security-critical situations to drivers.
✗	●●○	[P4]	Exploring threat models and defensive mechanisms of content creators in Pakistan.

"Privacy Option" denotes that the technology studied predominately violates (✗) or respects (✓) its users' privacy. Complementary, "Security Abilities" refers to the skills and expertise the technology demands from the users studied in the respective papers.

considerable knowledge and effort—there are fewer options available to users to protect their privacy effectively. The advertising industry shaped the technological landscape people use today, enabling far-reaching privacy violations. One example of this is the mobile ecosystem, where both Android and iOS phones contain unique identifiers that enable frictionless mapping of a phone user's activity [125, 12]. Data brokers aggregate data across applications, for example, sensitive location information. This enables the creation of surveillance technology for governments that may also be (ab)used by ordinary people [82]. It is almost impossible to circumvent these practices completely, and to attempt to do so, users must compromise on the technology they use and bring time and technical expertise to the table. The majority of users have to rely on the protections available to the masses. Cookie banners are a prominent example from the Web ecosystem. They were introduced as a response to the GDPR in an attempt to grant users more control over data collection. However, they shift the burden to the user while failing to be effective protections [S3, 358, 90].

1.3 Studying the State of Digital Sovereignty on the Internet

This thesis examines the critical role of security in achieving digital sovereignty. I narrow the scope down to the perspective of individuals as users of Internet technology to explore the overarching research question:

What security and privacy challenges limit individuals' digital sovereignty?

With a series of four studies, I explore the tension between technological systems, allowing for different degrees of digital sovereignty. I designed these studies to represent extreme cases of privacy-preserving or violating technology while demanding varying degrees of security-related expertise and skills from the users. This way, through the lens of people's *abilities*, I investigate how they assess threats, make security decisions, and navigate challenges with regard to security technology. I study these in the context of different technologies offering or lacking privacy preserving *options*.

Table 1.1 depicts an overview of the studies and contributions. The following sections provide details on study methods and how the individual studies contribute to a better understanding of my overarching research question.

1.3.1 Digital Sovereignty in the Context of Privacy-Preserving Technology

This section investigates self-hosting as a phenomenon that enables the largest degree of digital sovereignty. As a counter horizon to mainstream technology, self-hosting is the practice of providing a service yourself on the hardware you own or control. With two large-scale studies, I uncover challenges to self-hosting, quantify the phenomena, and describe the population of self-hosters and what makes them unique.

Study A: Self-Hosting Nextcloud Self-hosting allows for maximum control over both hardware and software. This enables hosters to take full control over their data, as they can decide on which servers it is processed and stored. Free and open source software (FOSS) is a popular choice for self-hosting as it additionally provides full transparency of the code that runs on the system. However, self-hosting puts lots of responsibility on the person who maintains the system, something that is not required when using mainstream proprietary services. Critically, self-hosters have to make a magnitude of security-critical decisions surrounding hosting and server management, requiring extensive domain knowledge.

Research focused on security practices of administrators in companies [200, 47, 97], and challenges for private people when administrating IoT devices in a smart home environment [58, 109, 146, 347, 50]. However, there is a critical lack of data studying the phenomenon of self-hosting in all its facets and how it impacts peoples' digital sovereignty. Therefore, I followed an inductive research approach to gather empirical data to explore the dimensions of self-hosting [P1]. Especially, I focused on the reasons why people self-host, how they operate, and how they secure their systems. I investigated self-hosting on the example of Nextcloud, which is a major open-source file-sharing service. In collaboration with the Nextcloud community, I ran a large-scale survey of Nextcloud instances ($N = 994$). Then, I conducted follow-up semi-structured interviews with select survey participants ($N = 41$). This methodology allowed me to tie the broad but coarse survey data to in-depth insights gained in the interviews, thereby exploring the possibilities and challenges self-hosting poses to individuals' digital sovereignty. In particular, the study makes the following contributions. The study:

- Categorizes motivational factors that lead people to self-hosting.
- Describes self-hosting as a socially embedded activity.
- Explores the security mindset and practices of self-hosters.
- Identifies expertise as a major challenge to self-hosting.

This study hints at a major challenge regarding individuals' digital sovereignty: While technology that offers the maximum number of privacy-preserving options is available

through self-hosting, the security of the operations poses a major challenge to individuals and institutions. Without robust security, however, the privacy guarantees that self-hosting offers do not hold in practice.

Study B: Quantifying Self-Hosting and Characterizing Self-Hosters To complement the mostly qualitative findings of study A, I designed a second study to quantify self-hosting on a large scale [P2]. Through a series of two large-scale surveys, I estimated the prevalence of self-hosting in a representative sample of the US population obtained through Prolific. Directly asking about whether a person self-hosts or not is not feasible, as it would lead to a high false positive rate. Thus, the survey elicits self-hosting status through software usage and a series of follow-up questions to determine how the service is provided. In the second survey, I compared the population of self-hosters against a demographically matched control group to identify which individual characteristics are associated with self-hosting. This allows me to reason about potential roadblocks and enabling factors for self-hosting. This study contributes to a better understanding of the prevalence of self-hosting and its role in achieving digital sovereignty on a larger scale. Specifically, it:

- Estimates an upper bound of 8.4% private self-hosters in the US population.
- Offers a comprehensive overview of self-hostable technology across five use cases (communication, file storage, synchronized password managing, websites, and smart home) and the server types self-hosters rely on.
- Provides a demographic description of self-hosters by age, sex, and ethnicity, identifying statistically over- and underrepresented groups.
- Finds that self-hosters are not more privacy aware than the general population. However, the results show that self-hosting correlates with having an IT background, IT administration skills, an affinity for technology interaction, and a "maker" self-identity.

This study hints at challenges to the digital sovereignty of individuals. While software options exist that enable individuals to be fully digitally sovereign, they are not widely used. Results show that self-hosters also use more services in general, so while some people claim to self-host to gain independence, in the grand scheme, one cannot say that self-hosters refrain from using conventional services. However, this might hint at severe usability issues that span the hosting ecosystem, including self-hostable software options.

1.3.2 Digital Sovereignty in the Context of Privacy-Violating Technology

As highlighted in the previous section, self-hosting is not a suitable option for the majority of Internet users. It is technically demanding and requires extensive domain knowledge. Thus, in this section, I take a look at systems that the majority of Internet users rely on. I focus on technologies that are criticized for privacy-threatening business practices and allow for little digital sovereignty. First, I present a study on partially

autonomous cars as an example of systems that demand minimal security abilities from their users as the systems are mature and moving towards autonomy. Second, I present a study on Pakistani content creators as a data point to study the digital sovereignty of a marginalized population while social media services demand average security abilities from their users.

Study C: Autonomous Vehicles Partially autonomous vehicles are driving computers that rely on 1000 to 3000 chips on average [114, 326]. Through this, a magnitude of sensitive data, such as location and driving habits, are available for collection and monetization. The car industry has been criticized for moving towards privacy-violating practices without giving drivers meaningful privacy control options [178]. This makes modern cars a suitable case to study the extreme end of closed systems that limit people's digital sovereignty. Similarly, in the realm of security, partially autonomous vehicles pose an interesting case, as they are an example of mature systems that are designed for high-risk scenarios while requiring minimal input from the driver. At the same time, increasing digitization opens up new attack vectors that remain invisible to drivers. To build trust and to enable users to exert control, prior work in the area of human-computer interaction researched how systems can become more intelligible [218]. A central aspect is to provide users with helpful and relevant information to make technology's inner functioning more perspicuous and to promote better decision-making. However, in the context of high-risk scenarios, people's ability to assess threats and make security-relevant decisions is understudied [218]. Thus, to support drivers with a digitally sovereign use of modern cars in security-critical situations researchers first need to understand if they can assess risks and which information they need to react appropriately. Building on prior work, I designed a mixed-methods MTurk survey ($N = 60$), which relies on scenarios and priming to elicit driver's information demand after safety- and security-critical incidents [P3]. I picked scenarios in which the car malfunctioned, but the error source can be attributed to either a technical malfunction or a malicious intrusion. In two conditions, the survey prompted participants to believe the error was caused by either of the two. A third unprompted control condition in which the error cause was not stated enabled me to uncover if participants think about security compromises in these contexts. This study makes several contributions toward understanding the digital sovereignty of drivers of partially autonomous vehicles. Specifically, it:

- Develops a taxonomy of information types relevant to safety and security-critical scenarios.
- Identifies factors that moderate drivers' information demand.
- Provides implications for designing human-car interaction in the context of technical malfunctions and malicious intrusions.
- Finds that drivers struggle to assess threats related to malicious intrusions, often failing to identify threats and determine appropriate responses.

Moreover, this study hints at a grand challenge for individuals' digital sovereignty in the context of security: Automation is not the sole solution to cope with people's lacking

security-related abilities. Despite the high level of automation, in which users hardly have to make any decisions, the results suggest that drivers want to take action in security-critical situations and require precise information to make appropriate decisions.

Study D: Threats and Security Practices of Pakistani Content Creators Social Media is a prime example of an industry that is widely known for its privacy violations [74, 226, 184]. While decentralized solutions exist [237], people who want to use platforms for growing an audience or building a business are incentivized to stick to big platforms such as Facebook, Instagram, or TikTok. Moreover, these large commercial platforms are international ventures, which makes them suitable testbeds for studying how technology that was developed with a focus on Western populations impacts the digital sovereignty of non-WEIRD (Western, Educated, Industrialized, Rich, and Democratic [162]) users. With a study of Pakistani content creators, I collect data points to better illuminate this research question. Pakistan is a suitable case as it represents an opposite pole to the West due to its religious and cultural background, ranking second to last in gender parity [119]. Especially for women, content creation is a way to earn a living in a culture that otherwise does not facilitate women working outside of the home [42, 255]. Prior work identified threats that apply to US-based content creators [345, 318, 241, 149]. Moreover there is a growing body of work that studies the influence of gender on security and privacy [68, 315, 314, 313, 375, 373], as well as studies in the global south highlighting how woman are especially vulnerable and structurally disadvantaged [363, 16, 315, 314, 261]. As there is no prior work on the security and privacy of content creation in the global south, I explore the inter-sectionalized marginalization of content creators in Pakistan across genders with qualitative research.

I designed a semi-structured interview study with 23 Pakistani content creators across three gender identities [P4]. By focusing on negative experiences, the interview protocol explores how the sociocultural context of Pakistan impacts threats and how defensive mechanisms provided by platforms are lacking, missing, or ineffective. This study contributes to a better understanding of gaps in the digital sovereignty of a non-WEIRD population in a context where they have to rely on privacy-threatening technology. In particular, it:

- Identifies the online and offline threat landscape faced by Pakistani content creators.
- Categorizes technical and behavioral defenses, security mindsets, and support structures that creators rely on.
- Maps defensive mechanisms, mindsets, and external support structures to threat categories, exposing blind spots in the security and privacy options offered by platforms.

Moreover, this study hints at a challenge to the digital sovereignty of marginalized populations: Their threat models are not adequately considered, which leads to gaps in defensive technologies. Thus, in addition to their widespread privacy violations, big technology platforms also fail to offer comprehensive security mechanisms to marginalized users.

1.4 Thesis Structure

This thesis builds on four peer-reviewed papers that are published at the A* venues USENIX Security and ACM CHI, where I am the first author. Each paper comprises one chapter in Part 1 and Part 2 of this thesis. The structure of this thesis is as follows:

Chapter 1: The introductory chapter motivates the research, outlines research questions, describes the contribution of each paper, and provides an outline of the structure of this thesis.

Chapter 2: This chapter presents background and related work.

Part 1: Digital Sovereignty in the Context of Privacy-Preserving Technology

Chapter 3: This chapter describes a mixed-methods study on self-hosting in the case of Nextcloud. The contents of this chapter have been published as parts of the paper: "To Cloud or not to Cloud: A Qualitative Study on Self-Hosters' Motivation, Operation, and Security Mindset." Lea Gröber, Rafael Mrowczynski, Nimisha Vijay, Daphne Muller, Adrian Dabrowski, Katharina Krombholz. *32nd USENIX Security Symposium* (USENIX Security 23) [P1].

Chapter 4: This chapter describes a large-scale survey study to quantify the prevalence of self-hosting and analyze the population of self-hosters. The contents of this chapter are based on the paper: "Towards Privacy and Security in Private Clouds: A Representative Survey." Lea Gröber, Simon Lenau, Rebecca Weil, Elena Groben, Michael Schilling, Katharina Krombholz. *33rd USENIX Security Symposium*(USENIX Security 24). [P2]

Part 2: Digital Sovereignty in the Context of Privacy-Violating Technology

Chapter 5: This chapter describes a mixed-methods online survey to elicit car drivers' information demand in security-critical scenarios. The contents of this chapter have been published as part of the paper: "Investigating Car Drivers' Information Demand after Safety and Security Critical Incidents." Lea Gröber, Matthias Fassl, Abhilash Gupta, Katharina Krombholz. *In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (ACM CHI). [P3]

Chapter 6: This chapter describes a qualitative study to explore the threat landscape and security practices of Pakistani content creators. The contents of this chapter have been published as parts of the paper: " "I chose to fight, be brave, and to deal with it": Threat Experiences and Security Practices of Pakistani Content Creators." Lea Gröber, Walled Arshad, Shanza, Angelica Goetzen, Elissa Redmiles, Maryam Mustafa, Katharina Krombholz. *33rd USENIX Security Symposium*(USENIX Security 24). [P4]

Part 3: Discussion and Conclusion

Chapter 7: This chapter discusses the results of chapters 3 - 6 in light of the overarching research question of security challenges for the digital sovereignty of individual Internet users.

Chapter 8: This chapter summarizes the findings, concludes this thesis, and presents future research directions.

2
Background and Related Work

2.1 Digital Sovereignty

The discourse around digital sovereignty takes place primarily in a political context. To understand the challenges and origins of the concept, I refer to Pohle and Thiel in this section [287]. First, I introduce the traditional notion of political sovereignty, then outline its challenges in the digital space, and finally describe why digital sovereignty has become a popular leitmotif for digital politics today.

In a political context, the term sovereignty was primarily used to describe the independence of states from one another and their control over their own territory [287]. It is important to note that a physical territory was considered a prerequisite for an authority to be able to exercise its sovereignty [145]. However, the importance of the concept of national sovereignty was challenged in the 1990s due to increasing globalization [322] and the growing impact of supranational institutions such as the EU [225], which had a significant influence on the economy and politics of individual states and thus undermined their absolute autonomy. Voices proposing global governance models became prominent, emphasizing the need for stronger international institutions to address global challenges [158]. These ideas influenced the early development and regulation of the Internet [287]. In this context, two concepts were decisive in establishing the exceptional status of the digital space and attempting to exempt it from traditional state-based sovereignty:

- **Cyber exceptionalism** describes the premise that the Internet and digital spaces are fundamentally different from the physical world and, therefore, occupy a special position that requires its own legal and regulatory frameworks [287]. An integral argument in establishing the "exceptional" status of cyberspace refers to its non-territoriality. Johnson and Post proposed that cyberspace is not limited by physical borders and that legal norms based on territorial sovereignty are therefore unsuitable for regulating the Internet [180]. Famously, Barlow declared that cyberspace is "*a world that is both everywhere and nowhere*", asserting its independence from government regulation and the physical world's legal systems [38]. This exemplifies a central line of thought following cyber exceptionalism: as the Internet gains importance, state sovereignty declines [185]. The proposed reasons for this were (1) the decentralized structure and global scale of the Internet, which national legislation is ill-equipped to deal with, (2) a high degree of innovation and rapidly changing business processes, which legislation cannot keep pace with, and (3) the difficulty of holding Internet users accountable for their actions [290]. Cyber-exceptionalism tends to go hand in hand with cyber-libertarianism [188], which characterized the beginnings of Internet commercialization in Silicon Valley [37, 353].

- **Multi-stakeholder Internet governance** is an approach to regulating and managing the Internet that does not require a central decision-making authority and therefore disregards states as sovereign entities [287]. The approach originates from the technical community, where processes and decisions for the maintenance and further development of the Internet are decided in a meritocratic manner. The core principles are inclusivity, openness, decentralization, transparency, and

consensus-based decision-making. Stakeholders in the decision-making process should be all those who are affected by the technology. Thus, it facilitates collaborations between governments, technical experts, civil society, academia, and private companies.

However, time has shown that neither of these concepts caused a fatal disruption in national sovereignty [287], although the practical difficulties of exercising sovereign authority in the digital domain persist [254]. Multi-stakeholder Internet governance has become firmly established in the global policy landscape [287], but it deals with conflicts stemming from coordination problems [230], its rejection of traditional international institutions dominated by governments [287], and a transition from predominantly technical issues to more pronounced political and social issues [230]. The core idea of cyber exceptionalism that the rising importance of the Internet weakens concepts of sovereignty that are tied to territory has not proven to be true, either [287]. On the contrary, the Internet developed in a direction that increasingly made it possible to influence and monitor data flows [287]. Fueled by the commercialization of the Internet, other actors besides states came onto the scene who were interested in a regulatable, less anonymous, and less horizontal architecture, which led, among other things, to the emergence of so-called walled gardens [92, 91]. Walled gardens describe closed ecosystems in which the service providers exercise control over applications and content and limit interoperability with external platforms [280]. Such models have significant implications for user autonomy and privacy but are the common denominator of the big technology companies of today [386]. They contrast with the open nature of the Internet's foundational principles, where users can freely access a wide range of information and services. In fact, platforms and intermediaries have become so central to content distribution that open protocols have diminished in importance [279, 331, 164]. Scholars argued that these actors have grown so powerful that they are hard to govern by traditional means of policy [33, 287, 118], and at the same time they control key areas of public interest such as communication and online markets, making powerful commercial actors quasi-sovereign [118, 287]. These developments, and incidents such as the Snowden revelations [348, 349, 98], have prompted scholars and policymakers alike to call for digital sovereignty [287].

2.1.1 The Role of the Individual

Digital sovereignty is a hotly debated topic in Europe, especially in Germany. The term is not clearly defined and is used in different narratives, depending on the actors and technologies involved. The German Informatics Society (Deutsche Gesellschaft für Informatik) has proposed a unifying definition (translated from German): *"Digital sovereignty is the sum of all abilities and options of individuals and institutions to be able to exercise their role(s) in the digital world in an independent, self-determined and secure manner"* [128, 61]. In political discourse, digital sovereignty is proposed as a means to secure economic prosperity [287, 208], establish security of national infrastructures [287, 208], protect the European way of life [208], and as an integral part of a modern, digitized state [208]. Those narratives have in common that they predominantly focus on states as the main actors [208]. Floridi believes that digital sovereignty should not

replace traditional analog national sovereignty but rather complement it [117]. A key challenge lies in establishing legitimacy: Digital sovereignty, like national sovereignty, must derive its power from the people, e.g., through a democratic process. However, it remains unclear how this principle should be implemented in practice, although the role of the individual in shaping the future is likely to be crucial [117]. For this reason, even if it sounds like a technical problem, digital sovereignty should not just be a matter for specialists [117]. Interestingly, especially voices from the technical community emphasize the importance of empowering individuals to make decisions and take action in an independent and self-determined manner [287, 208, 61]. In this way, the concept of digital sovereignty is not tied to the authority of a state but linked to individuals' abilities [287]. Moreover, it takes the technological options that are available to people into account. In the current technological landscape, those are characterized by business models that incentivize privacy violations [386], and bind users to closed ecosystems [280]. Narratives surrounding people's digital sovereignty are thus linked to strengthening their rights as consumers and democratic citizens, as well as enhancing data protection [287, 61, 208]. However, the discourse portrays individuals primarily in a passive role [208]: they are powerless in the face of monopolistic market forces [117]. In this dissertation, in addition to the effects of walled gardens and systems that require a minimum of user intervention, I explicitly address active behavior through which individuals exercise self-determined digital sovereignty, namely in the form of self-hosting.

2.2 Privacy Technologies

Privacy is a major structuring theme of this thesis, as it is a central challenge that digital sovereignty seeks to address for individuals. In the following, I provide an overview of privacy-enhancing technologies and justify why this thesis focuses on self-hosting as a form of privacy-preserving behavior. Then, I review research on user-facing privacy violations and people's perceptions of them.

2.2.1 Privacy-Enhancing Technologies

In 1997, Burkert described privacy-enhancing technologies (PETs) as "*technical and organizational concepts that aim at protecting personal identity*" [62]. Although PETs generally rely on information security technologies and concepts, such as encryption, Burkert strictly distinguishes them from data security technologies. This is because security primarily safeguards the processing of the data but is not concerned with whether the processing is permitted by the owner of the data. Security is, therefore, a necessary but not sufficient condition for data privacy [62]. The key goal of PETs is to eliminate the collection and usage of personal user data without loss of functionality [62]. If this is not possible, the usage of personal data should be minimized, and users should be granted control over how their data is being processed [361, 62, 342]. Traditionally, PETs grant users these individual control options in the form of *choice* and *consent* [342]. In this respect, PETs are technologies that enable users to become digitally sovereign.

Organizational concepts of PETs can be industry standards or guidelines for the

protection of privacy [342]. Examples are *privacy preference signals* that enable users to make global privacy decisions that are then shared with the services they use without further user interaction. A prominent milestone was the now obsolete protocol *Platform for Privacy Preferences* (P3P) that was developed with the World Wide Web Consortium and recommended in 2002 [367]. It standardized how websites communicate data processing practices, enabling users to quickly understand what happens to their data. However, it saw limited adoption by both browsers and websites. Nevertheless, privacy legislation such as the General Data Protection Regulation (GDPR) [113] and California Online Privacy Protection Act (CalOPPA) [172] affirmed the idea of privacy preference signals, and new standards are being developed today [368, 147, 11]. Technical concepts of PETs are specific tools that enable users to conceal their online identity and exert control over how and which data is collected, processed, and stored [342]. Examples include ad-blockers, cookie banners, end-to-end encrypted messengers, and technologies that provide user anonymity, such as the Tor browser.

PETs yield increased privacy controls for users, however, in reality this does not imply that user's privacy also increases. In 2001, Tavani et al. already sketched out this tension area [342], and recent developments surrounding cookie banners underpin their claim. Cookie banners depend on user consent, which requires informed and voluntary decision-making. However, in practice, consent is frequently influenced by coercive mechanisms, such as dark patterns or the threat of functionality loss [51, 269]. Even if consent is given freely, users may remain unaware of potential secondary uses of their data beyond the primary purpose to which they agreed [342]. Another limitation of PETs is the reliance on user awareness and initiative [342]. People must be informed of their existence and actively seek out and implement these tools to benefit from their functionalities. These challenges highlight why technical tools that put the burden on users alone are not enough to protect people's privacy. Policy and legal advances are needed to incentivize lasting change. In a world where privacy is inherently built into the fabric of the Internet, there would be less need for technical PETs [342].

Self-Hosting

In this thesis I focus on self-hosting as a behavior that enables people to maximize the level of control over how and where their data is stored and processed. Thus, it can be seen as the epitome of privacy-preserving behavior and as an expression of digital sovereignty. This section is an extension of the description of self-hosting of paper [P1].

Self-hosting refers to the process of running and maintaining services or software for personal or organizational use under the direct control of the user rather than relying on third-party shared services (whether paid or free).

Typically, this involves hosting services on-premises, meaning on the service owner's property. However, self-hosting can also include placing privately owned servers in third-party data centers (co-location) or renting servers from such facilities. The level of control varies along a spectrum, from dedicated servers that allow full hardware and software configuration to virtual private servers, which offer software control within a virtual machine.

For the purposes of this thesis, I define self-hosting in three key aspects:

- user control over the hardware, i.e., running on the user's own hardware or them renting said hardware
- control over the software, i.e., the operating system, the configuration
- a dedicated installation for the usage of that user or organization, i.e., including members of that organization in a broad sense, e.g., family members, students of a school, or customers of a printing service

To enable a broad exploration of the concept of self-hosting, the qualitative work [P1] also examines an edge case: Clients of Software-as-a-Service (SaaS) or managed servers are included if the service is a dedicated installation for a single customer on a rented server, with most of the maintenance outsourced to a third party. In such cases, the customer retains low-level access, such as through an SSH shell. In contrast, typical commercial cloud services operate as shared platforms serving millions of customers, where individual customers do not have SSH access or similar control. However, an entity that operates a service exclusively for others (clients) and not for their own use does not qualify as a self-hoster but as a service provider. For instance, a (commercial) SaaS provider running software primarily for clients rather than for their own organization falls outside the scope of this thesis.

Self-hosting is often carried out with open-source software (i.e., which source code is available and compilable to the finished product). Open-source software, also known as free software, is defined by its rights. One has the right to study the code without restrictions, the right to use the software for any purpose free of restrictions like licensing agreements, one has the right to distribute the software without any costs, and one has the right to modify the code and share the modifications. Some of the advantages of Open Source or Free Software, in comparison to proprietary software, are that this model avoids vendor lock-in and monopolization of the market, supports the autonomy of users, and serves users instead of a corporate business model, and the code allows for independent security checks. Examples of such projects are Firefox, Linux, Apache, Signal, Wikipedia, and Nextcloud. However, open-source software is not a necessary condition for self-hosting. The term also applies to hosting closed-source products, such as game servers.

2.2.2 Privacy Violating Technologies

A central challenge that digital sovereignty seeks to address for individual users is the prevalent privacy violations they face in today's technological landscape. In that context, privacy primarily refers to data protection [208]. But it also encompasses the ability to conceal one's online identity, for example, as a means to exercise free speech in oppressive regimes, thereby empowering individuals in their rights as democratic citizens [287]. Analogous to the definition of privacy-enhancing technologies (PETs) [62], the characteristics of technologies that disregard the privacy of their users can be derived. First of all, they are characterized by the excessive collection of data beyond what is necessary to operate the service [284]. In addition, there is usually a lack of transparency. Users are not sufficiently informed about how their data is collected, processed, and stored [240]. Data is passed on to third parties without user consent, and user behavior

is manipulated, for example, to gain consent [51, 269]. Finally, inadequate security can render a service vulnerable and lead to data leaks that violate users' privacy [319].

Prominent examples of privacy-violating technologies can be found in the advertisement industry. Through cookies, browser fingerprinting, and tracking pixels, among other things, users' online activities are tracked, often without clear disclosure. In the area of cell phones, applications have been criticized for unnecessarily requesting permissions that allow them to collect excessive amounts of sensitive data, e.g., by accessing contacts, location, or microphone. Moreover, there exists designated surveillance technology, such as spyware, that is built to run on users' devices and secretly monitor their behavior. In the following, I describe two domains of privacy-violating technology in more detail. This thesis focuses on semi-autonomous vehicles and social networks, as these two domains exemplify contrasting yet complementary aspects of the impact of privacy-violating technologies on individuals and society.

Partially-Autonomous Vehicles

Modern cars are evolving increasingly towards autonomy. To achieve this goal, engineers embed an ever-growing multitude of sensors to capture data relevant to autonomous driving. However, we are still far from having fully autonomous cars at scale. Currently, all models that offer some sort of autonomous driving or semi-autonomous features require strict human supervision. The car industry has recently been heavily criticized for its excessive and obscure data collection practices. In fact, experts from Mozilla examined cars for how they treat driver data and attested that modern cars are the "*worst category of products for privacy that [they] have ever reviewed*" [178]. They investigated 25 major car brands and critiqued that all of them collect excessive amounts of unnecessary personal data. Moreover, this data is used for reasons other than strictly necessary to operate the vehicle or communicate with the driver. Most car companies sell data to third parties, including government and law enforcement. The experts were not able to assess if minimum security standards are met as companies do not disclose information on security practices. To top things off, most brands give drivers no or only little options to control what happens with their data.

Social Networking Services

The major social networking services (SNSs) of today, like Facebook, Instagram, TikTok, and Twitter, are integral to daily communication and social interaction and are prime examples of technologies that engage in privacy-violating practices. The privacy risks of SNSs are well researched and documented [309, 329, 165, 350]. For individual users, there are risks with how the data they upload to a platform is being accessed, shared, and used [321]. Moreover, due to the network effects of SNSs, additional privacy risks emerge based on what other people share about themselves [329]. Due to the vast amount of data being uploaded, it becomes increasingly infeasible for people to manually identify what might be concerning their privacy [329]. Moreover, it is often not clear to people how the data they share and the ways they interact with platforms can be used to target them [106, 7]. In these ways, the concept of control and consent to manage privacy is being challenged in the domain of SNSs [350]. Moreover, the major platforms

are global enterprises expanding privacy violations to billions of people. Widespread consequences include mass surveillance and interference with individual decisions, e.g., in elections, enabled through the data-driven business models of SNSs [386].

2.2.2.1 The Privacy Paradox

Given the prevalent privacy violations and the fact that people continue to use these systems, one could assume that privacy is just not important to people or that they do not care about how companies treat their data. Indeed, scientists have intensively studied how people perceive privacy and their actual behavior in online spaces. Studies in e-commerce were early to point out a possible *privacy paradox* [57, 330]. Although participants had concerns about data misuse or expressed that privacy was important to them, this was not reflected in their behavior. For example, they were willing to disclose more data than previously stated, e.g., in consultations [330] or through loyalty cards [57]. Numerous studies have found evidence in favor of the privacy paradox, e.g., in the domains of SNSs [39, 171, 211, 306, 339]. However, data is inconclusive overall [194], and studies have found a link between privacy concerns on online activities [250, 383, 76]. As more and more perspectives and variables are being considered, enabling the drawing of more ecologically valid conclusions, scholars argue that the privacy paradox may dissolve [96]. This led Hoffmann et al. to propose the term *privacy cynicism* as a more accurate framing to describe people's resignation concerning omnipresent inescapable privacy threats [167].

2.3 Human-Centered Security

Against the backdrop of different privacy-preserving or violating technologies, this thesis makes several contributions to human-centered security. Security is a key concept to consider, as good security is a prerequisite for privacy, which in turn is an integral objective of the digital sovereignty of individuals. However, security is not only an interesting angle to digital sovereignty because it enables privacy. Security topics run through all aspects of human activity in online spaces and increasingly also through our offline world, which is characterized by digitalization. Thus, to enable self-determined technology use in general, it is critical to study how to facilitate secure use in specific. This not only encompasses the development of theoretically and proven correct encryption schemes, protocols, and algorithms but also how these can successfully be applied to real-world problems. The research area of human-centered security is concerned with enabling this by viewing security as a socio-technical problem space. The field established itself through three groundbreaking papers that focused on the poor usability of security technologies relieving users from being framed as "*the enemy*" [8, 387, 377]. Soon, the focus was extended to experts, such as software developers, to combat security vulnerabilities before they arise (e.g. [260, 134, 6, 259, 130]). Researchers quickly realized that even experts have considerable problems with (the development of) security technology. Nowadays, research no longer only focuses on "standard" users but invests considerable resources in researching a wide variety of user groups with different accessibility needs [369]. Moreover, the use of security technology is increasingly

being examined across different social and cultural contexts [43, 370]. The following paragraphs are based on parts of the related work and background sections of papers [P4] and [P1].

2.3.1 Security as a Socially Embedded Task

The research community found that social context plays a role in how people learn about and adopt security practices. One example of this is how people work together to make decisions on security and privacy topics [270]. Prior work has studied different social groups and relationships such as friends [295], acquaintances [203], intimate partners [278], and family members [257, 15, 203, 295]. Research found that people may take a collectivist approach to managing security with individuals becoming "*tech caregivers*" for their community [203, 63]. A study by Watson et al. uncovered how people from the same social group may collaborate to share and secure digital resources. This collaboration is facilitated by collective mental models regarding threat actors and individual accountability [371]. In a cross-cultural study on account security incident response, Redmiles et al. found that Facebook users felt a sense of belonging when seeking support from friends and relatives to address suspicious account login attempts [295]. The same study also highlights cultural differences in how people handle security challenges. People from more collectivist countries rely on their peers more frequently than people from individualist countries [295]. Similarly, research highlighted several cases of how social norms, e.g., about privacy, impact technology use and thus influence threat models [315, 314]. Especially in the Global South, several studies found instances of intermediated technology use to help protect community members [316, 15], and alternate modes of operation, such as shared phone usage [315]. However, there are limitations to relying on communities for collective security management. A central issue is that people are reluctant to share their security and privacy incidents in groups [371] but seek to learn from other's negative experiences [75]. In general, group discussions on security are seldom held, tend to be avoided, and are often superficial [371, 203, 202]. Research assumes this is due to missing incentives for individuals to participate, power imbalances due to lacking expertise and existing hierarchies, and challenges in protecting individuals' privacy when engaging in group discussions on security topics [21, 22, 202, 20, 203].

My work on content creators in Pakistan [P4] and on self-hosters [P1] showcases two instances of how people rely on online and offline communities for protection. In the face of missing or inadequately adjusted defenses provided by platforms, creators in Pakistan rely on their followers and other creators to respond to threats. In addition, creators commonly rely on family, friends, and peer groups for social support to cope with negative experiences. Moreover, by viewing self-hosting as a socially embedded task, my co-authors and I have characterized different modes of operation for providing self-hosted services. We found that people enter specific *operator constellations* to cope with lacking expertise and skills.

2.3.2 Vulnerable User Populations

In both the human-computer interaction and security communities, researchers have identified the need to study vulnerable populations to understand the structural threat landscape these people are exposed to and to investigate possible shortcomings in defenses [370, 242]. In a meta-study, Warford et al. [370] identified *contextual risk factors* that can be used to categorize at-risk populations. *Societal factors* describe how users are at risk due to their cultural and social embeddedness and the public roles they take on. Examples of people facing elevated risk due to taking legal or political action include but are not limited to, people involved with political campaigns in the US [81], activists [86, 196, 233, 340], and journalists [243, 244, 245]. Moreover, research focused on vulnerable, marginalized groups such as LGBTQ+ people [214, 49, 324], sex workers [41, 333], and populations that are at-risk due to low socioeconomic status (e.g., non-Western women [93, 314, 315, 364], people in developing regions [15, 247, 303, 363], and people in developed regions [296, 297, 79, 366]). Apart from that, Warford et al. [370] identified the *relationships* of people (e.g., intimate partner abuse [67, 122, 156, 238]), and the *personal circumstances* (e.g., underserved accessibility needs [175, 17, 234]) as potential risk factors.

My study on Pakistani content creators contributes to gaining a better understanding of at-risk populations [P4]. The intersectional marginalization of content creators in Pakistan, consisting of cultural discrimination against women and non-binary people, and higher risk exposure due to creating content for large audiences, is particularly well suited to explore candidate factors that might impact threats. In the study, my co-authors and I find that content creators who use their voices to make political statements or otherwise comment on societal norms such as religion or sexuality are exposed to elevated risk in the context of Pakistan, especially if they are women or non-binary. Further, my analysis on self-hosters contributes a representative analysis of a population that is believed to be strongly motivated by privacy [P2]. This work extends the recent stream of human-centered security work exploring the practices of different sub-populations by investigating the practices of the general population at large.

2.3.3 Ability for secure Use

A core problem for human-centered security is that people have different abilities in dealing with technology. This refers to both their expertise and computer skills and has been operationalized by Rajivan et al. [294]. According to them, at least four factors contribute to computer security skills and expertise: basic computer skills, advanced computer skills, security knowledge, and advanced security skills [294]. Studies support the suspicion that differences in user expertise lead to unequal, sometimes unpredictable, use of security- and privacy-enhancing technologies [294, 31, 173]. For example, a study by Büchi et al. suggests that general Internet skills enable users to protect their privacy while browsing the Internet [59]. Consequently, these skills play a central role in explaining the privacy behavior of users [59]. Thus, this thesis focuses on systems that demand different levels of expertise from their users, offering a broad view of the resulting challenges to their digital sovereignty. The following paragraphs explain

the levels of user expertise analogous to the classification I use in the Introduction to structure my work (compare Table 1.1).

- **Minimal Abilities: Human Out-of-the-Loop**

  One approach to addressing security issues is to automate them, thereby reducing or eliminating the need for user involvement in decision-making [83, 110]. This can be particularly advantageous in high-risk domains, such as autonomous vehicles, where human errors could have severe consequences. However, security automation presents challenges. One key issue is that automation removes control from users, undermining their sense of agency [110]. This loss of control can conflict with the principles of digital sovereignty, which emphasize self-determination as a fundamental aspect of individual autonomy. Moreover, hiding security features and automating workflows negatively impacts the mental models users develop about how technology functions [379, 110]. Without accurate or complete mental models, users may struggle to make appropriate decisions when technology requires intervention or when automated systems fail [110]. Consequently, there is an ongoing debate about how much users need to know to make self-determined and well-informed decisions regarding security [110, 83, 100]. Finding the right balance between automation and user involvement remains a major challenge, as both the practical benefits of automation and the users' fundamental need for control and agency must be taken into account. My work on drivers' information demand in safety- and security-critical situations in the context of partially autonomous vehicles contributes actionable design implications for the development of interfaces that deal with situations in which automation fails [P3].

- **Average Abilites: End Users**

  I refer to average security abilities to describe the skill set and expertise end users need to carry out everyday tasks on the Internet, such as accessing Websites or using social media. For these tasks, users are generally in the loop when it comes to security practices, such as creating and managing passwords, assessing the trustworthiness of content, and navigating online hate and harassment. Research has put tremendous effort into studying the perceptions and practices of general users regarding security. This thesis focuses on the security abilities of end users in two domains: social networking services (SNSs) and private hosting. SNSs pose a challenge for their users as the online landscape is increasingly characterized by hate and harassment [344]. Content creators specifically are exposed to elevated risks, as they cater to large audiences [345]. My paper on content creation highlights how protecting against these threats poses severe challenges, especially as creators are usually not trained professionals in the domain of information technology and defenses are inadequately adjusted to the sociocultural context of Pakistan [P4]. In the context of private hosting, research investigated how people use their home network for more than just Internet access. There is work on the complexity and challenges of administrating home networks [58, 109, 146, 347, 50], and design guidelines for better home network management tools [288]. Bly et al. [50] investigated the overhead, or *problem-time* people have to invest when

dealing with network devices. Similarly, my work on self-hosting highlights how people's expertise is a major roadblock. This is because self-hosters are a diverse population, and they may become administrators without necessarily having the relevant technical expertise to perform hosting-related tasks [P1].

- **Maximal Abilities: Experts** While my qualitative work indicated that people turn to self-hosting to protect their privacy regardless of their technical expertise [P1], the quantitative analysis yielded a different perspective of the population at large [P2]. My co-authors and I found that technical expertise discriminates self-hosters from the general population and not security or privacy attitudes. There are several studies that compare end users' and experts' mental models of threats and defensive mechanisms [S1, 200, 47, 89]. Especially relevant in the context of hosting is the deployment of HTTPS. Krombholz et al. [200] investigated end users' and administrators' comprehension of the mechanism and found differences in the level of abstraction and perceived security benefits. Similarly, my qualitative paper on self-hosting finds that, especially, non-experts under- or overestimate the level of protection of security mechanisms [P1]. However, being an expert does not imply that hosters have no gaps in their security mindsets [P1]. Similarly, in a study about experts' and non-experts' mental models on VPN [47], Binkhorst et al. found that even experts have misconceptions about the security aspects of VPNs. Still, self-hosting demands technical expertise that the average end user does not have, so even if people start without this knowledge, they have to educate themselves in the process. This is why I consider self-hosters to have above-average to expert technical skills, which study [P2] confirms.

Part I

Digital Sovereignty in the Context of Privacy-Preserving Technology

3

A Qualitative Study on Self-Hosters' Motivation, Operation, and Security Mindset

The contents of this chapter were published as part of the publication *"To Cloud or not to Cloud: A Qualitative Study on Self-Hosters' Motivation, Operation, and Security Mindset"* (USENIX Security 23) [P1]. This chapter uses the academic "we" to highlight the contributions my co-authors, Rafael Mrowczynski, Nimisha Vijay, Daphne A. Muller, Adrian Dabrowski, Katharina Krombholz, and me. The following table details the contributions of each author to this paper:

Author	**Contribution**
Lea Gröber	I developed the original idea of investigating self-hosting and established the connection to Nextcloud. I led all phases of the study, i.e., I developed the method, conducted all the interviews, analyzed the data, and wrote the paper.
Rafael Mrowczynski	Rafael contributed to the qualitative analysis and brought in the perspective of self-hosting as a socially embedded task. He wrote parts in the background and results section and reviewed the paper.
Nimisha Vijay	Nimisha and I conducted the qualitative analysis of the survey. Moreover, Nimisha contributed to the qualitative analysis of the interview data.
Daphne A. Muller	Daphne discussed my initial research ideas with me, coordinated the community survey, provided feedback on the interview guideline, and recruited participants from my shortlists. She also contributed input and feedback on drafts of the background and method sections.
Adrian Dabrowski	Adrian contributed to the qualitative analysis of the interview data. He helped with the framing of the storyline, contributed parts to the background and introduction section, and was overall involved with reviewing and polishing the paper.
Katharina Krombholz	As my academic advisor, Katharina guided key project decisions, provided feedback on my initial ideas and methods, reviewed the final paper before submission, and discussed the conclusions with the team.

Reference

Gröber, Lea and Mrowczynski, Rafael and Vijay, Nimisha and Muller, Daphne A and Dabrowski, Adrian and Krombholz, Katharina. (2023). To Cloud or not to Cloud: A Qualitative Study on Self-Hosters' Motivation, Operation, and Security Mindset. 32nd USENIX Security Symposium, 2491-2508.

Despite readily available cloud services, some people decide to self-host internal or external services for themselves or their organization. In doing so, a broad spectrum of commercial, institutional, and private self-hosters take responsibility for their data, security, and reliability of their operations.

Currently, little is known about what motivates these self-hosters, how they operate and secure their services, and which challenges they face. To improve the understanding of self-hosters' security mindsets and practices, we conducted a large-scale survey (N_S=994) with users of a popular self-hosting suite and in-depth follow-up interviews with selected commercial, non-profit, and private users (N_I=41).

We found exemplary behavior in all user groups; however, we also found a significant part of self-hosters who approach security in an unstructured way, regardless of social or organizational embeddedness. Vague catch-all concepts such as *firewalls* and *backups* dominate the landscape, without proper reflection on the threats they help mitigate. At times, self-hosters engage in creative tactics to compensate for a potential lack of expertise or experience.

3.1 Introduction

The year is 2023 A.D. The Internet is entirely occupied by commercial cloud services. Well, not entirely... One small minority of indomitable self-hosters still holds out against the invaders.[1] Cloud computing has been on the rise for the past decade, and is popular with both individuals and organizations for its scalability, affordability, and accessibility [154]. On the flip side, commercial clouds are criticized for posing privacy risks to consumers [170, 336, 337]. The associated concentration of user data also carries security risks, such as increased attractiveness for attackers due to the proximity of the data [252]. Tim Berners-Lee criticizes the current centralization of the Internet and its services by a few companies as the creation of *data silos* where users' data is locked away. Not only has the user considerably less control, but they also need to trust the service- and the data center operator [222]. Self-hosting is sometimes promoted as an opposition to this development [187], promising to protect and secure one's own data by regaining autonomy. The term *self-hosting* describes maintaining the hard- and/or software for internal and external services on your own as opposed to buying access to these services from a third party [252]. A wide range of self-hostable software covers file synchronization, streaming, calendars, password managers, messaging, and many more [1]. Additionally, self-hosting allows for a diverse set of deployment and configuration options, and, with respect to different threat models, security strategies.

Anecdotal evidence suggests that self-hosters commonly find themselves thrown in at the deep end of suddenly being responsible for an Internet-facing service [2]. In this context, self-hosters represent a special population, as they become administrators without necessarily having the relevant technical expertise nor experience. They are an intermediary group between end-users and professional administrators.

To shed light on security challenges within the complex self-hosting ecosystem, we investigate the security mindset and practices of people with varying levels of technical

[1] If this chapter were a French comic about Romans [131].

expertise who host in personal, organizational, or non-profit contexts. To do so, we combine a large-scale survey (N_S=994) with semi-structured interviews involving selected survey participants (N_I=41). All participants are users of Nextcloud, a well-known self-hostable cloud office suite that covers a wide range of functionality with a variety of apps. The Nextcloud community is a suitable test bed to study the self-hosting phenomenon, as it has a large and active community covering a broad variety of use cases. Hence, with a combination of qualitative and quantitative methods, we answer the following research questions:

RQ1: *What motivates people to self-host?* Uncovering reasons to self-host helps to understand self-hosters' goals and might explain why they make certain (security-relevant) decisions.

RQ2: *How do self-hosters operate?* Understanding admin constellations and social embeddedness will help to uncover unique roadblocks that self-hosters face, and the resources they rely on to overcome problems. Understanding the context of operations is necessary to improve adoption, support, and administrative tools.

RQ3: *What are perceived threats and how do self-hosters manage them?* Analyzing security practices, including attacker modeling, risk perception, and selection of defensive mechanisms is a crucial step to uncover structural gaps in self-hosters' security mindset.

RQ4: *How do self-hosters maintain their operations?* Understanding maintenance, as a facet of security practices, helps to explore self-hosters' security mindset.

RQ5: *In how far does the multidimensional space of self-hosting create tension?* Understanding what problems certain (combinations of) individual characteristics, such as knowledge or motivation, cause and how they affect security outcomes can help people make better decisions.

We found that a lack of it-expertise does not prevent people from self-hosting, especially if they are driven by normative values. To overcome their inexperience, they may enter special operational constellations, such as knowledge barter arrangements, or embed themselves in online communities. Certain motivational factors can impact how participants approach security. The results are meant to guide the development and deployment of helpful advice, information sources, and tools for the self-hosting community.

Replication Package We provide a full replication package and artifact repositories to support open science, reproducibility, and follow-up studies.[2]

3.2 Background

3.2.1 Self-Hosting

Self-hosting refers to running and maintaining services or software under one's own control for personal or organizational use, rather than relying on shared services from third parties. Most of the time, this means the services run on-premise, i.e., on the

[2]https://github.com/usrgroup/USENIX23-selfhosting

service owner's own property, but can also mean putting their own servers in a third-party data center (co-location) or renting servers there. Renting servers falls on a spectrum of various levels of control over hardware and software (e.g. dedicated servers enabling hard- and software configurations; virtual private servers enabling software choices within the virtual machine). It can also include cases where customers have dedicated installations on rented servers with limited access (shared hosting). These Software-as-a-Service (SaaS) instances are included as an edge case, but a typical cloud service without dedicated installations for a single user is not. The three cornerstones for self-hosting are: (i) user control over hardware, (ii) control over software including the operating system and configuration, (iii) and a dedicated installation for the user or organization. Self-hosting is not limited to open-source software and can include closed-source products like game servers.

3.2.2 Nextcloud

Open Source software, also known as Free Software, is software-defined by its rights. One has the right to study the code without restrictions, the right to use the software for any purpose free of restrictions like licensing agreements, one has the right to distribute the software without any costs, and one has the right to modify the code and share the modifications. Some of the advantages of Open Source or Free Software, in comparison to proprietary software, is that this model avoids vendor lock-in, monopolization of the market, supports the autonomy of users, and serves users instead of a corporate business model, and the code allows for independent security checks. Examples of such projects are Firefox, Linux, Apache, Signal, Wikipedia, and Nextcloud.

For this study, we research the self-hosting ecosystem on the example of Nextcloud [264]. Nextcloud developed from a mere file-syncing tool similar to Dropbox to a content collaboration platform with support for office documents, calendars, contacts, forms, and workflow management. They are installed on around 400,000 servers [263] and entail a large online community to tap into. This allows us to study the self-hosting population from a holistic point of view, as Nextcloud is adopted by private, commercial, non-profit, and governmental organizations. All of which we captured in the survey and most of them in the qualitative interviews.

3.2.3 Social Science and Sociology

In our analysis, we will draw on social-scientific concepts and ideas to understand the broader social contexts of self-hosting. In addition to an IT dimension, self-hosting also entails various forms of interactions between humans taking place under specific social-structural conditions. Therefore, we will briefly introduce this terminology.

Social embeddedness encapsulates the idea that all individual human actors (also self-hosters) are involved in various social relations [133, 289, 359]. We distinguish between two dimensions and forms of social embeddedness that we identified in the domain of self-hosting: (1) *Digitally mediated interaction* by which we mean all sorts of social interactions mediated by IT and especially by the self-hosting infrastructure. Examples are sharing photographs with family members, collaborating on a paper draft

existing as a text file in a self-hosted cloud, or processing student data on a self-hosted server of a university. Here, social embeddedness means the social constellation at large in which given digitally mediated interactions take place. (2) *IT operation*, focuses on interpersonal and broader social constellations in which the interviewee's activities, specifically aimed at the operation of the IT infrastructure, take place. These two dimensions can overlap since IT operators often use digital means of communication for coordinating their activities. As for the forms of social embeddedness, we will focus in our analysis in particular on the following:

Organizational embeddedness We understand organizations as clearly defined and coherently acting groups of humans that are meant to exist over longer periods of time. They are usually established to pursue some specified goals. They also have sets of explicit behavioral rules for their members, e.g., duties, explicit commands, membership fees, or general loyalty expectations. Very often, organizational rules include hierarchical relations between different members of an organization [268, 251]. However, even most formalized organizations are simultaneously full of informal relations. Some of those may improve professional interactions, while others can circumvent or even undermine official goals [262].

Collaborative networks is a concept that builds on the broader notion of "social networks" developed in sociology and adjacent social sciences [132, 291, 221, 64]. It denotes frequent interactions between social actors (individuals or groups) based on relevant characteristics of these actors who cooperate without forming an official organization. For example, a particular game programmer and a particular graphic designer frequently cooperate on different projects (i.e., they form a collaborative network) without establishing a formal organization (a small company) because they reciprocally value their particular skills and mutually trust in their abilities "to do a good job."

Knowledge Barter describes an exchange relationship between at least two social actors (usually individuals) who directly trade knowledge-based services (assistance) without using money as a transactional medium [159]. It is a derivative of the broader economic-sociological term *barter* used to denote moneyless exchanges of goods [292]. An often characteristic of knowledge-barter exchanges is the *delayed reciprocity*: actor A does not immediately reciprocate a helping act by actor B, but rather offers his or her assistance to B when the latter is really in need of it and vice versa – often described as "helping each other out" or "exchanging favors."

3.3 Related Work

We study security practices of self-hosters. While there is no directly related work on self-hosting security practices, as a counter horizon we review security, privacy, and human factors research on cloud computing.

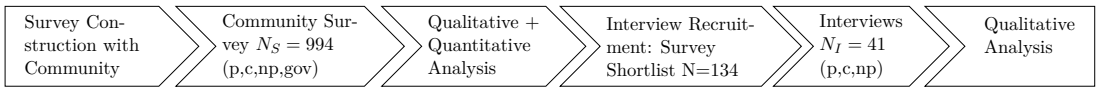


Figure 3.1: Overview of the study process, and user groups involved in each step (**p**ersonal, **c**ommercial, **n**on-**p**rofit, **gov**ernment).

Security and Privacy in the Cloud Monlar et al. [252] discuss the security implications for organizations that move their infrastructure from self-hosting to third-party clouds. Our participants share the concerns that cloud service providers make attractive targets for attackers. Similarly, there is multiple research on the security of commercial cloud computing discussing potential privacy violations [170, 336, 337]. We find that privacy and autonomy are central motivating factors that drive people to self-hosting. While there is research on how privacy in clouds can be achieved by client-side encryption [10, 85], Van et al. [362] argue that cryptography alone cannot solve privacy issues in cloud computing.

End User Perceptions on the Cloud Users that rely on a third-party cloud storage are a counter horizon to self-hosting. There is multiple research on cloud adoption and its influencing factors [14, 124, 129], user perceptions of cloud-services [77, 338, 34, 376, 365], and technology to assist users with data management in third-party clouds [190, 191, 55]. A common theme is that people lack awareness of which data is stored in clouds and that they have a need to take control. Tabassum et al. researched users' understanding of smart home devices and concerns regarding privacy-risking data practices [338]. They found that knowledge of smart homes did not impact their threat models and protection behavior. This is an interesting finding in the context of what motivates people to self-host. We found that IT knowledge alone does not predict if people will become self-hosters.

3.4 Methodology

Since little is known about the phenomenon of self-hosting, we carefully combined different qualitative and quantitative methods to explore the topic broadly. Our study consists of two parts: **(1)** a Nextcloud community **online survey** (N_S=994) covering demographic information about the instance, as well as motivations, use cases, and a coarse security assessment. **(2) semi-structured interviews** (N_I=41) with selected participants of the Nextcloud community survey, focusing on self-hosting as a socially-embedded activity, operator constellations, maintenance practices, threat-modeling, and defensive measures. Figure 3.1 provides an overview of the methodology. We compensated interviewees with 30€[3]; participants of the community survey were not compensated, as the survey was a joint effort with the community.

[3]Some waived compensation, as the study serves the open-source cause.

3.4.1 Study Population and Recruitment

Self-hosting is a broad concept with a broad variety of use cases, motivations, and approaches. Nextcloud is a suitable test bed to study the phenomenon, as it covers a variety of use cases, is open-source, and exhibits a large and active community. Although our findings are in detail Nextcloud specific, we expect generalizability for overarching concepts such as motivation to self-host, structural issues stemming from operator embeddedness, and security assessments of self-hostable solutions. We supposed in advance that there might be differences between personal and institutional usage of self-hosting. To study the phenomenon holistically and to identify areas of tension, we examine the following user groups: personal, commercial, non-profit, and government (the latter is only for the survey and not the interview).

We worked with the Nextcloud community to create a voluntary community survey; see Section 3.4.2 for details. Participants were also asked to indicate whether they would like to be contacted for possible further questions. Based on the survey's records, we shortlisted potential interview partners. We manually selected interesting participants covering a broad set of traits. That is, we accounted for (self-declared) security expertise, team size, reasons to use Nextcloud, and security concerns regarding their instance. We reached out to shortlisted candidates via email and invited them to online video interviews. During the course of the study, we updated the shortlist with complementing candidates according to our recruitment success, until we achieved coverage for the traits. The positive response rate per shortlist was 50% (personal), 26% (commercial), and 29% (non-profit/gov), and no one from governmental users.

3.4.2 Community Survey (N_S=994)

The survey was created in collaboration with the Nextcloud community. Community members started a discussion on the Nextcloud forum about how details of the community are unknown, and the idea arose to gather questions in a shared document [231]. This document was public, so everyone was able to collaborate. The community then reached out to Nextcloud employees who took over the operational aspects of the survey construction and distribution. In addition to the questions from the community, we added complementing questions about motivation, security perceptions, and operator constellation. Finally, Nextcloud's marketing department distributed the survey invitation via their newsletter, and a community member shared it in the forum [232]. After 1000 entries, we closed the survey, and a few open sessions increased the total returns to 1015, resulting N_S=994 after data cleanup. We collected data in September and October 2021. The survey enables us to get a bird-eye perspective of a self-hosting population. It served as a starting point for analysis and provided the basis to select a broad range of participants for interviews (see Section 3.4.3). The survey contained 21 questions in free-text and multiple-choice format focusing on Nextcloud instance-specific information (see replication package). Individual characteristics were subject to the interviews, not the survey. Questions group into three categories: (1) **technical details** such as server type (SaaS, Home Server, Dedicated Server, Virtual Private Server (VPS), Colocation), CPU architecture, Nextcloud version, security concerns (free text), (2) **operator constellation** including team size, and number of people with security background, (3) **details**

on use case such as number of users, population (drop-down: personal, commercial, saas, non-profit, governmental), apps installed (free text), reasons to use Nextcloud, additional self-hosted services. The survey also included some Nextcloud-specific questions about app usage and development requests that are out of scope for this paper.

3.4.3 Interviews (N_I=41)

We conducted 45 in-depth semi-structured interviews (total: 50 hours and 49 minutes; average: 67 minutes) to complement the survey's findings with rich qualitative data but removed four as they did not meet our definition of self-hosting (Section 3.2.1), i.e., they were predominantly hosting for third parties, not for themselves. For an overview of the selection process see Figure 3.2 in the Appendix. Talking to selected survey participants enabled us to tie the in-depth insights from the interviews to the large-scale but coarse picture that the community survey yields and vice versa. Thus, allowing us to explore interesting concepts that surfaced in the survey and fill in the gaps. Those 17 personal, 11 commercial, and 13 non-profit users give insight into the reasons that led them to self-host their service, the social ties in which they operate, and how they maintain and secure their instances. The analysis of the survey informed the development of the interview guide, enabling us to complement the instance-specific data of the survey with concrete technical challenges in a socially embedded context that surfaced during interviews. The resulting interview guide consisted of four parts containing questions that were tailored to the different user groups (personal, commercial, non-profit). See the replication package for the full interview guide.

In the first block, we talked about **reasons for adoption, and areas of application**. To open the conversation, we invited participants to tell us about their professional and educational backgrounds and history. Then we asked about their story of how they became Nextcloud users. Both questions gave us context and doubled as ice-breaker questions. We continued with the participants' privacy notions, how they use Nextcloud, and their technical setup. For organizational users (commercial, non-profit), we additionally asked how the self-hosted service is socially and technically embedded in daily operations, e.g. when working with clients. In the second block, we talked about **maintenance practices**. Participants reported their approach to maintenance, regarding different components of the software and hardware stack. Third, we inquired about **threat models and defensive mechanisms**. Participants told us about any past incidents, their approach to securing their instance, which defensive mechanisms they deployed, who they try to protect from, and where they think their system could be vulnerable. Additionally, we asked organizational users if they have any security policies or guidelines for their infrastructure. Lastly, we complemented missing **demographical information** if not mentioned in the first block. We recorded the age (bracket), gender, country, occupation, and technical & security background. For organizational users (commercial, non-profit) we also asked about the size of the entire organization, sector of operation, and size of the operational and security team.

3.4.3.1 Interview Pre-tests

We conducted two pre-tests to ensure the questions were suitable for IT-savvy and non-savvy participants and understood correctly.The first pre-tester hosts Nextcloud on a home server without a technical background. The second pre-tester studied computer science and hosts Nextcloud instances for commercial and personal use on virtual private servers and home servers. The pre-tests led to minor rephrasing and changes to the order of questions to improve the flow of the interview.

3.4.4 Data Analysis

The qualitative analysis was a multi-step process, involving a total of four coders with different backgrounds (two computer scientists, one designer, and one sociologist). We followed an iterative procedure combining the "top-down" approach of qualitative content analysis [239, 360, 204, 308, 325] with the "bottom-up" strategy inspired by "open coding" in Grounded Theory [66, 332, 224]. First, two coders constructed codebooks for the survey. Based on these, we conducted thematic analysis [56, 360], grouping codes into core themes and concepts. Second, all coders worked together and iteratively analyzed the interview data. For each research question, we tied the coarse findings of the survey to the detailed insights of the qualitative interview analysis. The rich interview data allowed us to explore, confirm, and extend the themes and categorizations of the survey. While analyzing, we always re-read the corresponding transcript segments to make sure the analysis is grounded in data. The following sections provide a detailed description for the survey and interviews.

3.4.4.1 Community Survey

For the analysis of the two qualitative questions about reasons to self-host, and security concerns, we only considered entries that contained an answer to at least one of the two questions resulting in 912 records (see Table 3.2).Two researchers with different backgrounds constructed a separate codebook for each question following the *open coding* approach. The lead author is a computer scientist with a focus on security and privacy who constructed the initial codebooks based on 10% of the dataset. The second author has a background in design and used the initial codebooks to independently code the same percentage of the dataset. The coders discussed their coding and adjusted the codebooks accordingly. Subsequently, they proceeded to iteratively code and discuss portions of the dataset until they reached saturation [356, 48]. Saturation was reached after three iterations, taking into account 27% of the dataset. We conducted two additional rounds of coding where no new high-level concepts emerged. The inter-coder reliability Krippendorff's alpha [198] was between 0.69 and 0.869 for each codebook version. The remaining dataset was split in half among the coders to be analyzed with the final codebooks. Afterward, the coders discussed if new concepts emerged or if any changes were needed. They agreed that the codebooks were stable and needed no further alterations. The final codebooks contain 35 codes and are provided in the supplementary material referenced in Section 3.1.

3.4.4.2 Interviews

We analyzed the interviews starting with some initial thematic codes derived from the survey findings and the interview guide, but enriched and specified them by open coding. For the initial codebook construction, we picked five interviews constituting the presumably most contrastive cases in our dataset. Two coders (computer scientist, sociologist) coded the interviews independently. They discussed their coding and merged the codebooks into one. We started a *documentary* analysis [52, 285] at this point, where we derived themes, concepts, and how the different self-hosters react to similar problems, based on case comparisons. We then proceeded to iteratively code selected interviews to test and contrast patterns in our analysis. That way we reached a stable codebook after coding 25% of the dataset. We proceeded that way, involving two additional coders (computer scientist, designer) until having coded 50% of the dataset and we agreed that we reached saturation. We jointly discussed the codings and identified six axial categories (knowledge, motivation, social embeddedness, it-operations, security mindset, use cases) with respective sub-themes. All coders worked together to write concise summaries of all interviews, containing quotes and references to the raw data, for the sub-themes of the axial categories. This type of analysis does not need an inter-coder agreement calculation, as all codings were jointly discussed and resolved resulting in a hypothetical agreement of 100%. The final codebook contains 585 codes and is provided in the supplementary material referenced in Section 3.1.

3.4.5 Ethical Considerations

This study got approval from the Universität des Saarlandes ethical review board. Self-hosting is a sensitive topic since it revolves around personal data. In particular, the interviews expose personal security and privacy choices and provisions. We made sure that participants were informed about data collection practices prior to taking part in the study. Additionally, before each interview, we thoroughly explained the process in order to obtain informed consent.

3.4.6 Limitations

Generalizability We recruited participants through the Nextcloud newsletter. We selected Nextcloud as a case study because of its widespread use and engaged community. While we have no indication that self-hosters of other projects feel and behave differently, we also cannot deny the possibility. Nonetheless, many of our participants also self-host other projects (see Section 3.5), indicating a large overlap and a similar mindset. Questions about product specifics (e.g., the type of update mechanism) are obviously not generalizable to other projects.

Selection Bias Community-based recruitment limits our view to successful installations. Furthermore, users would need to be invested enough in the topic that they subscribe to the newsletter and volunteer to the interviews. While we asked about installation problems, we were missing out on potentially fatal roadblocks to self-hosting.

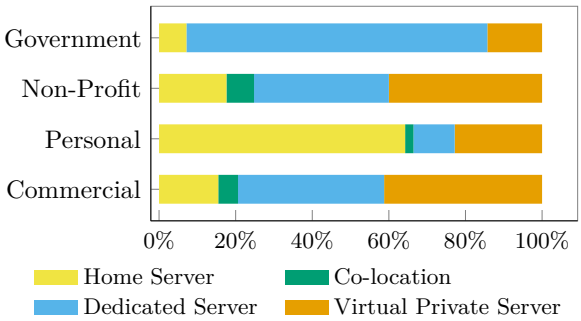


Figure 3.2: *Survey data:* Relative frequencies of reported server types across user groups.

Recall Bias With an interview and questionnaire methodology, self-reported experiences might be several years old. Future work could use more controlled lab or diary studies for details on current, e.g., installation problems, but would miss out on the mindset of the experienced user base.

Social Desirability Bias We mitigated social desirability bias by stressing that the study's goals are to improve Nextcloud and identify roadblocks to self-hosting. In recruitment, participants were not primed for security. During interviews most people did not hesitate to discuss security choices, and were seeking guidance or feedback (which we offered after the interview to avoid bias). Against this background, we assume that our interviewees were relatively open to talk about weaknesses and vulnerabilities of their instances.

3.5 Results

Direct survey$^{S\text{-}x,g}$ and interview$^{I\text{-}x,g}$ quotes are translated verbatim into English where necessary. x denotes the participant id within the dataset (survey or interview), and g adds the group (personal, commercial, non-profit, government).

3.5.1 Demopgraphics

Survey The survey data provides insights on admin constellation and other self-hosted services (compare Table 3.2). Refer to Figure 3.2 for an overview of server types. The majority of participants self-hosted at least one service in addition to Nextcloud. Use cases are broad, from other file storage, synchronization, and file transfer solutions, websites, home automation, communication and messaging tools, password managers, over mail servers, DNS servers, and software development version control, to game servers. The survey indicates that most instances are administrated by single admins (per user group: com 60.37%, p 87.09%, np 55.79%, gov 42.28%). For organizational self-hosting, admin teams between two and three people are common, however, the biggest team with 11 admins was reported by a personal self-hoster. It is not a given that admin teams have members with a security background (on average per user group:

ID	Use Case	Country	Sector of Operation	Occupation	IT Occupation	Hosting Occupation	IT Background	Sec Background	Servertype	Motivation	Social Emb.
13	Non-Profit	Germany	Research Institute	IT System Administrator	●	●	●	○	Private Data Center	A, U, C	Org:Team
17	Non-Profit	France	Freelancer Association	⌀	●	●	●	○	Virtual Private Server	N, C	Coll. Network
20	Non-Profit	Germany	University	Developer/ System Administrator	●	●	●	●	Private Data Center	U, P	Org:Team
21	Non-Profit	Germany	Art Preservation	Art Conservator	○	○	⊙	⊙	Virtual Private Server	U, A, P	Org:Team
23	Non-Profit	Germany	Bicycle Club	IT Consultant	●	●	●	⊙	Private Data Center	C, P	Org:Team
24	Non-Profit	Italy	EU Project	Technical Translator	●	○	⊙	○	Virtual Private Server	N	Org:Team
26	Non-Profit	Germany	Theater Club	Student	○	○	●	○	Virtual Private Server	C	Org:Sole
29	Non-Profit	Germany	Sports Club	CTO in Telecom Company	●	○	●	○	Virtual Private Server	P, A, C	Org:Sole
30	Non-Profit	Germany	School	Media Designer	○	○	⊙	○	Virtual Private Server	C, N, U	Org:Sole
32	Non-Profit	Spain	School	Teacher	○	○	●	⊙	Virtual Private Server	P, N	Individual
36	Non-Profit	Germany	Crisis Line	Professor Computer Science	●	○	●	○	Dedicated Server	U, N, A	Org:Sole
37	Non-Profit	Slovenia	Dataprotection Community	Web Developer	●	○	●	○	Commercial Server	P, C, A	Coll. Network
40	Non-Profit	Switzerland	Political Party	System Engineer	●	●	⊙	⊙	Commercial Server	P, S, A	Individual
2	Commercial	U.S.	Production	CEO	○	○	●	●	Commercial Server	C, A, S, P	Individual
12	Commercial	Germany	IT Consulting	IT Consultant	●	●	●	⊙	Dedicated Server	N, P, A	Org:Sole
14	Commercial	Germany	Law Firm	Lawyer	○	○	○	○	Commercial Server	U, A, P	Knowledge Barter
15	Commercial	Sweden	Journalism	Investigative Journalist	○	○	⊙	○	Co-location	A, P	Knowledge Barter
22	Commercial	France	Consulting	Public Policy Consultant	○	○	○	○	Commercial Server	U, N	Individual
25	Commercial	Canada	Consulting	Consultant	○	○	●	○	Home Server	U	Individual
31	Commercial	Netherlands	Production	System Administrator	●	●	●	⊙	Commercial Server	N	Org:Sole
34	Commercial	France	Travel Agency	Tour Guide	○	○	○	○	Shared Hosting	F, N	Org:Sole
35	Commercial	Germany	Media Design	Freelancer	●	○	⊙	○	Shared Hosting	P, U	Org:Sole
44	Commercial	Netherlands	IT Consulting	IT Support	●	●	⊙	⊙	Virtual Private Server	U, A, P, N	Org:Sole
45	Commercial	Netherlands	Architecture	IT Professional	●	●	●	○	Private Data Center	F, A, U	Org:Sole
1	Personal	New Zealand	⌀	IT Project Manager	●	○	●	⊙	Home Server	A, P, F	Individual
3	Personal	U.S.	⌀	Software Engineer	●	○	●	○	Virtual Private Server	P, A, N	Individual
4	Personal	U.S.	⌀	Networking Systems Engineer	●	●	●	●	Virtual Private Server	C, A, S	Individual
5	Personal	U.K.	⌀	Teacher	○	○	○	○	Home Server	F, A, N	Individual
6	Personal	Italy	⌀	Student	●	○	●	●	Home Server	F, P, A, S	Individual
7	Personal	Germany	⌀	Doctoral Student	○	○	⊙	○	Home Server	U, F	Individual
8	Personal	Germany	⌀	Cloud Architect	●	●	●	●	Home Server	N, P, S, U	Individual
9	Personal	Czech Republic	⌀	Data Specialist	●	○	●	○	Home Server	U, P, A	Individual
10	Personal	Germany	⌀	IT Administrator	●	●	●	○	Home Server	P, A, C, F	⌀
11	Personal	Hungary	⌀	DevOps Engineer	●	●	●	⊙	Virtual Private Server	P, S	Individual
19	Personal	Finland	⌀	Kernel Programmer	●	○	●	●	Home Server	U, A, P	Individual
28	Personal	U.S.	⌀	Software Engineer	●	○	●	●	Home Server	U, A	Individual
33	Personal	Germany	⌀	IT Consulting	●	●	●	●	Home Server	F, N	Individual
38	Personal	U.S.	⌀	Software Engineer	●	○	⊙	○	Virtual Private Server	N, U	Individual
39	Personal	U.S.	⌀	System Engineer	●	●	⊙	○	Virtual Private Server	U	Individual
41	Personal	Germany	⌀	Journalist	○	○	●	○	Home Server	F, N	Individual
43	Personal	France	⌀	Software Engineer	●	○	⊙	○	Home Server	U	Individual

Table 3.1: Interview demographics (four interviews were excluded as they did not match our criteria). ⌀=no answer given; Self-reported IT proficiency: *IT-related Occupation*, *Hosting-related Occupation*, *IT Background*, *Security Background* where ●=yes, ⊙=self-taught, ○=no. *Motivational factors* according to Section 3.5.2: N = Normative, P = Privacy, A = Autonomy, S = Security, C = Cost, U = Use Case, F = Personal Challenge or Fun. *Social Embeddedness* c.f. Section 3.5.3: Individual = individual operators with family & friends, Org:Sole = organizationally-embedded sole operators, Org:Team = team members within organizations, Coll. Network = collaborative networks

com 38.33%, p 49.59%, np 50.07%, gov 80.09%). For single admins, less than half of the people report having a security background (per user group: com 42.18%, p 32.76%, np 31.16%, gov 16.66%). The majority of participants self-hosts at least one service in addition to Nextcloud (per user group: com 66.98%, p 65.89%, np 67.39%, gov 71.42%).

Interviews We interviewed 40 men and one woman. Table 3.1 provides a detailed overview. Participants come from 16 countries across Europe, North America, and Oceania. The participants' professional background is broad and ranges from non-technical occupations (e.g., teachers, journalists, lawyers), to generally IT-related (e.g., developers, data specialists), to hosting-related occupations (e.g., system administrators, system engineers, IT-support). Likewise, the participants' educational background has a similar broad spread and partly reflects different educational opportunities available at the time caused of the wide age range. For 10 participants, high school is their highest level of education, four have completed an occupational apprenticeship, and 27 have a

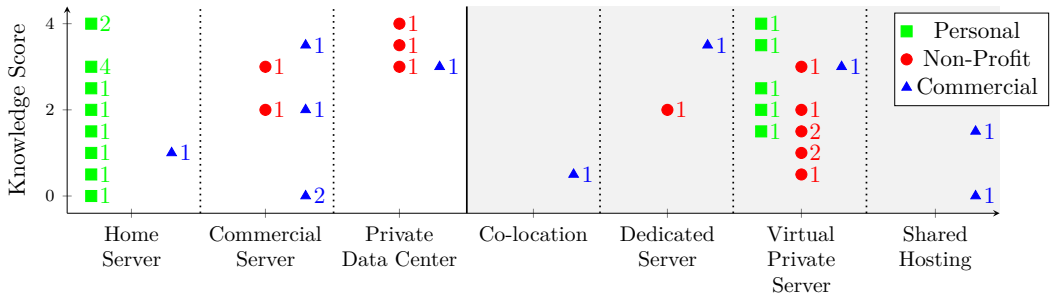
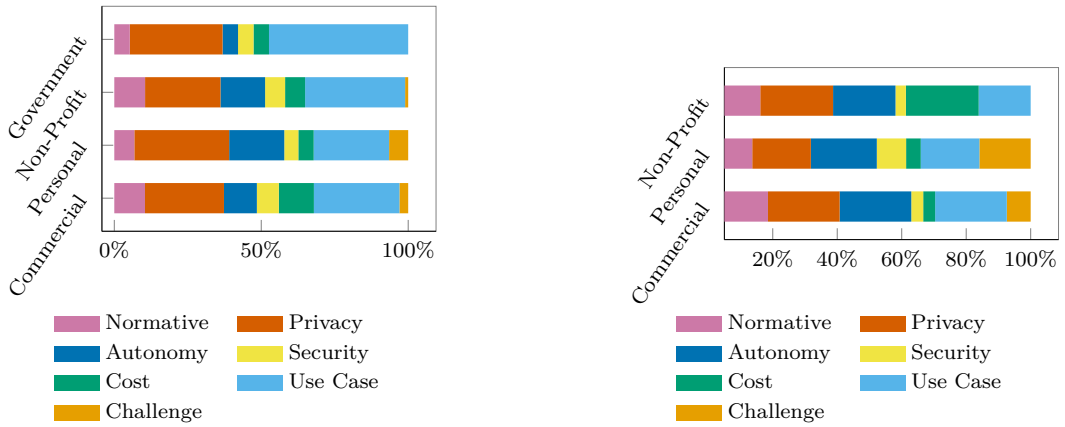


Figure 3.3: Visualization of the interview population by mapping IT expertise against server types. Knowledge score, as a points sum of self-reported *IT background, security background, IT-related occupation,* and *hosting-related occupation* data; where yes=1, self-taught=0.5, no=0 (compare Table 3.1). Server types: on-premise facilities in white, third-party hosting in gray.

university education (BS, MS, PhD, diploma, state examination). While 26 participants report having an educational IT background, 11 people claim to be self-taught, and four say they have no technical background. The distribution (edu/self/none) across user groups is as follows: personal (12/4/1), commercial (5/3/3), and non-profit (9/4/0). Moreover, we asked participants about their security background. Eight participants reported having a security background either obtained through education or extensive work experience, Nine reported being self-taught, and 24 said they had no background in security. The distribution across use cases is as follows: personal (6/2/9), commercial (1/3/7), and non-profit (1/4/8). For organizational use of self-hosting, we cover a broad spectrum of different industries (commercial: travel agencies, law firms, journalists, etc.; non-profit: research institutes, universities, schools, political parties, etc.). Participants relied on a variety of server types to provide their operations. Figure 3.3 provides an overview of the interview population in relating participants' it-knowledge to their set-up choices. Server types are diverse and ranged from under-the-table Raspberry Pis [120][I-5p], repurposed or upcycled hardware[I-28p], to private data centres[I-20n], and hosting on third-party clouds[I-2c].

3.5.2 Motivation

Across all groups, **normative** driven self-hosters practice self-hosting because they see it as the right thing to do. Based on the interview data, we identify two (not necessarily mutually exclusive) sub-types of normative motivation for self-hosting: (1) by general socio-political values (e.g., for the society); (2) by professional-ethical values (e.g., for themselves and their occupation). A common theme for (1) is the appreciation of *"privacy [...] as a fundamental right. I believe that privacy is what protects us from totalitarian states. That means that by exercising or using my right to privacy, I am, in a way, strengthening democracy."*[I-8p] We find (2) for professionals who strongly rely on trustful relations between individual practitioners and their clients (attorney, journalist): *"I deem it utterly unacceptable from the legal point of view when attorneys use official [third party] cloud systems like iCloud or OneDrive and store client data there because they cannot control at all to what extent confidentiality can be ensured there."*[I-14c] Prominent motivational factors are **privacy** and **autonomy**, which often



(a) Survey data

(b) Interview data

Figure 3.4: Relative frequencies of reported motivational factors across user groups.

co-occur: *"I like the idea of having my data being completely private, physically on my own storage devices that I own and I can manage so that I don't have to use other means of encryption or something."*[I-6p] Privacy is about establishing ownership over, and protection of sensitive data. Autonomy refers to the need for independence from third parties (usually commercial vendors), and a need to exert control over technical set-up and configuration. Tightly connected to this, some participants turn to self-hosting to improve **security** in contrast to relying on commercial cloud solutions. Those who mentioned security as a motivation reported having security expertise either through education or self-taught and expressed a need for transparency or a lack of trust in third-party vendors. Across all groups, participants turned to self-hosting to save **costs**, usually to avoid the explicit cost of buying in the market, but it can also reduce administrative costs within organizations. Saving costs, however, can also interfere with privacy considerations, e.g., when non-profit organizations have to opt for cheaper server types. Unlike people who are strongly normative driven, for others, the decision to self-host is a pragmatic one. These people are mostly driven to meet a specific **use case**. Last but not least, participants self-host for the fun of it. They enjoy the **personal challenge** or want to learn something about hosting. Figure 3.4a provides an overview of motivational factors.

Normative Driven (tech extrinsic motivation)

The decision to self-host can be driven by normative considerations of right and wrong human conduct (normative ideas). Hence, normatively motivated self-hosters practice self-hosting because they see it as the right thing to do.

Our interview data allow us to identify two sub-types of normative motivation for self-hosting: (1) by general socio-political values (e.g., for the society); (2) by professional-ethical values (e.g., for themselves and their occupation). As with the more abstract motivational categories, these two sub-categories do not have to be mutually exclusive, but should be analytically distinguished because they are based on

two different normative ideas.

Self-hosters can be motivated by **socio-political values**, such as *"freedom"*[S-397p], *"liberty"*[S-723p], or decentralization. These values are attractive for people with different political views though they tend to understand them differently within their political orientation. Accordingly, self-hosting is also interesting for those who gravitate towards the left side of the political spectrum and for those who position themselves on the right wing[I-5p].

A common theme is the appreciation of privacy as a fundamental right and pillar of free speech and democracy. The focus here is not on the protection of specific data, but more on privacy. *"So privacy I think is important as a fundamental right. I believe that privacy is what protects us from totalitarian states. That means that by exercising or using my right to privacy, I am in a way strengthening democracy."*[I-8p] To illustrate their point, participants regularly refer to privacy violations by *"Big Tech"*[S-805p] or *"GAFAM*[4]*"*[S-785p] and general business practices fueled by data collection as a *counterhorizon*[5].

Some interviewees have a sentiment that services should be paid for with money, not with data: *"I simply have no desire to use any kind of ad-driven or ad-financed server. I [...] have simply internalized: service costs money, and I pay money. [...]I don't like to pay with my data."*[I-36n]. This can also spur a need for autonomy. Others frame these privacy-violating practices as a threat to society.

Participants value open-source technology for its community-driven, inclusive approach to knowledge exchange. For normative-driven people, the fact that open-source technology is mostly free of charge is secondary.

Professional-ethical values are relevant for professionals (e.g., medical doctors, lawyers, tax advisors, chartered accountants, psychotherapists, etc.) who strongly rely on trustful relations between individual practitioners and their clients; hence, require strict confidentiality. Journalists have very similar requirements when it comes to contacts with informants, although those are not strictly their clients. As a consequence, these professions have developed elaborate systems of norms that regulate their members' behavior inter alia of handling confidential information obtained or generated.

Our interview data demonstrate that professional-ethical considerations can explicitly motivate the practice of self-hosting: *"Due to my professional status, I mean, I'm an attorney, I am subordinated to the (principle of) confidentiality, and it also applies to the handling of data. Well, under normal circumstances, I cannot write an e-mail to my clients unless he absolves me from the confidentiality obligation for that particular reason. It's the same when it comes to the issue where I store my data. And I deem it utterly unacceptable from the legal point of view when attorneys use official [third party] cloud systems like [...] and store client data there because they cannot control at all to what extent confidentiality can be ensured there."*[I-14c]

A similar rationale occurs in the interview with a journalist: *"That means that if*

[4]Google, Amazon, Facebook/Meta, Apple, Microsoft

[5]Counterhorizon [52, 285] describes the utterances by research participants reporting practices, behavioral patterns, or habits of others that they consider being misguided in normative terms. Counterhorizons are a common means of expression used for emphasizing the speaker's point (and own righteous behavior) by contrasting with others' wrongful behavior.

the information is stored, the information that the police wants, that I, for good reason, don't want to share with the police because that's not my job. I would lose the trust of my sources and the public if they thought I was a collaborator with the police. [...] However, our CEO is in a different situation than I am. [...] If she gets a search warrant from the prosecutor, she might stand in front of the gate and say, "I don't like this," but she can't chain herself to the door. [...] As a result, I keep the pictures unblurred in my Nextcloud server, where you and I, the source, and myself, as well as other journalists, can work on them. However, the original material stays there, but what is going to be published is what I share with the ⟨B⟩ group, where all the producers and the cameramen need to have access to be able to do the story."[I-15c] (⟨B⟩ is another widely used videoconferencing and filesharing software).

The primary understanding of privacy constitutes the difference between socio-politically motivated self-hosters and those who are driven by professional-ethical considerations. The former think of privacy in general terms and dislike the idea that powerful state or corporate actors can get access to citizens' data. The latter is focused on data of specific people: the data they share with their clients, patients, sources, etc. However, the mutually non-exclusive character of different normative motivations leaves the possibility open that professionally motivated self-hosters also appreciate this form of data handling because of its broader socio-political implications.

Privacy Driven (tech extrinsic motivation) In both interviews and the survey, across all user groups, *privacy* is a prominent motivation, especially in the sense of data protection. Participants distinguish the need to protect personal and organizational data and self-host to establish ownership. This implies knowing where the data is stored, and controlling who can access and share it. *"From a personal perspective, I like the idea that I'm trying to take ownership of the data that I have, and some of it is my behavior. What music am I listening to? What's on my to-do list for today? [...] All the little bits about me. I don't know what gets collected, but it's nice to say that «This isn't collected because I put it in my own little box and nobody gets it but me.»"*[I-2c] Also, some participants expressed the fear of losing access to their data when relying on free-of-charge third-party services: *"[...] If you're using a free service, like most people do, you never know when that's going to be taken away. They just might close your account, and then you can say goodbye to your data."*[I-19p] Depending on the use case and the type of data, taking adequate steps for data protection can also be legally binding. Our EU-based attorney[I-14c] and a journalist[I-15c] argued that American services are not GDPR compliant and thus self-hosting is without alternative: *"And I think it is simply not legally permitted for attorneys to use official [public] cloud systems such as iCloud or OneDrive and store client data there because they cannot control the extent to which confidentiality can really be guaranteed there. And via which servers, via which interfaces, and in which part of the world the data is stored."*[I-14c]

The wish to operate one's self-hosting infrastructure under a trustworthy jurisdiction is an important aspect of privacy-driven SH motivation we discerned in some of our research participants. For Interviewee 15, it is crucial that his server is located in his (Scandinavian) home country where law-enforcement agencies really adhere to the rule of law (that protects him as an investigative journalist), and he knows legal professionals

are able to defend his rights if needed. Hence, he was reluctant to accept a hosting offer from a company based in another EU member state that was a part of the "Soviet bloc" until 1989, because he feared that corruption, believed by him to be still widespread there, could undermine the official EU-compliant legislation (e.g., GDPR) and make his data vulnerable to a privacy breach.

Autonomy Driven (tech extrinsic) We were able to identify two major themes based on the survey and interview data. First, participants seek autonomy in the sense of independence from third parties, i.e., to avoid strong dependencies because of vendor lock-ins or becoming susceptible to pricing changes, policies, or terms and conditions. *"Privacy is a small part of it. I would say a bigger part is not wanting to be reliant on services provided by companies."*[I-28p] They justify this need by a lack of trust in third-party companies, a disagreement with their business practice or philosophy (possible link to *normative* driven), and a general need for transparency thereof; e.g., interviewee 21 requires independence from third-party content policing: *"[some of our artists] work in the artistic-feminist field. They just get kicked off Vimeo, they get kicked off YouTube. And for this reason, for example, we run a PeerTube instance [...], and that's why we have more autonomy."*[I-21n] Second, participants want autonomy in their actions. Here, especially people with IT backgrounds who want control over server locality, set-up, and configurations[S-8n]. However, IT proficiency is not a necessity. Strong needs for privacy[I-14c I-15c] or normative ideas[I-5p] can lead to the same autonomy need.

Exercising full control over hard- and software highlights the aforementioned lack of trust. However, new tensions arise when participants delegate some control over hardware and software, e.g., when hosting on a VPS instead of on-premise. Managing their data on a local server at home gives participants the feeling of autonomy and comfort: *"I like the self-hosted. I like the idea that it's sitting here."*[I-5p] For some, this autonomy and control over the server location are even directly linked to security, which is described in detail in the following section. *"I like the idea of having my data being completely private, physically on my own storage devices that I own and I can manage so that I don't have to use other means of encryption or something."*[I-6p]

Security driven (tech extrinsic motivation) People who explain their reason to self-host with security considerations usually also express a need for transparency or a lack of trust in third-party vendors. *"It's basically because I feel it more secure when I'm operating my Nextcloud instance, and instead of trusting in third parties, I'm storing some sensitive personal data on my Nextcloud instance, and these data are so confidential that I simply cannot trust others to host it and I cannot trust others to secure and store it."*[I-11p] In this sense, the motivational factor *security* is related to *autonomy*. These people believe that full control over hardware and software allows them to make deliberate security decisions to protect their data. This manifests in the ability or wish to deploy specific security mechanisms that they otherwise would not have access to, such as *"server-side data encryption."*[S-210p] In the interviews, this motivational factor was most prominent in the personal use case. Generally, those who mentioned *security* as a motivation reported to have security expertise either through education or self-taught. However, this does not imply a strategic approach towards

security, as we show in Section 3.5.5. Some participants turn to self-hosting because they believe staying off big commercial clouds is inherently more secure, e.g., because big vendors who manage large amounts of data are an attractive target for attackers. Other participants argue, that they cannot compete with the knowledge and people power big tech companies can invest in security, thus making their instances inherently less secure than commercial alternatives.

Cost Driven (tech extrinsic motivation) Across all use cases, participants turned to self-hosting to save cost, although during the interviews, this factor was most prominent within non-profit organizations (see Figure 3.4b). The cost-saving aspect has different facets. First, there is the explicit cost of buying in the market (software, maintenance, subscriptions, support). By relying on open-source technology and forgoing support contracts (which can be one way for open-source projects to monetize), people keep the monetary expenses of self-hosting low. Second, within organizations, there are administrative costs that can be minimized when turning to self-hosted solutions, e.g., not having to do public tenders. On the flip side, self-hosting comes with high costs, which are not a major concern to people in this category: *"A lot of the self-hosting comes right down to money because there are many things and places out there that want me to do, let's say, $5 a month or $30 a month. It all adds up. As a technologically proficient person, I've always been able to do it myself."*[I-4p] According to our participants, the cost of self-hosting is made up of (i) work effort (set-up, technical maintenance, administrative work (e.g., adding users), (ii) time spent to educate oneself, (iii) money spent on hardware, software, and premises, (iv) electricity costs. *"I like to put down the electricity on my home, for the cost, for security, but for climate. [...] I use more electricity with my home computer than the computer at the hosting company."*[I-43p] Cost pressure is a factor that is in constant competition with other motivations. An example of this is Participant 21, who works for a non-profit digital art conservatory and hosts their service on a major public cloud. In this case, the privacy needs are overpowered by cost pressure: *"For example, Nextcloud is attractive to many people who want to do it on-premise. For reasons of data protection and so on. And of course, I think that's great too. But for me, it's more of an afterthought, because we don't have any premises at all, which means that it runs on a public cloud service and then you have to use the cloud storage, otherwise it's unaffordable."*[I-21n]

Use Case Driven (tech intrinsic motivation) Unlike people who are strongly normative driven, for others the decision to self-host is a pragmatic one. Organizations self-host in an effort to digitize their operations, while preserving their *privacy* and *autonomy*[I-14c]. Similarly, personal users are looking for solutions to manage, sync, and share their data in a multi-device set-up[I-39p]. They explain the decision to host based on quality criteria the service they decided for offers. In that line, they perceive Nextcloud as an easy-to-use, easy-to-deploy, stable, secure, high-quality software product that fits their use case.

Interest Driven (tech intrinsic motivation) For some people, self-hosting is an end in itself, which is like a personal challenge for them. In the process, they do not necessarily

have a concrete use case in mind: *"When it was running, I thought, great, now you've got it running, but somehow you're missing the 'use case'. I then gave it up again."*[I-41p] Participants reported being driven by curiosity and the desire to acquire new skills. They enjoy putting together their own set-up and solving problems. People who are driven by fun in this *"tinkering"*[S-595p] were least likely to mention the usability of Nextcloud as a reason to host the service. In contrast, the fact that self-hosting is perceived as something that takes time, knowledge, and effort is concerning or even repelling people of all use cases, both with and without a technical background. The personal challenge motivational factor is most prominent in the *personal* self-hosters. Interestingly, there can be a spill-over effect when people start *"playing with it"*[I-45c] in private, and then adopt it in an organizational context where the motivation is more *usecase* and *autonomy* driven[I-45c].

Key Takeaways: Based on the survey and interviews, we categorized seven motivational factors that led people to self-hosting. People can exhibit multiple motivational factors simultaneously. Thereby, tensions can arise where one factor can outweigh another.

3.5.3 Operator Constellations

In this section we describe self-hosting operations as a socially embedded activity. In the interviews, we learned that self-hosting is practiced in different constellations of social actors. The two major dimensions are *digitally mediated (social) interactions* and *IT operations*. We use the social embeddedness of technical operations as the primary structuring category to identify several types of self-hosters and tie them to the survey results on operator constellations (compare Table 3.2).

Individual operators with family and friends 87.1% of personal self-hosters run their Nextcloud instances on their own without any significant assistance from other individuals. They use them for private purposes and often also host data of their family members, friends, and acquaintances, but they are the only person responsible for the entire operation of the self-hosting infrastructure. Hence, they are usually socially embedded in their digitally mediated interactions (e.g., sharing family photographs, coordinating activities via a self-hosted calendar app, etc.), but they act on their own in the domain of IT operations. The only rudimentary form of social embeddedness in the latter domain is the participation in online forums from where *individual operators* extract needed pieces of IT expertise.

Most individual operators have profound IT knowledge and practical skills providing them with the self-confidence for self-hosting, though only 32.8% of individual operators report having a background in security (through education or self-taught). However, a lack of expertise can lead to a permanent struggle with technical problems. For example, interviewee 5 is strongly driven towards self-hosting by a normative (specifically political) motivation: as a person with a communitarian-socialist mindset, he does not want to use the cloud services of big capitalist companies. But as a humanities teacher, he commands very little IT knowledge. As a result, he is not able to overcome more

complex technical problems: *"Basically, networking is what defeats me, I just don't get it. I set up my own server, which is here, that's it."*[I-5p]

Organizationally embedded sole operators

They are partly similar to the individual operators since they are the only individuals running the respective self-hosting instance. However, as members of organizations, they take certain (formal) responsibility for the functionality of the IT infrastructure including its self-hosting components. Sole and part-time administrators are typically found in small organizations, with other duties on top of it. According to the survey, 60.4% of commercial, 55.8% of non-profit, and 42.3% of governmental Nextcloud instances are administrated by sole operators. Such constellations are particularly common in civic organizations like clubs or associations, where this task is allocated to an (assumed) domain-competent member. This subjective assumption is not necessarily accurate, and competence is relative to other members.

The major similarity between *individual operators* and *sole operators in organizations*: they make their choices without any substantial interference by others; they determine all key aspects of the self-hosting infrastructure on their own; they also outline all further maintenance strategies including security measures. And if they err, there is virtually none to stop or even to warn them. However, online forums were frequently mentioned as an important source of critical information. Thus, we propose that there is something like virtual social embeddedness besides interpersonal social embeddedness.

Key Takeaways: *Sole operators*, both individual and those in organizations, enjoy the highest degree of leeway in determining their entire IT infrastructure, but they may face serious challenges when they hit the limits of their technical expertise.

Team members within organizations

Bigger organizations usually deploy entire teams for IT operations (up to 39.6% of commercial, 44.2% of non-profit, and 57.3% of governmental Nextcloud instances). Hence, self-hosters acting in such organizational contexts are often embedded in a group of people with a certain division of tasks, responsibilities, and expertise. This adds another layer to their social embeddedness besides their general membership in the organization. Members of IT teams usually command extensive expertise which is often the main reason why they became an IT team member in the first place.

Table 3.2: Survey demographics on admin constellation, other self-hosted services in percentages per user group.

		p	com	np	gov
Group size	[#participants]	656	95	131	13
Single admin	[%]	87.1	60.4	55.8	42.3
Single admin w/ security bg.	[%]	32.8	42.2	31.2	16.7
Admin teams security bg.	[%]	49.6	38.3	50.1	80.1
Host additional services	[%]	65.9	67.0	67.4	71.4

We identified two sub-types of specialized teams in which self-hosters can be embedded:

1. Teams with no internal specialization where each individual member is potentially responsible for all IT-related tasks within the organization. This redundancy can provide for continuity in situations when individual team members become temporarily incapacitated. But it may also lead to a confusion of responsibilities, inter alia, in terms of security practices.
2. Teams with internal specialization where each individual member focuses on a limited sub-set of tasks in accordance with their expertise. Individuals maintaining the organizational self-hosting infrastructure in such constellations may partly resemble sole operators with regard to their independence and autonomy. However, differences arise from their embeddedness in a more complex organization, as detailed below.

Organizational embeddedness has both limiting and enabling implications. A bigger and more complex organization with a specialized IT team can restrict individual choices since other organization members may have a stake in (fundamental) decisions whether to cloud or not to cloud as well as more detailed technological choices. In other words, organizationally embedded self-hosters have to account for the diverse needs, preferences, and interests of different stakeholders.

The enabling aspects of organizational embeddedness are exemplified by Interviewee 20, who operates several Nextcloud instances for a night-school program at a German university. Here, he is a member of a small team that can be extended, if needed, by a few student assistants. In addition to that, he can also team up with the employees of the university's central IT services: *"We also have very frequent interactions with the computing center of the university. They also have a Nextcloud instance for the entire university [...]. Whenever they discover a problem, sure, we work then together on a solution. And it is the same the other way around when we discover something."*[I-20n] The third layer of his organizational embeddedness results from institutionalized cooperation between his university's IT specialists and their counterparts at another higher-education institution with an explicitly technical focus. Hence, he is able to easily mobilize additional expertise and workforce available within other segments of a complex organizational structure. This constellation enables him to implement self-hosting solutions that are beyond the reach of individual self-hosters.

Key Takeaways: Organizational team members are more constrained in their choices than self-hosters, but they can more easily receive support from their colleagues whose IT expertise complements their own knowledge and skills.

Collaborative networks refer to the cooperation of multiple individuals based on their personal traits rather than organizational structures. These networks emerge bottom-up as each member decides whether to initiate, continue, or renew cooperation with specific partners. Some participants may take on a pivotal role and lead the network, while others follow their lead.

This social mode of IT operation is vividly exemplified by Interviewee 37: he belongs to a cyber-activist community leaning towards leftist anarchism. He perceives free and open source software and self-hosting as key technologies that make independence from

big capitalist companies possible (normative/socio-political motivation). Together with several other tech enthusiasts gravitating around a non-profit radio station, he operates different online services, including a Nextcloud instance for a community of like-minded users. The services are operated by a team that does not constitute any kind of formal organization: *"People freelance [...] and currently, I'd say that there are almost ten people that come and go, but in equivalent full-time, I would say like five people."*[I-17n] Hence, they form a social network connected by their common socio-political cause. The interviewee describes knowledgeable members of the team training novice admins as part of the community activities aiming at the dissemination of IT knowledge.

As an operator of self-hosting infrastructure, being part of a collaborative network has both benefits and drawbacks similar to those of organizational embeddedness. On one hand, the network can limit choices, but on the other hand, it can provide access to additional skills and workforce. However, the constraints of being in a network are less rigid due to its voluntary nature. Individuals who feel too confined can easily leave since there are no formal exit barriers typically found in organizations.

Key Takeaways: Collaborative networks demonstrate that pooling of self-hosting expertise can also occur without a formalized organizational framework. An important prerequisite for this form of collective operation appears to be a shared motivational idea that coalesces people commanding the required skills.

Knowledge barter This term denotes long-term, non-monetary exchanges of knowledge-based services as regular reciprocal favors. It is a form of social embeddedness that involves individuals with a relatively low level of own IT skills, but a strong motivation to self-host on their own.

This relationship is best illustrated by the case of Interviewee 14, a German defense attorney with a very strong professional-ethical (normative) motivation to self-host. Since he lacks sufficient IT expertise for operating a self-hosted instance at his law firm's premises entirely on his own, he relies on crucial technical support from an *"IT nerd"*[I-14c] (as the interviewee repeatedly calls this person) who happens to also be a client of him in need of criminal defense for alleged digital crimes. The fact that law-enforcement agencies were unable to penetrate "IT nerd's" own systems makes him a credible IT expertise provider in the eyes of the attorney. Such knowledge-barter relationships require a high level of general interpersonal trust because the IT-savvy partner acquires access to the most sensitive parts of the counterpart's computer infrastructure: *"This alternative support is the difficult part for me. It's trust-based because I have to let him in very deep into my system."*[I-14c] Hence, Interviewee 14 is aware of the general issue, but specific cyber-security risks of such a relationship were not discussed in detail by him.

Key Takeaways: *Knowledge bartering* can be a way to overcome one's own lack of expertise, but at the same time, it poses far-reaching security risks.

3.5.4 Maintenance Practices

We asked interviewees about their approach to maintaining their service. Participants broadly regarded updates as a crucial step in providing their operations reliably and securely[I-20n]. While some participants have a structured approach to maintenance[I-40n], others update sporadically[I-30n], or admit neglecting it[I-5p].

Structured We identified different building blocks that suggest a structured approach to maintenance. Participants in this category reported on at least one of these, often in response to negative experiences with updating (e.g., data loss, downtime, functionality loss). **(1)** Participants **defined update cycles** (e.g., weekly to bi-annually) that may vary between different software components. Additionally, these participants frequently reported waiting for a stable version of Nextcloud before updating. A common approach to *"save some work"*[I-21n] is to slack *"a couple of minor releases"*[I-21n] behind the current release, especially when new major releases come out[I-1p]. Some participants reported treating critical security updates differently, immediately updating once they receive notice[I-13n]. Some participants stick to **(2) defined update procedures**, e.g., in the form of self-made technical checklists[I-2c]. Update procedures include making snapshots to recover from failed updates[I-35c] and prior checking forums for reported issues[I-21n]. If the setup allows, some participants carry out the updates step by step, starting with the least critical system (e.g., in case of multiple Nextcloud instances[I-20n]). However, not everyone with this option utilizes it: *"I upgrade all of them at once, but this might not be a good strategy. Recalling what has recently happened. Maybe, I'll just upgrade one at a time and see how it works"*[I-19p]. **(3) Testing** Organizational users reported on testing strategies. Sometimes they have dedicated test instances for development and updates. Additionally, they might define a set of use cases and manually test functionality after updates[I-20n].

Best effort Across user groups, people reported on maintenance behavior that we describe as *best effort*. They do updates sporadically when they have time or get a notification. *"So [updates are] a bit of high life as it comes."*[I-14c] For people who work in distributed admin teams or rely on third parties for help, this may lead to a diffusion of responsibilities: *"Otherwise, our approach is: You check from time to time whether there are security patches, or you check: When is the next big update? Or if it's super important, then someone will get in touch. So if it's really, really important, we also have other admins who let us know."*[I-13n]

No Strategy Some personal and organizational participants report not regularly updating, e.g., not having updated the OS since initial set-up[I-36n]. The choice of infrastructure can block participants' update abilities. One participant opted for a shared hosting set-up and is now stuck with an outdated database which prevents him from updating Nextcloud[I-34c]. Some participants report missing updates regardless of their technical proficiency. Here, one participant identifies his knowledge gaps as a major roadblock: *"No, my maintenance is very poor. I think, I probably reached the limit of my comfortable knowledge with setting it up. [...] and so I'm slightly on a wing*

and a prayer. I do all the updates, the stable updates, and I keep Debian updated when I remember and just hope for the best, which isn't good, not good at all."[I-5p]

Key Takeaways: Maintenance practices are inconsistent. For both organizational and private self-hosters, there are examples of structured and unstructured approaches, with elaborate strategies predominantly found in the organizational context.

3.5.5 Security Mindset

In the survey and in the interviews, different perspectives on cybersecurity emerged. Most participants expressed either a *fatalistic* or *pragmatic* security mindset, which are two sides of the same coin. There is a broad understanding that *"security is [a] prerequisite for everything else"*[S-957g], and that without good security, self-hosting is a lost cause[S-879p]. Similarly, they usually share the view that no software[I-17n] or system can ever be 100% secure[I-20n]. People with a *fatalistic* mindset conclude that therefore a skilled attacker can break into any system, so they *"wouldn't even try [defending]"*[I-44c]. In contrast to that, people with a *pragmatic* mindset acknowledge threats, but conclude that security is achievable[I-2c] when following state-of-the-art security recommendations. These mindsets are relative to attacker models, e.g., people can be *pragmatic* when defending against untargeted external attacks, and *fatalistic* with respect to state actors. We were not able to identify a candidate factor that correlates with the security mindset, e.g., security knowledge does not seem to influence if people are *fatalistic* or *pragmatic*. There is also a third group of people who did not comment on security because they lack the confidence due to a lack of expertise, they completely outsourced security to a third party, or they neglected the topic based on their self-perceived unworthiness as a hacker target[I-15c].

3.5.5.1 Attacker Models

In the following, we describe concrete and unspecific attacker models that surfaced during the interviews.

Targeted State Actor A lawyer[I-14c] and an investigative journalist[I-15c] in our dataset explicitly framed state actors as their most important attacker model. Both have concrete institutions and their capabilities in mind against which they want to protect. Their threat model is based on public knowledge of how these intuitions are legally allowed to operate, and of their own and their colleagues' experiences in dealing with and defending against them. They especially define the threat of them gaining physical access via search warrants. Neither of the two is tech-savvy, so they rely on a *knowledge-barter* (see Section 3.5.3) constellation to secure their operations. However, both have a *pragmatic* mindset with regard to defending against the state actors they defined. This is because they believe self-hosting is ultimately the only way to protect their data, and they have trust in the capabilities of their security operators: *"[the operator] is a former client of mine. And no law enforcement agency in the world had managed to penetrate [their] systems"*[I-14c]. Across all user groups, people are aware of state actors (technical or legal, like above). While most don't view themselves as targets, others

explicitly state that even if they were, they would not stand a chance: *"it would be game over against a national security service. I don't think someone at my level can defend against that, so I wouldn't even try"*[I-44c]. Interestingly, these participants don't refer to concrete capabilities or attack vectors. They seem to view state actors as omnipotent, omniscient adversaries.

Targeted External Attacker Only organizations identified targeted attacks from non-state actors, such as business competitors, opponents to their cause, or personal enemies. Attacker's goals varied: While rivaling artists use hacking as a form of dialogue[I-21n], globally operating energy corporations seek to spy on and sabotage climate activists[I-17n]. Similarly to *state actors*, participants predominately have a *pessimistic* mindset about successfully defending instances: *"Any kind of attacker that can spend on one person that is skilled/motivated for some months would be able to access data. So this is my rough estimation, which is based on nothing"*[I-17n].

Untargeted External Attacker The most prominent attacker model across user groups was untargeted external attacks. Participants frequently referred to automated bots and *"Script Kiddies"*[S-447p], who *"poke around the Internet for the fun of it"*[I-5p]. However, they can also work as a first-stage reconnaissance to select easy targets for ransom or extortion. Most participants rank this to be the top threat they need to address. Although people usually were *pragmatic* in defending against these automated untargeted attackers, especially people with low technical expertise struggled in identifying adequate means of protection: *"My security is probably woeful"*[I-5p]. Additional mismatch emerges when the security mindset is borrowed from the end-user domain, e.g., a personal self-hoster who thinks his Ubuntu server is safe because *"ransomware usually targets Microsoft, not Linux"*[I-5p].

Internal Attacker Few participants also mention a need to protect from internal attackers, such as malicious admins[S-822n]. In the case of off-premise instances, participants often identify the hosting provider as a potential attacker, accessing their data[I-32n]. One participant describes users as a potential threat with regard to data theft[I-40n]. There is a broad understanding that users are not trustworthy[I-4p], but they think it is their incompetence that makes them a risk, not malicious intent: *"[I know my users], so it's unlikely that there would be malicious intent"*[I-44c]. Personal self-hosters do not report users as potential attackers, possibly because their user base mostly consists of friends and family.

Unspecific Attacker Model Across all user groups, participants elicited vague ideas of who could be an adversary to their system. They made unspecific claims that *"everything is a threat"*[I-40n], or that they are not protecting from anyone specifically[I-22c]. Both people with and without IT background or security expertise lack explicit attacker models. Likewise, this attitude is found across all use cases, and particularly concerning in contexts where one would expect elevated threat models, such as schools[I-32n].

Key Takeaways: We identified four attacker models in the interview data. However, only a few participants explicitly analyzed threats and threat actors prior to deciding on defensive mechanisms. The majority had unclear perceptions of attackers' capabilities.

3.5.5.2 Self-Hosters Security Perceptions

78% of survey participants explicitly stated that the security of their Nextcloud instance is a concern to them. They are concerned for a variety of reasons. First, because they think they are an attractive target to attackers based on who they are (e.g., government institution, independent media organization, lawyers), or because of the kind of data they possess (e.g., sensitive private or business data, client data). In particular, personal self-hosters are worried about their *"digital identity"*[I-14c]. As a consequence of a breach, organizations anticipate reputation damage and losing customers' trust[I-13n]. Personal users additionally worry about letting down their family and friends. Second, security in the sense of reliability concerns them as service is a critical infrastructure within their organization. Any downtime or loss of access would negatively impact the organization's day-to-day operations. If the use of the instance is perceived as non-critical *"it's a hobby project"*[S-893p], this can have the opposite effect. Third, taking adequate security measures might be a legal obligation for certain organizations, e.g., if they process personal data that is under GDPR protection. Participants said, them failing to secure their instance makes them liable to prosecution[I-14c].

Perceived Risks Both in the survey and interviews participants expressed their concerns about a variety of risks that they associate with self-hosting.

1. **Hosting on-premise.** Participants referred to the possibility of physical theft or confiscation of data, e.g., in the case of dealing with state actors. Family and friends who also had physical access were not a concern, because of a trustful relationship. Also, participants identified a need to maintain hardware components, e.g., to avoid data loss due to aging hard drives.
2. **Hosting on the public Internet.** One of the most pressing security concerns are Internet-facing instances which participants perceive as the primary entryway for attacks: *"it needs to be accessible easily which is (but does not have to) sort of contradictory to being secure"*[S-810c]. Similarly, the secure configuration of software components such as web servers, databases, and all attached services is understood as the first line of defense. Simultaneously, the potentially complex interactions between software components leave ample room for mistakes: *"I learn what I can [...], but server security feels like a bottomless pit"*[S-118p]. Participants acknowledge the importance of maintaining the set-up, posing a security risk if updates are not rolled out regularly throughout the software stack. This gets complicated if services demand different versions of dependencies, or apps within a service block the update process because they are not compatible with the service's latest version, as can be the case with Nextcloud. Also, participants are concerned with

the rapid update cycle of Nextcloud, feeling overwhelmed when trying to keep up[S-655p].

3. **Software.** Participants are aware of risks, that are generally associated with any software product, such as vulnerabilities in the code, and corresponding 0-days: *"I am afraid of 0-days at all levels of my Nextcloud/Linux system as state-sponsored attackers have access and vulnerabilities for all types of infrastructures and software"*[S-149p]. Moreover, they worry about supply chain attacks, especially when it comes to using pre-configured yet unsigned docker containers[I-19p]. In the case of Nextcloud, some worry about the underlying substrate as *"PHP has a reputation for security problems"*[S-235p]. Many participants view third-party apps for Nextcloud as one possible entryway into their system[S-628p].
4. **Admin capabilities.** With self-hosting, as opposed to relying on hosting providers, participants mainly identified two risks: First, knowledge gaps with respect to general server setup, the configuration in general, and security expertise in specific: *"[I] only have pro-amateur know-how"*[S-555p], *"I am no expert, so it could leak any moment"*[S-123p]. Second, they acknowledge a lack of resources, e.g., time and team size to properly secure the instance: *"As much as I don't trust Silicon Valley with my data. I always have to think they have more people working on security than I could have."*[S-15p].
5. **Users.** Users are broadly viewed as a risk to the system. Participants usually see them in a passive role where they fall victim to malware[S-586p], viruses[S-656c], and ransomware[S-710c]. *"I am more afraid of the users being stupid than the box being hacked"*[S-1002p].

Trust Anchors In both the survey and interviews, participants named several factors that alleviate their security concerns. We distinguish these *trust anchors* from actively deployed defensive mechanisms like 2FA, HTTPS, or security training for end users, as they are things out of participants' control or tools they use to assess security. Open-source software is a central trust anchor, because of its transparency, especially if a large community is involved. Here, participants also rely on social proofs to manage their security needs: *"I rely on the community average needs [for security]"*[S-2c]. In this context extensive documentation, including guidelines also builds confidence: *"[I] rely on well-documented software that I can trust"*[S-29p]. Participants are aware about yet undetected software vulnerabilities, but Nextcloud's bug bounty program helps to establish trust. When using third-party hosting providers, users are less worried depending on the server's applicable jurisdiction, e.g. EU[I-32n]. Participants often rely on security audits as feedback channels, e.g., automated scanners[I-5p], or more seldomly on penetration tests[I-31c]. We found that in particular non-professionals widely acknowledge audits as useful: *"That would be what I'd pay for, is a security audit."*[I-5p]. If these certify a good score, it relieves participants' security concerns. However, this can also lead to a false sense of security, e.g., if people rely on outdated or incomplete scanners[I-25c]. Having security knowledge was a trust factor for some[I-28p]. Others shifted admin responsibilities, e.g., by relying on external maintenance, such as NextcloudPi[S-370p], however, often security remains a concern.

Key Takeaways: The security of their operations is a concern for the majority of participants. Regardless of technical expertise, they are creative with identifying potential risks spanning across hardware, software, network, and human factors. Measures that create transparency, and social proofs are important trust anchors.

3.5.5.3 Perceptions of Defensive Mechanisms

During the interviews, we discussed concrete mechanisms participants use to secure their operations. Across all user groups, participants report following security advice[I-5p], best practices[I-2c], and documentation[I-21n]. One participant wrote their own security mechanisms to protect against and monitor suspicious activity and explained why: *"I want to know what's going on in the software. I only trust software as far as I can see what's happening."*[I-44c] While an overview of all defensive mechanisms is presented in Appendix Figure A.1, we describe selected, controversial ones below. **Firewalls** are very popular with our participants. They use them to separate subnets[I-40n] from each other, and to restrict access from outside to selected ports[I-2c], giving them a secure feeling: *"I don't have to pay attention to what services are running and what ports they have open"*[I-35c]. One person combined them with self-written intrusion detection[I-44c]. While firewalls are *"the most important thing"*[I-1p] for some, others leave them out completely: *"I believe that the firewall issue is simply misunderstood in the vast majority of cases. A firewall [...] only does port filtering in 99% of the installed instances. [...] But that doesn't make any difference if you simply check what else is running on the instance and simply shut down these services"*[I-33p]. This participant prefers a simple set-up because he had a negative experience with a firewall appliance that broke down and shut him out of his instance when he was abroad. **End-to-end encryption (E2EE)** is requested by many participants, but Nextcloud's implementation is not feature complete [265]. Self-hosters without their own hardware see this as a way to protect their data from unauthorized access by hosting providers[I-32n]. As a consequence of E2EE, participants have concerns about complexity of key management, increased computing load on the server, reduced recovery options in the case of data loss, and users not understanding the mechanism and its implications[I-40n]. Interestingly, participants often have misconceptions about the security benefits E2EE provides over other measures (e.g., over HTTPS[I-5p], server-side, and hard-drive encryption[I-15c]). **Backups** are considered to be very important by most participants. They often make participants feel safe, even if they are aware of their poor security strategy. Interestingly, this is also the case if participants do not perform regular backups[I-5p]. **Two-factor authentication** is an example of how participants deploy different levels of security on user or instance level[I-13n], e.g., enforcing 2FA for administrators or users with access to sensitive data[I-33p]. In this context, participants report on challenges explaining the concept of 2FA to non-tech-savvy users[I-1p]. In Nextcloud 2FA is implemented in apps (add-ons). Participants reported being locked out of their instance when the 2FA app did not work after an update[I-2c]: *"If I have to turn off two-factor authentication, I don't think that it means tomorrow somebody's going to get hacked but it means that people have trouble logging in, because all of a sudden their method for logging in changed and that is when people take shortcuts that leads to security risks."*[I-2c] Because of these issues, some participants reported being reluctant to adopt 2FA.

Key Takeaways: Participants find it difficult to select suitable security mechanisms. There is a tendency to pick supposed "catch-all" defenses, while the actual effectiveness and security guarantees are often unclear.

3.6 Discussion

We discuss security mindset and practices in relation to administrator constellations and identify high-level gaps in participants' reasoning. Moreover, we relate individual characteristics to participants' server-type choices. Last, we discuss areas of tension and outline recommendations.

3.6.1 Gaps in Security Mindsets (RQ 3,4)

Participants have contradicting perceptions about the security of self-hosting. Some turn to self-hosting because they believe staying off big commercial clouds is inherently more secure, e.g., because data proximity makes big vendors an attractive target for attackers. Researchers likewise identified this threat in the context of cloud computing [252]. Other participants argue that they cannot compete with the knowledge and resources of big tech companies, thus making their instances inherently less secure than commercial alternatives. While it is difficult to compare security across instances and organizations, research on end user and expert perceptions of threats reveals different levels of abstraction and comprehension [200, 47]. We found gaps and inconsistencies in the security mindsets of both personal and organizational self-hosters. Thereby, neither the technical expertise nor administrator constellations imply a structured approach to security (a.k.a. threat modeling). This suggests, that even experts who are socially embedded into organizations struggle with a systematic approach to security. Understanding gaps in security mindsets will allow academics and practitioners to develop tooling and targeted information sources to help users in securing their instances. Based on our findings, we discuss four major gaps in participants' security mindsets:

(1) Attacker models are often unclear or non-existent. Most participants do not actively model attackers, even if they have the technical expertise. This is true for both single operators and people working in teams in an organizational context. When asked, participants were not confident in identifying potential attackers and which capabilities they would have. Unawareness of possible attackers or their capabilities is detrimental, as it is an essential step in modeling threats and implementing effective security mechanisms.

(2) The data suggests that most participants find it difficult to **prioritize risks**. While participants are in general comfortable identifying potential risks, they tend to lose track in the face of the multitude of potential vulnerabilities. For some, this gives the impression that all efforts are wasted and security cannot be achieved. This is especially an issue for self-hosters who cannot draw on additional resources or who have a limited time budget to secure their operations.

(3) Participants struggled with **identifying defensive mechanism** that are suitable for their use case and set-up. The mapping of perceived risks to defensive mechanisms is

especially hard for them, all the more if an understanding of potential attacker models is missing. Most participants were somewhat confident in naming risks they thought could apply to their operations. However, when it came to defensive mechanisms, some take the view of not having enough expertise to judge those. A few (even commercial users) turn to knowledge barter relationships to manage the situation. Others try to find help in online communities. In general, most participants have a *more is better* mindset regarding the deployment of security mechanisms (e.g., wanting E2EE, but not knowing how it would protect them). Only two experts chose an approach to keep their set-up and defensive mechanisms simple (e.g., not deploying a firewall, but making sure ports are closed). Moreover, misconceptions of security benefits can lead to adopting inadequate security practices. For example, it was a frequent notion that data is safe because there are backups, which was occasionally also used to justify a lack of security.

(4) Maintenance, most notably regular updates, are not performed by all participants. Both personal and organizational self-hosters lack adequate maintenance practices, with some not having done updates since installation. This might indicate that some people see security more as a one-time action item, while others view it as a continuous effort.

3.6.2 Impact of Individual Characteristics and Social-Embeddedness (RQ 1,2)

We found that expertise alone is not enough to predict the server type that people opt for. One might expect that people who have less IT knowledge would prefer managed servers that give them less control, but have dedicated people working on security, thus balancing privacy needs and work effort. However, this is not the case. People who are strongly normative or challenge-driven might go for the on-premise setup even if it potentially causes great struggles for them due to lacking expertise. Similarly, cost constraints may overrule privacy needs and steer people toward hosting providers. Our data suggests, that the server type choices can negatively impact security outcomes, e.g. when cheap server types block software updates. This highlights how *motivation*, *operation*, and *security practices* are connected and we conclude, that we need to take people's motivation and use cases into account when making server-type suggestions.

A major roadblock for people is a lack of expertise and resources, e.g., time, across all user groups. This is in line with research on the complexity and challenges of administrating home networks [58, 109, 146, 347, 50], especially when having to configure network devices [50]. Depending on their social embeddedness, self-hosters choose different ways to overcome their inexperience. *Individual operators* might tap into online communities, while *organizationally embedded operators* sometimes enter *knowledge barter* relationships. No participant voiced any concerns or reported negative experiences regarding these two forms of support. Especially the Nextcloud community was universally described as friendly and helpful, unlike other online communities known for toxic interactions between users [45, 71, 112]. Our data hints at a gender skew towards men in the self-hosting population, although future work needs to validate this.

Another way to overcome lacking expertise and resources, is to go for a server type that requires less maintenance (e.g., with a full-managed hosting provider), or

rely on ready-made software solutions to ease the burden of maintenance (e.g., pre-configured or managed docker containers). However, balancing knowledge requirements and automation requires tradeoffs: (1) The goal that people have in mind might not be compliant with outsourcing hardware/software maintenance, (2) Relying on solutions that make adoption easier, might later complicate maintenance. Future work can explore how concrete set-up choices influence roadblocks people have to cope with when adopting self-hosted services, and possible impacts on security practices, such as update experiences.

3.6.3 Areas of Tension (RQ 5)

Self-hosting can involve a wide range of motivations, use cases, set-ups, admin capabilities, and social embedding. Therefore, no one-size-fits-all solution exists and personalized information sources are necessary. Participants especially need help with securing their operations. Here, even people who have concrete attacker models, talk about problems identifying attackers' capabilities and realistic threat models. Additionally, people often struggle to identify adequate defensive mechanisms. Information sources tailored to their specific use-case and set-up could assist self-hosters in identifying potential attackers, corresponding risks, and defenses. While security scanning tools are valued, participants may require assistance in selecting a reliable one. Automation is often seen as a promising solution to improve security, but it cannot be the sole solution, since self-hosters are responsible for both set-up and maintenance. Yet, tools like set-up wizards might mitigate the major roadblock that is admin capabilities. However, people actively avoid solutions like ready-made docker containers to reduce complexity, both for ease of maintenance and to lower security risks. Attempting to keep things simple and transparent can also pose risks, as exemplified by the participant who writes all security tools themselves.

3.7 Conclusion

This study explores and connects three dimensions of self-hosting: *motivation*, *operation* in the form of self-hosters' social embeddedness, and *security* mindset. A need for privacy, autonomy, and security together with the belief that self-hosting is the right thing to do are prominent motivational factors. Yet, the decision to self-host is frequently a pragmatic one influenced by cost considerations and the availability of high-quality self-hostable solutions. Motivational factors don't exist in a vacuum but are enabled or constrained by the resources participants can rely on. For instance, participants report varying levels of technical proficiency and work in different *operator constellations*, some of which are deliberately entered to cope with lacking expertise, especially when self-hosting is a hard requirement for them due to their professional-ethical values. Strongly normative-driven self-hosters might opt for server types that allow for a maximum level of hardware and software control, although their lacking expertise turns hosting into a cumbersome task. Others find themselves in the conflict between their need for privacy and cost constraints, causing them to rent third-party servers from large tech companies. Security is often approached in an unstructured fashion. Only a few — even commercial

and organizational users with a dedicated admin team — invest in a threat and attacker analysis. Without such an analysis, security features are chosen more spontaneously than reflected.

4

A Representative Survey on the Prevalence of Private Hosting and Administrator Characteristics

The contents of this chapter were published as part of the publication *"Towards Privacy and Security in Private Clouds: A Representative Survey on the Prevalence of Private Hosting and Administrator Characteristics"* (USENIX Security 24) [P2]. This chapter uses the academic "we" to highlight the contributions of my co-authors and me. The paper was the result of a collaboration with the psychology department and was based on Elena Groben's master's thesis, which I co-advised. The following table details the contributions of each author to this paper:

Author	**Contribution**
Lea Gröber	I developed the research idea and methods with Michael, Rebecca, and Elena. I supervised Elena throughout the project and evaluated the focus groups with her. Especially, I contributed the technical knowledge of self-hosting for the construction of the survey instrument. I decided on the analysis methods in collaboration with Rebecca, Michael, and Simon. Finally, I wrote the majority of the paper.
Simon Lenau	Simon reviewed and advised on the survey instruments and paper framing. He ran the statistical analyses we agreed on. He contributed tables and a section on lessons learned to the paper.
Rebecca Weil	Rebecca advised throughout the project. She contributed to the study idea, methods, framing, and oversaw data collection. She reviewed and wrote parts of the paper in the methods and results sections.
Elena Groben	Elena worked on parts of the project as her master's thesis. She contributed to the study idea, development of methods, and data collection.
Michael Schilling	Michael had the initial idea for the project. He provided critical feedback through all stages of the study and contributed to the framing, methods, and analysis plan. He reviewed the final paper before submission.
Katharina Krombholz	Katharina gave feedback on the initial idea and methods. She reviewed the survey instrument and the final paper before submission.

Reference

Gröber, Lea and Lenau, Simon and Weil, Rebecca and Groben, Elena and Schilling, Michael and Krombholz, Katharina. (2024). Towards Privacy and Security in Private Clouds: A Representative Survey on the Prevalence of Private Hosting and Administrator Characteristics. 33rd USENIX Security Symposium, 6057–6074.

Instead of relying on Software-as-a-Service solutions, some people self-host services from within their homes. In doing so they enhance their privacy but also assume responsibility for the security of their operations. However, little is currently known about how widespread private self-hosting is, which use cases are prominent, and what characteristics set self-hosters apart from the general population. In this work, we present two large-scale surveys: (1) we estimate the prevalence of private self-hosting in the U.S. across five use cases (communication, file storage, synchronized password managing, websites, and smart home) based on a representative survey on Prolific ($\mathbf{n = 1505}$), and (2) we run a follow-up survey on Prolific ($\mathbf{n = 589}$) to contrast individual characteristics of identified self-hosters to people of the same demographics who do not show the behavior.

We estimate that 8.4% of the US population are private self-hosters. Websites are the most common self-hosted use-case, predominately running on home servers. All other use cases were equally frequent. Although past research identified privacy as a leading motivation for private self-hosting, we find that self-hosters are not more privacy-sensitive than the general population. Instead, we find that IT administration skills, IT background, affinity for technology interaction, and maker self-identity positively correlate with self-hosting behavior.

4.1 Introduction

In the past decade, both end users and companies have migrated to public clouds[248, 116, 335, 154, 376]. Due to an abundance of Software-as-a-Service (SaaS) offerings, clouds are not only used for file storage but for a broad set of use cases effectively putting users' data in the hands of third parties. As an alternative, some people set up and maintain their own services on hardware they control. This behavior has distinct security and privacy implications for these so-called "self-hosters": (1) There is a privacy benefit in taking control of data by hosting services on controlled premises [170, 336, 337]. Nonetheless, this benefit is negated if the self-hosted service lacks proper configuration, maintenance, and backup procedures, thereby rendering it vulnerable to attacks or data loss. Accordingly, (2) self-hosters need to take responsibility for securing their operations, a task that requires technical knowledge that not all people who self-host have [P1]. Some see a security advantage in self-hosting because they believe small instances are unattractive targets [P1], while cloud providers are more prone to be attacked due to the centralization of user data [252]. Yet, big providers usually have the means to invest in experts to secure their operations. Having the means still does not imply that commercial clouds are secure [275, 169, 385] or that individuals without these means fail to secure their self-hosted services [272].

Currently, assessing the security and privacy implications of self-hosting is challenging given the diverse interplay of configurations, use cases, and operator capabilities that have not yet been thoroughly studied. Additionally, there is a lack of comprehensive data on the scale of self-hosting, and how many people are impacted by its security and privacy implications. Moreover, to design context-specific solutions that make securing private data easy, we first need to understand the distinct disposition of the self-hoster population.

Accordingly, we propose the following research questions as first steps towards gauging the security impact of self-hosting, and laying the foundations for designing effective solutions:

RQ1: *How widespread is self-hosting for private use cases? What kind of people are self-hosting?* Estimating the prevalence of self-hosting is a first step towards understanding its security and privacy impact. Characterizing self-hosters' demographics lays the foundations for constructing personas that are representative of the population and valuable for any kind of security-focused design efforts.

RQ2: *Which tools are self-hosted and how?* Understanding technologies and server-type choices enables us to concentrate research efforts on common security-relevant and potentially high-risk use cases.

RQ3: *In which characteristics do self-hosters differ from the average U.S. population?* Investigating what distinguishes self-hosters from the general U.S. population contributes to a better understanding of enabling and constraining individual characteristics relevant to system administration work, and hints at roadblocks in the hosting ecosystem.

We addressed these questions by using a sequential two-survey design. First, we identified self-hosters in a demographically representative U.S. sample with $n = 1505$ participants and analyzed which services they self-hosted. Second, we ran a follow-up survey with $n = 589$ participants and compared the identified self-hosters against a socio-demographically matched control group with regard to 14 characteristics suggesting a relationship with self-hosting. We selected characteristics that we hypothesized would predict self-hosting behavior, based on qualitative insights from focus groups we ran prior to survey construction. This work makes the following core contributions:

1. We estimate the prevalence of self-hosting for private use in the U.S. with 8.4 %, CI [7, 9.6]. Based on this we suggest that self-hosting is not a niche phenomenon, and research efforts are worth investing to understand and support the community in securing their data.

2. We compile demographic data and individual characteristics that describe the self-hoster population. This information may inform future design studies focused on developing security-enhancing solutions.

3. We provide an overview of prominent use cases and server-type choices. Understanding common self-hosting practices helps identify critical use cases and informs future research and development efforts.

4. We highlight individual characteristics that are positively and negatively associated with self-hosting behavior. This yields insights about which factors could be roadblocks to self-hosting and helps concentrate efforts to support future administrators.

5. We offer a reflection on our study design and share takeaways from preparing and running large-scale representative studies on technical topics.

Replication Package We provide a full replication package including survey instruments, anonymized data, and evaluation artifacts.[1]

4.2 Related Work and Self-Hosting Definition

Self-Hosting For the scope of this paper, we define self-hosting as (1) user control over the hardware, i.e., running on the user's own hardware or renting said hardware, (2) control over the software, i.e., the operating system, the configuration, (3) the self-hosted service needs to be available over a network, and (4) responsibility for the service, e.g. not relying on a third-party for set-up and maintenance.

The research community looked at self-hosting from the angle of people administrating their own home networks [58, 109, 146, 347, 50]. This line of work focuses on the growing complexity of home networks due to the increasing number of IoT devices. Problems arise when people have to deal with hardware and software failures and when expectations of usefulness are defied (e.g., mismatches between what a person expects to be able to do and specific device capabilities) [50]. Moreover, recent work investigated private and organizational self-hosting in the case of a popular file-sharing and content collaboration platform. In the previous chapter [P1], we found that privacy, autonomy, and security are prominent motivational factors when people decide to self-host services for private use. These self-hosters are diverse in terms of technical background and face challenges when it comes to maintenance and security choices.

Security Practices of Administrators The tasks of system administrators are manifold and research looked into different facets of their work. Various studies examined root causes and usability challenges for transport layer security misconfigurations [115, 200, 201]. Similarly, Kraemer et al. [197] shed light on the influence of human error on network administration and information security. Li et al. [216] uncover challenges administrators face when updating servers. Dietrich et al. [97] investigated system operators' perspectives on security misconfigurations. They find that mitigations often exist, but are not put to effective use, thus highlighting the importance of a human-centric research approach. Moreover, recent research started investigating the social-embeddedness of system administration work. Kaur et al. [186] uncovered structural challenges marginalized genders face when working in IT administration. Gröber et al. [P1] categorized five different constellations in which operators work together to maintain IT infrastructure across organizational and private use. They found that individual characteristics such as operators' level of expertise and use case requirements influence how system administration work is carried out.

Representative Studies in Security Representative user studies in the area of security and privacy enable researchers to draw generalizable results but are costly to conduct. Redmiles et al. conducted a series of representative studies, such as a U.S.-census-representative survey via Survey Sampling International panel to investigate how demographics, knowledge, and beliefs correlate with security advice and behavior of end

[1] Find survey instruments in Appendix B.1; other artifacts will be made available after publication.

users [296]. Further, Redmiles et al. ran a census-representative telephone survey to understand how different socioeconomic status correlates with security incidents [297]. Finally, Redmiles et al. investigated if data gathered on Amazon Mechanical Turk (MTurk) on topics on security and privacy generalize to a broader population [298]. To this end, the authors compared an MTurk sample with a census-representative web panel and a probabilistic telephone sample. They found that MTurk responses regarding security and privacy tasks are actually more representative of the U.S. population than the web panel. This study has been replicated by Tang et al. on MTurk and Prolific and was compared against a probabilistic survey obtained through a service provider [341]. All studies above focus on the U.S. population. Recent research broadened the scope to German citizens [307], and large-scale international studies comparing representative samples from 12 countries on four continents [163]. Both studies obtained their sample through online panelists carried out by a service provider.

4.3 Methods Overview

This work comprises two quantitative surveys that we conducted consecutively. Thereby the self-hoster classification of Survey 1 informed the sampling of Survey 2. This way we were able to recruit a demographically matched control group for the self-hosters and study which characteristics distinguish them from non-self-hosters. Our team of researchers constructed both surveys simultaneously in an iterative process in which we refined survey items (e.g., wording) and measurement details (e.g., scale anchors) to minimize misunderstandings or complications for participants. We outline the final wordings and survey flows in appendices B.1 and B.2. Both surveys also underwent technical pre-testing to ensure data collection quality and a smooth experience for participants. We preregistered the study design including hypotheses.[2] Hypotheses for Survey 2 were grounded in qualitative data of two focus groups we ran prior to survey construction. Thus, we explain the focus groups as a methodological aspect of Survey 2 in Section 4.7.1.1. In the following, we present both studies in sequential order including details on methodology, results, and discussion.

4.4 Prevalence Survey 1 - Methods

To allow for valid estimation of the prevalence of self-hosting behavior in the USA, we aimed to recruit a representative sample of the U.S. population and administer a survey based on which responses we could identify self-hosters. Representativeness in the context of our study refers to meeting the census distribution of age, sex, and ethnicity and, due to sampling via Prolific, is restricted to individuals who have internet access via a computer or mobile device. We assume the influence of this restriction to be minimal because access to such devices exceeded rates of 94% for heads of households up to 64 years old in 2018. For heads of households aged 65 and older, the rate was 80% [235]. With a continuous positive trend in recent years, the influence today can be

[2] https://osf.io/4apwe/?view_only=b08a9b2d7b6d4f288b57f8382b26e41f

assumed to be even less [311]. Our measurements, sampling method, and identification strategy are detailed in the following.

4.4.1 Measurements

We introduced our survey as a part of research about software and application use in private and professional contexts to avoid biasing participants toward the subject of self-hosting. As a basis for an operational identification strategy of self-hosting behavior, we presented participants with 5 use case categories (i.e., file storage - synchronization, file transfer; Web sites - CMS, blogging; communication - messaging, voice/video telephony; synchronized password managing; smart home) and lists of 5 self-hostable and 6 non-self-hostable tools within each category (see Table 4.3), presented in random order. We included tools based on their popularity on GitHub for self-hostable tools [54, 151] and on Google Trends (https://trends.google.com) for non-self-hostable tools. Participants also had the opportunity to add other tools not mentioned in the pre-selected lists. We asked them to indicate for each tool whether they use it in a private or work context.

For each self-hostable tool, and for each tool entered via the "other" option that was being used in a private context, we asked participants whether the tool was set up for them or whether they had set it up themselves on a server. All participants indicating that they had set up the tool themselves were subsequently asked on which type of server they had set it up (i.e., home server, virtual private server, dedicated server, or other). This was identical for all use case categories, except the smart home category. Here, we asked participants if they had enabled remote access from outside the local network (i.e., accessible via the internet, accessible via a virtual private network, only on the local network).

To reduce the number of false positives (i.e., participants incorrectly identified as self-hosters based on their response in the tool selection) we introduced a definition of private self-hosting in the second part of Survey 1. We distinguished it from commercial cloud services and asked participants if they had come into contact with self-hosting before, providing us with a short description of how they had come into contact via a free text field. We used this information later on as a sanity check for the identification of self-hosters based on their tool selection. To further ensure data quality, we used two attention checks in the survey to identify inattentive participants [274].

In the last part of Survey 1, we collected demographic data, final comments on the survey, and announced the follow-up survey (Survey 2) that we might contact some of the participants again about. The median response time of the survey was 499 seconds, and participants were compensated with 1.18 USD, corresponding to an hourly wage of 8.51 USD.

4.4.2 Sampling

Based on the data available to us at the time of sampling in the U.S. Decennial Census of Population and Housing [354] we defined target sub-samples, representing the U.S. population in terms of sex, age, and ethnicity. Because census data only contain information about the distribution of sexes (i.e., male, female) but not gender, we sampled based on the information available to us. However, because we assessed

self-reported gender (i.e., man, including trans man/trans male; woman, including trans woman/trans female; non-binary; self-described) in our surveys, we report this information in the result sections. Moreover, we used the brackets 18-28; 28-38; 38-48; 48-58, and > 58 for the target age and Asian, Black, Mixed, Other, and White for the target ethnicity.

For data collection we used Prolific (www.prolific.co), defining 50 target sub-samples for U.S. residents as a result of the cross combination with sex, age, and ethnicity, utilizing the Prolific pre-screening functionality. Replicating Prolific's in-house representative sampling solution determined the simplification of ethnicity groups in the U.S. census [293, 341]. Prolific allows a relatively cost-effective and fast way to collect data that has been shown to be representative of the U.S. population, at least with respect to some assessed items [341]. As such it might be considered a more feasible approach for non-commercial research purposes as compared to other face-to-face omnibus representative sampling procedures (e.g., ipsos.com). Nevertheless, we have to assume a general bias in our sample, potentially relevant to our research questions, in the sense that it only contains people who have internet access via a computer or mobile device to answer Prolific surveys. Accordingly, we treat our estimation of the prevalence of self-hosting behavior in the U.S. population as an upper bound. The total targeted sample size was $n = 1500$, based on the availability in the Prolific participants pool. Data collection took place between August and September 2022.

4.4.3 Data Cleaning and Preparation

We only included participants in the sample who had completed the entire survey and passed both attention checks. Due to our aim to achieve a representative sample, we excluded participants who did not meet these criteria and we re-sampled in line with the requirements of our target sub-samples. Due to technical issues, some participants were observed twice. For these participants, we kept only the first complete data set. Once data collection was complete, three researchers coded participants' final survey comments (codes: "issue", "no issue") to see whether any comments raised doubts about data quality, e.g. the participant mentioned having accidentally indicated to self-host a tool. In case of disagreement, they discussed their ratings and came to a consensus. Four participants were excluded from the sample.

4.4.4 Survey Weighting

The representative sampling procedure closely aligned the distribution of sex, age, and ethnicity in our sample to the distribution in the population (see Table B.1). Still, slight deviations occurred due to integer sample size limitations and data cleaning (see section 4.4.3). To resolve these small discrepancies and accurately estimate self-hosting prevalence, we applied the Generalized Regression (GREG) Estimator [94] to obtain calibration weights for our survey, as is best practice in other research areas and official statistics [70, 213, 283, 304]. The weights were determined such that the estimated weighted proportions exactly meet the census proportions in Table B.1. The weights are $w_h = N_h \big/ n_h^{(1)}$ for each element in the h-th socio-demographic group defined by

sex, age, and ethnicity. N_h and $n_h^{(1)}$ denote the group sizes in the census and our first survey, respectively. This ensures that our results accurately reflect the diversity of socio-demographic groups in the population, mirroring their exact proportions in the U.S. population. For example, this counterbalances variations among different groups in participant exclusions based on attention checks.

4.4.5 Operational Classification of Self-Hosters

We identified participants as self-hosters when they selected or added at least one self-hostable tool in one of the use case categories as privately used, indicated that they had set it up themselves on a server, and confirmed that they had come into contact with self-hosting before taking part in our survey. Alternatively, they were identified as self-hosters when they described in the open response format (i.e., Are there any other tools that you self-host and that have not been mentioned above?) that they use self-hosted tools or network services, set up and maintain services themselves, and are in control of the infrastructure. In addition, they had to confirm that they had come into contact with self-hosting before taking part in our survey. As the latter strategy relied on open responses, four raters coded the open responses independently checking for the criteria described above. In case of disagreement, they discussed their ratings and came to a consensus. A final sanity check was applied to all self-hoster classifications. Open responses of all identified self-hosters were coded by four coders for any indications raising doubts about the classification (e.g., a non-self-hostable tool was mentioned by the participant as being self-hosted). In case of disagreement, they discussed their ratings and came to a consensus. The classification process produced three outcomes. Participants were either classified as self-hosters, as non-self hosters, or as having an unclear self-hosting status (i.e., due to conflicting information about their self-hosting behavior).

4.5 Prevalence Survey 1 - Results

We achieved a representative sample ($n = 1505$) of the United States, with respect to age, sex, and ethnicity, in line with the U.S. Decennial Census of Population and Housing Census [354] and the corresponding demographic data available on Prolific. A comparison of the respective shares of age, sex, and ethnicity for our survey and the population can be found in Table B.1 in the Appendix.

4.5.1 RQ 1: Prevalence of Private Self-Hosting

One of the main goals of Survey 1 was the determination of self-hosting prevalence in a representative US sample. In total, we identified $n = 124$ self-hosters in our sample, indicating a prevalence of 8.4 %, CI [7, 9.6] of self-hosting behavior in the US. We also identified $n = 1355$ non-self-hosters and $n = 26$ received an unclear self-hoster status. Table 4.1 shows the self-hosting prevalence (i.e., estimated occurrence of self-hosters in the population) broken down by age, ethnicity, sex and gender in column 2. For example, we estimated a self-hosting prevalence of 11.6 % in the age group 48-58. We

Table 4.1: Estimated self-hosting prevalence and group comparison by age, ethnicity, sex and gender (in %)

		Distribution in subgroups	
	Prevalence	**Self-Hosters**	**Non-Self-Hosters**
Age			
18 – 28	8.3 ± 3.2	18.4 ± 6.3	18.6 ± 0.6
28 – 38	10.9 ± 3.7	22.3 ± 6.6	16.5 ± 0.6
38 – 48	7.9 ± 3.2	17.1 ± 6.3	18.2 ± 0.6
48 – 58	11.6 ± 3.6	**26.1 ± 7.1**	**18.1 ± 0.7**
58	4.9 ± 2.1	**16.2 ± 6.2**	**28.6 ± 0.6**
Ethnicity			
asian	8.0 ± 6.0	4.8 ± 3.5	5.0 ± 0.3
black	10.0 ± 4.1	14.3 ± 5.5	11.7 ± 0.5
mixed	5.9 ± 7.1	1.5 ± 1.8	2.2 ± 0.2
other	7.4 ± 5.3	5.7 ± 3.9	6.5 ± 0.4
white	8.3 ± 1.6	73.8 ± 7.1	74.7 ± 0.6
Sex			
female	2.9 ± 1.2	**17.8 ± 6.6**	**55.1 ± 0.8**
male	14.3 ± 2.5	**82.2 ± 6.6**	**44.9 ± 0.8**
Gender			
woman	3.1 ± 1.2	**18.6 ± 6.7**	**53.6 ± 1.0**
man	14.1 ± 2.6	**79.9 ± 6.9**	**44.2 ± 1.0**
non-binary	0.0	**0.0**	**1.3 ± 0.6**
self-described	11.0 ±20.3	0.8 ± 1.6	0.6 ± 0.4
not stated	18.5 ±33.1	0.7 ± 1.5	0.3 ± 0.3

Boldened shares indicate significant differences between estimated self-hoster and non-self-hoster shares

± indicates the lower and upper bounds of the 95% confidence intervals

compared the share of age, ethnicity, sex and gender characteristics between the group of self-hosters (column 3) and non-self-hosters (column 4) to determine any significant differences between the two groups with regard to these characteristics (e.g., are self-hosters more likely than non-self-hosters to fall in age group 48-58?). Compared to non-self-hosters, the age group of 48 to 58 year old people is significantly more frequent, while people older than 58 years are less frequent in the self-hosters sample. Moreover, men are more while women and non-binary people are less frequent among identified self-hosters. All other assessed demographic characteristics do not differ significantly between self-hosters and non-self-hosters.

4.5.2 RQ 2: Tool Usage Patterns

To better understand self-hosting behavior, we looked at the distribution of the different use cases and tools selected from the range of presented tools by participants identified

as self-hosters. Tools additionally listed by these participants under a respective use case were, due to relatively small numbers, summarized as 'Other'. Participants who we identified as self-hosters solely based on their responses in other open response formats (i.e., Are there any other tools that you self-host and that have not been mentioned above?) are not included in the table due to inconsistent information about use cases, tool names, and server types. Table 4.2 shows that from our predefined use cases, websites are most frequent among self-hosters. Communication, file storage, synchronized password managing, and smart home are less frequent use cases and do not differ significantly from each other in their frequency. For websites, the most frequently used tool is WordPress, which in the majority of cases is hosted on a home server. For the smart home use case, Home Assistant is most frequently used and in the majority of cases not accessible from the internet. All tools that were indicated as being privately used by self-hosters from our pre-selection of tools, can be found in Table 4.2.

In addition, we not only looked at the tools that are self-hosted by participants but also at the usage of non-self-hosted tools (i.e., self-hostable and non-self-hostable) for the same use cases in self-hosters and non-self-hosters. Interestingly, Table 4.3 shows that, across all use cases, self-hosters seem to use more tools in general. That is, they not only use more self-hostbale tools but also more of the non-self-hostable tools (e.g., Microsoft Teams, Google Drive and Home, LastPass) as compared to non-self-hosters.

4.6 Prevalence Survey 1 - Discussion

The goal of Survey 1 was to determine the prevalence of self-hosting in the U.S. to get a better understanding of how widespread the phenomenon might be. To this end, we used a representative sampling method (additionally corrected by weighting), which in turn allows us to gauge the upper bound of the occurrence of self-hosting in the U.S. population. Based on the results of Survey 1 we estimate the occurrence of self-hosting with 8.4 %, CI [7, 9.6]. As such, self-hosting should not be considered a niche phenomenon.

The results of Survey 1 also indicated that self-hosters are more likely male and belong to the age group of 48-58 year-old people, and less likely to the age group of people older than 58. Speculatively, a possible connection between the age of the participants and self-hosting could be the time of the emergence of relevant technologies and the life phases in which the people were at that time. People in the 48-58 age group were born between 1964 and 1974. This means that they were in the late adolescent phase of their lives at the time of the advent of the Internet and might have been open to innovations. At that time, however, the Internet was more technically demanding. One needed more technical knowledge and cloud computing in the current sense did not exist back then. If one wanted a certain service or functionality, “self-hosting" was the default. This is one attempt to reason about this finding but the present research approach does not allow us to verify such claims. Still, our finding serves as a basis for future research to explore causal demographic variables to explain self-hosting behavior.

Our results also revealed that a large proportion of self-hosting behavior is hosting WordPress websites on a home server. This is a potential high-risk use case, as people are running internet-connected services from their homes. Further research might look

Table 4.2: Usage and hosting type shares of self-hostable tools (in % of all self-hosters)

Tool	Self-Hosted by		Server type							
			Home		VPS		Dedicated		Other / unknown	
Communication	20.2	± 9.1								
Teamspeak	12.0	± 7.3	64.5	±22.7	11.6	±15.1	23.9	±20.4	0.0	
Mumble	8.0	± 6.2	63.3	±28.6	18.1	±22.7	18.5	±23.2	0.0	
Rocket.Chat	2.7	± 3.6	32.0	±52.2	68.0	±52.2	0.0		0.0	
Jitsi Meet	2.6	± 3.6	65.8	±54.0	0.0		34.2	±54.0	0.0	
Mattermost	1.3	± 2.5	32.1	±52.2	67.9	±52.2	0.0		0.0	
Other	1.4	± 2.7	0.0		50.1	±69.3	0.0		49.9	±69.3
File Storage	17.7	± 8.8								
Nextcloud	4.1	± 4.6	74.9	±43.5	0.0		25.1	±43.5	0.0	
ownCloud	2.8	± 3.8	49.0	±69.3	51.0	±69.3	0.0		0.0	
SparkleShare	2.7	± 3.7	100.0		0.0		0.0		0.0	
Synthing	2.6	± 3.6	65.7	±54.1	34.3	±54.1	0.0		0.0	
Seafile	1.4	± 2.7	100.0		0.0		0.0		0.0	
Other	6.7	± 5.7	59.5	±43.0	20.2	±35.4	0.0		20.2	±35.4
Synchr. PW Managing	16.3	± 8.4								
Vault-/bitwarden	6.8	± 5.7	73.3	±22.4	13.7	±17.6	13.0	±16.9	0.0	
sysPass	2.7	± 3.7	33.0	±53.1	67.0	±53.1	0.0		0.0	
Passbolt	1.3	± 2.6	75.2	±42.2	24.8	±42.2	0.0		0.0	
Teampass	1.3	± 2.6	50.3	±69.3	49.7	±69.3	0.0		0.0	
Other	6.8	± 5.7	89.0	±20.4	0.0		0.0		11.0	±20.4
Websites	51.4	±11.4								
WordPress	46.0	±11.4	47.9	±11.9	26.2	±10.4	21.5	± 9.7	4.4	± 4.8
Ghost	2.7	± 3.7	49.9	±49.0	50.1	±49.0	0.0		0.0	
Cockpit	1.3	± 2.6	51.4	±69.2	48.6	±69.2	0.0		0.0	
Other	4.0	± 4.5	33.2	±37.7	33.4	±37.7	0.0		33.4	±37.7
			Accessible from Internet							
			yes		**via VPN**		**no**		**Other / unknown**	
Smart Home	22.8	± 9.6								
Home Assistant	21.5	± 9.4	25.3	±13.4	0.0		59.4	±15.2		
Node RED	2.6	± 3.6	25.8	±43.5	0.0		24.5	±41.8		
WebThings Gateway	1.4	± 2.6	0.0		100.0		0.0			

± indicates the lower and upper bounds of the 95% confidence intervals

into this use case to investigate security configurations and assist people in making secure decisions. Notably, our analyses also showed that being a self-hoster does not necessarily mean solely turning to self-hostable tools and avoiding commercial alternatives. Quite the contrary, across all use cases, self-hosters seem to use more self-hostable and commercial tools in general as compared to non-self-hosters. An avenue for future research is to explore usage dependencies between and across tools and use cases when people self-host.

Table 4.3: User share per pre-defined tool (in %)

	Tool	Usage		
		Self-Hosters	Non Self-Hosters	
Communication	Zoom	77.5 ± 7.3	75.6 ± 2.3	
	Discord	**67.8 ± 8.2**	**35.6 ± 2.3**	
	Microsoft Teams	**48.5 ± 8.8**	**38.8 ± 2.5**	
	Whatsapp	37.0 ± 8.5	37.8 ± 2.5	
	Telegram	**27.4 ± 7.9**	**15.7 ± 1.9**	
	Signal	**16.9 ± 6.6**	**7.2 ± 1.4**	
	Mumble	**10.4 ± 5.4**	**1.1 ± 0.6**	✘
	Teamspeak	**10.4 ± 5.4**	**2.6 ± 0.8**	✘
	Jitsi Meet	**5.7 ± 4.1**	**1.1 ± 0.6**	✘
	Rocket.Chat	4.9 ± 3.8	1.2 ± 0.6	✘
	Mattermost	3.2 ± 3.1	0.8 ± 0.5	✘
File Storage	Google Drive	**92.0 ± 4.7**	**78.6 ± 2.1**	
	Dropbox	**67.0 ± 8.3**	**53.0 ± 2.6**	
	Microsoft OneDrive	59.6 ± 8.7	51.1 ± 2.6	
	iCloud	46.7 ± 8.8	49.0 ± 2.6	
	MEGA	**24.2 ± 7.5**	**7.6 ± 1.3**	
	Box	12.2 ± 5.8	7.1 ± 1.4	
	Nextcloud	**8.2 ± 4.8**	**1.0 ± 0.5**	✘
	ownCloud	**5.8 ± 4.1**	**1.0 ± 0.5**	✘
	SparkleShare	**4.9 ± 3.8**	**0.9 ± 0.5**	✘
	Seafile	4.9 ± 3.8	1.0 ± 0.5	✘
	Syncthing	4.0 ± 3.5	1.2 ± 0.6	✘
Smart Home	Amazon Alexa	47.7 ± 8.7	38.7 ± 2.6	
	Google Home	**46.0 ± 8.8**	**25.1 ± 2.3**	
	SmartThings	**20.0 ± 6.9**	**6.9 ± 1.3**	
	Home Assistant	**13.7 ± 6.1**	**3.2 ± 0.9**	✘
	Apple HomeKit	7.2 ± 4.6	4.2 ± 1.1	
	Node RED	4.0 ± 3.5	1.0 ± 0.5	✘
	WebThings Gateway	2.5 ± 2.8	0.9 ± 0.5	✘
	Bosch Smart Home	2.4 ± 2.7	1.2 ± 0.6	
	Domoticz	1.7 ± 2.3	0.8 ± 0.5	✘
	Gladys	1.7 ± 2.3	1.0 ± 0.5	✘
	Vivint Home	1.7 ± 2.3	1.8 ± 0.7	
Synchr. PW Managing	iCloud Keychain	24.9 ± 7.5	20.4 ± 2.1	
	LastPass	**21.9 ± 7.3**	**9.1 ± 1.5**	
	1Password	**13.9 ± 6.1**	**3.6 ± 1.0**	
	Vault-/Bitwarden	**12.9 ± 5.9**	**3.5 ± 1.0**	✘
	Roboform	**9.8 ± 5.2**	**2.7 ± 0.9**	
	Keeper	5.7 ± 4.1	2.2 ± 0.8	
	Dashlane	4.9 ± 3.8	2.3 ± 0.8	
	Padloc	**4.8 ± 3.8**	**0.9 ± 0.5**	✘
	Passbolt	4.1 ± 3.5	1.0 ± 0.5	✘
	Teampass	4.1 ± 3.5	1.6 ± 0.7	✘
	sysPass	4.1 ± 3.5	1.1 ± 0.6	✘
Websites	WordPress	**48.4 ± 8.8**	**19.7 ± 2.1**	✘
	Blogger	**16.2 ± 6.5**	**8.1 ± 1.4**	
	Wix	10.5 ± 5.4	7.0 ± 1.4	
	Squarespace	8.9 ± 5.0	7.6 ± 1.4	
	Weebly	7.3 ± 4.6	4.1 ± 1.1	
	Ghost	**4.9 ± 3.8**	**0.8 ± 0.5**	✘
	Webflow	4.8 ± 3.7	1.6 ± 0.7	
	Strapi	2.5 ± 2.7	0.7 ± 0.5	✘
	Cockpit	2.4 ± 2.7	0.6 ± 0.4	✘
	Jimdo	2.4 ± 2.7	0.7 ± 0.4	
	Wagtail	2.4 ± 2.7	0.7 ± 0.4	✘

Boldened shares indicate significant differences between estimated self-hoster and non-self-hoster shares
✘: Tool is self-hostable
± indicates the lower and upper bounds of the 95% confidence intervals

4.7 Characteristics Survey 2 - Methods

To be able to better describe the group of self-hosters and understand individual characteristics that correlate with self-hosting behavior, we invited all self-hosters identified with the help of Survey 1 to take part in Survey 2 and matched them with an invited control group of non-self-hosters identified in Survey 1.

4.7.1 Measurements

This section presents details of the survey instrument, as well as hypothesis construction based on qualitative insights from focus groups.

4.7.1.1 Focus Groups to Identify Predictors

To identify candidate characteristics relevant to self-hosting behavior we conducted two focus groups with self-hosters (three participants) and non-self-hosters (ten participants) respectively. We provide an overview of the procedure, analysis, and results below.

One researcher moderated both sessions which took about 90 mins each. The researcher followed a protocol (Appendix B.4) but allowed and encouraged participants to discuss topics freely. To provide a common ground for everyone, the sessions started with an introduction to self-hosting which is especially vital for the non-self-hoster group. As an ice-breaker question, we asked participants about their prior experiences with self-hosting. Only the self-hoster group was then asked about their personal definition of self-hosting, contrasting it with the definition we offered. Afterward, the researcher opened the main discussion which was structured into six key questions: We explored reasons that would discourage individuals from using cloud services (Q1), and why they might be inclined to engage in self-hosting (Q2). Then, participants reflected on situations that might have influenced their decision to self-host (Q3). Moreover, we discussed other possible aspects and domains of life that could be relevant to self-hosting such as personality traits and individual characteristics (Q4). Last participants reflected on possible reasons not to self-host even though one would want to (Q5), and why some individuals would reject to self-host altogether (Q6). Afterward, we asked only the self-hoster group to identify technical skills they consider essential for self-hosting.

Thematic Analysis Two researchers (computer scientist and psychologist) first grouped the audio data by questions, then listened repeatedly while applying open coding [66, 332, 224] independent from each other. The researchers discussed their coding, resolving all mismatches by revisiting the audio data and updating the codes. During this iterative process, it became evident that coding across questions yielded a better fit with the data compared to strictly adhering to the question structures. For instance, the theme "privacy" was observed both as a lack of privacy, which acted as a deterrent from using cloud services and as a desire to attain privacy, which served as a motivation for self-hosting. Consequently, the coders developed the codebook to capture such overarching concepts. Once the coders reached an agreement, they organized the resulting codes in a mindmap, grouping them into topics and illustrating connections between them. Finally, they used the mindmap as the basis for a discussion to identify themes. In doing so, the

coders listened to audio data again, this time identifying and transcribing relevant quotes. They identified ten core themes, for which they found supporting data in both focus groups: *work effort* (the amount of work, time and resources it requires to self-host), *security* (security advantages and disadvantages of self-hosting), *technology interest and skills* (aptitude for and ability to acquire know-how to self-host), *tinkering and DIY* (aptitude for self-taught learning and doing things yourself), *interpersonal aspects* (different personal factors influencing the motivation to self-host), *money* (financial aspects involved in self-hosting, required spending's and saving money), *privacy* (privacy concerns and needs that can be addressed by self-hosting), *usefulness of self-hosting* (fulfilling unique needs that are not fulfilled by other services), *control* (self-hosting as a means to gain control over one's own life), *openness to new things* (trying out things and setting trends). Many of these core themes concur with recent research findings, identifying privacy, security, and autonomy needs, saving costs, usefulness, and enjoying learning something new as motivational factors for self-hosting [P1]. The authors also showed that a specific skill set and expertise is needed (or needs to be brought in) for self-hosting.

4.7.1.2 Scale Measurements and Hypotheses

We mapped scale measurements of individual characteristics to all core themes that could be captured by such measures and for which we found validated and reliable scales in the literature (i.e., *security, privacy, technology interest and skills, openness to new things, tinkering and DIY, money, work effort, control*). This allowed us to empirically test if the characteristics that emerged from the qualitative analysis of the focus groups can predict self-hosting behavior.

Security To assess participants' security concerns with respect to the protection of their personal information, we used the 4-item security concern scale (e.g., "I worry about wrong information being linked to my identity due to security breaches") [9]. Responses were given on a 5-point answering scale ranging from strongly disagree (1) to strongly agree (5) (see Appendix B.2.12 for details). Because results from the focus groups revealed two possible directions for the relationship between security concerns and self-hosting (i.e., providing more security and increasing security risks), we predicted a non-directional relationship between security concerns and self-hosting behavior.

Privacy Participants' concerns regarding the availability of private information on the Internet were assessed with the 4-item privacy concern scale (e.g., "I am concerned that a person can find private information about me on the Internet") [99]. Responses were given on a 5-point answering scale ranging from strongly disagree (1) to strongly agree (5). Details can be found in B.2.11. Based on the results of the focus groups we predicted a positive relation between privacy concerns and self-hosting behavior. This would indicate that self-hosting is accompanied by a higher concern for information privacy.

Table 4.4: Consistency between operational and self-classification as self-hoster

	Self-classification		
	Non-Self-Hoster	Self-Hoster	Overall
Operational classification			
Non-Self-Hoster	355 (60.3 %)	122 (20.7 %)	477 (81.0 %)
Self-Hoster	32 (5.4 %)	80 (13.6 %)	112 (19.0 %)
Overall	387 (65.7 %)	202 (34.3 %)	589 (100.0 %)

Technology interest and skills: Affinity for technology interaction (ATI) To measure participants' aptitude for interacting with technical systems we used the 9-item ATI scale (e.g., "I try to understand how a technical system exactly works") [121]. Responses were given on a 6-point answering scale ranging from completely disagree (1) to completely agree (6) (see Appendix B.2.2 for details). Based on the results of the focus groups we predicted a positive relation between ATI and self-hosting behavior because self-hosters might show a higher interest in technical systems .

Openness to new things: Personal innovativeness in the domain of information technology (PIIT) We assessed participants' interest in trying out and experimenting with new technologies with the 4-item PIIT scale (e.g., "If I heard about a new information technology, I would look for ways to experiment with it") [13]. Responses were given on a 6-point answering scale ranging from strongly disagree (1) to strongly agree (6) (see Appendix B.2.10 for details). Based on the results of the focus groups we predicted a positive relation between PIIT and self-hosting behavior because self-hosters might be more open to trying out and experimenting with technologies.

Technology interest and skills: Computer self-efficacy To assess participants' confidence in performing computer-related activities we used the advanced (e.g., "Using a computer's task manager"; BITS-Ad) and expert ("Analyzing computer error log files"; BITS-Ex) subscales of the Brief Inventory of Technology Self-efficacy (BITS) [374], asking people to indicate their level of confidence performing each activity. Responses were given on a 6-point answering scale ranging from not at all confident (1) to completely confident (6) (see Appendix B.2.4 for details). Based on the results of the focus groups we predicted a positive relation between BITS-Ad and BITS-Ex and self-hosting behavior because self-hosters might show more advanced and expert technology skills.

Technology interest and skills: Self-hosting skills Self-reported skills to set up and administrate a server were assessed with a self-developed 6-item scale. We first presented participants with a job description of a system administrator (compare Appendix B.2.13). Subsequently, we asked them to rate their abilities in different domains (computer networks, operating systems, servers [virtual or physical], software, system security, and system administration). Responses were given on a 6-point answering scale ranging from poor (1) to excellent (6). Based on the results of the focus groups we predicted

a positive relation between self-hosting skills and self-hosting behavior. This should be the case if self-hosting requires a certain basic skill set to be able to perform the behavior.

Technology interest and skills: IT background To assess IT background we used an item (i.e., “Are you studying or have you been working in any of the following areas: information, technology, computer science, electronic data processing, electrical engineering, communications technology, or similar?”) introduced by Elbitar and colleagues [111] (see Appendix B.2.9 for details). Responses were given on a dichotomous answering scale (yes/no). Based on the results of the focus groups we predicted a positive relation between IT background and self-hosting behavior. An IT background might provide people with the necessary skill set to be able to perform the behavior.

Tinkering and DIY: Maker activities To measure how much time participants typically spend with domestic activities (e.g., baking), DIY activities (e.g., woodworking) and arts and crafts (e.g., ceramics), we adapted 18 different activities from Collier and Wayment [80] and asked participants to indicate their time spent with each activity on a scale from “none” (1) to “I spend large amount of time doing activity” (5) (see Appendix B.2.6 for details). Based on the results of the focus groups we predicted a positive relation between Maker activities and self-hosting behavior. This should be the case if self-hosting goes along with other tinkering and DIY activities.

Tinkering and DIY: Maker self-identity Participants’ self-identity as a maker or DIY person was assessed by presenting them with a short description of what is meant by do-it-yourself (e.g., "sometimes this can be called crafting, sometimes it refers to hobbies. Typically it leads to making something tangible. That is, what you can create with your own two hands.") and asking them to indicate on a scale from not at all (1) to extremely so (5) how much they identify as a maker or DIY person (see Appendix B.2.5 for details). This procedure was adapted from Collier and Wayment [80]. Based on the results of the focus groups we predicted a positive relation between Maker self-identity and self-hosting behavior. This should be the case if self-hosters perceive themselves as DIY persons.

Money: Frugality Participants’ economical consumer lifestyle was assessed with the 9-item frugality scale (e.g., “I believe in being careful how I spend my money”) [210]. Responses were given on a 6-point answering scale ranging from definitely disagree (1) to definitely agree (6) (see Appendix B.2.7 for details). Based on the results of the focus groups we predicted a positive relation between frugality and self-hosting behavior because self-hosters might ponder more about what to spend their money on.

Work effort: Grit The extent of participants’ consistency of interest (GRIT-Co) and perseverance of effort (GRIT-Pe) was assessed with the 8-item GRIT-S scale (e.g., “I often set a goal but later choose to pursue a different one”; “I am diligent.”) [104]. Responses were given on a 5-point answering scale ranging from strongly disagree (1) to strongly agree (5) (see Appendix B.2.8 for details). Based on the results of the

focus groups we predicted a positive relation between GRIT-Co and GRIT-Pe and self-hosting behavior. This might be the case if self-hosting behavior goes along with being consistent in an area of interest and putting effort into reaching a goal.

Control: Autonomy To assess participants' valuation of self-direction and freedom of choice in their daily activities and undertakings we adapted eight autonomy items (e.g. "I feel I'm doing what really interests me") from the Basic Psychological Need Satisfaction and Need Frustration Scale (BPNSNF) [69]. Responses were given on a 5-point answering scale ranging from not important to me at all (1) to very important to me (5) (see Appendix B.2.3 for details). Based on the results of the focus groups we predicted a positive relation between Autonomy and self-hosting behavior because self-hosting behavior might go along with the importance that is ascribed to having control over one's life.

Similar to Survey 1, to ensure data quality, we used two attention checks in the survey to identify inattentive participants [274].

4.7.1.3 Self-Report Identification

Because the identification of (non)self-hosters in Survey 1 was based on operational criteria (i.e., participants' response behavior in closed and open questions), we employed a self-report identification procedure in Survey 2 to complement the measurement from the previous survey. To this end, we presented participants with a definition and examples of self-hosting and then asked them whether they would describe themselves as a self-hoster (i.e. whether they currently self-host or have recently self-hosted at least one tool/service in their personal life). We emphasized that their answer would not affect survey length or compensation to minimize externally motivated answering behavior.

4.7.2 Sample Selection Process

To allow for a meaningful comparison between self-hosters and non-self-hosters with reliable group estimates, self-hosters were matched with non-self-hosters in an approximate (influenced by the availability and responsiveness of participants in the pool) ratio of 1:3 [334] with respect to age, ethnicity, and sex, keeping the influence of these demographics as constant as possible in both groups. For data collection, we used again Prolific. Data collection took place between September 2022 and November 2023. 98.68% of the sample was completed in December 2022. However, to increase the chances of including the maximum number of self-hosters, the survey was opened up to November 2023.

4.7.3 Data Cleaning and Preparation

We only included participants in the sample who had completed the entire survey and passed both attention checks. Participants who did not meet these criteria were excluded and we re-sampled in line with the requirements of our sample selection process. Once data collection was completed three researchers coded participants' final survey comments to see whether any comments raised doubts about data quality

Table 4.5: Logistic regression model for self-hoster status: stepwise selection

	Model 00	Model 01	Model 02	Model 03	Model 04	Model 05	Model 06	Model 07	Model 08	Model 09
Intercept	0.1	0.6	0.6	−0.2	−0.3	−0.9	−0.8	−0.6	0.3	0.5
ATI	0.3	0.2	0.5	0.4	0.4	0.7 **	0.7 **	0.7 **	0.7 **	0.7 **
DIY-self	0.5 *	0.5 *	0.5 *	0.5 *	0.5 *	0.6 **	0.6 **	0.6 **	0.5 **	0.5 **
GRIT-pe	−1.5 ***	−1.6 ***	−1.5 ***	−1.6 ***	−1.6 ***	−1.4 ***	−1.7 ***	−1.7 ***	−1.8 ***	−1.8 ***
IT background	0.9 *	0.9 *	0.8 *	0.8 *	0.9 *	1.0 **	1.0 **	1.0 **	0.9 *	0.9 *
Skills	1.1 ***	1.1 ***	1.1 ***	1.1 ***	1.2 ***	1.4 ***	1.4 ***	1.4 ***	1.4 ***	1.4 ***
Privacy	0.2	0.2	0.2	0.1	0.2	0.2	0.2	0.1	0.1	
DIY activities	−0.9	−0.9	−0.8	−0.8	−0.8	−0.9	−1.0	−1.0		
Security	0.0	−0.1	0.0	−0.1	−0.1	−0.1	−0.1			
GRIT-co	−0.4	−0.3	−0.4	−0.4	−0.4	−0.4				
BITS-ad	1.0 *	1.0 *	1.0 *	1.0 *	1.0 **					
BITS-ex	0.1	0.1	0.1	0.1						
Autonomy	−0.3	−0.3	−0.2							
PIIT	0.3	0.3								
Frugality	−0.7									
Deviance	131.4	132.7	133.6	133.9	134.0	138.3	140.0	140.1	141.4	141.4
AIC	167.3	164.6	161.8	159.3	157.0	154.7	152.8	151.0	149.7	148.2
# of observ.	432	432	432	432	432	432	432	432	432	432

$^{***}p < 0.001$ $^{**}p < 0.01$ $^{*}p < 0.05$
Results after controlling for the effects of Gender, Age and Ethnicity

(e.g., participant indicated having trouble with filling out the scale items). In case of disagreement, they discussed their ratings and came to a consensus. One participant was excluded.

4.7.4 Survey Weighting

To be able to make valid claims about the population for our findings in Survey 2, we adjusted for deviations between our sample and the population with respect to the distribution of sex, age, and ethnicity as well as for the over-representation of self-hosters, both resulting from our sampling design (see section 4.7.2). The calibration weights are determined such that the estimated weighted proportions across all socio-demographic groups containing self-hosters are the same in Survey 2 as they were in the representative Survey 1, and that the estimated share of self-hosters within each of these groups also corresponds to the prevalence estimated in Survey 1 (see tables B.1 and B.2). As in section 4.4.4, we used the Generalized Regression estimator [94] for this purpose. The weights for Survey 2 are therefore $w_h = \widehat{N}_h^{(1)} \big/ n_h^{(2)}$ for each element in the h-th group defined by sex, age, ethnicity, and self-hoster status. $\widehat{N}_h^{(1)}$ denotes the *weighted* size of this group in Survey 1 and $n_h^{(2)}$ the sample size in Survey 2.

4.8 Characteristics Survey 2 - Results

Our initial sample in Survey 2 consisted of $n = 112$ self-hosters (i.e., a retention rate of 90.32%) and $n = 477$ non-self-hosters (self-selected from a pool of $n = 1355$), identified in Survey 1. Sample demographics can be found in Table B.3 in the Appendix.

Because we asked participants in Survey 2 to indicate whether they described themselves as self-hoster, we used this indication to compare it with the classification

based on operational criteria from Survey 1. Table 4.4 shows the consistency between classifications of Survey 1 and Survey 2. The results reported hereafter follow a conservative approach and are thus, solely based on participants whose classification concurred.

4.8.1 RQ3: Individual Characteristics

We inspected all scales and if applicable subscales, with respect to their internal consistencies [182]. We calculated Cronbach's alpha (α) for all scales requiring simple mean scores [121, 374, 80, 210, 111, 13, 9] and McDonald's omega (ω) for all scales requiring mean scores weighted with factor loadings [69, 104, 99]. Internal consistency ranged from $\alpha = .82$ to .95 and $\omega = .77$ to .92, indicating overall acceptable to good reliability for the scales. To investigate which of our selected predictors explain self-hosting behavior best, we entered all predictors into a backward stepwise regression analysis (see Table 4.5). The model that offers the best fit to our data, as indicated by the lowest Akaike information criterion (AIC) [18], contains affinity for technology interaction (ATI), maker self-identity (DIY-self), perseverance of effort (GRIT-Pe), IT background and self-reported self-hosting skills as significant predictors. ATI, DIY-self, IT background, and skills showed significant positive relations to self-hosting behavior, indicating that participants who belong to the group of self-hosters, show a higher aptitude for interacting with technical systems, identify themselves stronger as makers, report more frequently having an IT background and report better self-hosting skills than participants who belong to the group of non-self-hosters. GRIT-Pe showed significant negative relations with self-hosting behavior, indicating that self-hosters report less perseverance in their efforts as compared to non-self-hosters.

4.9 Characteristics Survey 2 - Discussion

The goal of Survey 2 was to get a better understanding of self-hosters with respect to their individual characteristics. Our results showed, in line with our predictions that self-hosters (as compared to non-self-hosters) show greater interest in technical systems, have more often a skill set that allows them to perform the behavior, and more frequently have an IT background. In Survey 1, we learned that self-hosters in general use more tools, including SaaS solutions. Both findings suggest that self-hosting goes hand in hand with a strong technical background. Speculatively, technical people who use a broad set of tools also adopt self-hostable solutions. Our observations may also point to major roadblocks in the self-hosting ecosystem, allowing only skilled people to stay in it.

Moreover, we found that self-hosters perceive themselves more often as DIY persons, although we did not find evidence that self-hosters' DIY activities differ from non-self-hosters. Contrary to our prediction, self-hosters seem to show less perseverance in their efforts as compared to non-self-hosters. What might appear as a counter-intuitive finding at first glance, could be explained by research showing that having grit not only helps people to achieve difficult goals [103] but also can have a flip side, making it hard for people to let go [23] and to persist when moving on might be the better choice

[223]. Accordingly, that self-hosters show less perseverance of effort also indicates that they might have an easier time letting go of goals that are not worth pursuing. Further research should explore whether and how being more flexible in goal pursuit might aid or result from self-hosting behavior.

Unexpectedly, we did not find any evidence that self-hosters differ from non-self-hosters with respect to their security or privacy concerns, their computer self-efficacy, their openness to new technologies, their economical consumer lifestyle, or their valuation for autonomy. This is especially surprising because self-hosters named privacy, autonomy, and security as motivational factors [P1]. However, our results do not imply that security or privacy concerns play no role when it comes to self-hosting. Rather, these factors do not explain the behavior beyond the predictors discussed above.

At present, we can only speculate that, although carefully considered, selected scale measures might not exactly represent the core themes identified in the focus groups (e.g., did the frugality scale capture all financial aspects involved in self-hosting?), or identified core themes might not apply to all self-hosters but potentially only to specific a subgroup (i.e., concurring themes in Gröber et al.'s research were found for Nextcloud users [P1]). It is also possible that identified core themes are at least partly influenced by focus group participants' ideas and conceptions about self-hosters and that these conceptions do not perfectly match the actual characteristics of self-hosters.

4.10 Discussion and Future Research Directions

Security and Usability of Infrastructure This work focuses on people who are currently self-hosting. Thus, we must assume survivorship bias with regard to the perspective of people who would like to become self-hosters, or people who tried and failed. Based on our findings, having technical skills (or believing to have technical skills), and IT background are indicators for self-hosting. This possession of technical skills could lead to people "surviving" when self-hosting. However, it may actually indicate severe technical roadblocks or usability challenges. Thus, future research should investigate what is currently preventing people from self-hosting. In doing so, studies should maximize external validity to provide a realistic view of entry burdens such as set-up procedures, infrastructure decisions, and secure configurations. We argue further, that research may focus on usability and security challenges of hosting infrastructure in the long run. In general, conducting research tailored to assist private hosters, who may have fewer resources and background knowledge, will benefit the greater population of IT administrators if security and usability challenges are streamlined.

Investigate Socio-Technical Influences Infrastructure does not exist in a vacuum, but is directly impacted by the people who administrate it and the social environment they are embedded in [P1]. We identified different individual characteristics that (may) predict self-hosting behavior. Future research should investigate the interplay of individual characteristics enabling or constraining different stages of the hosting process. The social-embeddedness may actually be a determining factor of self-hosting success. Long-term studies could help to link infrastructure configurations with socio-technical parameters such as individual characteristics of administrators, and relate those to

observable security outcomes. This way we obtain valid assessments of the security of self-hosted systems. Moreover, this work identifies demographic data that describes the self-hosting community. This data can now be used to better reach the respective target group or to tap into underrepresented groups. For example, we found that gender minorities are less likely to self-host. Future research might investigate this beyond organizational embedded administrators [186].

Community-Driven Design While we cannot directly offer implications for design based on our findings, the demographic data we collected provides valuable perspectives for future design efforts. Specifically, it is useful for building personas representative of the self-hosting population, while being mindful of underrepresented groups. We suggest aiming for a community-driven and participatory methodology when designing solutions or tooling for self-hosters. That is because there are various use cases and demographical traits to take into consideration, which makes one-fits-all solutions unlikely.

4.11 Ethical Considerations

Both studies received approval by Saarland University's ethical review board. Before collecting any data we obtained informed consent of participants for both studies. We told participants that the survey was anonymous and that all data would be treated confidentially. Moreover, we clarified that their participation was voluntary, they had the right to withdraw at any point. We disclosed our identity and offered a contact mail for any questions. Moreover, we were transparent about the overall study process including an optional follow-up survey. The collected data was stored on CISPA's private servers. To protect participant's privacy we anonymized the study data we made available to the public.

4.12 Limitations

With our sampling and weighting method, we recruited a representative sample of the United States with respect to age, sex, and ethnicity. However, our sample might still be biased by our approach to recruiting participants only via Prolific. However, recent research showed that sampling on Prolific does allow to generalize results at least with respect to certain topics [341] and outperforms other means of online data collection [102, 281]. Accordingly, in terms of overall feasibility, our approach maximizes the currently available resources and instruments.

Our sampling method and thus, our results are not immune against self-selection bias [157]. However, we took utmost care when announcing and in the instructions of the survey to conceal the actual purpose of our research. Accordingly, we cannot entirely rule out self-selection bias due to the general topic (e.g., software and application usage) but we minimized self-selection based on the topic of self-hosting.

The results of our research largely rest on our definition of self-hosting, which also determined the selection of use cases and tools. Accordingly, our research might underestimate other instances in which people administrate their own infrastructure and

services but did not see their behavior reflected in our definition of self-hosting. We paid close attention while coding the open responses in Survey 1 to include all possible cases that went beyond our pre-selection of use cases (e.g., gaming, media server) and are confident to have included these in our classification. Yet, the self-report identification in Survey 2 (i.e., whether people would describe themselves as self-hoster, taking our definition into account) might have led to an exclusion of actual self-hosters who do not agree with our definition. Because it is not possible to conduct the present research without at least a working definition of self-hosting, future research should explore other use cases and potentially more inclusive or more narrow sub-definitions of self-hosting.

As our results rest on correlations we cannot make any claims about causal relationships between self-hosting behavior and individual characteristics [217]. More specifically, we do not know whether interest in technical systems, self-hosting skills, and having an IT background is a precondition to start self-hosting or follows from self-hosting behavior. Likewise, perceiving oneself as a DIY person might be a necessary prerequisite or might simply result from the experience of self-hosting. Similarly, having an easier time letting go of goals might be beneficial for self-hosting in a fast-moving technology ecosystem or might be a result of having experienced the need to adapt and shift goals quickly when practicing self-hosting.

Nevertheless, the present research provided an estimation of self-hosting behavior in the US population and identified characteristics that set self-hosters apart from people who do not self-host. As such the present research serves as important groundwork for future research, looking into the causal relationships between the identified characteristics and self-hosting behavior.

4.13 Large-Scale Studies on Technical Topics

Identifying a Sub-Population Based on Operational Criteria and Self-Report

One of the main goals of the present research was to identify a group of people based on their specific behavior (i.e., private self-hosting). We spent time and effort to (1) define the behavior of interest as exactly as possible, and (2) derive measurable indicators of usage from this definition. This allowed us to use operational criteria for identification without the necessity for participants to self-identify. Using operational criteria can be advantageous to avoid answering behavior based on demand characteristics [35] and to ensure that the behavior intended to be captured is represented entirely in what is actually measured. However, this method is not immune to participants' misconceptions influencing their response behavior. To illustrate, without relevant background knowledge, downloading and installing an application might be misconceived as administrating it on their own hardware. Accordingly, the right wording of items and questions is paramount in minimizing the number of false positives for identification. Yet, exact wording cannot rule out misconceptions entirely.

A remedy against misconceptions is educating participants about the behavior of interest. Under the premise that participants read the information attentively, understand the definition, and can apply it to their own behavior, false positives in identification might be reduced. Yet, such an approach is more susceptible to demand characteristics. In our research, we combined both methods to keep participants'

misconceptions and demand characteristics at a minimum. We suggest that this might be good practice for investigating certain behaviors with surveys to balance out trade-offs.

Representative Sampling and Data Processing In order to assess how commonly individuals host their own services and understand how this behavior relates to personal traits, an important task was to facilitate drawing conclusions from volunteer web surveys (i.e., Prolific) that are reasonably generalizable to the population. As part of our two web surveys, we put a lot of effort in increasing the precision of our results by (1) carefully selecting a sample that reflects the overall population in terms of age, gender, and ethnicity, and (2) mitigating potential biases and selectivity in responses through thorough data cleaning and the application of calibration weighting.

We are convinced that such efforts to capture the population's full diversity in the sample and minimize the impact of potential selectivity and bias (e.g. due to low quality responses) contribute significantly to advancing human subjects research in the area of security and privacy.

4.14 Conclusion

Self-hosters take control over their data by managing it on their own vs. relying on proprietary SaaS providers. In doing so, they take responsibility not only for maintenance but also for the security configurations of their infrastructure. One might assume, that people who take on this extra burden are probably more privacy-concerned than their peers. However, our results suggest that this is not the case. People who have higher privacy or security concerns are just as likely to be self-hosting as people who have lower concerns. In other words: neither privacy nor security concerns are predictors of self-hosting behavior. Instead, we find that characteristics related to technology interest, hosting skills, and maker self-identity are positively correlated with self-hosting. Based on this we assume that self-hosting currently requires deep technical knowledge and presents barriers to non-technical people. Still, we find that it is not a niche phenomenon with an estimate of 8.4% of the U.S. population running self-hostable services. Taken together, our findings provide a solid basis to better understand the security and privacy impact of self-hosting and instigate further research and development efforts to advance the field.

Part II

Digital Sovereignty in the Context of Privacy-Violating Technology

5

Investigating Car Drivers' Information Demand after Safety and Security Critical Incidents

The contents of this chapter were published as part of the publication *"Investigating Car Drivers' Information Demand after Safety and Security Critical Incidents"* (ACM CHI 21) [P3]. This chapter uses the academic "we" to highlight the contributions of my co-authors Matthias Fassl, Abhilash Gupta, Katharina Krombholz, and me. The paper is based on my Master thesis and resulted from a significant extension. The following table details the contributions of each author to this paper:

Author	**Contribution**
Lea Gröber	I developed the research idea, designed the study, collected the data, evaluated the qualitative and quantitative data, and wrote the paper.
Matthias Fassl	Matthias contributed to the framing, suggested using the correspondence analysis, and revised the paper.
Abhilash Gupta	Abhilash contributed to the qualitative analysis and wrote parts of the technical background section of the paper.
Katharina Krombholz	As my academic advisor, Katharina guided key project decisions, provided feedback on my initial ideas and methods, reviewed the final paper before submission, and discussed the conclusions with the team.

Reference

Gröber, Lea and Fassl, Matthias and Gupta, Abhilash and Krombholz, Katharina. (2021). Investigating Car Drivers' Information Demand after Safety and Security Critical Incidents. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 1 - 17.

Modern cars include a vast array of computer systems designed to remove the burden on drivers and enhance safety. As cars are evolving towards autonomy and taking over control, e.g. in the form of autopilots, it becomes harder for drivers to pinpoint the root causes of a car's malfunctioning. Drivers may need additional information to assess these ambiguous situations correctly. However, it is yet unclear which information is relevant and helpful to drivers in such situations. Hence, we conducted a mixed-methods online survey ($N = 60$) on Amazon MTurk where we exposed participants to two security- and safety-critical situations with one of three different explanations. We applied Thematic and Correspondence Analysis to understand which factors in these situations moderate drivers' information demand. We identified a fundamental information demand across scenarios that is expanded by error-specific information types. Moreover, we found that it is necessary to communicate error sources, since drivers might not be able to identify them correctly otherwise. Thereby, malicious intrusions are typically perceived as more critical than technical malfunctions.

5.1 Introduction

In recent years, modern cars' automation levels increased from driver assistance to partial automation – thereby making car-integrated technology unprecedentedly complex. These cars increase driving safety while also reducing the burden on drivers. To date, modern automation features require constant supervision which drivers struggle to provide over a longer period of time [205]. However, even if they do pay attention, the reactions of the car may be hard to predict or explain, e.g. in case of an accident. In older cars, the blame was usually on the driver or some technical malfunction. Modern cars' behavior is becoming increasingly opaque due to a rising level of autonomy, while opening up new attack surfaces [27, 271, 310, 378, 380, 381], and exposing drivers to unknown threats. Hence, drivers increasingly rely on proper in-car risk communication. If drivers would receive relevant information they could: (1) explain the car's behavior, which also builds trust and confidence in the technology, (2) resolve liability issues, i.e., blame the correct party for the accident, and (3) take appropriate actions to avoid such accidents in the future.

However, to provide drivers with helpful explanations and warnings, we first need to understand the drivers' information demand in safety- and security-critical incidents. The impact and design of explanations and warnings have been extensively studied with respect to security warnings in browsers [19, 218, 219, 299, 300, 302]. However, the domain of partially-autonomous vehicles constitutes a special case, as it involves potentially life-threatening situations. Lim and Dey investigated the demand for intelligibility in context-aware applications [220]. However, they focused on desktop applications and explicitly did not cover any level of autonomy or high risk situations. Recent work of Smith et al. focused on high risk situations. The authors explored pilot reactions to attacks on avionic systems [328]. However, their emphasis was more on reactions and not on information demand in security critical situations.

To investigate drivers' information demand, we conducted a mixed-methods online survey on Amazon Mechanical Turk with $N = 60$ participants. At this point we want to distinguish the "drivers" in our study from real-world drivers. That is, in our study

participants react to hypothetical scenarios relieving them from any driving-related tasks. We exposed participants to safety- and security-critical situations. We carefully selected ambiguous malfunctions for these situations which could be explained by either a malicious intrusion or a technical defect: (1) a car with activated autopilot hits construction barrels on the highway, and (2) a car does not unlock upon the first click of the key. The survey provided one of the following explanations for these situations: (a) a malicious intrusion (security breach), (b) a technical malfunction, or (c) no explanation. Exposing participants to different explanations expands the exploratory space, as the context of the critical situations shifts according to the car's explanations. Adding a condition in which the car does not provide an explanation gives us insights about the participants' own interpretation of the error cause.

Afterward, we used open-ended questions to elicit the participants' information demand and quantitative questions to assess their trust, satisfaction, and operational intent. We used Thematic Analysis [56] to evaluate the qualitative data. Additionally, we applied Correspondence Analysis [135] to understand which factors in these safety- and security-critical situations moderate drivers' information demand. The quantitative data was analyzed with statistical tests to verify qualitative results.

We found a basic need for information across scenarios, which is expanded depending on perceived error causes. Technical malfunctions and malicious intrusions have little overlap resulting in more car or situation specific information demand. Malicious intrusions were consistently perceived as critical, even if other perceived error sources in the same scenario were not. There exists a gap between highly critical situations and less critical situations in terms of trust, satisfaction, and operational intent ratings. Additionally, we identified the need to communicate error sources, as participants are not aware of malicious intrusions. They also have trouble to assess and react to highly critical situations.

5.2 Related Work

In the following, we discuss related work on intelligibility in human-computer interaction, trust and explainability in autonomous systems.

Intelligibility in Human-Computer Interaction

Our work is heavily influenced by Lim and Dey's paper "Assessing Demand for Intelligibility in Context Aware Applications" [218]. The authors conducted two experiments to elicit users' demand for information and to verify their findings. The first experiment was an online study carried out on Amazon Mechanical Turk [351]. Participants had to answer qualitative questions regarding the behavior of one of four context-aware applications. Additionally, Lim and Dey assessed participants' satisfaction ratings regarding their experience with the application. The second experiment assessed whether or not users' satisfaction levels rise if they are presented with the type of information they demand. The authors found, among other things, that users want any available information in critical situations, while at the same time they are hardly satisfied with the information they get. The authors, however, did not include autonomous vehicles

or systems of any kind in their study. In the following year, Lim and Dey published a toolkit to support intelligibility in context-aware applications [219]. The toolkit was designed to assist developers with incorporating different intelligibility types into applications. However, since the context of driving a car is inherently different from using a desktop application, further work is necessary to investigate this specific use case and technology. Our study aims to close this gap in the literature. Bellotti et al. came up with four principles to support intelligibility and accountability in context-aware systems [44]. They identified a need to inform the user of a system's capabilities, provide feedback, ensure identity and action disclosure, and grant the user control over the system. While these principles are an excellent point of reference, their broad character does not allow for concrete design decisions. Our study provides actionable insights that help to improve car-driver communication in line with these principles. Research has also addressed the information needs of users in other areas. For example, McGuinness et al. conducted an interview study that identified themes influencing the willingness of users to use and trust an adaptive agent [246, 127]. In addition, Gregor et al. did a meta-review and identified what kind of explanations users of knowledge-based systems demand [144]. Jakobi et al. investigated long-term information demands in do-it-yourself smart home systems, identifying changing information demands over time [176]. Again, the results of these works cannot be directly transferred to the domain of modern cars. Therefore, our study will provide valuable contributions to complete the picture of the information needs of users in different contexts.

Explainable Artificial Intelligence in Autonomous Systems

Recent research focused on making the actions and internal processes of systems with varying levels of autonomy understandable, and communicating them to users [155, 209, 168, 317, 101, 181]. For instance, Hastie et al. [155] introduced a multimodal interface (MIRIAM) for remote autonomous systems. MIRIAM is intended to increase the transparency of the system and thus strengthen the operator's confidence in the system. The interface allows one to pose *why* and *what* questions to the system. Langley et al. [209] framed the concept of explainable agency for intelligent autonomous systems and claim that an agent needs to convey its internal reasoning to the user, and which actions it executed, among other things. These approaches form a good basis when it comes to conveying knowledge to users. However, depending on the situation, the drivers may not be receptive to different types of information. Therefore, it is important to investigate the information needs of drivers in order to provide them with adequate information adapted to the situation.

Trust in Automation

Among other things, our work studies how we can maintain a trustful communication between vehicles and drivers in the context of critical situations. Therefore, we discuss research about trust in automation [166, 212, 249, 152, 227, 305] in the following. Madhavan and Wiegmann found that the process of forming trust in a machine differs from trusting humans [227]. This is because humans initially treat other people with caution [305]. The trust relationship is built slowly, as long as the other person does not

make any mistakes. In contrast to this, people usually assume that machines function flawlessly. Hence, they encounter them with a trust advance [227, 212]. With every mistake the machine makes, this trust is then corrected downwards [107]. However, this effect only occurs if the person has had no previous contact with the machine [227].

5.3 Technical Background

Remote keyless system (RKS) technology was first introduced in cars in the 1980's [207]. Since then, the underlying technology of RKS has continuously evolved after each version was demonstrated to be exploitable. At the time of writing this, it uses encrypted rolling codes. However, this is also vulnerable to exploits as shown in various demonstrations [183, 123, 139, 140, 141, 143], the last one as recently as November 2020 [143]. This is most likely due to incorrect implementation of protocols or reliance on flawed protocols. The most common of these exploits are relay attacks (which repeat the signal from the driver's key to the car from a large distance using relays) and replay attacks which capture and block valid signals from the driver's key fob and use these signals later on. Vulnerabilities in cars are not limited only to car keys [381]. Researchers have already gained control of a moving car while sitting in the back seat [136] as well as from kilometers away [137]. They gained control of the steering wheel, brakes, windshield wiper, air conditioner, and the dashboard system. Recently, researchers found that they could fool Tesla's autopilot program into believing "phantom" signs. They were able to trick a Tesla to stop, by flashing a stop sign for a second on a billboard next to the road [142]. Apart from malicious intrusions, the computer systems of a car may suffer from technical malfunctions. The video used in our study shows an example of when the autopilot failed to recognize objects in its path and crashed through construction barrels [384].

As the number of computerized features in cars increases, so does the potential for exploits and malfunctions to be life-threatening [310, 384]. While car systems currently do not communicate warning messages about third party interference to the driver, scientists are working on solutions to detect malicious intrusions in vehicles to safeguard their internal functioning and ensure that such exploit attempts are thwarted [271, 380, 126, 256, 73, 105]. Such mechanisms can possibly be further developed to alert drivers about third-party intrusions.

For this study we chose scenarios inspired by technical malfunctions and exploits that either occurred in the real world or were demonstrated to be feasible by scientists. However, to the best of our knowledge, there is currently no mechanism to alert drivers of an ongoing attack, even if it were detected. For this study we assume the car is capable of such a detection and notification to the driver, to investigate which information people need in critical situations.

5.4 Methodology

Our study is designed to elicit drivers' information demand depending on different critical situations. Hence, our study lays the foundation to improve in-car risk communication

to drivers and to provide helpful information at appropriate times. Accordingly, we identified the following research questions:

RQ1: What information do drivers demand for safety- and security-critical incidents?

RQ2: Which factors moderate information demand after critical incidents?

RQ3: Which error sources for safety- and security critical incidents do drivers think of?

Since it would be unethical to put participants into critical situations we use an online survey with scenarios to investigate their attitudes, trust, satisfaction, and information demands. To cover a broad spectrum of situations, we selected a high-critical scenario (crashing against construction barriers) and a low-critical scenario (key malfunction). We specifically chose ambiguous scenarios in which the cause of vehicle malfunctions is not obvious. Since we confront participants with hypothetical scenarios, the participants ("drivers") are relieved from all driving-related tasks. This constraint is further strengthened as the car in the scenarios is not moving at the time we elicit participants' information demand. After each scenario, participants fill out a questionnaire with qualitative and quantitative questions. We apply *Correspondence Analysis (CA)* to investigate which factors moderate drivers' information demand. We describe each of the identified correlations in detail using qualitative data from the free text response questions.

5.4.1 Online Survey

All participants ($N = 60$) are exposed to two scenarios (C: crashing against construction barriers and K: key malfunction) in a randomly chosen sequence. Each scenario contains (1) an introductory text, (2) a description of the situation, and (3) the vehicle's explanation. The vehicle explains its behavior with one of the following explanations (randomly assigned per participant and used for both scenarios): **malicious intrusion (MI, $N = 17$)** by third parties, a **technical malfunction (TM, $N = 19$)**, or with **no explanation (NO, $N = 24$)**. Hence, scenarios (C and K) are studied within subjects and explanations (MI, TM, NO) are studied between subjects. After each scenario, participants fill out a questionnaire about their experience. The supplementary material provides the scenarios as presented to the participants.

Introductory text The introductory text embeds each scenario in the setting of partially-autonomous vehicles by describing the vehicle's capabilities and limitations. This introduction directly addresses the participant to make the setting more tangible, thus making it easier for the participant to immerse into the situation. The text describes the car's functionality according to the claims on Tesla's website [343]. In particular, that it can automatically steer, accelerate, and brake within its lane. However, the text explicitly states that active driver supervision is required at all times. We did not want to study a specific brand of vehicle, but used the Tesla description for a realistic abstraction of such a vehicle. Hence, we did not specify the brand of the car in the survey. From the video illustrating scenario C, one cannot infer which car it is.

Additionally, we omitted a description of the center console or visual representation of the error message to minimize the influence of factors beyond our main focus.

We chose scenarios in which the vehicle communicates a malfunction that could have been caused by a functional error or a malicious attacker. We hypothesize that in such cases the driver cannot identify the source of the malfunction without further context.

Scenario C: Crashing against construction barriers This scenario asks the participant to imagine driving on the highway with an activated autopilot. The description explicitly emphasizes that this requires active driver supervision. Just like the introductory text of the setting, this description is designed to be as tangible and immersive as possible. Hence, it contains elements that should make it easy for the study participant to imagine herself in the situation. For example, instead of simply saying that the driver was briefly inattentive, the text provides a vivid description of why this is the case: *"You receive a text message from your best friend to which you reply immediately."* This not only ensures that the study participants can better identify with the situation but also establishes a common ground and thus leaves less room for interpretation and misunderstandings.

The actual situation is presented in a 22-second video [384]. It shows the collision of a vehicle with construction site barrels from the driver's perspective. This scenario is based on an actual event: A dashcam recorded this situation in a Tesla while the car's autopilot failed. In the video, the vehicle drives towards the end of a highway lane that is closed due to construction. For an unknown reason, it does not recognize the construction site barrier. The driver reacts too late and only intervenes after the vehicle has hit 10 barrels[1]. Figure 5.1 depicts the entire description of the scenario as presented to the participants. After this video, another tangible description clarifies that the driver, not the vehicle, activated the brake.

The vehicle in this scenario responds in one of three ways to the incident: explaining its behavior with a malicious intrusion (MI), attributing it to a technical malfunction (TM), or not explaining at all (NO):

MI *You look at the car's center console and learn that your car's behavior was caused by a hacker. They temporarily took control of the vehicle and steered it into the construction barrels.*

TM *You look at the car's center console and learn that your car's behavior was caused by a sensor malfunction. The front sensors did not recognize the construction barrels, causing the incident.*

NO [*no explanation is offered*]

After participants experienced the scenario and the vehicle's explanation, the survey continues with the open and closed questions shown in Table 5.1.

[1]According to the video description the driver fell asleep behind the wheel of his Tesla. Although the driver acknowledges that the accident was mainly his fault, the vehicle is also held accountable: *"Automatic Emergency Braking totally failed me on the one time we needed it most. With all the phantom braking events we have experienced in the 2½ months we've owned it, it does seem like it would panic when it saw this coming."* [384] Note: The owner of the video has since taken it offline. Please contact the authors of this paper if you have further questions regarding it.

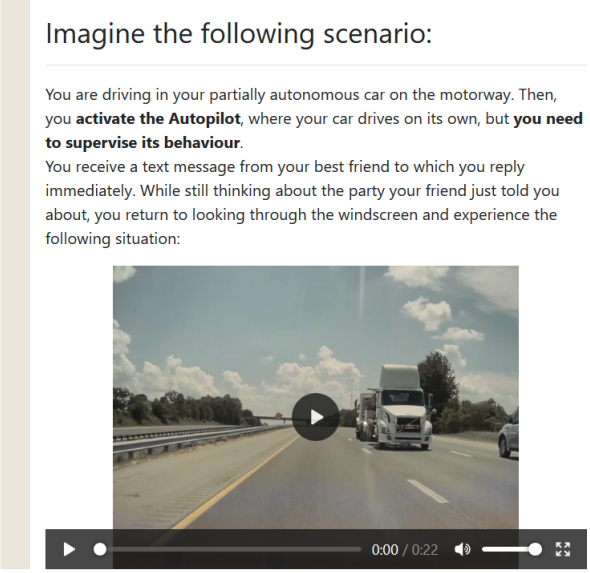


Figure 5.1: Scenario C: a vehicle crashing into construction barrels as presented in the survey. The supplementary material provides the complete version of the survey.

Scenario K: Key malfunction In this scenario, the driver wants to unlock her vehicle with a remote key fob. However, it does not respond the first time and the driver needs to press the "unlock" button again to unlock the vehicle. We chose this scenario because the problem of cloning keys has been present for many years [192, 138, 139, 123, 183] and was recently prominent in the media again when a Tesla was stolen from a driveway [323]. Furthermore, this scenario is less critical than the other one, as it usually only causes material damage without threatening the lives of the vehicle's occupants. The attack mentioned by the malicious intrusion explanation refers to key fobs and vehicles that synchronize using rolling codes [183].

The vehicle in this scenario responds in one of three ways to the incident: explaining its behavior with a malicious intrusion (*MI*), attributing it to a technical malfunction (*TM*), or not explaining at all (*NO*):

MI *When you look at the car's center console you learn that someone may have cloned your key and can now use it to unlock your vehicle.*

TM *When you look at the car's center console you learn that the battery charge of your key is weak and that you need to replace it soon.*

NO [*no explanation is offered*]

After participants experienced the scenario and the vehicle's explanation, the survey continues with the open and closed-ended questions shown in Table 5.1.

Questionnaire After each scenario, participants answered four qualitative open-ended questions regarding their (1) perception of the scenario, (2) next actions, (3) feelings

Table 5.1: Questionnaire after each scenario containing qualitative and quantitative questions.

Measure	Scenario Question	Answer Type
Perception/ Attention	What do you think happened in this scenario?	free text
Action	What will you do next as the driver?	free text
Driver Feeling	How do you feel about the vehicles response?	free text
Information Demand	What information should the car provide about the situation?	free text
Vehicle Satisfaction	I am satisfied with the vehicles behaviour in this specific situation.	7-point Likert
Trust in Automation Scale	According to Jian et al. [179] adjusted to partially-autonomous vehicles	7-point Likert
Operational Intent	After experiencing this incident, I would buy a vehicle of this kind again.	7-point Likert
	I would sue the manufacturer of the partially-autonomous vehicle.	7-point Likert
	I would continue to use the partially-autonomous vehicle.	7-point Likert
	I would warn my family and friends about the partially-autonomous vehicle.	7-point Likert

about the scenario, and (4) demand for information. Likewise, participants answered four quantitative closed-ended questions regarding their (1) satisfaction with the vehicle's response, (2) trust in automation (using Jian et al.'s [179] scale), and (3) operational intent. All quantitative questions asked for a response on a 7-point Likert scale. Table 5.1 lists all qualitative and quantitative questions asked after each scenario.

5.4.2 Pilot Study

The goal of the pilot tests was to test and improve the comprehension of questions and scenario descriptions. We conducted a total of 6 pilot tests in which we asked participants to think-aloud while completing the survey. This not only allowed optimization of the texts and questions, but also revealed layout flaws. We iteratively conducted pilot tests and directly incorporated the results into the survey after each round of testing. We continued until we had covered every condition once and the participants completed the survey without any problems. Based on the outcome of the pilot tests we adjusted phrasing of the open-ended questions and conditions. For example, we slightly rephrased some questions to clarify the direct reference to the scenario. Additionally, we added gray bars to the left of each paragraph to provide visual guidance. The final analysis does not include the results of the pilot tests. We recruited pilot test participants from our university achieving an even distribution of women and men, computer-science students, and administrative employees, aged 23-45.

5.4.3 Recruitment and Participants

We recruited $N = 60$ study participants from Amazon Mechanical Turk (MTurk) [351]. Participants were randomly assigned to conditions (MI, TM, NO), resulting in an uneven distribution among conditions. We carefully balanced sample size considerations for our mixed methods study. We performed power calculations to estimate the number of participants for the quantitative analysis. For a statistical power of 0.8 and $\alpha = .05$ we estimated 60-80 participants for a medium effect size. With regard to the qualitative analysis, we are confident that the number of participants is sufficient as we reached saturation (see Section 5.4.4).

We chose MTurk because it enables us to effectively investigate the information demand of a broad set of people, as opposed to e.g. lab studies. Additionally, we wanted a culturally homogeneous sample that is known to be suitable for security research. Prior work by Redmiles et al. suggests that MTurk responses regarding security and privacy experiences, advice sources, and knowledge are more representative of the U.S. population than are responses from a census-representative panel [298]. To participate in our study, MTurk workers needed to own a car and be located in the US. We selected car ownership as a criterion to ensure that participants have experience with regular cars. None of the participants owned a partially-autonomous car. 13 participants reported having previous experience in driving or riding cars with autonomous driving features. Driving experience varied between 7 and 52 years (median 21, mean 23.94). Additionally, we required a HIT Approval Rate[2] for all Requesters' HITs greater than 95%, and that they have more than 100 approved HITs. In the pilot test, participants completed the survey in about 20 minutes, so we compensated participants with $3 for the completion of the survey. However, participants invested more time than anticipated (26 minutes on average) which resulted in a wage below the US federal minimum ($7.25). To remedy this situation, we gave a $0.50 bonus to all participants. A total of 23 woman and 27 men took part in our study with ages ranging between 24 and 73 (median 27, mean 40.67). Table 5.2 shows a detailed overview of our participants' demographics.

Table 5.2 contains an overview of the participants demographic split by conditions.

5.4.4 Coding Procedure

We used open coding according to Strauss and Corbin [332] to evaluate the qualitative data. In total, we created 6 codebooks, one for each scenario and condition.

Two researchers independently coded answers to open-ended questions for each scenario and condition in two iterations. Initially, one of them skimmed the first half of each dataset and constructed an initial version of the codebook. The draft version of the codebook captured the dataset's concepts and topics. Afterward, both researchers used this codebook to code the second half of the data. They tagged pieces of the answers with labels, at once summarizing, categorizing, and describing the data [66]. After this first iteration of coding, the two researchers discussed their codes and adapted the codebook accordingly. During the second iteration, the two researchers coded the first half of the dataset. In cases where the discussion of the second iteration also led to a change in the codebook, the researchers coded the second half of the dataset again. This resulted in an *inter-rater reliability* Krippendorff's α [199] between 0.77 and 0.91 for each codebook. We achieved saturation after the first coding iteration for 5 codebooks. In the Crash, *NO* condition one participant reported to repair the sensor which added a new concept to this codebook in the second coding iteration.

[2] "A Human Intelligence Task, or HIT, is a question that needs an answer. A HIT represents a single, self-contained, virtual task that a Worker can work on, submit an answer, and collect a reward for completing. HITs are created by Requester customers in order to be completed by Worker customers." [352]

	TM	MI	NO	Total
Gender				
Women	7	9	7	23
Men	17	8	12	37
Age				
Min	25	24	24	24
Max	73	59	69	73
Median	37	36	38	37
Mean	40.73	38.94	42.36	40.67
Education				
Min	4	4	4	4
Max	21	13	17	21
Median	13	10	10	10
Mean	11.00	10.17	9.89	10.35
Driving Experience				
Min	9	7	8	7
Max	48	44	52	52
Median	20	22	22	21
Mean	23.43	23.70	24.78	23.94
ATI Scale				
Min	1.55	2	3.55	1.55
Max	6	6	5.88	6
Median	4.66	4.55	5	4.66
Mean	4.53	4.34	4.80	4.56

Table 5.2: Participant demographics. Education reported according to OPM educational level (273). Driving experience in years. Affinity for technology interaction(ATI) scale (121) results on a scale from 1-6. Higher values indicate a tendency to actively participate in intensive technology interaction (346).

5.4.5 Analysis

We conducted *Thematic Analysis* [56] to identify topics, correlations and themes in the coded data. Further, we applied *Correspondence Analysis* [135] to explore the relationship between different situational factors and the occurrence of information demand codes. The information demand codes are a result of the open coding procedure from the previous section. The different situational factors that might moderate information demand are a result of the Thematic Analysis.

5.4.6 Ethical Considerations

Our university's ethical review board (ERB) evaluated and approved this research project. To enable informed consent, we explained the study objective to the participants, stated that participation is voluntary and that they may abort the survey at any time. Further,

Table 5.3: Consolidated codebook of participants' information demand across all scenarios and conditions

Code	Description	Example Quotes
Question Types		
What	What happened in the situation?	"It should provide a report of what happened [...]", "It should say that there is a breach in the system or something of that regard [...]"
Why	Why did the malfunction occur?	"It should also have some kind of an explanation as to why it didn't brake"
Who	Who is the attacker?	"Who is responsible."
When	When did the attack happen?	"when it was hacked"
How	How did the attack happen?	"Any information about how it was exploited by the hacker"
Car		
Status	Status information of malfunctioning parts	"Status of the remote; low battery indicator; weak signal strength maybe"
Diagnostic Report	Car's diagnostic report, e.g. error codes, damage	"The car should provide a report about the damage - when the car collided with a hazard, how fast it was traveling, any potential damage to look out for", "if the autopilot is still working"
Decision Process	Car's internal decision process leading to accident	"How it interpreted the situation and any negative reactions to the incident."
Parameters	Car's parameters during accident, e.g. velocity etc.	"show a graph of some sort of how long/distance it continued driving from the first hit of a cone to when the vehicle eventually stopped along with the speed, if it slowed down at all etc"
General Usage	General information about how to use the car	"How many clicks is necessary to unlock the car"
Situation		
Preventative Measures	How can such accidents be avoided in the future?	'[...] what steps I can take next to keep this from happening again."
Message	Message, Warning, or Alert about Incident	"It can send a message that it only got a partial signal the first time.", "It would be great if the car could give a warning [...]"
Troubleshooting	Information to resolve the issue manually and clues to find the attacker	"It should tell you when it was copied and that way you could try to figure out who it was"
Recommendations	Recommendation how to react to the situation	"Do the sensors need to be checked?"
Quality of Warning		
Visual	Demand visual message	"The car should save the visual evidence if it has a built in dash cam"
Audio	Demand audible message	"The car should have an automatic warning system [...] like a voice warning"
Before/During	Demand message prior to or during the incident	"The car should have sounded a warning of the cones approaching."
None		
None	No information demand	"Nothing at all, unless there's a reason why the fob truly needed more presses [...]", "Nothing really, it seems self-explanatory to me"

we did not collect any personally identifiable information (PII). At the end of the survey, we provided links to our webpage and contact information in case participants had further questions.

5.5 Results

We first used Thematic Analysis to identify concepts, topics and connections in the coded data. Those insights then form the basis for a subsequent Correspondence Analysis of situation-related information needs. Table 5.3 shows an overview of all information types elicited in the study with explanations and quotations. Table C.1 in the Appendix shows a comparison of all codebooks grouped by themes. All quotes in this section are unaltered including spelling mistakes.

5.5.1 Results of Thematic Analysis

The results of the Thematic Analysis are grouped by high level codes about perceived error causes. For each, we report on similarities and differences, as well as specifics of both scenarios. In addition, we talk about differences in the perception of both scenarios and about first trends in the need for information.

5.5.1.1 Malicious Intrusion as Perceived Error Source

Across both scenarios nobody thought of security breaches as possible error sources, unless they were primed for it in the *MI* condition. This suggests that security breaches are a concept that is not deeply rooted in people's minds when it comes to driving. However, if the participants were then made aware of a malicious intrusion, this was perceived differently depending on the situation.
In the key scenario people were unsure about how the attack was carried out: *"I don't know how someone would go about cloning a key."* (P34). They tried to make sense of the situation, each coming up with different explanations to what might have happened, being more or less close to the actual attack we had in mind: *"Someone accessed the computer in the vehicle and made a clone of the entry system."* (P37), *"[...] there was someone standing nearby with some sort of rf reader and intercepted the authentication code used to unlock the vehicle. the first click to unlock the vehicle did nothing because it went to the interceptor, the second time unlocked because it was gong to the car, not the rf reader"* (P20). 2 participants thought that the malfunction was actually a security functionality, impeding attacks: *"it's actually good that it doesn't unlock right away because it will take longer for the hacker to access it"* (P37). The participants reacted to the key scenario with mixed feelings. 5 participants perceived the scenario as *scary* (P30). One participant expressed that *"Hacking is a real concern."* (P21). However, at the same time 9 participants perceived the warning of the car in a positive way, e.g. stating that they were *satisfied* (P21,63), and *happy* (P65). This is likely the case because the car warned its driver prior to a potential theft: *"Since the car wasn't stolen/missing after a stranger cloned my key, it seems the security features are working for the time being"* (P29). Nevertheless trust issues may remain as one participant stated *"I would be paranoid about the issue really being resolved once it had been corrected."* (P34).
In the crash scenario participants understood that the accident was caused by a malicious intrusion, but were uncertain about what exactly happened: *"somehow the hacker was able to disable the safety features of the vehicle [...]"* (P20). Additionally, participants had different ideas about hacker capabilities. For example, 1 participant stated that *"The car should have stopped much quicker. Even if a hacker impacted the steering, the brakes should have been activated because of collision warnings."* (P27), and a second one stated: *"[...] I would figure the systems would only allow things like that if they were being manually overridden from withing the car itself"* (P39). To add to this, 7 participants stated that they would continue to drive manually, while only 2 reported to call the police or get the car towed. 12 participants had predominantly negative feelings with regard to the car's behavior, stating they were *irritated* (P29), and *angry* (P52). In contrast to the key scenario, participants did not appreciate the car's explanation of the situation. This is likely because the accident already happened and the car failed to notify the driver early. 2 participants explicitly classified the situation as potentially lethal: *"[...] It is scary to think that someone can hack the system and potentially kill you. [...]"* (P30). With regard to responsibility, 15 participants held *hackers* (P4) accountable for the accident. Out of those, 5 participants acknowledged that the driver is guilty of not paying attention to the road and 1 person blames *"[...] the people who designed the vehicle [...] "* (P34).

5.5.1.2 Technical Malfunction as Perceived Error Source

Technical malfunctions were the most named error sources in the *NO* and *TM* conditions across both scenarios. Thereby, participants demonstrated a good understanding of and intuition for technical error sources. In general, the concepts and themes mentioned in the *NO* and *TM* conditions broadly overlapped in both scenarios.
13 participants identified a *"bad battery"* (P24) as a potential error in the key scenario, *NO* condition. In the *TM* condition everybody who correctly understood the scenario identified the key's battery as the error source. 3 participants misinterpreted the scenario to be about an electric vehicle. We excluded them from the analysis of this scenario. In *TM* condition, the majority had neutral to slightly positive feelings about the car, describing its behavior as *helpful* (P19), and *acceptable* (P9). 5 participants had negative feelings, expressing they would *"prefer that it [the car] warns me before the battery gets so low that i may stop working correctly"* (P33). In the *NO* condition the majority had neutral to slightly negative feelings, stating that batteries being low are a *common occurrence* (P22) and that they *"didn't like how it [the car] was unresponsive"* (P24).
In the crash scenario 15 participants thought the accident was caused by a technical malfunction in the *NO* condition. In the *TM* condition, 22 participants correctly attributed the error to a *sensor malfunction* (P32). In both conditions the majority was extremely unsatisfied with the car's reaction, stating they were *"upset and frightened"* (P19), *"surprised and a little panicked"* (P32), or *scared* (P23). The message in the *TM* condition had no positive effect on the overall impression of the situation. Similar to the *MI* condition, participants demanded to *"be forewarned if a sensor is failing or has failed"* (P69). In terms of liability, 7 participants in the *TM* and 9 participants in the *NO* condition held the driver accountable for the accident: *"The driver (me) was not paying adequate attention to the situation [...]"* (P28). 8 participants in the *TM* and 3 participants in the *NO* condition blamed poor design: *"Apparently the sensors weren't programmed to recognize the particular obstacles [...]"* (P15), *"Shouldn't the vehicle be able to recognize the signs warning of the lane ending in the first place."* (P33). 1 participant wanted the car to *"acknowledge that it made a mistake"* (P39).

5.5.1.3 Human Error as Perceived Error Sources

Apart from technical malfunctions and malicious intrusions, participants also identified other potential error sources. Especially if no explanation for the car's behavior was offered, participants blamed the malfunction on themselves across both scenarios. In the key scenario, 6 participants in the *NO* condition made statements like *"Sometimes [...] you don't press it [the key] correctly so you need to do it again"* (P22), or *"I didn't press the button hard enough"* (P49). Nobody mentioned human failure as the error source in the *TM* or *MI* condition in the key scenario. In the crash scenario, on the other hand, the concept came up more frequently. This is likely the case, as the description pointed out that the driver was inattentive. Here, 9 participants of the *NO* condition stated that they *"stopped paying attention in a situation where I should have been supervising"* (P25). 6 participants (*TM*) and 5 participants (*MI*) made similar statements.

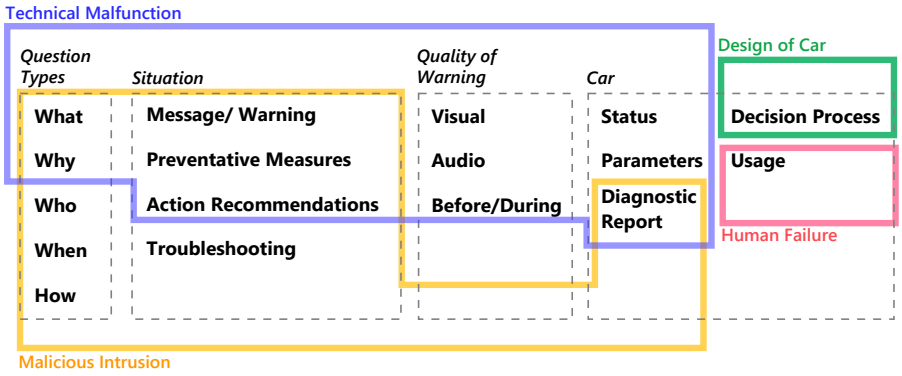


Figure 5.2: Overview of how information demand (in the background) corresponds to the perceived cause of error (colored boxes) across both scenarios.

5.5.1.4 Design of the Car as Perceived Error Source

Some participants thought the malfunctions were not actual malfunctions, but intended by design, e.g. in the key scenario, or limits of the cars functionality in the crash scenario. When no explanation was given in the key scenario, 5 participants explained the car's malfunction with statements like: *"The car was programmed to unlock at two clicks [...]"* (P48). Note, that out of those 3 participants understood the malfunction as a security feature: *"I feel safer with this and know that my car would not open for just any one just for the remote that I have"* (P59). Nobody in the *MI* or *TM* conditions thought the malfunction was intended by design. In the crash scenario 3 participants in the *NO* and 6 participants in the *TM* condition thought that the malfunction was due to limited functionality of the autopilot: *"The car was apparently only programmed for any side abstruction"* (P74), or *"I think the vehicle got confused. It knew there was a road there but wasn't aware of the construction."* (P56). Nobody in the *MI* condition thought the malfunction was due to gaps in functionality.

5.5.1.5 Perceived Criticality of Scenario

The qualitative findings suggest that participants perceive the key scenario less critical than the crash scenario. This is indicated by different impressions participants have about both scenarios as well as their reported actions to the scenarios. However, the *MI* condition of the key scenario constitutes a special case, as 14 participants reported on calling the police or contacting the manufacturer. Thus, this condition was also perceived as more critical. The overall gap between key and crash scenario was also evident in the quantitative data. The scores of trust, operational intent, and satisfaction in the key scenario were significantly higher than in the crash scenario. For each dependent variable we ran MANOVA. The test provides information about Wald-type statistic (WTS), the ANOVA-type statistic (ATS) and re-sampling versions of these test statistics [320]. Using WTS and ATS, there was a significant effect of the scenarios on the trust scale rating, satisfaction rating, and operational intent rating, df=1, $p<0.001$.

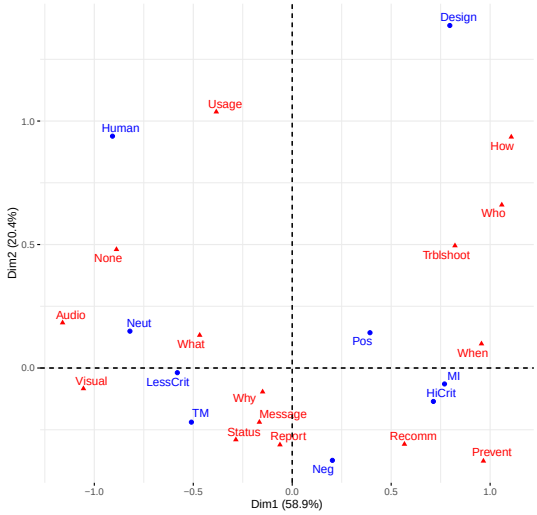


Figure 5.3: *Key* scenario

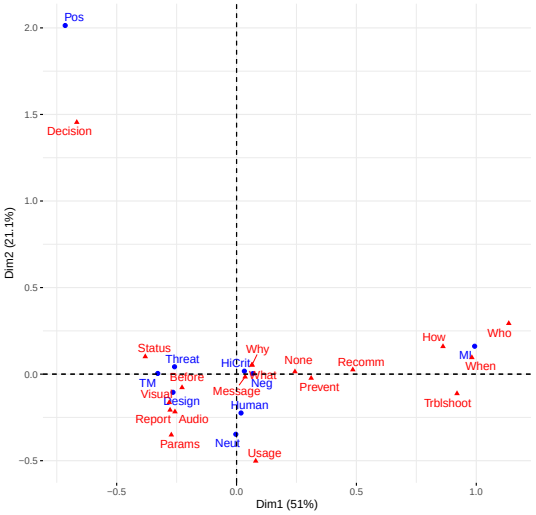


Figure 5.4: *Crash* scenario

Figure 5.5: Asymmetric biplots of *moderating factors* (blue dots) and *information demand codes* (red triangles). The dimensions correspond to the eigenvalues that cover the largest percentage of variance.

5.5.1.6 First Trends in the Need for Information

A fundamental need for information of the study participants became apparent across scenarios and conditions. It includes the questions why the malfunction occurred and what happened in the situation. In addition, the need to receive a warning or message from the car was widely expressed. Depending on the perceived severity of the situation and on the error source, participants asked for additional information. Figure 5.2 illustrates the information types, grouped by perceived error source across all scenarios. For technical malfunctions participants generally cared more about information on the car and were more specific about the kind of warning they wanted. If participants thought a malicious intrusion was the error source, they were more interested in information about the situation including different question types about the attack. The perceived error source human failure resulted in interest about how to properly use the car or the key. Last but not least, if participants thought the malfunction was due to limited functionality they wanted to know more about the car's internal decision process.
In the following section, we present the results from our Correspondence Analysis to identify more fine-grained trends in information demand.

5.5.2 Correspondence Analysis

Based on the Thematic Analysis, we identified different situational factors that might moderate information demand (*moderating factors*):

- **Highly Critical Situations (HiCrit)** Participant perceived the situation as highly critical.

- **Less Critical Situations (LessCrit)** Participant perceived the situation as less critical.
- **Technical Malfunction (TM)** Participant identified a technical malfunction as the error cause.
- **Malicious Intrusion (MI)** Participant identified a malicious intrusion as the error cause.
- **Human Failure (Human)** Participant identified a human failure as the error cause.
- **Design of Car (Design)** Participant perceived the malfunction as an intended design choice.
- **Life Threatening (Threat)** Participant perceived the situation as life-threatening.
- **Positive Impression (Pos)** Participant had a positive impression of the car's response to the situation.
- **Negative Impression (Neg)** Participant had a negative impression of the car's response to the situation.
- **Neutral Impression (Neut)** Participant had a neutral impression of the car's response to the situation.

We applied Correspondence Analysis between *moderating factors* and *information demand codes* (shown in Table 5.3) to explore their relationship. We explored this relationship using biplots (shown in Figure 5.5) and checked each conclusion against the raw data. These biplots visualize the relationship between the moderating factors (blue dots) and information demand (red triangles). Put simply, the further away labels are from the origin, the more discriminating they are, and smaller angles between a moderating factor and an information demand label (connected through the origin) indicate an association of the two.

Relative inertias indicate for each cell of the contingency table (refer to Tables C.7 and C.4 in the Appendix) the relative contribution to the total value of the chi-square statistic. The higher the value of a cell, the higher the association of the respective row and column categories. We report relationships with relative inertia larger than 0.01 only if they are grounded in the qualitative data.

Appendix C.2 provides tables with exact results of the Correspondence Analysis. The following paragraphs report on information demand trends depending on moderating factors.

5.5.2.1 Perceived Error Cause

Participants identifying a **malicious intrusion (MI)** as the error cause in a scenario, was globally a strongly discriminating factor for information demand. Participants were more interested in information about the situation than in information about the car. In the crash scenario *MI* was the only factor that had an impact on information demand.

It led to increased demand for the information types *Who, When, How, Preventative Measures, Troubleshooting,* and *Action Recommendations*. In the key scenario *MI* is a strongly discriminating factor. Similar to the crash scenario there is increased demand for *Who, When, Preventative, Recommendations*. One participant asked *"how and when it [the key] was cloned"* (P37). This person thought that the car's computer was hacked to clone the key and that the second click on the key was a security mechanism. Because of this *How* is also closely associated with the moderating factor *Design*.

Globally, the perceived error cause **technical malfunction (*TM*)** resulted in participants being more interested in information about the car and being more specific about the quality of warning they want. In the crash scenario *TM* led to increased demand for the type *Before*: *"The car should have indicated there was something in it's path."* (P8). A positive impression of the car's response to the situation is an outlier in the *crash* scenario. It is positively associated with *Decision*. This is the case because only two participants demanded information about the internal reasoning of the car in the situation and one of them misunderstood the situation at least partially: *"It's ability to recognize a collision and pull to the side of the road is good." (P58)*. In the key scenario *TM* let to increased demand for *Status* information and less demand for *When* and *Troubleshooting*.

Globally, the perceived error cause **human failure (Human)** is not a strongly discriminating factor. It is more closely associated with technical malfunctions. In the *crash* scenario and *NO* condition one participant wanted information about *"[...] changing lanes"* (P42) which led to increased demand for the type *Usage*. *Human failure* is a discriminating factor in the *key* scenario with increased demand of the types *Usage* and *None*. Participants typically wanted to know *"how many clicks [are] necessary to unlock the car"* (P48).

2 participants perceived the malfunction as an intended **Design** feature of the car in the *crash* scenario. They were interested in *"how it [the car] interpreted the situation and any negative reactions to the incident."* (P58). Additionally, they had the impression that the car pondered on what would be the best reaction to the situation *"[...] the autopilot made the best decision of the situation. Staying on the lane was safer than swerving to either lane. [...]"* (P35), and humanizing it *"I would hope that it knows it made a mistake [...]"* (P38). In the *key* scenario the design factor is more closely associated with the perceived error cause *malicious intrusion*. This is because participants perceived the double clicking of the key as an addition security feature: *"In the manual there should be an explanation of why I would have to push the button twice and if there is a trouble shooter for this"* (P59). People demanded more information about *Who, How, Usage, Troubleshooting*.

5.5.2.2 Criticality of the situation

Across both scenarios participants associated **highly critical situations (HiCrit)** more closely with technical malfunctions and malicious intrusions than with human failure or design issues. Highly critical situations did not spark increased information demand with any particular information type in the crash scenario. Participants were broadly interested in all information types they were aware of. In the *Key* scenario, participants associated highly critical situations more strongly with malicious intrusions

than with other perceived error causes. This resulted in an increased information demand of *When, Preventative Measures* and *Action Recommendations*. There was decreased demand for *Status* and *None*.

Globally, situations which participants perceived as **less critical or not critical (LessCrit)** are a strongly discriminating factor for information demand. In those situations participants had increased demand for the information types *Status, Usage* and *None* and decreased demand for *Before*. There is no evidence in the codebook that the *crash* scenario was perceived less critical. This is why LessCrit is not present as a moderating factor in Figure 5.4. In the *Key* scenario less critical situations are more strongly associated with technical malfunctions. There is increased information demand for the types *Status and None*. The demand for the information types *When, Preventative Measures, Troubleshooting,* and *Action Recommendations* decreased.

5.6 Discussion and Implications

First, we discuss the combined results of perceived error causes and situational characteristics to highlight the lessons learned. Based on those we derive actionable implications for design that may form the baseline for a meaningful communication of technical error sources and malicious intrusions.

Our analysis of information demand indicates differences between (1) scenarios perceived as less or highly critical, and (2) between the error causes technical malfunction (TM) and malicious intrusion (MI). While the fundamental need for information remained similar across the conditions and scenarios, participants demanded additional information depending how they perceived the error source. Furthermore, we observe a more differentiated splitting of information codes between TM and MI in the key scenario, which is perceived as less critical, than in the crash scenario, which is perceived as highly critical. A key finding is that while the participants could imagine many different causes of vehicle failures from technical malfunctions, human error, to deliberate design decisions, no one mentioned malicious intrusions. However, we argue that complete threat models cannot be expected from drivers, and that the gaps we have identified must be taken into account to support drivers when they need it most. For this reason, the following discussion pays special attention to the specifics of MI error causes and differentiates them from the needs in terms of TM error causes.

5.6.1 Lessons Learned about Perceived Error Causes

The experimental setup not only allowed an investigation of given error causes, but also revealed which error sources and threats the participants were aware of. We found that, if possible, participants tried to find simple explanations for the presented scenarios. This resulted in an increased attribution to human failure in the NO condition across both scenarios. Additionally, participants commonly named technical malfunctions as potential error sources in all conditions. This may be due to the fact that cars used to cope with mostly technical malfunctions in the past, and cars suffering from malicious intrusions are at the moment still the rare exception. Interestingly, participants tried to explain malfunctions as intended design choices, e.g., to enhance the security of the

vehicle. This was mentioned in the context of clicking twice to unlock the car in the key scenario and could potentially be borrowed from experiences in the online world, such as Two-Factor Authentication. However, this at least indicates fuzzy concepts with regard to enhancing security which is not an uncommon concept. E.g. Distler et al. found that software displaying security mechanisms to its users is better received than equal software that does not [100]. Apart from this, participants demonstrated no sensitivity to security. Unless primed for it, nobody thought of security breaches in the context of car accidents. This indicates the need for guidance in security critical situations and is reflected by the increased demand for the information type *Action Recommendation*. Moreover our results indicate that simply stating the error cause is not enough in highly critical situations. While this is sufficient to place the scenario in the correct context, it is not sufficient to help participants assess and react to the situation correctly. This is true for both *TM* and *MI*.

5.6.2 Lessons Learned about Situational Factors

We found that the more critically the situation is perceived, the greater the need for information but the worse the situation is perceived in terms of trust, satisfaction and intentions to act, regardless of whether the car offers an explanation or not. This coincides with the results of Lim and Dey [218], who also found that people in critical situations have a broad need for information and are difficult to satisfy even if their information demand is met. We believe that in the context of vehicles, however, it is precisely these extreme situations that require special attention, since people need the best possible information, especially in critical, potentially life-threatening scenarios. For less critical situations, on the other hand, it was already possible in this study to satisfy the participants' need for information. The majority of participants was satisfied with the explanation that the battery was empty and needed to be replaced in the key scenario. This was reflected in a neutral to positive impression and a reduced need for information. This scenario also illustrates the key issue our study tackles: how critically a situation is perceived depends on its context. Since the participants do not consider *MI* as a source of error, they are dependent on a classification of the situation. However, the context and thus the need for information changes through the classification. If *MI* is then identified as the error source in the key scenario, the participants classify the scenario as critical, also reflected by dropping satisfaction and operational intent scores. At the same time, they are grateful for the indication of a possible security breach.

5.6.3 Implications for Design

All implications for design are based on the the findings of the Thematic and Correspondence Analysis and have to be validated in future work. We identified a fundamental information demand across all conditions and scenarios. Communicating the error cause (*Why*), explaining *What* happened in the situation and alerting the driver (*Warning/Message*) help to describe the situation and make the driver aware of malfunctions or threats. Depending on the situation and malfunction drivers wanted additional information. In case of *MI*, participants were typically interested in situational and

attack specific information. The following design recommendations can help drivers to understand security critical situations and to act accordingly:

1. **Provide Precise Action Recommendations.** Across both scenarios, the study participants wanted concrete recommendations for action. These could relate, for example, to what should be done next in a concrete situation, such as calling the police. However, it can also be higher level recommendations on how to remedy the security breach. Information types: *Action Recommendation*

2. **Explicitly Communicate Threats.** In the *crash* scenario many of the study participants misjudged the current threat situation. This led to most of the participants simply driving on, which could be potentially life-threatening in this scenario. Therefore it is important to communicate threats realistically and understandably. Information types: *What, Why, Message/Warning*

3. **Communicate Preventative Measures.** Participants across both scenarios demanded information on how to prevent security breaches of this kind in the future. This information not only contributes to a better understanding of the situation, but at the same time educates the driver and makes her sensitive to the subject. Information types: *Preventative Measures*

4. **Provide Information About the Attack Vector.** The majority of participants wanted to know what happened and how this was possible. It is important to communicate the information at a level that the drivers can understand. Here it may be necessary to adapt the information to the level of expertise of the driver or to have her select the degree of detail she wants to know about. Information types: *Who, When, How*

5. **Provide Investigative Cues.** Many participants wished for hints that could help them identify the attacker, e.g., the time and place of the attack or whether someone has already gained access to the vehicle. Information types: *Diagnostic Report, Troubleshooting, Who, When, How*

People in the *TM* and *NO* conditions, however, were interested in different types of information. They typically demanded more information about the car and were more specific about the quality of warnings they expected. The following design recommendations could serve as a baseline to design suitable communication structures for technical malfunctions:

1. **Provide Visual and Auditory Alerts.** The study participants demanded visual and auditory signals, which ideally draw their attention to the defective part before malfunctions occur. This can be, for example, a *beep* sound to attract the driver's attention, or a flash of the key before the battery charge becomes too low. Information types: *Message/ Warning, Visual, Audio, Before/During*

2. **Provide Status Information of Malfunctioning Parts.** Study participants most often inquired about the status of the faulty parts. In the *key malfunction* scenario, they wanted information about the battery status and how long the

charge would last. In the *crash* scenario they demanded information about the state of the front sensor. Information types: *Status*

3. **Provide Diagnostic Report.** Some of the study participants requested a diagnostic report from the vehicle. The report should contain information about the malfunction and the damage report. Additionally the car's parameters during the accident such as velocity can be supplemented. Information types: *Diagnostic Report, Parameters*

4. **Communicate Next Actions.** Similar to the *MI* condition, the study participants wanted actionable recommendations. However, they focus more on what the driver needs to do in order to repair the defect and relate less to the specific circumstances of the situation. Information types: *Action Recommendation*

5. **Explain Car's Internal Decision Process.** Some participants thought the malfunction was due to the autopilot's lacking functionality. They were interested in how the car perceived the situation and what caused it to misbehave. Information types: *Decision Process, Why*

Our results are consistent with the four principles supporting intelligibility and accountability in context-aware systems [44]. Here, the identified information types of our study complement the principles with details for the domain of semi-autonomous vehicles, with malicious intrusions identified as special cases. As the work of Jacobi et al. suggests [176], information needs can change over time. This certainly needs to be considered for the domain of cars in general. However, we argue that (highly) critical situations are a special case because they are rarely experienced. In order to establish a practical relevance, we would like distinguish ourselves from the NIST Cybersecurity Framework [267]. It focuses primarily on building and maintaining critical infrastructure and thus provides a set of activities to achieve specific cybersecurity outcomes. However, it is not tailored to achieve good computer-human communication.

5.7 Limitations

Our study has several limitations, some of which originate from the study design, while others are intrinsic to the measures we use to gather our data. First, we drew our sample from MTurk. Hence, it is not representative of the population of American car owners, since MTurk users are usually between 18 and 48 years old and have some level of college education [298]. However, as we were particularly interested in the specifics of malicious intrusions, MTurk serves a suitable population for our purpose [298]. Second, we confront participant with hypothetical scenarios. This means that the participants have to imagine experiencing the described situation. Although we tried to make the scenarios as tangible as possible, they cannot offer the same quality as a personal experience. Due to the hypothetical character of the study, participants are relieved from any driving-related tasks. Additionally, at the point in time of the scenarios, when we asked for the driver's information demand, the car is not moving (anymore). Hence, the "drivers" - in that sense - are no longer drivers as they are not constrained by

typical driving tasks when we elicited information demands. Although participants reported demand for visual and auditory alerts in response to a detected malfunction while driving, we acknowledge this general limitation to our scope of work. Since only 13 participants reported having previous experiences in driving or riding cars with autonomous driving features, this lack of experience among participants likely biased our results as well. Nevertheless, there is evidence that hypothetical surveys are able to identify tendencies [65], which can be verified in future work. Yet, we do not claim our results to be exhaustive, especially with regard to the information types we elicited.

Third, as our study follows a mixed-methods approach, we have to deal with a sample size trade-off. This means that it might be necessary to recruit more participants in order to detect small effects; however, this conflicts with qualitative data analysis. We did power calculations to estimate our sample size for an estimated medium effect between scenarios, resulting in $N = 60$. Although we usually reached saturation within the first round of coding, $N = 60$ was still manageable in terms of qualitative evaluation.

5.8 Conclusion

We identified 18 information types ranging from situational aspects to question types over car and warning specifics. Some of these information types form a basic need for information across scenarios. Depending on the perceived error cause, people may demand more situational or car specific information. The findings could be used to display relevant, sought-after information in appropriate contexts and inform design decisions for human-car interaction.

Moreover, we found that malicious intrusions were consistently perceived as critical, even if other perceived error sources in the same scenario were not. Critical situations sparked increased information demand, while at the same time making it hard to satisfy people. However we believe that particularly these situations require our attention, since people need the best possible information, especially in critical, potentially life-threatening scenarios.

Last but not least, we found the need to properly communicate error sources. Participants did not identify malicious intrusions in any scenario, unless being primed for it. If primed, they were not able to assess the situation correctly and act accordingly. In case of technical malfunction a similar effect surfaced in the crash scenario were people were unsure of the Autopilots capabilities and reasoning. We also found that simply prompting the error source is not sufficient in highly critical situations as this solely places the malfunction in the correct context leaving many open questions. We argue that complete threat models cannot be expected from drivers, and that the gaps we have identified must be taken into account to support drivers when they need it most.

Our study is an important first step to improve in-car risk communication to drivers and to provide helpful information at appropriate times. Depending on the situation this may build drivers' confidence and trust in the safety and security of the car, while it also may improve their decision-making capabilities in critical situations. Future work could test and validate our results, e.g. in providing drivers with relevant information and measuring how they assess and react to different situations. If feasible, it would

be beneficial to test our results in a more realistic setting, though it is not ethical to have participants experience highly critical situations first hand. More realistic study set-ups could also investigate drivers' information demand while being constrained with driving-related tasks, which we did not cover. Last, but not least our insights could be used to develop or enhance interface solutions for cars to meaningfully communicate error sources in the future.

6

"I chose to fight, be brave, and to deal with it": Threat Experiences and Security Practices of Pakistani Content Creators

The contents of this chapter were published as part of the publication *"I chose to fight, be brave, and to deal with it": Threat Experiences and Security Practices of Pakistani Content Creators"* (USENIX Security 24) [P4]. This chapter uses the academic "we" to highlight the contributions of my co-authors, Waleed Arshad, Shanza, Angelica Goetzen, Elissa M. Redmiles, Maryam Mustafa, Katharina Krombholz, and me. The following table details the contributions of each author to this paper:

Author	**Contribution**
Lea Gröber	I trained Waleed and Shanza in qualitative methods. We worked together to pre-test and refine my interview guideline. I oversaw the recruitment and interview phase. Waleed, Shanza, Angelica, and I evaluated the qualitative data. I wrote the majority of the paper.
Waleed Arshad	Waleed contributed to pre-testing and refining the interview guideline. He recruited participants, conducted interviews, and evaluated parts of the qualitative data. Waleed contributed content to the results section.
Shanza	Shanza contributed to pre-testing and refining the interview guideline. She recruited participants, conducted interviews, and evaluated parts of the qualitative data. Shanza contributed content to the background section.
Angelica Goetzen	Angelica evaluated parts of the qualitative data. She contributed to the related work section.
Elissa M Redmiles	Elissa had the initial idea to study the domain of content creation. She advised, contributed to the framing, and reviewed the final paper before submission.
Maryam Mustafa	Maryam contributed to the interview guideline. She advised, contributed to the framing and background section, and reviewed the final paper before submission.
Katharina Krombholz	As my academic advisor, Katharina guided key project decisions and provided feedback on my initial ideas and methods. She contributed to the framing and background section and reviewed the final paper before submission.

Reference

Gröber, Lea and Arshad, Waleed and Shanza and Goetzen, Angelica and Redmiles, Elissa M and Mustafa, Maryam and Krombholz, Katharina. (2024). "I chose to fight, be brave, and to deal with it": Threat Experiences and Security Practices of Pakistani Content Creators. 33rd USENIX Security Symposium, 19–36.

Content creators are exposed to elevated risks compared to the general Internet user. This study explores the threat landscape that creators in Pakistan are exposed to, how they protect themselves, and which support structures they rely on. We conducted a semi-structured interview study with 23 creators from diverse backgrounds who create content on various topics. Our data suggests that online threats frequently spill over into the offline world, especially for gender minorities. Creating content on sensitive topics like politics, religion, and human rights is associated with elevated risks. We find that defensive mechanisms and external support structures are non-existent, lacking, or inadequately adjusted to the socio-cultural context of Pakistan.

Disclaimer: This chapter contains quotes describing harmful experiences relating to sexual and physical assault, eating disorders, and extreme threats of violence.

6.1 Introduction

Online content creators express themselves, reach broad audiences, raise awareness, or build careers [215, 60] using services such as TikTok, Instagram, and YouTube. They cater to large audiences, share a sizable volume of private data and are consequently exposed to elevated risks [345]. Prior research explored the threats to US-based content creators [345, 318] and the ways in which sensitive topics – such as sex work – affect the security and privacy of creators [241, 149]. While it is estimated that approximately an equal number of men and women engage in content creation work [195], a survey by Thomas et al. [345] found that U.S. content creators who identify as women are more likely to experience sexual harassment, excessive negative reviews, stalking, spreading of rumors, and surveillance than those who identify as men. Furthermore, outside of content creation, a growing body of work establishes how gender impacts the digital security and privacy experiences [68, 315, 314, 313, 375]. In the global south [363, 16, 315, 314], women are structurally disadvantaged and face unique threats, often tied to their socioeconomic status and literacy [315, 16, 261].

In this paper, we present a study with 23 semi-structured interviews investigating the intersectional marginalization for content creators across genders in Pakistan. Pakistan is a particularly interesting country in this regard: It ranks second to last in terms of gender parity [119]. Due to cultural and religious factors, the access to the regular labor markets is limited for women [42] as is their presence in public spaces in comparison to men [255]. To navigate these constraints many women choose to migrate to online spaces for forming social connections and for work and business opportunities [382]. Prior work reveals the vital importance of online communities in countries like Pakistan for women to discuss taboo narratives, find work, explore identities and form connections [382, 26].

In contrast to general social media users, content creators are especially vulnerable and constitute a special case in two ways: (1) They cater to a large audience, and depending on the type of content they produce can (be perceived to) threaten social norms and structures.[1] (2) They *have to* stay online and maintain public profiles and personas to keep their business/activism going and thus cannot rely on affordances available to other users who have private profiles and are not dependent on making

[1] Blasphemy laws in Pakistan carry severe penalties up to death [236].

money as content creators.

Anecdotal evidence reports instances of women content creators in Pakistan facing severe harassment [53], group assault [87], or getting killed [327, 228]. The prominent case of Qandeel Baloch demonstrates how content creation can be a powerful way for women from low socioeconomic backgrounds in Pakistan to emancipate themselves, but who also face lethal consequences when family, friends or extended relatives start to feel threatened or dishonoured by their presence in online spaces, particularly when their private data (identity) is leaked [228]. Additionally, content creators - both men and women - who engage with sensitive topics such as religion, sexuality or politics face harassment [53, 174], may be forced to leave the country temporarily [3], and have even been killed for their work [177]. These stories showcase the elevated risk to which Pakistani content creators are exposed and underscore the critical need for research into how to best protect these marginalized content creators.

Our qualitative study investigates the security context of digital content creation in Pakistan including the threat landscape, how creators navigate concrete negative experiences, defensive mechanisms, and what support structures they rely on. Accordingly, we defined the following research questions:

RQ1: What does the threat landscape for content creators in Pakistan look like?

RQ2: What defensive mechanisms do Pakistani content creators implement to stay safe and secure? Which resources do they rely on?

RQ3: What gaps are present in existing security and privacy measures? Which interventions would be needed that are specific to the Pakistani social media ecosystem?

We find that the online and offline threat landscape is tightly connected, with online threats frequently manifesting in offline harm. The degree of both online and offline threats are severe and in the socio-political context of Pakistan can become lethal, especially if content revolves around religion, sexuality, or politics. Online threats are accounted for with a mixture of *technical* and *behavioral defenses*. While defenses exist for prominent threat categories like *toxic content*, participants are dissatisfied with options to deplatform attackers. However, certain threat categories such as *impersonation* lack any defensive mechanisms, while exposing victims to severe threats. Our findings inform the further development of inclusive safety tooling for social media platforms tailored to different populations and risks that are specific to those populations. Moreover, our findings contribute to the improvement of information sources for at-risk populations.

6.2 Background and Related Work

6.2.1 Sociocultural Background of Pakistan

Pakistan is a particularly interesting country to study the challenges of publicly exposed figures such as content creators and frontiers of security and privacy measures given its complex patriarchal and religious landscape. In the recent 2022 gender gap report by the World Economic Forum Pakistan places second to last (Afghanistan is last) in terms of gender parity [119]. In Pakistan, offline and online privacy behaviors are heavily influenced by religious and conservative values and patriarchal norms. Notions

of community, family honor, and piety as well as Islamic values heavily influence legal, political, and social norms[148]. Pakistan is a culturally and linguistically diverse country with significant region-specific differences within the country. It is also deeply class-based where wealth is not equally distributed, and the reality of high-income citizens is very different from those of low-income citizens [4]. There exists an educational divide between rural and urban citizens, and the majority of women have no formal education [276]. Mobile phones are equally available to both rural and urban households, however, rural households are three times less likely to have access to a computer or internet [276]. Islam is the state religion with approximately 95-98% of the population identifying as Muslim. The Islamic principles of *Purdah* and gender segregation often bleed into digital spaces as well. *Purdah* is broadly defined as the segregation of genders, and involves both modesty of the heart and the eye [153]. The practice of Islam in Pakistan values modesty, the segregation of genders and the covering of women's bodies [153]. These values also impact women's access to formal labor markets, public spaces, and digital spaces [229]. Prior work highlights the ways in which women in Pakistan navigate constraints on mobility, social networks, access to labor markets, and social support by leveraging digital spaces [382, 258]. Online spaces are a particularly critical pathway for women to access vital resources and gain financial independence. However, prior work also highlights the ways in which these spaces fail women in Pakistan [261].

Since 2013, the Pakistani national identity cards have a third gender category [277], and since 2018, Pakistan's Transgender Persons (Protection of Rights) Act theoretically strengthened the rights of transgender people. Although the transgender community – locally referred to as the *Khwaja Sira* – are officially recognized, they continue to face severe discrimination, are often excluded from the conventional job market, and are seen as beggars in public [266]. Despite legal recognition and a Supreme Court ruling allowing transgender people to be identified as a third gender, systemic social support is still lacking. Prominent representatives of the *Khwaja Sira* community argue that the Western LGBTQ+ acronym does not adequately capture their unique experiences, and they emphasize the separation of gender identity and sexual orientation [189]. In Pakistan, there are two main social attitudes towards the *Khwaja Sira*: conservative groups marginalize them for not conforming to a binary gender, while the liberal population often understands trans rights but through Western ideals [24].

6.2.2 Security and Privacy of Content Creators

In 2022, Thomas et al. [345] conducted a survey with 135 U.S. content creators from different platforms (esp. YouTube, Instagram, and TikTok) with a focus on quantifying the extent of negative experiences, reactions to attacks, and perceived gaps in protective solutions provided by platforms. They found that content creators in the U.S. are structurally exposed to risk, with one in three being the target of attacks and 95% experiencing hate or harassment at least once. In response, many content creators chose to ignore attacks, while others engaged in self-censorship, or leaving platforms altogether. A recent interview study by Samermit et al. [318] explored threats relevant to U.S. creators, and their protective practices. They found diverse threat models apply for creators, and that defenses are often adopted only after negative experiences. While these

works offer insight into experiences of content creators in the US, a full understanding of content creator safety and security requires an understanding of the experiences of non-WEIRD (Western, Educated, Industrialized, Rich, and Democratic [162]) creators. Centering the most marginalized users allows us to ensure that we fully capture the risks faced by content creators in designing interventions; as McDonald et al. [242] explain, "often, the privacy risks of vulnerable populations are not fully considered in the design of systems because those risks and potential harms are not fully understood (nor necessarily prioritized) by those responsible for research and design." For instance, recent work has begun to explore the unique safety and security experiences of digital sex workers. These content creators exist at the intersection between digital content creation and sex work; they utilize social media to create monetized adult content as well as connect with other creators and engage in community support [149, 150, 108, 357]. In addition to negative experiences faced by content creators at large [345], digital sex workers are also subject to social stigma and strict content moderation policies [40], producing additional negative outcomes like being de-platformed from social media sites [149, 36]. In general, algorithmic fairness is an issue for marginalized creators [72, 193, 95]. For instance, creators with disabilities face challenges with demonetization [193] and transfeminine TikTok creators need to navigate visibility traps [95].

6.3 Methodology

We conducted a qualitative semi-structured interview study ($n = 23$) to explore content creators' security and privacy perceptions, needs, and practices in the socio-cultural context of Pakistan. The following sections detail our recruitment, interview procedure, pilot tests, and analysis.

6.3.1 Recruitment and Participants

Our target population are content creators matching the following criteria: (1) living and working in Pakistan, (2) generating some form of income from their content (e.g., monetary or through compensation with physical goods). We recruited Pakistani content creators across platforms (YouTube, Instagram, TikTok, Facebook, Twitter, Snapchat, Spotify) and targeted a diverse sample across several dimensions (income level, literacy, gender, follower count, content type) to ensure we captured a range of relevant experiences. For recruitment, we manually compiled short lists of creators of different genders based on profile data[2] ensuring broad coverage of the above demographics. Recruitment followed an exploratory, iterative approach. We used the results of the ongoing qualitative analysis to expand the shortlists with candidates that could bring new perspectives to the study. For example, we found that content creators frequently engage in self-censorship to protect themselves. To complement our perspective, we then specifically targeted people who post content on sensitive topics that other content creators avoid, such as gender and politics, but also people who only post seemingly uncritical content, such as food or pets. Moreover, after each interview, we asked the

[2]In Pakistan, few people openly disclose their gender online. We validated our assumption with participants' self-reported gender data.

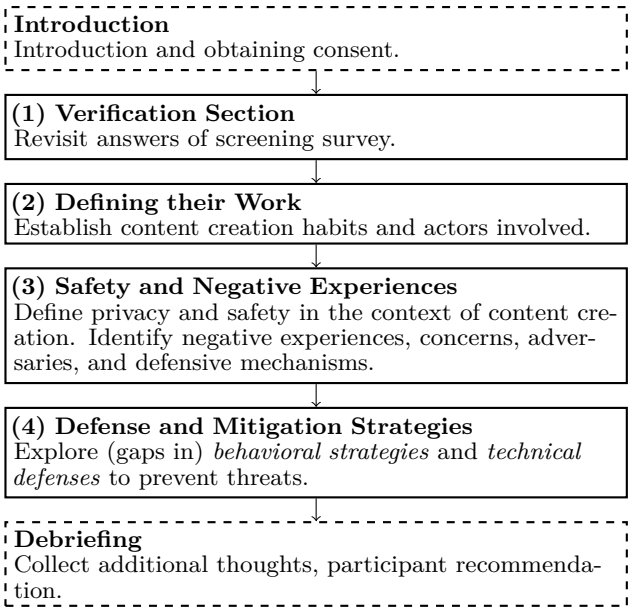


Figure 6.1: Overview of the semi-structured interview guideline.

participant if they knew someone whom we should interview for this study. We reached out through content creators' designated contact channels and asked them to fill out a three-minute screening survey to ensure they fit our criteria. The screening survey collected personal information (like age, location information, gender, disabilities, education, and income), and platform information (where they create content, compensation for content creation, description of content topic, posting habit, and follower count). The positive response rates per shortlist were: 21.4% *woman*, 30% *man*, 100% *Khwaja Sira*.

6.3.2 Interviews

Two interviewers conducted 23 one-hour semi-structured interviews in Pakistan from December 2022 to May 2023. This resulted in a total of 1261 minutes, with a median of 56 minutes. We compensated each participant with PKR 5000 (USD 19.02), resulting in a median hourly wage of PKR 5258 or USD 20.37. To comply with socio-cultural norms, the interviewers were a man and a woman Pakistani co-author who respectively spoke to the men and women participants. Khwaja Sira chose their interviewer. This practice was intended to create a room where participants could feel safe and talk openly. We conducted interviews in English and Urdu, depending on the participants' preferences. Interviews took place in person or online via Zoom.

We developed the interview guideline with our exploratory research questions in mind. We also took the results of previous studies into account [345], building on them while exploring the topic broadly through semi-structured interviews. Figure 6.1 provides an overview of the interview flow and contents. For the full interview guideline refer to the Appendix Section D.2.

We describe the interview sections below and provide the interview guideline in

the supplementary material. Each interview started with a fixed introduction that established the purpose and procedure of the interview and allowed the participant to ask organizational questions. Afterward, we obtained the participant's consent to record the interview and proceeded with the interview guideline. The interview was structured as follows:

Verification Section: We revisit questions from the screening survey about the scope of content creation. This serves both as opening questions to get participants talking and to re-verify that creators fall within our target population and are, indeed, the creators we reached out to. No participants were screened out.

Defining Their Work: This section establishes our participants' context and content creation habits. In the beginning, participants define their job title (e.g., *content creator* vs. *influencer*) by which we address them throughout the interview. Further, they describe posting habits and the type of content they create. We provide an overview of participants' background information in Section 6.4.1.

Negative Experiences: We ask participants to describe their personal definition of safety and privacy in relation to content creation. Then, we ask about their biggest digital and physical security and safety concerns. To make the conversation more tangible, we focus on concrete negative experiences that participants encountered in connection to their content creation work. Moreover, we ask who they identified as potential adversaries, and how they protected themselves in these scenarios. Based on this, we categorize the negative experience of content creators in Pakistan in Section 6.4.2.

Defense and Mitigation Strategies: We explore how participants aim to prevent these threats from occurring via *behavioral* (e.g., adjustment of content types, or leaving platforms) and *technical* defenses (e.g., platform features, or authentication schemes). Then, we explore any additional support structures participants relied on to cope with negative experiences. Finally, we explore (gaps in) participants' socio-technical approaches to mitigating harm from the negative experiences they have endured (see Section 6.4.3).

6.3.3 Pilot Testing

To pilot test the interview guideline, one researcher took on the persona of a Pakistani content creator and we conducted two pilot tests, one with each interviewer. We tested the flow of the interview guideline in terms of the overall structure and order of questions. In response, we restructured minor parts of the interview guideline to remove redundancies. As the study progressed, we further developed the interview guideline by incorporating participant feedback. Changes included minor rewording to make questions more open-ended and changes to the question order. No substantial changes were made to the interview guideline.

6.3.4 Analysis

We transcribed the interview recordings manually and using a GDPR-compliant transcription service (Amberscript). Urdu parts were first transcribed and then translated into English by the researchers of our team. Four researchers were involved in the coding process: the two co-authors from Pakistan who conducted the interviews and two co-authors with German and U.S. backgrounds, respectively. Three of them are computer scientists; one is a social scientist. We conducted a thematic analysis where data collection (interviewing) and analysis was an iterative process. Our goal was to collect a set of threat patterns which are associated with in-depth data about individual experiences. Researchers used a combination of "top-down" qualitative content analysis [239, 360, 204, 308, 325] (informed by previous frameworks [344]) and "bottom-up" analysis inspired by "open coding" [66, 332, 224] allowing for emerging themes. Three researchers jointly established a first codebook taking into account a male and female interview. We followed an iterative approach where at least two researchers coded one interview independently. Then they discussed and resolved all disagreements and updated the codebook accordingly. As we coded, we wrote summaries and memos to collect and systematize potential themes. Once coding was complete, we jointly discussed themes by revisiting memos and grouping codes in the codebook. Thereby we identified five axial categories (negative experiences, concerns, attackers, defense practices, support). We reached stability in threat patterns after having interviewed 16 men and women. We continued interviewing but focused recruitment efforts on Khwaja Sira creators, to explore if (individuals of) this gender experienced additional threat patterns (because we found that gender is likely an influencing factor on negative experiences). However, we did not find additional threat patterns in the last six interviews (two Khwaja Sira and four men/women), although we learned about extreme instances of already known threats. Therefore, we closed data collection and concluded reaching thematic saturation [160, 161] with respect to threat patterns after 16 interviews. The final codebook contains 495 codes.

6.3.5 Limitations

Interview studies are limited by self-reported data, which may lead to under- or over-reporting. To address this, we designed the interview guideline to focus on specific experiences and provided prompts to aid memory recall. Participants sometimes felt uncomfortable reporting negative incidents, and we respected their decision. Therefore, we acknowledge that our findings may not fully capture extreme negative experiences. To mitigate social desirability bias, we reassured participants that we were interested in their experiences as content creators and would not judge their actions or responses to threats. Given the strict social gender norms in Pakistan (Section 6.2.1), we conducted interviews in a same-gender setting (selected gender for Khwaja Sira) to create a safe space for discussing sensitive topics. Our convenience sample does not necessarily represent the wider population of Pakistani content creators. While qualitative studies like ours do not strictly require representation, we made deliberate efforts to recruit diverse participants. We defined recruitment criteria based on previous research and anecdotal evidence to identify potential high-risk populations [277, 148], resulting in

Table 6.1: Platforms where participants create content.

Online Community	N	%
Instagram	23	100%
TikTok	15	65%
YouTube	14	61%
Facebook	6	26%
Twitter	2	9%
Snapchat	1	4%
Spotify	1	4%

a diverse sample that includes Khwaja Sira creators. However, future work should quantitatively validate our findings. Moreover, our sample does not include participants who decided to withdraw from creating content, or who faced lethal consequences.

Ethical Considerations We obtained approval for this study by Saarland University's ethical review board. This research touches upon critical negative experiences, thus potentially bringing up past trauma. To obtain informed consent, we thoroughly explained the process of the study, including how we recorded, anonymized, and stored the collected data according to GDPR. We paid special attention to anonymizing participant quotes and present only aggregated demographical information to avoid potential harm due to de-anonymization.

6.4 Results

Findings are illustrated with participants quotes (Gx); x denotes the participant id within the gender group G (**M**an, **W**oman, **K**hawaja **S**ira).

6.4.1 Sample Descriptives

We recruited 12 women, 9 men, and 2 Khwaja Sira (see Section 6.2.1). Participants are young with 83% ages 18-24 and 17% ages 25-34; and educated with highest degree high school (48%) or University (48%). One person was still pursuing a high school degree. 56% of interviews contained answers in both English and Urdu, 26% were solely conducted in English, and 17% in Urdu alone. 48% of participants have 10k - 50k total followers across platforms, while the biggest creator has between 300k - 400k followers. Everyone earned some form of income from content creation, but only 36% were comfortable sharing annual income details. 48% of participants reported earning upwards of 100k PKR annually. Table 6.2 reports demographics; Table 6.1 presents a breakdown of the platforms participants create content on. All participants use Instagram, followed by TikTok (65%), and Youtube (61%). Participants post various content topics: from fashion and lifestyle to awareness of socio-political issues. Some participants posted multiple types of content which fall into a common broader category, e.g., fashion, lifestyle, etc. Most commonly participants create lifestyle content

Table 6.2: The demographic of the 23 creators of this study.

Demographic	Group	N	%
Gender	Woman	12	52%
	Man	9	39%
	Non-Binary	0	0%
	Khwaja Sira	2	9%
	Prefer not to say	0	0%
Age	18-24	19	83%
	25-34	4	17%
Income from Creation in PKR	<50k	1	4%
	50k - 100k	1	4%
	100k - 150k	2	9%
	150k -200k	2	9%
	>200k	7	30%
	Prefer not to say	10	44%
Education	Below high school	1	4%
	High school	11	48%
	University	11	48%
Total Followers	10k - 50k	11	48%
	50k - 100k	3	13%
	100k - 200k	5	22%
	200k - 300k	3	13%
	>300k	1	4%

(39%), followed by fashion (35%), comedy (17%), and music, life stories, vlogs, and socio-political and human rights issues (13% each). It is important to note that socio-political issues include topics on gender, sexuality, and religious minorities, politics, and human rights include topics like women and trans rights. Two of our participants posted dance videos. The remaining categories include makeup, food, mentoring, photography, pets/animals, family content, and documentaries (4.3% each).

6.4.2 Threat Landscape and Negative Experiences

We structured the threat landscape of Pakistani creators following Thomas et al.'s taxonomy of online hate and harassment attacks [344] and expanded it towards offline threats. Threats were identified based on the axial categories *negative experiences*, *concerns*, and *attackers*. Figure 6.2 provides an overview of online and offline threat categories, including associated attacks that emerged in the interviews. Throughout our results, we compare our findings to prior work on U.S. creators [318, 345] and the general Pakistani population [32], however, we do not provide comparative statements for all findings due to differences in analysis lense.

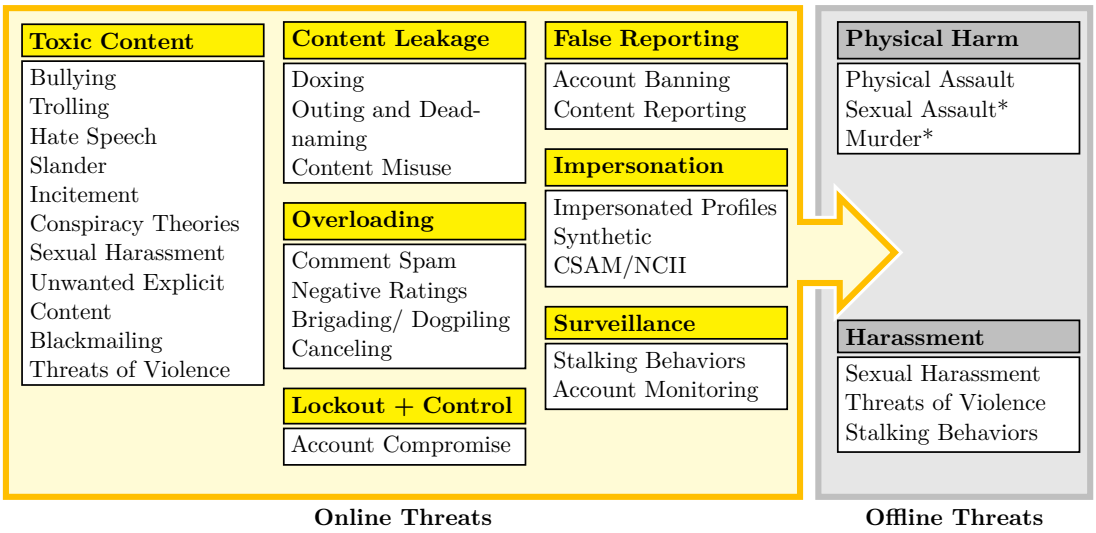


Figure 6.2: Spillover effect from online (yellow) to offline threats (gray) for content creation in Pakistan, grouped by categories. A * denotes threats explicitly mentioned (but not experienced); all other threats were experienced by at least one participant.

6.4.2.1 Toxic Content

This category includes online attacks that are intended to be seen by the creator, often with the goal to intimidate, harass, influence, or silence them. In line with our work, research on U.S. creators found these attacks impact emotional safety [318].

Bullying and Trolling. The boundaries between *bullying* and *trolling* are fluid, and attacks are prevalent across genders [w:11 | m:6 | k:2]. Women report instances of body (W7, W8, W11) and skin tone (W1) shaming, with one participant developing an eating disorder as an outcome (W7). Similarly, men are also shamed for their bodies, such as for not growing facial hair (M3). Participants are also targeted if they are in positions of power; for instance, W9 works as a civil servant and experienced bullying and harassment due to attackers' belief that she does not deserve a good life while the local population lives in poverty. In general, *marginalization* is a risk factor [370] that applies to U.S. creators, too. Attackers focus on identity characteristics such as gender or religion, with intersectional identities (e.g., woman, overweight) compounding risk [318].

Hate Speech. In our sample, all genders report experiencing *hate speech*, although it was most prevalent among woman and Khwaja Sira [w:10 | m:3 | k:2]. Extreme cases of hate speech and harassment are associated with specific content topics, such as religion (W9), politics (K1), and discussion of social issues such as rights for women and/or transgender people (W9, K1, K2). As W9 puts it, activism can be dangerous, as people start targeting you based on the minorities you speak up for: *"When I talk about Ahmadis [religious sect] or religious minorities in general, people would just tag me an Ahmadi. [...] If I talk about women, they would just spew slurs at me just*

randomly. They would curse me, saying I'm a feminist. It's a curse word itself. March is just upon us and Women's Day is here. The whole cycle just starts every single year. The other kind of harassment that I face is when I talk about transgenders. They would just say that I am promoting gay marriages just because I'm talking about the protection of transgenders" (W9). Outside of religious, political, and social issues-centric content, even pictures with friends in a mixed-gendered group can lead to harassment: *"I remember posting a group picture with my friends and they were, five or six girls and there were two guys in the picture [...]. For some reason, I received a lot of backlash on that post. 'Why are you promoting Western ideals in Pakistan? [...] She's Western propaganda.' Even though it was literally just a couple of friends standing together and there was nothing problematic about it"* (W3).

Slander, Incitement, and Conspiracy Theories. *Slander* is a distinct type of toxic content meant to harm the reputation of creators. It is particularly dangerous in Pakistan, as people – especially women – need to comply with strict cultural norms of *pardah/modesty* and *honor* [355]. Among our participants, all genders [w:3 | m:3 | k:2] fell victim to the spread of disinformation – not only those who produce content associated with heightened risk (e.g., religion, politics, social issues), but also those of non-sensitive content types. For instance, W12 owns a pet account, and a competing account that did not grow as quickly as hers spread disinformation that she abused her animals. Slander may lead to *incitement* [w:2 | m:3 | k:2], either when creators decide to remain associated with someone who got canceled (W2), or if their content is controversial or sensitive (K1). *Incitement* can easily lead to online (W2, W9, M4, M5) and offline (K1) harassment, for instance K1 reports of online articles speculating about her gender that led to religious sects showing up at her home (K1). The boundaries between *slander* and *conspiracy theories* are fluid, and the potential for harm is great if the creators reputation is damaged. We identified four instances of *conspiracy theories* [w:1 | m:1 | k:2] that were all trying to explain creators behavior with hidden political agendas: *"If these individuals are sitting and saying that I'm an American agent working on a conspiracy to promote LGBT rights in Pakistan"* (K1).

Sexual Harassment. Among our participants, only women and Khwaja Sira reported experiencing online sexual harassment [w:2 | m:0 | k:2]. One prominent experience involved religious figures leaving sexualized comments on the (non-sexual) picture of a woman creator. Being targeted by religious figures can escalate the incident into the religious realm where the victim becomes a target for a broader audience that often is willing to resort to offline violence [228].

Unwanted Explicit Content. The majority of women in our sample receive unsolicited messages [w:9 | m:2 | k:1], often containing explicit content. In one incident, a friend sent the unwanted content (M2), but usually attackers are unknown as the messages are sent from throw-away accounts: *"This guy was DMing me really sexual stuff, pictures, unsolicited pictures of himself [...]. I blocked him from so many different accounts and he would just keep making new accounts to do that"* (W7).

Blackmailing. Few participants report being *blackmailed* with the majority identifying as men [w:1 | m:3 | k:1]. Attackers can come from within their inner circle; for example, former friends threatened to leak a participant's identity (W11). Other attackers are anonymous and threaten participants for attention or personal gain: *"The most extreme level of it is when someone reaches out to you and basically asks you for [...] a shout out, [...] a reply, it can be anything like that and if you don't do it, then they threaten people that you love"* (M9).

Threats of Violence. Some participants report receiving *threats of violence* [w:2 | m:2 | k:2] through comments and direct messages (M2). Attackers are usually anonymous, although participants sometimes link them to known opponents of them (e.g., influential industrialists (K1)). Threats can also be a result of *hate* and *canceling* campaigns (M4). The goal is to intimidate or silence the creator, and may target the creator's family and close ones as well (M4). Participants report that content on topics like religion and gender lead to more extreme (sexualized) threats of violence (K1,K2, M2, W9). In general, creators think that *"mostly girls [...] receive perverted comments"*(M2). In our data set only woman and Khwaja Sira received comments regarding sexualized violence: "*Whenever I talk about religion, these kinds of things happen. I believe that [...] the concept of religion that has been propagated by different Madrasas is the ultraconservative version of a religion that I'm talking about. They do not see women or random creators talking about religion. They believe that religion is something that only they can talk about. Just because a person [...] is in westernized attire, [they are] not Muslim enough to talk about the religion. That's when I got some rape threats"* (W9). Participants did not report any threats of violence that were realized offline.

6.4.2.2 Content Leakage

Content leakage summarizes attacks in which private data or content is leaked by a third party. This was done to threaten, embarrass, or shame our participants.

Doxing. Most woman creators were concerned about or experienced *doxing*, where personal information is exposed to a greater audience [w:5 | m:2 | k:2]. Attackers can come from the inner circle (W11), or engage in social engineering to obtain personal information such as phone numbers (W4). Through *doxing* creators and their loved ones become targets for offline attacks (e.g., *stalking behaviors*, *physical harm*), but also because people surrounding them might feel threatened or dishonoured. In the context of Pakistan, leakage of personal information such as family details are therefore severe threats, and deanonymization has led to the murder of a woman creator in the past [228]. Brands can also be a threat as they obtain participants' information to send PR packages; one participant had a brand give out her home address to a fan who was a relative of the brand's owner (W7).

Deadnaming. Another form of leaking personal information and disrespecting the victim is deadnaming a Khwaja Sira, which was reported in one instance: *"Dead naming me is a line for me [...]. When people [call me by my old name] to belittle me that I*

think is a line that I do not like being passed because <old name> is a very personal name." (K1). Here, *deadnaming* was used to harass the creator, however this practice reveals sensitive gender information that might not be intended for a broader audience. In case of K1, gender-identity is part of her content.

Content Misuse. We expanded this category to also include unauthorized content use, which was a prominent concern across participants [w:5 | m:3 | k:1]. It is different from content leakage, as in this case the content was previously posted on a public profile. Unauthorized content use by third parties primarily leads to financial harm [149], but may also be used for *impersonation* that in some cases can result in *physical harm.* In Pakistan this is a severe threat to woman, as it damages their reputation: *"Being a Pakistani girl with a public account, there is always a little thing in your heart that anything can happen anytime, your pictures can be edited and transformed and a lot of other things can happen so social media is one of the most unsafe places, which we all know but still we use it"* (W1).

6.4.2.3 Overloading

Overloading refers to online attacks or interactions that overwhelm the creator by spamming communication channels (e.g., comments or notifications), usually with the aim to silence or influence them. Non-malicious intentions, such as getting the creator's attention or showing fan affection were also reported (W12), but can equally be perceived as harassment. Attackers are groups of people, often entire communities, that coordinate to target a victim.

Comment Spam and Negative Ratings. Especially when it comes to religious content (M5), our participants face excessive negative engagement like hate comments and messages [w:7 | m:3 | k:1]. Comment spam can also be a way for people to get the attention of creators. Complementary, W12 describes out of the norm positive attention from fans as a form of harassment.

Canceling, Brigading, and Dogpiling. These attacks [w:3 | m:2 | k:1] are similar because they are all carried out by coordinated online communities, e.g., through comment spams. The goal is what makes them different: *Canceling* aims to silence and de-plotform creators (W2, M4, M7), *brigading* tries to disrupt the discourse (W2, W8, W9), and *dogpiling* aims to get victims to recant their views (K2).

6.4.2.4 Lockout and Control

These attacks aim to silence creators or de-platform them alltoghether by breaking into or leveraging privileged access to a targets' accounts or devices [345]. In our sample, attackers were never identified, but third-party reports highlight the risk of attackers from the inner circle of the victim [315, 238].

Account compromise. The threat of getting hacked is known to the majority of participants, and they deploy Two-Factor-Authentication (2FA) as a defense [w:12 | m:7 | k:2]. However, the concepts of attack vectors and recovery seem to be obscure: *"I couldn't access my account and I don't know why that happened"* (W12). M7 told us about the account deletion of a fellow creator, allegedly because the attacker did not like her content: *"Her account got deleted recently. And the kid who hacked it said that because he didn't like the content [...] he deleted it. If it's so simple for a kid like that to do something, I would say Instagram is not really a safe account. And you can't even appeal to anyone to help"* (M7). In our sample, three creators (W12, M9, K2) have fallen victim to hacks that resulted in account lockouts. None named details on who the attacker was or how their account got compromised. All were able to recover their account with the help of the platform's support system and external digital rights organization. Only after the hacks and on the platform's request, two of them switched to 2FA. Participants also expressed their concerns of account or content deletion, although nobody in our sample experienced it firsthand.

6.4.2.5 False Reporting

A platform's reporting function is intended to protect individuals and the community at large. However, due to its semi-automated handling, reporting can be misused to silence, de-platform, and financially harm creators or, more generally, demonstrate power. Attackers usually stay anonymous. We extended the taxonomy [344] to fit creators.

Account Banning and Content Reporting Participants were as concerned about attacks in this scenario as they were about hacking attempts. However, they were dissatisfied with the lack of defensive options. One participant experienced account loss through banning twice on TikTok (W8) but was able to recover from it. Another participant reported a similar incident of a fellow creator: *"Some random person [...] reported her account and got her banned from Instagram. There's nothing that she could do to get her account back. She had around 20,000 followers. The person who reported her account also told her, 'I don't agree with your content, and I think that you're annoying [so I took your platform], and I want you to apologize to me, then I'll give it back to you.' "* (W3). Similarly, some participants had their content wrongly taken down from platforms [w:1 | m:1 | k:0]. Attackers coordinated malicious reporting attacks to silence creators: *"Even though my videos did not have any sort of explicit content [dance], due to people reporting it, a lot of my videos got taken down. They still do to this day"* (W8). Generally, participants wished for transparency on how banning works from the platforms (W1, W3, W12).

6.4.2.6 Impersonation

This category involves using, altering, or artificially recreating content to impersonate a creator online. These kinds of attacks can do severe reputational damage. As a result, M9 reported offline harm such as physical assault and lethal threats.

Table 6.3: Mapping of defensive mechanisms, mindsets, and external support to threat categories (c.f. Figure 6.2).

Threat Category	Defenses				Mindsets				External Support				
	Technical	Data Policies	Self-Censorship	Offline	Ignore	Comply	Fight	Fatalism	Information Sources	Platforms	Authorities	Social Suppport	Therapy
Toxic Content	●		●		●	●	●	●	●	●	●	●	●
Content Leakage + Theft	●	●					●					●	
Overloading	●					●	●					●	●
Lockout + Control	●								●	●	●	●	
False Reporting								●	●	●		●	
Impersonation					●		●		●	●	●	●	
Surveillance		●						●	●		●	●	
Physical Harm		●	●	●			●	●	●			●	
Offline Harassment		●	●	●			●	●	●			●	

●: observed with one or more participants

Impersonated Profiles. Half of the woman creators in our sample were impersonated online [w:6 | m:1 | k:0]. Common platforms are dating profiles (W2, M9), and social media platforms (W1, W10). Attackers are unknown to participants, and motives are often unclear. Attackers deceived people to send them money (W2) or extract information from family members (W10). There was one case that led to offline harm: *"People started making fake Tinder and Bumble profiles. I don't know if it was something personal against me."* (M9). Even several years after the incident, the participant faces severe threats, as family members of the matches track him down and threaten to kill him. They are motivated by violations of strict cultural norms of *pardah/modesty* and *honor* [355].

Synthetic CSAM and NCII. Another form of impersonation, that was experienced by one woman in our sample, is synthetic non-consensual intimate imagery (NCII) or synthetic child sexual abuse material (CSAM). Participant W7 had synthetically-created pictures of her shared in an online community forum: *"I think to me the most unpleasant experience was the Reddit picture [...] where someone did something really weird on my picture. That was uncomfortable for me because that was more of a sexual nature. [...] I was 16. [...] someone sent me a screenshot of [the image]. I reported that and it got taken down because I [was] a minor "* (W7). Further, technologies such as Deepfakes generate content that is harder to distinguish from reality, opening the door to harass the victim based on societal norms of decency.

6.4.2.7 Surveillance

Participants identified two types of attackers: (1) online contacts, such as fans or opponents, and (2) offline contacts, such as friends, family, or work-related acquaintances.

Stalking Behaviors. Stalking is a common experience among creators [w:4 | m:2 | k:2]. One participant reported being digitally stalked: *"There was a case with a house help of mine that got access to my accounts and then ended up stalking me from different accounts. That was a very big concern for me and a very unpleasant experience"* (W6). Stalking, especially if exerted by people close to the victim, directly or indirectly reduces the victim's physical or psychological integrity. Incidents of stalking may co-occur with physical threats, as was the case for this participant: *"I've had physical concerns in the sense that randomly, people 1-2 times have sent me messages: 'This was your car*

and this was the numberplate. We saw you take a u-turn from here.' " (W5). Thereby stalking of creators is an example of a post-digital security issue [79], as it blurs the line between the online and offline worlds.

Account Monitoring. Similarly, female participants reported that they are concerned and annoyed that parents and other relatives monitor their account (W2, W3, W6, W9, W11, W12). This is in part due to socio-cultural norms of decency, as W9 puts it: *"People in our society are not accustomed to looking at pictures or videos of a woman on social media."* and W3: *"I know that a lot of people from my family have approached my mom in a way, saying you have absolutely no control over her daughter and she should be ashamed of what I was doing."*

6.4.2.8 (Offline) Physical Harm

The various online threats above can manifest into offline attacks causing physical harm to creators. Participants are especially concerned about influential Pakistani figures abusing their power and network to harm them. For U.S. creators physical-world harm was a top concern, however only few had personal experiences [318]. In our study 11 participants reported negative offline experiences, suggesting that they might be more prominent in the context of Pakistan.

Physical and Sexual Assault, Murder. The majority of participants were aware and concerned about the possibility of becoming targets in the real world. Three participants experienced physical assault [w:1 | m:1 | k:1]. M9 went through an extreme case of assault as a consequence of being impersonated on a dating profile: *"I was getting into my car and behind my car two cars came and parked. A person came out from one of the car and he held me by the collar pushed me against the wall and he pointed a gun to my head and he was you have done this and that to my sister."* K2 reported getting kidnapped. Mostly Khwaja Sira were concerned about sexual assault; none of our participants experienced it. Murder is real threat as is exemplified by M9 and past events [228].

6.4.2.9 (Offline) Harassment.

This category contains offline attacks that cause the victim to feel intimidated, dehumanized, or belittled.

Sexual Harassment. Sexual harassment like catcalling (W9) was not experienced by men [w:1 | m:0 | k:2]. Khwaja Sira faced severe instances of offline sexual harassment through cis-woman: *"As a trans woman interacting with a cis woman, the privilege lies with the cis women. Harassment at the hands of cisgender women, I have faced a lot, which can both be what I call sexual harassment and then harassment of a private nature where they ask me very, very intrusive and disgusting questions in the presence of other people"* (K1). She explains that in Pakistani patriarchy, women cannot harass cis-men, so they harass transgender people.

Table 6.4: Overview of threats experienced by participants according to the coded interview data. "E" denotes that the participant reported a personal negative experience in this category. "T" denotes that they did not report a negative experience, but explicitly stated that they are concerned about an attack in this threat category.

Threat Category		W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	M1	M2	M3	M4	M5	M6	M7	M8	M9	K1	K2
Toxic Content	Bullying	E	E	E		E	E	E	E	E	E	E	E	E	E	E	E			E			E	E
	Trolling	E	E	E													E		E				E	
	Hate Speech	E	E	E	T	E	E	T	E	E	E	E	E	E			E	E					E	E
	Sexual Harassment							E	T	E													E	E
	Slander		E							E			E	T			E			E		E	E	E
	Conspiracy Theories		E														E						E	E
	Unwanted Explicit Content			E	E	E		E	E	E	E	E	E		E					E				E
	Blackmailing											E					E				E	E		E
	Threats of Violence			E						E				T	T		E	E					E	E
	Incitement		E							E							E	E				E	E	E
Content Leakage	Doxing	T		T	E	T	T	E	E	E	T	E	T				E	E					E	E
	Outing and Deadnaming																						E	
	Content Misuse	T	E	E			T	E		E		E	T	E						E		E		E
Overloading	Comment Spam	E	E			E			E		E		E				E	E		E				E
	Negative Ratings								E								E							
	Brigading/ Dogpiling		E						E	E														E
	Canceling		E														E			E				
False Reporting	Account Banning	E		T					E				T		T									E
	Content Reporting	T							E													E		
Impersonation	Impersonation	T	E	E	E	E		E			E											E		
	Synthetic CSAM/NCII							E					T											
Surveillance	Stalking Behaviors		E			E	E				E						E	T				E	E	E
	Account monitoring		T	T			T			T		T	T											E
Lockout + Control	Account compromise			T		T			E				E			T		T	T	T		E		E
Physical Harm	Physical Assault	E		T		T		T		T				T								E	T	E
	Sexual Assault									T							T						T	T
	Murder					T				T							T	T				T	T	T
Offline Harassment	Sexual Harassment							T		E													E	E
	Stalking Behaviors			E		E		E			E		T	T			T	T	E		E	E	E	E
	Threats of Violence									E				T			T	T				E	T	E

Threats of Violence. Another form of intimidation is threats of violence, such as threatening sexual assault (W9) or death (W9, M9, K2). Many participants are concerned about the possibility of somebody showing up at their address to threaten them. Creators are sometimes threatened to influence their posting behavior or delete content: *"Lucky for me it was not someone influential otherwise they would have forced me to delete it rather than asking politely and might have reached my house and telling me are you deleting it or we make you delete it"* (M1). Others will threaten to kill creators to restore honor (M9).

Stalking Behaviors. All genders experienced offline stalking behaviors [w:4 | m:3 | k:2]. While participants are generally open to take pictures and meet fans in public, some are concerned about having pictures and videos taken without their consent because of being public figures (W3). Similarly, creators reported instances where fans went to their house or current location to take pictures or meet them (W7). In contrast to the previous category, here the intention is not to threaten the creator, but to be near them. The close-knit communities within Pakistan simplify the process of obtaining creators' personal information for attackers (K1). This is especially the case in rural areas, and when the creator has great visibility [345], e.g., because of being Khwaja Sira (K2).

Table 6.4 provides an overview of the negative experiences participants reported on during the interview.

6.4.3 Coping with Threats

We categorize *defensive mechanisms* and *mindsets* of creators when handling threats described above in Section 6.4.2 and the *external support* they rely on. The results

of this section contain a thematic analysis of the axial category *defense practices* and *support*. Table 6.3 relates the findings of this section to the previously described *threat categories*.

6.4.3.1 Defensive Mechanisms

Technical Defenses. All participants rely on **technical** defense mechanisms to protect their accounts, content, and manage their community. To battle *account lockout*, few participants use *authentication-related* defenses, such as password complexity rules (W1, M7), password managers (W1, W9), and login monitoring (F6). While some participants follow outdated security practices, such as changing passwords regularly (W1, W5, W6, M9), the majority use enhanced authentication schemes such as 2FA [w:12 | m:7 | k:2]. 2FA in particular is perceived as a strong security mechanism, however some participants only adopted it when the platform pushed it on them after account compromise (M9, W12), or they saw suspicious login attempts (W5). Similarly, U.S. creators adopt protective practices in response to attacks [318]. Moreover, most participants [w:12 | m:6 | k:2] rely on *community management* tools to prevent and *combat toxic content* and *overloading*. Mechanisms that restrict who can send messages or reply to content (W7) are used to silence attackers and perceived as effective. Yet, the majority of participants were not satisfied with the available comment moderation, blocking, and reporting tools platforms provide to prevent hate and harassment and de-platform attackers. Similar to how young Pakistanis collaborate to counteract impersonated profiles [32], creators rely on their community to mass-report attackers in the case of hate speech and impersonation. Some participants [w:5 | m:1 | k:0] use *content control* methods to protect against *content theft*. For example, creators sometimes use watermarks on their content (W11). However, these do not stop the content theft itself, but rather ensures that the source will be known (if the watermark is not removed entirely). Other creators completely separate content into private accounts to prevent it from being misused (W1, W2, W3, W12, M9). To prevent content being deleted after a hack, one creator curates backup accounts to which they can switch in case of a hack (W1). This purpose is also communicated to fans and they are encouraged to follow the backup account.

Data Policies. Most participants [w:12 | m:6 | k:2] report following personal data control policies to combat *content leakage*, *surveillance*, and offline threats. The underlying principle involves keeping personal data points private, such as details about family and friends, location data, daily activities, or means of transport. Some go as far as hiding their identity online or in real life (W9, W11, M9), or leaking false information to distract people (K1). These practices are universally perceived as effective strategies to preserve online and offline privacy among our participants.

Self-Censorship. Similarly to a study on U.S. creators [318], the majority [w:9 | m:9 | k:2] of our participants engage in self-censorship to protect from becoming the targets of online and offline hate and harassment. Self-censorship is a learned pattern of behavior that results from one's own or other people's experiences. Thereby, implicit or explicit social norms influence what may be said and done. While one participant acknowledges

self-censorship as an effective strategy to prevent hate (W2), our data suggests that it does not prevent creators from harassment, particularly related to religion. Against this background, self-censorship is a fragile precaution: *"I also avoid political and religious discussions online because I know for a fact that will never end well. [...] The most precautionary thing is to extremely filter out everything that I say and do"* (W3).

Offline Protections. Several participants [w:1 | m:2 | k:2] report relying on offline defenses to protect from *physical harm* and *harassment*. They carry pepper spray and rely on physical protection from other people (W9). Khwaja Sira face increased risk of offline harassment and report being selective about the people they meet (K2) and avoid attending parties in response to their digital visibility: *"I'm a digital creator who was getting very famous, and if I'm inviting 20 people and 25 show up, at least eight of them are going to show off saying [...] we ran into her. [...] They will go post my pictures and then that becomes a whole shitstorm. Like they would say this person talks about Islam or Sufism and now you see them in short cloths"* (K1). In that way, for creators offline privacy invasions can feed into online harassment.

6.4.3.2 Coping Mindsets

Based on how participants reacted to negative experiences we derived four *coping mindsets*:

Ignore. When coping with online harassment participants had the sentiment that this is something that one simply has to *accept* and *ignore* [w:12 | m:8 | k:2]. A similar theme was identified by Samermit et al. when U.S. creators claimed to develop a "thick skin" [318]. In line with their work, we find that creators do not want to give haters and trolls attention, hoping that this way the attack will die down. We identified an additional theme of Pakistani creators being concerned that by standing up for themselves they might confront someone influential who could target them in the real-world (M1). *Ignoring* is often informed by a sense of helplessness when creators feel they cannot defend against a threat. This resetting of safety expectations when facing elevated risk applies to U.S. creators, too [318]. We find this theme especially for *toxic content* type threats when creator's way of coping is to *"just get used to it"* (W7). Similarly, most could not imagine a world without online hate and harassment attacks. However, this does not mean that they do not deploy defensive mechanisms. Acceptance of risks was felt throughout our participants; for instance regarding impersonation, one participant noted they lacked adequate tools to deal with such attacks and therefore became desensitized: *"Impersonation happens a lot. Initially, I would get very scared about it. I would report the account and I would ask my friends [to report] them also. But then slowly, as the number of accounts increased, I got desensitized."* (W3).

Comply. One creator reports *complying with* attackers to stop hate and harassment, especially when facing cancelling. M4 apologized and changed his behavior and content. *"I uploaded three apology videos. [...] Did it once and deleted that account because I*

was getting too many death threats." In that way, this coping mindset is connected to *self-censorship*.

Fight. Many creators [w:4 | m:3 | k:2] report fighting online hate and harassment, especially if other coping mechanisms like *ignoring* or *complying* fail: *"I try to ignore it as much as I can, but if I can't and if I lose [my mind], then I'll have to answer to that person back"* (M3). They speak up for themselves (M3, K2), call people out (W2, W5), respond to comments (K2), and sometimes harass people back - the latter also offline (K1). We found that fighting back is a coping strategy that all participants who create sensitive or controversial topics rely upon (K1, K2, W9). We theorize that creators who actively decide to put out controversial content are willing to fight for their voice. But also creators on seemingly uncritical topics easily become targets. When asked whether she ever considered leaving the platform due to the hate she faces, W12 responded: *"I think that would be the weakest I can do. I chose to fight, be brave, and to deal with it."*

Fatalism. Across genders [w:7 | m:7 | k:1] participants developed a fatalistic mindset towards online and offline threats: *"If it's harassment, that's something that I still haven't figured out how to deal with [...] Though, I don't know how I will [ever] figure out how to deal with that"* (W3). In response to online hate participants left platforms (W5, W9), or even fled the country (M9). The attacker who almost killed M9 advised him to leave the country even after being convinced that M9 was innocent: *"He's like, 'Today was me tomorrow it could be some other guy from that group who tries to do something like this.' I was like, 'What do you suggest I should do?' He's like 'I would honestly say, stay off the grid. If you can [...] leave the country for a while.' "* (M9). W8 stopped creating content on TikTok because her account kept getting reported.

6.4.3.3 External Support

Information Resources. Two participants (W9, K2) explicitly mentioned education as a critical step to maintaining their security and privacy: *"Everything is just changing so much and technology is advancing at a speed of light. I'm just trying to understand it first, and then I'll take some measures about it."* (W9). Participants got their information about how to defend against threats from fellow creators [w:7 | m:7 | k:0], people in their personal lives [w:10 | m:7 | k:2], search engines [w:2 | m:1 | k:1], platform guidelines [w:3 | m:2 | k:1], advertisements (M8), and organizations (e.g., Digital Rights Foundation) (K1, K2).

Platforms. Creators appreciate platforms having physical offices (M8) and help centers (W12). They rely on *platform support* to combat *impersonation* (W4), and recover from *account compromise* after being hacked (W12). Also, platforms are the only way to recover from account loss due to *false reporting* attacks. In the case of *impersonation* participants stress that they *"don't know any other way to deal with it"* (W4). Creators criticize the slow response time of some platforms (Instagram) when taking down *content misuse*, and highlight that others (YouTube) are doing a good job (W9). Wanted features are tools to check for re-posted and impersonated content (W2, W3, W4, W7), controls

over who can view and screenshot content (W1, W3, M9), measures to counter spread of *slander* (W2), and involving real people to resolve conflict (W2, W6, W12, M2, M9), e.g. in the form of a help line (W2, W4, W6, W7, W9, M2). Moreover, they would appreciate transparency over how banning works (W1, W3, W12) and the formation of support communities (W5, W7, W9, W11, M7).

Authorities. Few participants reported *involving authorities* when dealing with online or offline threats. Creators rarely reached out to *government agencies*, and if they did it was not very effective (e.g., they got no response (W5)). W2 summarizes why: *"I think in Pakistan especially, like nobody really directly goes to authorities. It's usually through contacts. [...] I don't recall there being sort of easy access to authorities where it's like a helpline or something. Maybe there is one, but [...] it's not being advertised enough. Because I think there's also a general distrust in society, like in terms of authority. We don't really trust them because it's very hard to find authority that takes your word and actually brings justice"* (W2). Even creators who generally fight back against hate and harassment are reluctant to report attackers officially: *"[If] I lodge cases against them, I'm pretty certain things against me are going to go ballistic. It's a double edged sword. I will take an action, but that action is also probably going to cost me my life"* (K1). None of our participants mentioned involving authorities in response to offline threats. Other participants reported being in contact with activist agencies (K1), or reaching out to lawyers (W2, K2) and experts (K2).

Social Support. To cope with negative experiences, especially hate and harassment, participant rely on *social support*. They commonly refer to family and friends for support and to get a reality check (M3), although W3 tends to hide negative experiences from her close family because *"they already won't approve of <content creation>, and they'll tell me to stop"*. Another popular form of support is to talk to fellow creators and receive advice on how they navigate negative experiences (W4).

Therapy. Last but not least, a few participants go to therapy to cope with the outcome of online hate and harassment. The vast majority of our participants report on how content creation and the threats they face negatively impacts their mental health. They question their self-worth (M4), and face depression (M4) and burnout (W12, K1). One participant reported becoming suicidal after facing severe online and offline harassment (M4).

6.5 Discussion

6.5.1 Risk Factors (RQ1)

We discuss factors that correlate with experiencing heightened risk based on the relative frequency of reported threat types (compare Table 6.4), and the reasoning participants offered. These factors have to be quantitatively validated in future work.

Gender. Our data suggest that women and Khwaja Sira face heightened threats and are more likely to experience sexual harassment than men (average # of reported negative experiences: [w:7.33 | m:5 | k:17]); future work needs to validate this hypothesis. Women are pressured to adhere to social norms and face heightened risks in online spaces when identifiable as women (W12). Thomas et al. [345] found that woman creators in the U.S. statistically faced more threats than other genders; they did not find effects for transgender persons. In this study the threats Khwaja Sira faced were severe: Among others, they got kidnapped (K2), were sexually harassed online and offline (K1, K2), and experienced human-trafficing attempts (K2). Sexual harassment is an extremely sensitive topic. Victims are reluctant to speak up because of getting stigmatized, i.e., tainting their honor. Also, victims need to be cautious when acting against attackers. Confrontational responses like calling out people, especially if they are powerful (e.g., religious leaders) can become lethal [228]. Thus, technical solutions need to focus on victim protections; solutions such as blocking that draw attention on the actions of victims might not be suitable as it can provoke attackers.

Content Topic. Participants who create content on topics that are controversial face more negative experiences than those who do not. In contrast, in the U.S. content topics were not correlated with higher risk [345]. However, recent research identified content on sex work and nudity leads to de-platforming [30, 29] or shadowbanning [28]. In our study, examples of critical content topics are social issues and activism (W9), women's rights (W9), transgender rights (K1, K2), and sex and sexuality (K1, K2). Moreover, many participants report avoiding content on religion (W2), politics (M2), and influential people (W9) to avoid harm. On the contrary, creators in our sample whose content centers on topics like music (W4) and pets (W12) face fewer threats. Yet, seemingly harmless content such as pictures in a mixed-gendered group (W3), dance videos (W8), or wearing shorts (K1) can be *"stigmatized in [Pakistan's] society"* (W8), and bear the potential to be weaponized, especially against woman and Khwaja Sira. Thus, platforms need to take local cultural norms into consideration when deciding what classifies as "harmful" content and must not make assumptions based on Western views.

Platform and Audience. Some participants see a connection between the platform and the hate they receive. They suspect that the audiences differ between platforms, and that platforms with more low-literacy users (they name TikTok) harbor more potential for hate and harassment. Moreover, they think certain interaction features and affordances that allow direct contact with creators enable harassment. However, results are inconclusive, and this factor must be explored further in future work. Some creators reported that they started attracting hate only after reaching a bigger audience (W8). This is in line with Thomas et al. [345], who reports a correlation between hate and audience size. Moreover, Samermit et al. [318] describe *prominence* (in accordance to Warford et al. [370]) as a risk factor for creators in the US, e.g. when massive popularity and virality further amplify risk.

6.5.2 Gaps in Defenses and Support (RQ2+3)

Flawed Defenses. *Toxic content* is a prominent threat that a majority of our participants experience. It is also ranked a top priority by experts, and they advise *blocking*, *muting*, *reporting*, and *moderating* [372] which our participants also rely on. Platforms invest great effort into combating *toxic content*, and we find that participants use and appreciate moderation tools that help filter comments. However, they report that it is especially difficult to de-platform malicious actors, and criticise the lack in transparency and efficacy of reporting mechanisms. In line with this, other research discusses how reporting in Urdu is harder than reporting in English [355]. Work on U.S. creators found that involving platforms can be slow and opaque if the creator does not have a human contact at the platform [318]. In our sample creators often felt left alone with the responsibility to defend themselves. However, some active forms of defense, like speaking up and addressing hate, can have lethal consequences in Pakistan's tight social net when perpetrators in turn feel attacked and use their network to reach creators offline. Creators deliberate this when speaking up against hate; some may have a cause that they deem worth fighting for, e.g., activism for minorities, women and transgender rights. However, a policy brief by UNESCO and the Digital Rights Foundation [355] discusses that posts on sensitive topics lead to less engagement, either because audiences self-censor too or because of algorithmic biases [25]. In general, all creators were aware of the risks associated with sensitive topics such as religion and politics, and thus most of our participants refer to self-censorship.

Missing Protections. Participants lacked any defenses against threats of *false reporting* and *impersonation*. A recent study by Wei et al. [372] in which experts ranked threat categories internet users should prioritize, threats that fall under the category *impersonation* and *false reporting* were assigned medium ranking and rarely identified as top threats. In contrast, we found that *impersonation* attacks likely lead to offline threats and are thus highly critical in Pakistan. In our data set, we had one participant who almost got killed due to being impersonated on an online dating app (M9). Regarding Pakistan, *impersonation* poses a high risk for (1) the reputation of the victim, but also (2) for the reputation of the person who falls for the scam, especially in the context of impersonated dating profiles. M9 still receives online and offline threats from family members of the scam victim, even several years after the incident. Threats are so severe, that even former attackers advise him to leave the country. This points out the potential for geographic biases when selecting generalized security advice [372] and highlights the need for targeted advice, especially for populations facing elevated threats. Critically, even for the general internet user experts identified a lack of effective advice to defend against *impersonation* [372]. Future research needs to work towards effective and feasible defensive mechanisms taking into account the needs of marginalized populations such as creators in Pakistan. Moreover, creators are lacking effective protections against *stalking behaviors*; they rely on *data policies* that are implicit rules to control which data they release. However, it is difficult to foresee which data points might result in *stalking behaviors* and retracting data that is published once, is hard or impossible. In this regard creators are different from young adult internet users in Pakistan [32], as they cannot rely on the same affordances (e.g. having private profiles).

Lacking Support. The socio-cultural context of Pakistan impacts availability of support structures, especially as content creation is unregulated and lacks protective legislation [355, 32]. First, there are gaps when it comes to official support structures for online harm. Participants were reluctant to approach authorities, or unaware of agencies that deal with online hate and harassment. Underscoring the problem, even the one participant in our sample who filed a report never heard back from the authorities. This issue is similar for marginalized content creators in the West such as sexual content creators who similarly report being unable to turn to authorities due to stigma and lack of respect [149]. For U.S. content creators, law enforcement can be helpful, but they are not always taken seriously [318]. Second, participants report a stigma that is attached to content creation in Pakistan, especially for women: in our sample they were concerned about their parents and close relatives monitoring their accounts. This in turn might constrain the *social support* resources women can rely on, if they do not feel they can approach their parents when being attacked. Similarly, young adults in Pakistan are reluctant to report cybercrimes because they do not want their families to worry, or women want to avoid being victim blamed [32].

6.5.3 Towards Solutions

Threat Modeling. There is a need to explore the design for flexible threat priorities where socio-economic context determines what are severe threats, as opposed to viewing threats as static across social contexts. Comparing our results (e.g. impersonation attacks leading to offline harm) to previous work focusing on Western populations where experts ranked threat categories that internet users should prioritize [372], we suggest that different socio-economic contexts influence the prioritisation of threats.

Technical Defenses. Creators in Pakistan are especially dependent on well-designed technological defenses because of missing social and societal protections (e.g. reluctance to approach authorities, involve families). Defenses need to take cultural norms into account when (1) detecting harmful content (e.g. pictures of mixed-gendered groups, comments by religious figures), (2) protecting victims (defensive mechanisms should not leave traces), as Pakistan's tightly-woven social-net allows attackers to reach creators offline, or mobilize a crowd to attack victims based on false claims (e.g. religious offense).

Tailored Information Sources. Creators expressed a desire for contextualized information sources. We suggest to consider factors that impact creators' threat experiences when generating advice. Based on our findings we identified gender, content topic, platform, (size of) audience, and broader social embeddedness including cultural background.

6.6 Conclusion

Creators in Pakistan and the US [345] face online threats in the same categories, however there are differences in risk surface and expected harms: (1) what classifies as harmful content depends on the cultural context, and (2) we observe that online threats can

be expected to lead to offline harm, which was not reported for U.S. creators [318]. Here, threats like impersonation that are rated "moderate" by experts [372] lead to potentially lethal attacks. Hence, we hypothesize that although creators across the world face the same threat categories, the prioritization of threats changes across cultures. Finally, compared to young adults in Pakistan [32], the threats creators face are even amplified. Furthermore, countermeasures such as taking their profiles offline do not work for creators which suggests that they need protection mechanims for their specific needs.

Part III

Discussion and Conclusion

7 Discussion

Digital sovereignty is concerned, among other things, with empowering people to make self-determined decisions and actions in the context of digital technologies. A central aspect of this is the security and privacy needs of users, over which they should retain control. The past chapters presented four studies, through which I examined the current technology landscape in terms of systems that grant their users varying degrees of digital sovereignty. In the following, I discuss what we can learn from my findings about security and privacy challenges that limit the digital sovereignty of users.

7.1 Challenges for Privacy-Preserving Technology

I investigated self-hosting as a prime example of behavior that—if done correctly—grants people the highest degree of digital sovereignty. Through self-hosting, people can take full control over where and how data is stored and processed. From an academic perspective, there has been a critical lack of data on self-hosting. In part, this may be because the concept of self-hosting emerged recently as a counterpoint to Software-as-a-Service solutions. The move to third-party cloud computing is a recent phenomenon, and "self-hosting" used to be the default practice for decades. Still, human factors in system administration are relatively understudied, while they might be critical to success or failure, as my research suggests. In the following, I discuss key challenges for digital sovereignty that emerge from the findings of chapters 3 and 4.

The secure operation of infrastructure is too complicated

Some individuals have an intrinsic drive to achieve digital sovereignty. Chapter 3 highlighted such people and groups, including journalists and lawyers whose professions demand exceptional standards of data protection. Climate activists also stand out, seeking to safeguard their organizations from the influence of powerful oil companies and

government surveillance. Additionally, there are private individuals who value privacy as a fundamental principle and use digitally sovereign technologies to exercise their right to freedom of expression—thereby contributing to the promotion of democracy. Many of these individuals turn to self-hosting as their only viable option, citing the pervasive privacy violations inherent in mainstream technologies. This illustrates how, for some, the objectives of self-hosting align seamlessly with the core values of digital sovereignty.

However, my studies find that technical expertise is associated with self-hosting [P2], and people perceive it as a major barrier [P1]. This is because people need a comprehensive understanding of hardware and software components, as well as computer networks, to set up their own system. However, setting up a system once is not enough, as it needs to be regularly maintained and updated. My study with the Nextcloud community has shown that coming up with a suitable security strategy, in particular, causes major problems for self-hosters. Specifically, the paper found that they may struggle with (1) defining attacker models, (2) prioritizing risks, and (3) identifying appropriate defensive mechanisms. This is not only true for private people. Lacking security practices were reported across big and small commercial and non-profit organizations, too. Interestingly, the data did not suggest any factors, such as peoples' technological expertise, that might explain whether a structured or unstructured approach to security was used. That is, the study found bad and exceptional behavior everywhere regardless of use case, expertise, and social embeddedness. However, the privacy benefits that people expect from self-hosting are lost if the system is not properly secured. In this regard, the security of operations is a direct prerequisite for the digital sovereignty of these people. The actual security impact of self-hosting is not clear, and there are opposing views on it. Some participants stated they expect an advantage in staying off public clouds, as they assume to make less interesting targets for attackers if they self-host [P1]. On the flip side, participants acknowledge their lack of expertise and time to keep their instances secure as compared to big vendors who have the resources to spend on dedicated security experts. However, recent data breaches, such as those involving the cloud service provider Snowflake [78], which leaked business data, highlight the risks of centralizing user data. This centralization makes large providers particularly attractive targets for attacks. Currently, self-hosters heavily rely on their social network to cope with security challenges [P1]. For example, in cases where people turn to self-hosting to fulfill their professional-ethical obligations to data protection, my research found that they may enter knowledge-bartering relationships to deal with their lacking expertise. While this seemed to be successful for the participants in this study, it has its own drawbacks due to the dependency relationship and not being an option for everyone.

There are few easy-to-use and secure alternatives that do not require self-hosting

As outlined in the previous paragraph, the secure operation of infrastructure is a major obstacle for people who want to gain digital sovereignty through self-hosting. This makes self-hosting a poor choice for the majority of users. However, the aim does not have to be for everyone to take control of every aspect of the system. Still, there should be solutions for those who have a legitimate interest. But what about the majority of users?

There are alternatives that don't require the masses to self-host but are committed to protecting privacy and empowering users. We are currently experiencing a rise in decentralized technologies where users are not concentrated on one service but can interact with each other across many servers. A prominent example is Mastodon [237], a microblogging platform that experienced a massive surge in users after Twitter was sold. The idea is to provide users with a similar user experience they are used to from current big technology vendors. Users sign up with a server that federates with the entire network; thus, the service is decentralized in nature. The servers, however, are being hosted by community members, where I am confident the same high-level challenges apply that surfaced in my qualitative study on self-hosting [P1].

It is not made easy for people to prioritize digital sovereignty

8.4% of the US population self-hosts at least one service for personal use. For the website use case alone, this amounts to at least 20 million websites. That may sound like a lot in absolute terms, but the majority of user traffic still runs through a small number of providers [253]. However, of the 8.4% self-hosters in the States, a significant proportion probably aren't doing this to establish digital sovereignty. This assumption is based on the results of the qualitative study conducted in the Nextcloud community, which found that a large proportion of hosters follow a pragmatic approach. They host because they have the necessary skills, and a viable alternative is available with Nextcloud [P1]. The results of the quantitative study support this perspective, as it shows that self-hosters generally use more tools (both self-hostable and proprietary) than the general population. So, it seems that self-hosting is not an option for most people, and even for tech-savvy self-hosters, it is not a means of achieving digital sovereignty. Now, one might assume that digital sovereignty is not an issue of concern for the masses. However, I want to point out that this type of research does not allow any causal conclusions to be drawn. The observed phenomena could also be an effect of the fact that there are few functional and easy-to-use alternatives that enable users to be digitally sovereign. Proprietary solutions generally have low entry barriers and are used by the majority of people. This can lead to network effects, whereby people are incentivized to use already established services if they want to collaborate with other people. Moreover, a study on obstacles to the adoption of secure communication tools revealed that users prioritize utility over security [5]. When given a choice between two tools—one that is secure and usable and another that is insecure and usable—they tend to choose the tool offering greater utility, even if security is an important goal for them. Thus, similar to the controversy surrounding the privacy paradox (compare Section 2.2.2.1), it might turn out that people actually have no viable options to become digitally sovereign [167].

7.2 Challenges for Privacy-Violating Technology

At the moment, privacy-violating technologies are the norm. But that does not need to be the case. The goal should be for the industry to move away from surveillance capitalism towards sustainable business models that respect people's privacy and autonomy. This is both a policy and political problem, but also a problem of user interaction. Alternatives

need to be easy to use and secure. Based on the findings of Chapters 5 and 6, I outline the challenges that are developing for the digital sovereignty of users in the context of current privacy-infringing technologies.

It is unclear what users need to be digitally sovereign when dealing with closed systems

Digital sovereignty is a spectrum. People do not need full control over every aspect, as for example self-hosting allows people to exercise. There are systems where this is not sensible or even possible. For instance, it is not feasible for people to build their own cars. Especially with such security-critical systems, people are dependent on companies that enable them to use the technology in a digitally sovereign way. Key questions of digital sovereignty in this context include: Where must individuals exercise control, and where is transparency needed to enable informed decisions? Striking the right balance between automation and user involvement remains a significant challenge. It requires weighing the practical advantages of automation against the fundamental human need for control and agency. My research on drivers' information needs in safety- and security-critical situations within the realm of partially autonomous vehicles, offers actionable design insights for creating interfaces that effectively manage scenarios where automation fails [P3]. The findings emphasize that automation alone cannot address users' lack in security-related skills. Despite high levels of automation that minimize the need for user decisions, the study reveals that drivers prefer to take an active role in security-critical situations and require precise, actionable information to make sound decisions.

Claims to digital sovereignty differ across population groups

As the center of Internet commercialization revolves around advertising, leading technology companies are incentivized to gather large amounts of user data. Data collection and processing practices are opaque, and due to the rise of cloud computing data locality is hard to determine. Moreover, there are efforts to lock users into a company's ecosystem, restricting how a service can be used. To counteract these challenges, data protection laws have been put into place in some regions. One example of this is the General Data Protection Regulation (GDPR) that the European Union put into action to protect citizens' data. While these laws strengthen the technological self-determination of users, there is still a long way to go as big technology vendors have grown so powerful that it has become difficult to govern them by traditional political means [287]. According to the United Nations [88], in 2021 29% of countries had either no or only draft data protection legislation in place. Especially in Africa (39%) and Asia (43%) adoption is lacking. This leaves a tremendous amount of the world population vulnerable to the practices of global tech players. Moreover, these countries are disadvantaged in terms of resources (both ownership of technological infrastructure and access to experts and research institutions), and financial capital [286]. As Renata Pinto puts it: "*The world's offline populations are the disputed territory of tech empires, because whoever gets them locked into their digital feudalism, holds the key to the future*" [286]. This form of *digital colonialism* curtails the technological self-determination of those marginalized popula-

tions. Critically, these population do not benefit from the same security guarantees as the Western users, as my study on content creators in Pakistan highlights [P4]. The study finds that the threat landscape shifts due to the different socio-political context of Pakistan. Defensive mechanisms are non-existent, lacking, or inadequately adjusted for this context, thereby eroding the foundation of technological self-determination of the users of these platforms.

The digital sovereignty of users is not a primary goal for companies

Privacy and security are linked, but only to a limited extent. While robust security is a fundamental prerequisite for enabling data protection, it does not necessarily lead there. A service can be the most secure in the world, yet disregard the privacy of its users. Still, enabling secure use is a necessary first step towards a digitally-sovereign use of technology and often depends on factors such as the system's usability and the resources invested in making it securely accessible. Here, companies have the edge because they have more resources at their disposal. However, as highlighted by the example of content creators in Pakistan [P4], this advantage does not necessarily benefit all system users equally. If the population, their capabilities, or their threat scenarios differ from those of the "standard" user for whom the system was designed, users may find themselves with little or no protection—even within corporate systems with substantial resources. This may also occur when it is not economically viable for a company to allocate resources to enhance the experience for users who only make up a small market segment. Thus, it cannot be assumed that security challenges for end users are fully addressed simply by entrusting them to large companies.

8
Conclusion

Digital sovereignty is concerned with empowering people to make self-determined decisions and actions in the context of digital technologies. A central aspect of this is users' security and privacy needs, over which they should retain control. This dissertation examines security and privacy challenges limiting users' digital sovereignty. It explores these challenges through four studies that examine varying degrees of digital sovereignty across self-hosted systems, mainstream technologies, and corporate platforms. First, I consider self-hosted systems as the extreme point that can grant users maximum digital sovereignty. With a comprehensive qualitative study, I first examine why people self-host, how they administrate their systems, and how they secure them. In doing so, the study uncovers various dimensions of self-hosting and highlights areas of tension. In the next step, I present a representative quantitative study that estimates how prevalent self-hosting is among private people and what sets them apart from the population at large. The two studies indicate that although there are technological possibilities to become digitally sovereign, these are hardly used and that securing the systems, in particular, presents serious difficulties. As a counter-horizon, I examine technologies that the majority of people use, but which restrict the digital sovereignty of their users. Through a study on autonomous driving vehicles, I exemplarily examine a system that is advanced in its development and requires minimal user input but also gives users little opportunity to influence the system, e.g., on the much-criticized data collection. I study what information vehicle occupants need to be able to act appropriately in safety-critical situations and thus enable digitally sovereign handling of the vehicle. Finally, I present a study that uses social media to examine the impact of global corporations on the digital sovereignty of non-Western populations. I study the threat landscape of content creators in Pakistan and reveal how the technical safeguards provided by corporations are not sufficient to protect vulnerable communities. These studies suggest that automation is not the sole solution, as people want to make informed decisions, especially in critical situations, and that large corporations are not the saviors either, as, in addition to the prevalent privacy violations, they do not grant equal protection to all users.

Digital sovereignty remains out of reach for the majority of users. Approaches like self-hosting are not viable solutions for most people, as they demand significant technical expertise with a special emphasis on security practices. Beyond this expertise, another major barrier to self-hosting is the often poor usability of self-hostable software options. Beyond self-hosting, mainstream technology engages in prevalent intransparent privacy violations. Critically, the research community currently has no unified understanding of how and where systems need to provide information to their users to enable them to make self-determined decisions. This is generally a hard task, not only because use cases are vast but also because users are heterogeneous. Again, currently, we lack a profound understanding of which factors are suitable to describe a population and the associated threat models.

The recent trend towards decentralized platforms is a promising development for people's digital sovereignty. Future work could explore how people perceive decentralized social media and what barriers prevent people from using it safely. Self-hosting is likely to play a greater role as individual instances are provided by members of the community, e.g., in decentralized social networking sites. This calls for research to better understand the challenges and best practices of server administration, increasingly in

distributed teams. My research suggests that people find it particularly difficult to develop appropriate security strategies when they are self-hosting. Research should work towards providing more actionable advice in the area of system administration grounded in data. Moreover, research should explore how online communities can be leveraged to work towards more inclusive technologies. For example, while working with the Nextcloud community, I learned that many people without coding skills are eager to contribute to the open-source movement. However, the current open-source ecosystem is heavily focused on programming, which makes it difficult for people who don't have this background to contribute. Research could explore how open-source workflows can be opened up to design work and user studies. This could help address the persistent gaps in usability. Additionally, research needs to continue exploring how people's social embeddedness and cultural contexts influence their vulnerability to threats and the way these threats evolve across different sociocultural contexts. Developing a comprehensive theory of the factors that impact security outcomes will be crucial for informing both practical advice and design efforts.

Achieving digital sovereignty requires a paradigm shift that prioritizes privacy, inclusivity, and usable security while providing utility. Addressing these challenges requires a joint effort from researchers, online communities, and policymakers to disincentivize privacy-threatening business models. That way, we can create a digital landscape where individuals truly have the power to make self-determined choices and control their digital lives.

Part IV

Appendix

A

Self-Hosting Motivation, Operation, and Security Mindset

A.1 Community Survey

An asterisk (*) indicates mandatory questions.

1. How likely is it that you would recommend Nextcloud on a scale from 0 (not at all) to 10 (very likely)? *
2. What's your reason for this score and what could we do to improve it?
3. Which Nextcloud version are you running?
4. Which of the following describes your use case best?
   personal, non-profit, commercial, government, saas
5. Which apps do you have installed?
6. Which Nextcloud apps do you enjoy the most?
7. Which Nextcloud apps have you lost interest in?
8. Do you self-host any services in addition to Nextcloud?
9. How many users are on your server?
10. Why did you decide to host a Nextcloud instance?
11. In which area would you like to see more development from the community?
    Media hub, Digital office, Communication and social features, Connecting content from different apps, Performance, User experience, End to end encryption
12. Are there other areas that were not listed above that you would like to see more development in from the community?
13. Do you host an office suite alongside Nextcloud? None, Collabora CODE app, Collabora dedicated suite, OnlyOffice, Other: please specify
14. Do you use Talk, and if so, do you host the High-Performance Backend?
    newline I don't use Talk, I use Talk without the High-Performance Backend, I use Talk with the High-Performance Backend
15. Which CPU architecture are you using?
    x86 (Intel or AMD), ARMv32, ARMv64 (RPi, Pine Rock, etc.), something else (RiscV,

etc.)
16. Which kind of server are you providing Nextcloud on? Home server, VPS, Dedicated Server, Colocation, SaaS Provider
17. How many people are responsible for maintaining your Nextcloud instance?
18. How many of these people have a security background? If any, please specify which security training they had.
19. Is the security of your Nextcloud instance a concern for you? If yes, please give a reason why this is the case.
20. Please provide your name and e-mail if you would like us to be able to contact you if we have additional questions

A.2 Interview Guideline

Questions in red are for organizational self-hosters. **Intro.**

Hello. Thank you for your interest in participating in our study!

My name is $NAME$ and I am a researcher at $INSTITUTE$. If you have any questions about me or the study, we can talk about them first, then I'll ask your consent to start the recording and we'll start.

First, let's quickly go over how today's study is going to work. I'm going to ask you questions about your experiences with self-hosting in general and Nextcloud in particular. I expect that our conversation will take approximately one hour.

You can feel free to let me know if you don't want to answer a question, and we'll move on to the next question or we can stop the study, just let me know. Are you ok to start the recording?

I would like to begin with a few baseline questions about your reasons and contexts you use Nextcloud.

Motivation, Use Cases, and Social Embeddedness

1. I'd like to get to know you a bit to get started. Please tell me about your educational and professional background.
2. How did you get into self-hosting?
   a. [if privacy mentioned] What does privacy mean to you?
   b. [∀ answer] Is this a requirement for your company?
3. Which other products are you using simultaneously?
4. What needs to happen to make you switch to Nextcloud entirely?
5. Please tell me in which contexts you use Nextcloud.
   (a) For which purposes are you using it?
   (b) Who else is using your Nextcloud instance?
   (c) Who are the clients?
   (d) How many clients?
   (e) Do the clients have special requirements towards you?
   (f) Can you characterize your users/clients? (demography)
6. Who hosts the Nextcloud instance you are using?
7. Do you host a redundant Nextcloud for testing?
   (a) *If self-hosted:*
   (a) Are you the only one maintaining it?
   (b) Please tell me about your set-up.
      i. Which kind of server is Nextcloud running on?
      ii. Which operating system are you using?
      iii. Are you using any kind of virtualization to run Nextcloud (VMs, Docker)?
      iv. How is your Nextcloud reachable to its users?

   (c) How did you end up with this specific set-up?
      i. Did you encounter problems with:
         a) Hardware, b) Operating System, c) Virtualization, d) Software, e) Network?
4. Are you hosting any other services yourself?
   (a) Please tell me about them.
   (b) On which infrastructure does _____ run?

Thank you very much. Next up, I'd like to talk about how you first set-up your instance and any issues you might have encountered.

Maintaining the System

1. How/What do you do to maintain Nextcloud/your set-up?
2. Which issues did you face while maintaining/working with Nextcloud?
   (a) How did you fix those?
   (b) Did you adapt any strategies to prevent something like that from happening again?
      i. Please tell me about those.
   (c) Do you believe the security of your system was at risk at any point?
      i. Which steps did you take to maintain the security of your system?
3. How could maintenance of Nextcloud be made easier for you?

Attacks and Threat Models

Awesome. Next up, I'd like to talk about security.

1. Which security/privacy breaches or issues did you encounter with your self-hosted system?
   (a) Who was the attacker?
   (b) How was the attack executed?
   (c) Which damage was done to your system?
   (d) How did you respond to the breach?
   (e) How are you planning to prevent such an attack from happening in the future?
   (f) In how far would switching to a commercial solution prevent such an attack from happening?
      i. Why did(n't) you change back to a commercial solution?
2. Who is responsible for securing your Nextcloud instance?
3. What is your approach to keeping your Nextcloud instance secure?
4. What are your requirements and guidelines for security?
5. Do you believe you came up with a secure set-up of your self-hosted system?
   (a) Which security mechanisms did you deploy?
      i. How is (security mechanism) protecting your instance?
6. Who are you protecting from?
7. Why would someone be interested in attacking your self-hosted system?
8. What else could you do to enforce security?
   (a) Why didn't you do this?
9. At which point could your system be vulnerable? (software, hardware, network, user)
   (a) How could the vulnerability be used to attack your system?
   (b) Which measures are you currently taking to prevent such an attack from happening?
10. What could be done to help you secure your self-hosted system?
11. Which factors or constraints are currently preventing you from keeping your self-hosted system secure?
12. Did any security feature ever cause breakage? (for Nextcloud or any other self-hosted service)

Demography

1. How old are you?
2. What is your gender?
3. What is the highest level of school you have completed or the highest degree you have received?
4. What is your occupation?

5. Do you have a technical background? Please specify.
6. Do you have any background in security? Please specify.
7. Size of your company
8. Sector of operation
9. Size security team
10. Size operations team

A.3 Security Mechanisms

Figure A.1 gives an overview of security mechanisms participants used based on the interview codebook.

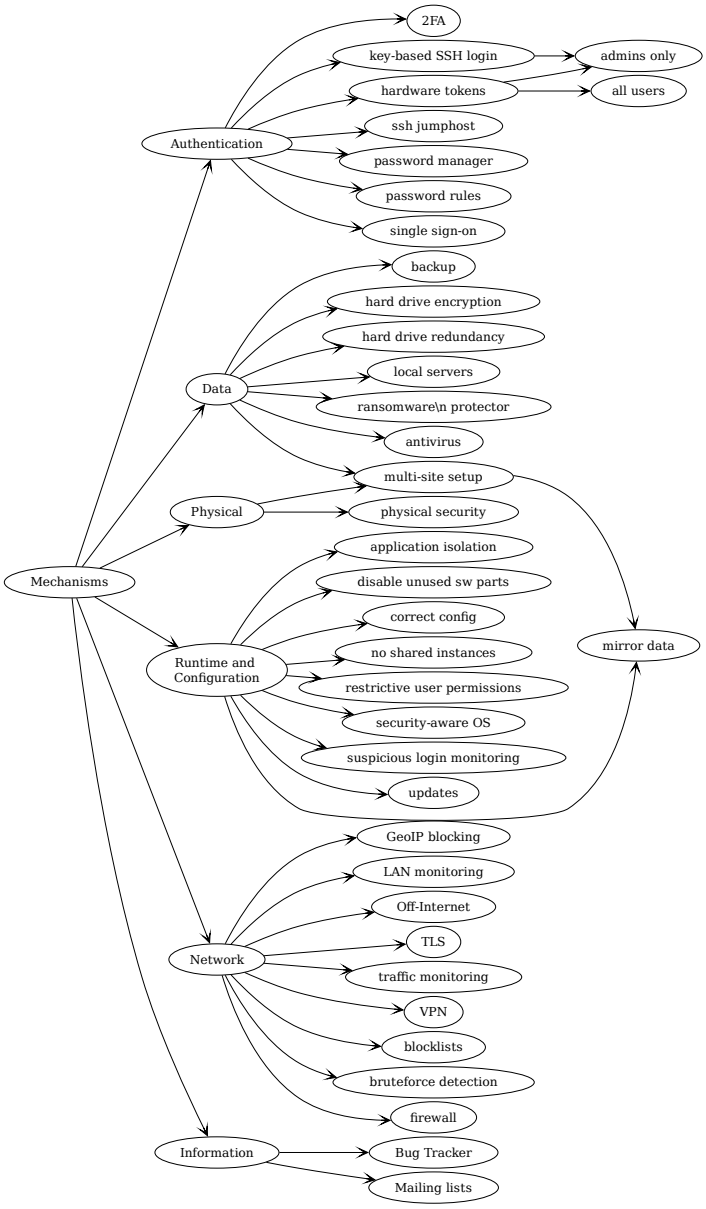


Figure A.1: Excerpt of interview code book for security mechanisms deployed by the participants

B

Quantifying the Self-Hosting Population

B.1 Flowchart and central questions of the prevalence survey

B.1.1 Survey consent

What is this survey about? Thank you for supporting our research! In this survey, we want to find out which software and applications people use, both in private and work contexts. You will be presented with various software tools and applications and asked to select the ones you use. This study will take approximately 6 min. Even if it is possible to fill out the survey with your mobile phone, we kindly ask you to fill out the survey on your laptop/desktop.

What data will be collected? In addition to your self-reported software usage, we will also collect demographic information (such as ethnicity, age, and gender, coarse region of residence), whether you complete the survey on your cell phone or desktop and your IP address (for technical reasons only, will be deleted immediately after the data collection is completed).

Data handling and confidentiality Your data will be used for research purposes only and will be treated confidentially. All records are completely anonymous. In accordance with the guidelines of the (redacted), data will be stored for 10 years. The data resulting from your participation may be made available to other researchers in the future for research purposes not detailed in this study description. In these cases, the data will not contain any identifying information that could link it in any way to you or your participation in a study.

What if I change my mind about taking part? Your participation is voluntary, you

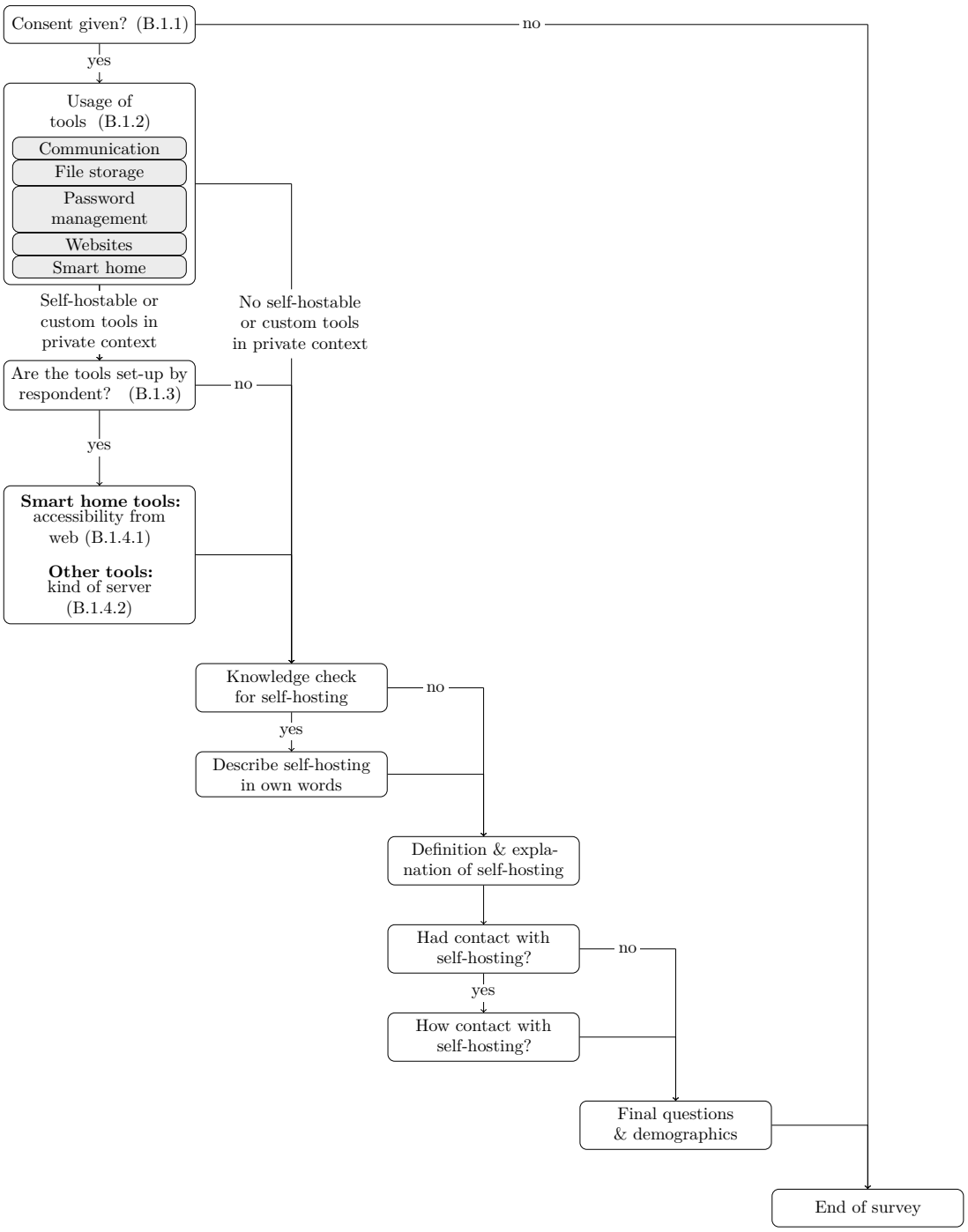


Figure B.1: Flowchart of the prevalence survey

have the right to withdraw your consent at any time during the study. Please note, however, that we require complete information and will therefore ask you to return your submission in case of withdrawal. Until the data collection is completed, you have the option to have your data deleted. From then on, the data is anonymized and deletion of the data is no longer possible.

What will happen to the results of the study? The results of the study will be summarised in a thesis, conference presentations and scientific journal publications. The anonymized data might also be published on the Open Science Framework.

Who is responsible for the study and who can I contact for more information? The study is conducted by (redacted). If you have any questions, require more information about this study or wish to make a complaint about the conduct of this study, you can contact us via Prolific or write an email to (redacted). We are keen to answer questions and to resolve ambiguities!

Do you agree to the terms and conditions listed above?

- I agree, I wish to participate.
- I do not agree I do not wish to participate.

B.1.2 Tools

B.1.2.1 Pre-selected tools

> For each of the tools in the list, please indicate whether you use it in a private context and/or in a work context (respectively for your studies).
>
> If you use a tool in both contexts, check both boxes.
>
> If you use a tool in neither context leave the boxes blank.
>
> If you don't use any of these tools, please select the "I do not use any of these tools" box for the respective context.

- **Communication**: Discord, Jitsi Meet, Mattermost, Microsoft Teams, Mumble, Rocket Chat, Signal, Teamspeak, Telegram, Whatsapp, Zoom
- **File Storage**: Box, Dropbox, Google Drive, MEGA, Microsoft OneDrive, Nextcloud, Seafile, SparkleShare, Synthing, iCloud, ownCloud
- **Smart Home**: Amazon Alexa, Apple HomeKit, Bosch Smart Home, Domoticz, Gladys, Google Home, Home Assistant, Node RED, SmartThings, Vivint Home, WebThings Gateway
- **Synchr. PW Managing**: 1Password, Dashlane, Keeper, LastPass, Padloc, Passbolt, Roboform, Teampass, Vaultwarden/Bitwarden, iCloud Keychain, sysPass
- **Websites**: Blogger, Cockpit, Ghost, Jimdo, Squarespace, Strapi, Wagtail, Webflow, Weebly, Wix, WordPress

B.1.2.2 Custom tools

Do you use any other tools for `<purpose of the respective category>` in work context or private context that are not mentioned above?

If so, enter the name of the tool below.

If you do not, please leave the text entry box blank.

B.1.3 Check whether respondent set-up self-hostable or custom tools themselves

B.1.3.1 Tools not falling in the smart home category

For some of the tools you have indicated that you use them in a private context. Because there are different ways to get access to these tools, we would like to ask you for some additional information. Please select the statement that applies to you.

`<tool>`

- `<tool>` was set up for me. I have created an account with my email address and a password to simply log in.
- A friend or family member set up `<tool>` and gave me a user account.
- I have installed and set up `<tool>` on a server myself (e.g. on a Raspberry Pi, Dedicated Server, VPS, etc.)
- I do not know which of the statements apply to me.
- None of the statements, I do it as follows:

B.1.3.2 Tools falling in the smart home category

For some of the tools you have indicated that you use them in a private context. Because there are different ways to get access to these tools, we would like to ask you for some additional information. Please select the statement that applies to you.

`<tool>`

- I am only a user of `<tool>` . A friend or family member has set it up and takes care of it.
- I have set up `<tool>` and I maintain it myself.
- I do not know which of the statements apply to me.
- None of the statements, I do it as follows:

B.1.4 Internet accessibility or server type

B.1.4.1 Internet accessibility (for all tools falling in the smart home category)

`<tool>` : Have you enabled remote access from outside your local network?

- Yes, <tool> is directly accessible from the internet.
- Yes, <tool> is accessible from the internet via a virtual private network (VPN).
- No, I am only using <tool> on my local network.
- I do not know which of the statements apply to me.
- None of the statements, I do it as follows:

B.1.4.2 Server type (for all tools not falling in the smart home category)

<tool> : On what kind of server?

- In my apartment / house (home server)
- Virtual Private Server (VPS)
- Dedicated Server

B.2 Flowchart and central questions of the characteristics survey

B.2.1 Survey consent

What is this survey about? Thank you for further supporting our research! Based on your given information in the first survey named "Software usage in private or work context" you are eligible to also participate in this follow-up survey. Your participation is therefore of great value to us and we would be grateful if you fill everything out to the end. In this survey, you will answer some questions about yourself across a spectrum of topics. These questions do not include any personal information or identifiers. This study will take approximately 13 minutes.

What further data will be collected? We will also collect demographic information (such as ethnicity, age, and gender, coarse region of residence) and whether you complete the survey on your cell phone or desktop.

Data handling and confidentiality Your data will be used for research purposes only and will be treated confidentially. All records are completely anonymous. In accordance with the guidelines of (redacted), data will be stored for 10 years. The data resulting from your participation may be made available to other researchers in the future for research purposes not detailed in this study description. In these cases, the data will not contain any identifying information that could link it in any way to you or your participation in a study.

What if I change my mind about taking part? Your participation is voluntary, you have the right to withdraw your consent at any time during the study. Please note, however, that we require complete information and will therefore ask you to return your submission in case of withdrawal. Until the data collection is completed, you have the option to have your data deleted. From then on, the data is anonymized and deletion of the data is no longer possible.

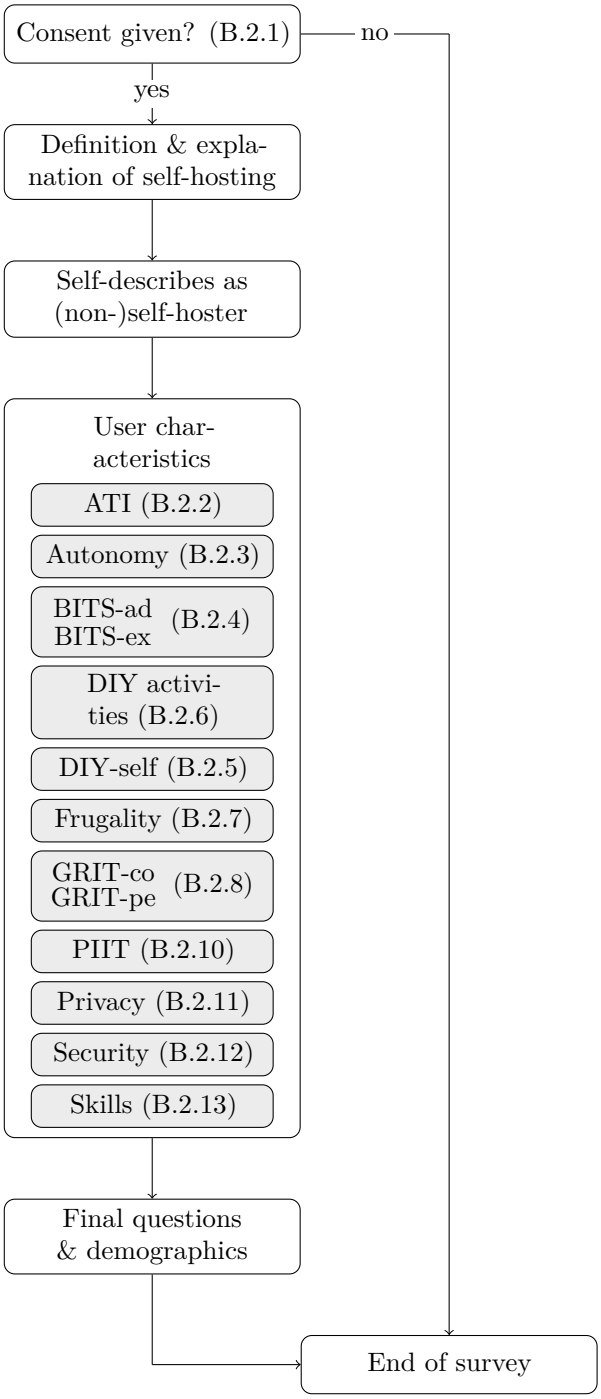


Figure B.2: Flowchart of the characteristics survey

What will happen to the results of the study? The results of the study will be summarised in a thesis, conference presentations and scientific journal publications. The anonymized data might also be published on the Open Science Framework.

Who is responsible for the study and who can I contact for more information? The study is conducted by (redacted). If you have any questions, require more information about this study or wish to make a complaint about the conduct of this study, you can contact us via Prolific or write an email to (redacted). We are keen to answer questions and to resolve ambiguities!

Do you agree to the terms and conditions listed above?

- I agree, I wish to participate.
- I do not agree I do not wish to participate.

B.2.2 ATI: Affinity for technology interaction

Instructions:

Indicate your agreement – Technical systems

Here, we ask you about your interaction with technical systems. The term 'technical systems' refers to apps and other software applications, as well as entire digital devices (e.g. mobile phone, computer, TV, car navigation).

On a scale of 1 to 6, please indicate the degree to which you disagree/agreee with the following statements.

Scale items:

- I like to occupy myself in greater detail with technical systems.
- I like testing the functions of new technical systems.
- I predominantly deal with technical systems because I have to.
- When I have a new technical system in front of me, I try it out intensively.
- I enjoy spending time becoming acquainted with a new technical system.
- It is enough for me that a technical system works; I don't care how or why.
- I try to understand how a technical system exactly works.
- It is enough for me to know the basic functions of a technical system.
- I try to make full use of the capabilities of a technical system.

Response scale:

1 Completely disagree
2 Largely disagree
3 Slightly disagree

4 Slightly agree
5 Largely agree
6 Completely agree

B.2.3 Autonomy: Need valuation: Basic psychological needs (BPN)

Instructions:

Indicate the degree of importance

On a scale of 1 to 5, indicate how important the following statements are to you in general.

Each level of the scale is assigned a number and the numbers increase from left to right. The higher the number, the more important to you.

For me, it is important that ...

Scale items:

- ...I feel a sense of choice and freedom in the things I undertake.
- ...I feel that my decisions reflect what I really want.
- ...I feel my choices express who I really am.
- ...I feel I am doing what really interests me.
- ...I do not feel like 'I have to' most of the things.
- ...I do not feel forced to do many things I wouldn't choose to do.
- ...I do not feel pressured to do too many things
- ...my daily activities do not feel like a chain of obligations

Response scale:

1 Not important to me at all
2
3
4
5 Very important to me

B.2.4 BITS: Brief Inventory of technology self-efficacy

Instructions:

Rate your confidence – Tech-related activities

Below, several tech-related activities are listed.

For each of the following statements on a scale of 1 to 6, please indicate your level of confidence that you can do the activity. Each level of the scale is assigned a number and the numbers increase from left to right. The higher the number, the higher the confidence.

Scale items:

[Subscale: Advanced]

- Creating a personal homepage
- Using advanced functions in office software
- Setting up a router
- Using a computer's task manager
- Setting up multiple computer monitors
- Troubleshooting computer problems

[Subscale: Expert]

- Using programming languages to write code
- Analyzing computer error log files
- Editing a computer's registry
- Designing professional websites
- Overclocking a computer
- Configuring a large computer network

Response scale:

1 Not at all confident
2
3
4
5
6 Completely confident

B.2.5 DIY-self: Maker Self-Identity

Instructions:

Indicate your degree of identification – DIY-Person

In this question, we would like to know how much you identify yourself as a Maker or DIY-Person.

To give you a better idea what we mean by that, we give you a description below:

Sometimes called a 'Maker' or sometimes 'do-it-yourself' (DIY); sometimes this can be called crafting, sometimes it refers to hobbies. Typically it leads to making something tangible. That is, what you can create with your own two hands. For example, you may like to garden, work on engines, build furniture, restore antiques, knit, quilt, make scrapbooks, take pictures, paint, fix up old houses, or cook for your friends or family. Please think about any kinds of activities that you have EVER done in your life that you enjoy and that you have ended up with some kind of tangible product.

Scale items:

- On a scale of 1 to 5, how much do you identify yourself as a "Maker" or DIY person?

Response scale:

1 Not at all
2
3
4
5 Extremely so

B.2.6 DIY activities: Maker Activites

Instructions:

Indicate your time spend – Activities

Below, several acitivities are listed.

On a scale of 1 to 5, for each of the activites indicate how much time you spend doing the activity.

Each level of the scale is assigned a number and the numbers increase from left to right. The higher the number, the more time you spend doing the acitivity.

Scale items:

[Subscale: Domestic Activities]

- Scrapbooking
- Baking
- Cooking
- Gardening/growing plants or flowers
- Fishing/hunting

[Subscale: DIY]

- Woodworking
- Electronics
- Metal working
- Fixing mechanical things (cars, machinery)

[Subscale: Arts and Crafts]

- Drawing/painting
- Ceramics
- Knitting/crocheting
- Computer graphics and web design
- Photography/film/movies
- Quilting
- Making jewelry
- Blogging/personal web page maintenance
- Sewing

Response scale:

1 None
2
3
4
5 I spend large amount of time doing activity

B.2.7 Frugality

Instructions:

Indicate your agreement – Consumer lifestyle

The following items refer to your individual, general consumer lifestyle.

On a scale of 1 to 6, please indicate the degree to which you disagree/agree with the following statements.

Scale items:

- If you take good care of your possessions, you will definitely save money in the long run.
- There are many things that are normally thrown away that are still quite useful.
- Making better use of my resources makes me feel good.
- If you can reuse an item you already have, there's no sense in buying something new.
- I believe in being careful how I spend my money.
- I discipline myself to get the most from my money.
- I am willing to wait on a purchase I want so that I can save money.
- There are things I resist buying today so I can save for tomorrow.

Response scale:

1 Definitely disagree
2 Mostly disagree
3 Slightly disagree
4 Slightly agree
5 Mostly agree
6 Definitely agree

B.2.8 GRIT

Instructions:

Indicate your agreement – Effort and interest

These statements are about effort and interest in goals and projects.

On a scale of 1 to 5, please indicate the degree to which you disagree/agree with the following statements.

Scale items:

[Subscale: GRIT-co: Consistency of Interest]

- I often set a goal but later choose to pursue a different one.
- I have been obsessed with a certain idea or project for a short time but later lost interest.
- I have difficulty maintaining my focus on projects that take more than a few months to complete.
- New ideas and projects sometimes distract me from previous ones.

[Subscale: GRIT-pe: Perseverance of Effort]

- I finish whatever I begin
- Setbacks don't discourage me.
- I am diligent.
- I am a hard worker.

Response scale:

1 Strongly disagree
2 Disagree
3 Neither agree nor disagree
4 Agree
5 Strongly agree

B.2.9 IT background

Scale items:

- Are you studying or have you been working in any of the following areas: information technology, computer science, electronic data processing, electrical engineering, communications technology, or similar?

Response scale:

1 Yes
2 No

B.2.10 PIIT: Personal Innovativeness in the Domain of Information Technology

Instructions:

Indicate your agreement – New information technologies

The following statements are intended to find out how you approach new information technologies.

On a scale of 1 to 7, please indicate the degree to which you disagree/agree with the following statements.

Scale items:

- If I heard about a new information technology, I would look for ways to experiment with it.
- Among my peers, I am usually the first to try out new information technologies.
- In general, I am hesistant to try out new information technologies

Response scale:

1 Strongly disagree
2 Disagree
3 Somewhat disagree
4 Neither agree nor disagree
5 Somewhat agree
6 Agree
7 Strongly agree

B.2.11 Privacy: Privacy Concerns

Instructions:

Indicate your agreement – Submitting information on the internet

The following statements refer to your concerns in general regarding submitting information on the internet.

On a scale of 1 to 5, please indicate the degree to which you disagree/agree with the following statements.

Scale items:

- I am concerned that the information I submit on the Internet could be misused.
- I am concerned that a person can find private information about me on the Internet.
- I am concerned about submitting information on the Internet, because of what others might do with it.
- I am concerned about submitting information on the Internet, because it could be used in a way I did not foresee.

Response scale:

1 Strongly disagree
2 Disagree
3 Neither agree nor disagree
4 Agree
5 Strongly agree

B.2.12 Security: Security Concerns

Instructions:

Indicate your agreement – Information security

The following statements refer to your concerns regarding information security.

On a scale of 1 to 5, please indicate the degree to which you disagree/agree with the following statements.

Scale items:

- I am concerned that databases that contain my personal information are not protected from unauthorized access.
- I worry about wrong information being linked to my identity due to security breaches.
- I worry about such information getting missing due to lack of adequate security measures.
- I believe stronger security measures are required to ensure the correctness of such personal information.

Response scale:

1 Strongly disagree
2 Disagree
3 Neither agree nor disagree
4 Agree
5 Strongly agree

B.2.13 Skills: Self-Hosting skill scale

Instructions:

In this part we would like to know how you rate your knowledge and skills in several technical areas. Thereby, the specific skills that we are interested in are mainly needed by people who work as system administrators.

To give you a better idea, we describe the work of such a system administrator below:

Install, configure, and maintain an organization's local area network (LAN), wide area network (WAN), data communications network, operating systems, and physical and virtual servers. Perform system monitoring and verify the integrity and availability of hardware, network, and server resources and systems. Review system and application logs and verify completion of scheduled jobs, including system backups. Analyze network and server resource consumption and control user access. Install and upgrade software and maintain software licenses. May assist in network modeling, analysis, planning, and coordination between network and data communications hardware and software.

Based on this job description on a scale of 1 to 5, how would you rate your skills and abilities in the following technical areas?

Each level of the scale is assigned a number and the numbers increase from left to right. The higher the number, the higher the skills.

Scale items:

- Computer Networks
- Operating Systems
- Servers (virtual or physical)
- Software
- System Security
- System Administration

Response scale:

1 Poor
2 Fair
3 Good
4 Very good
5 Excellent

B.3 Sample composition and self-hosting prevalence by socio-demographic groups

Table B.1: Share of sex × age × ethnicity groups in the population and our survey (in %)

Sex × Age	Asian		Black		Mixed		Other		White		Overall	
	Survey	Population	Survey	Population	Survey	Population	Survey	Population	Survey	Population	Survey	Population
Female												
18 – 28	0.5	0.5	1.3	1.3	0.3	0.3	0.9	0.8	6.0	6.1	9.0	9.1
28 – 38	0.6	0.6	1.2	1.2	0.3	0.2	0.8	0.8	5.8	5.8	8.6	8.5
38 – 48	0.5	0.5	1.1	1.2	0.2	0.2	0.7	0.6	6.6	6.6	9.1	9.2
48 – 58	0.5	0.4	1.2	1.2	0.1	0.2	0.5	0.4	7.4	7.3	9.7	9.5
58+	0.6	0.6	1.5	1.5	0.2	0.2	0.3	0.4	12.4	12.5	15.0	15.1
Overall	2.7	2.6	6.3	6.4	1.1	1.1	3.1	3.1	38.2	38.3	51.4	51.5
Male												
18 – 28	0.5	0.5	1.3	1.3	0.3	0.3	1.0	1.0	6.4	6.3	9.5	9.4
28 – 38	0.5	0.5	1.1	1.1	0.2	0.2	0.9	0.9	5.8	5.9	8.6	8.6
38 – 48	0.5	0.5	1.1	1.1	0.2	0.2	0.7	0.7	6.6	6.6	9.0	9.1
48 – 58	0.4	0.4	1.0	1.0	0.1	0.1	0.5	0.5	7.0	7.1	9.0	9.1
58+	0.5	0.4	1.1	1.1	0.1	0.1	0.3	0.3	10.4	10.4	12.4	12.4
Overall	2.3	2.3	5.6	5.6	1.0	1.0	3.3	3.3	36.3	36.4	48.6	48.5
Overall												
18 – 28	0.9	1.0	2.7	2.6	0.7	0.6	1.9	1.8	12.4	12.4	18.5	18.5
28 – 38	1.1	1.1	2.3	2.2	0.5	0.5	1.7	1.7	11.6	11.6	17.2	17.1
38 – 48	1.0	1.0	2.2	2.3	0.4	0.4	1.3	1.3	13.2	13.3	18.1	18.3
48 – 58	0.9	0.8	2.2	2.2	0.3	0.3	0.9	0.9	14.5	14.4	18.7	18.6
58+	1.1	1.0	2.5	2.5	0.3	0.3	0.7	0.7	22.9	22.9	27.4	27.5
Overall	5.0	4.9	11.9	12.0	2.1	2.1	6.4	6.4	74.6	74.7	100.0	100.0

Table B.2: Estimated self-hosting prevalence by sex, age and ethnicity (in %)

Sex × Age	Asian	Black	Mixed	Other	White	Over-all
female						
18 – 28	0.0	5.0 ± 9.8	0.0	0.0	2.2 ± 3.1	2.2 ± 2.5
28 – 38	0.0	12.5 ± 16.7	0.0	0.0	3.4 ± 3.9	3.9 ± 3.4
38 – 48	12.5 ± 24.5	0.0	0.0	0.0	6.1 ± 4.8	5.1 ± 3.7
48 – 58	0.0	0.0	0.0	14.3 ± 28.0	2.7 ± 3.0	2.7 ± 2.6
58+	0.0	0.0	0.0	0.0	1.6 ± 1.8	1.3 ± 1.5
Overall	2.6 ± 5.0	3.2 ± 3.5	0.0	2.0 ± 4.0	3.0 ± 1.4	2.9 ± 1.2
male						
18 – 28	14.3 ± 28.0	5.0 ± 9.8	40.0 ± 48.0	0.0	17.2 ± 7.7	14.3 ± 5.7
28 – 38	37.5 ± 35.9	41.2 ± 24.1	0.0	23.1 ± 23.8	11.9 ± 7.0	18.1 ± 6.6
38 – 48	14.3 ± 28.0	0.0	0.0	11.1 ± 21.8	12.6 ± 6.7	10.8 ± 5.4
48 – 58	0.0	26.7 ± 23.2	0.0	16.7 ± 32.7	22.1 ± 8.0	21.1 ± 7.0
58+	0.0	20.0 ± 21.0	0.0	20.0 ± 39.2	8.4 ± 4.4	9.3 ± 4.2
Overall	14.3 ± 11.5	18.0 ± 7.9	12.3 ± 14.7	12.6 ± 9.7	14.0 ± 2.9	14.3 ± 2.5
Overall						
18 – 28	7.1 ± 14.0	5.0 ± 6.9	19.7 ± 23.6	0.0	9.7 ± 4.2	8.3 ± 3.2
28 – 38	17.5 ± 16.8	26.9 ± 14.7	0.0	12.2 ± 12.6	7.6 ± 4.0	10.9 ± 3.7
38 – 48	13.3 ± 18.5	0.0	0.0	5.5 ± 10.7	9.3 ± 4.1	7.9 ± 3.2
48 – 58	0.0	12.5 ± 10.8	0.0	15.4 ± 21.3	12.2 ± 4.2	11.6 ± 3.6
58+	0.0	8.0 ± 8.4	0.0	9.2 ± 18.1	4.7 ± 2.2	4.9 ± 2.1
Overall	8.0 ± 6.0	10.0 ± 4.1	5.9 ± 7.1	7.4 ± 5.3	8.3 ± 1.6	8.4 ± 1.4

± indicates the lower and upper bounds of the 95% confidence intervals

Table B.3: Share of sex × age × ethnicity groups in our second survey (in %)

Sex × Age	Asian	Black	Mixed	Other	White	Overall
Female						
18 – 28	0.7	1.1	0.0	0.0	2.1	3.9
28 – 38	0.0	2.9	0.0	0.0	1.8	4.7
38 – 48	0.7	1.1	0.0	0.0	4.9	6.7
48 – 58	0.0	0.7	0.0	0.0	4.4	5.0
58 – 150	0.0	0.0	0.0	0.0	4.2	4.2
Overall	1.3	5.9	0.0	0.0	17.4	24.6
Male						
18 – 28	0.2	0.7	0.7	0.2	9.6	11.2
28 – 38	1.1	2.3	0.0	1.8	9.8	15.0
38 – 48	0.7	1.0	0.3	1.6	13.8	17.4
48 – 58	0.7	1.8	0.0	1.0	15.4	18.9
58 – 150	0.0	2.0	0.0	0.7	10.4	13.0
Overall	2.6	7.6	1.0	5.2	59.0	75.4
Overall						
18 – 28	0.8	1.8	0.7	0.2	11.7	15.1
28 – 38	1.1	5.2	0.0	1.8	11.5	19.7
38 – 48	1.3	2.1	0.3	1.6	18.7	24.1
48 – 58	0.7	2.4	0.0	1.0	19.8	23.9
58 – 150	0.0	2.0	0.0	0.7	14.6	17.2
Overall	3.9	13.5	1.0	5.2	76.4	100.0

B.4 Focus Groups Protocol

We detail the outline of the focus group protocol below. The order of the questions (from top to bottom) aligns with the chronological order in which they were posed during the focus groups.

Opening/ Introductory Questions

1. Have you ever heard of self-hosting before?
2. Where did you hear about it or in which context?
3. Do you engage in self-hosting yourself?
4. What do you host (which use case) and on which server structure?

Transition Question

This question was exclusively addressed to participants in the computer scientist focus group.

1. How would you define self-hosting?

Key Questions

1. What are the motives/reasons for rejecting non-self-hosted cloud services? ("push" factors)
2. What are motives/reasons to self-host? ("pull" factors)
3. What might have been relevant situations that led to self- hosting?
4. Can you think of other aspects or domains of life that are related to self-hosting (e.g., personality; individual characteristics)?
5. What are possible reasons not to self-host even though one would want to?
6. What are the motives/reasons to refuse self-hosting?

Additional Questions

This question was exclusively addressed to participants in the computer scientist focus group.

1. What are specific skills needed for self-hosting?

C

Information Demand in Critical Situation for Partially-Autonomous Vehicles

C.1 Codebooks

Table C.1 shows an comparative overview of the six codebooks. Each codebook contains high level codes about participants' perceptions of what happened in the scenario, feelings about the car's response to the scenario, reported next actions after the scenario, and information demand in the scenario.

C.2 Results of Correspondence Analysis

We provide the contingency tables of the key (Table C.2) and crash (Table C.5) scenario. Also we provide the results of the Correspondence Analysis for both scenarios including chi-square distances and relative inertias.

Table C.1: Comparative overview of the six codebooks. Mean inter-rater reliability of each codebook reported with Krippendorff's α (199).

Key - NO K's α: 0.842	Key - TM K's α: 0.795	Key - MI K's α: 0.815	Crash - NO K's α: 0.906	Crash - TM K's α: 0.771	Crash - MI K's α: 0.819
Perceptions of what happened					
Technical Malfunction	Technical Malfunction	Technical Malfunction	Technical Malfunction	Technical malfunction	Technical Malfunction
Human Failure	Human Failure	Hack/ Intrusion	Human Failure	Human Failure	Hack/ Intrusion
Security/ Safety Mechanism	Correct Description	Security/ Safety Mechanism	Technical Limits	Technical Limits	Human Failure
Design of Car	Incorrect Description	Correct Description	Correct Description	Correct Description	Minor Damage
Correct Description		Incorrect Description	Incorrect Description	Incorrect Description	Accident Avoided
		Uncertainty			
Feelings about car's response to situation					
Not Vehicle's Fault	Not Vehicle's Fault	Vehicle behaved appropriately	Negative Feeling	Negative Feeling	Negative Feeling
Positive Feeling	Positive Feeling	Positive Feeling	Surprised Feeling	Safety Hazard	Potentially lethal
Neutral Feeling	Neutral Feeling	Neutral Feeling	Neutral Feeling	Positive Feeling	Neutral Feeling
Negative Feeling	Negative Feeling	Negative Feeling	Positive Feeling	Surprised Feeling	Lost Trust
Safety Feeling	Safety Feeling	Safety Feeling	Safety Hazard	Worked properly	General wariness of AI
Improvement is needed	Car should react earlier	Insecurity Feeling	Driver should be attentive	Need for Fallback Mechanism	Need for Improvement
			Improvement is needed	Driver should be attentive	
Next Actions					
Key/ Battery Repair	Key/ Battery Repair	Key/ Lock Repair	Continue with Autopilot	Deactivate Autopilot	Call Manufacturer
Supervise Car/ Key	Use Key Analogously	Contact Manufacturer/ Dealer	Take over Manual Control	Take over Manual Control	Pull Over Car
Testing of Key	Use 2nd Key of Car	Contact Police	Inspect Car	Assess Car Damage	Turn off Autopilot
Adaption of One's Behaviour	Continue with Actions		Repair	Get Professional Help	Continue Driving Manually
Continue with Actions			Take care not to repeat HF	Report to Police	Never Use Autopilot Again
Check Manual			Don't use Autopilot	Report to Insurance	Monitor Autopilot
No Action				Report to Manufacturer	Fix Accident Scene
				Reflect on One's Responsibility	Call Police
				Get Rid of Car	Check Car for Damage
				Understand Car Mechanics/Tech	Get Car Towed
					Never Drive this Car Again
Information Demand					
Message/ Warning	Message/ Warning	Message/ Warning	Status of Malfunctioning Parts	Status of Malfunctioning Parts	What happened
Why	Status of Key	What happened?	Car's Diagnostic Report	Car's Diagnostic Report	Why it happened
Status of Key	Action Recommendations	Why did it happen?	Why it happened	Why it happened	Who is attacker
No Information Demand	Audio Reponse	How was it possible?	What happened	What happened	Preventative Meassures
Logs + Analytics	Visual Response	When did attack happen?	Car's Decision Process	Preventative Measures	Data for Fix
Usage Instructions	No Information Demand	Investigative Clues	Car's Parameters During Accident	Car's Parameters During Accident	Message/ Warning
Action Recommendations		Error Codes	Warning/ Message	Message/ Warning	When it happened
Troubleshooting		Status of Key	Audio Warning	Visual Warning	Action Recommendations
		Action Recommendations	Visual Warning	Audio Warning	Damage Report
		Preventative Measures	No Information Demand	Warn Before/ During Accident	No Information Demand
			Usage Instructions		
			Preventative Measures		
			Action Recommendations		
			Warn Before/During Accident		

Table C.2: Contingency table table of moderating factors and information demand codes of all conditions in the Key scenario.

	What	Why	Who	When	How	Status	Report	Decision	Param	Usage	Prevent	Message	Trblshoot	Recomm	Visual	Audio	Before	None
HiCrit	0	5	1	6	1	3	1	0	0	0	3	4	3	5	0	0	0	0
LessCrit	4	7	0	0	0	16	1	0	0	2	0	12	1	2	1	1	0	11
TM	2	5	0	0	0	12	1	0	0	1	0	7	0	2	1	0	0	5
MI	0	5	2	7	1	3	1	0	0	0	3	4	3	5	0	0	0	0
Human	0	3	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	4
Design	1	0	1	2	1	0	0	0	0	1	0	0	3	0	0	0	0	1
Threat	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pos	1	2	2	6	1	7	0	0	0	1	1	6	4	5	0	0	0	3
Neg	2	8	0	4	0	10	1	0	0	0	2	6	3	5	0	0	0	0
Neut	2	5	0	0	0	4	1	0	0	0	0	4	0	0	1	1	0	8

Table C.3: Chi-Square Distances of moderating factors and information demand codes of all conditions in the Key scenario.

	What	Why	Who	When	How	Status	Report	Usage	Prevent	Message	Trblshoot	Recomm	Visual	Audio	None	Total
HiCrit	1,352	0,054	0,155	3,597	0,669	1,649	0,155	0,676	3,889	0,147	0,614	1,949	0,338	0,225	3,606	19,077
LessCrit	0,979	0,167	1,225	5,106	0,817	2,024	0,041	0,49	1,838	1,179	1,76	1,718	0,245	0,857	3,05	21,496
TM	0,151	0,001	0,761	3,169	0,507	3,626	0,075	0,075	1,141	0,44	2,155	0,357	1,01	0,254	0,22	13,942
MI	1,437	0,009	2,287	5,365	0,567	1,951	0,11	0,718	3,43	0,256	0,457	1,574	0,359	0,239	3,831	22,592
Human	0,338	3,114	0,169	0,704	0,113	1,549	0,169	4,086	0,254	1,211	0,479	0,676	0,085	0,056	10,651	23,654
Design	0,789	1,408	2,945	1,424	5,241	1,937	0,211	2,945	0,317	1,514	9,634	0,845	0,106	0,07	0,014	29,4
Pos	0,255	2,221	1,679	1,919	0,37	0,04	0,824	0,038	0,045	0,002	1,188	0,881	0,412	0,275	0,442	10,591
Neg	0,041	0,858	0,866	0,042	0,577	0,534	0,021	0,866	0,378	0,007	0,121	0,68	0,433	0,289	4,62	10,334
Neut	0,74	0,489	0,549	2,289	0,366	0,213	0,37	0,549	0,824	0,001	1,556	2,197	1,916	3,645	8,776	24,479
Total	6,082	8,322	10,636	23,615	9,227	13,524	1,977	10,443	12,116	4,758	17,965	10,878	4,903	5,91	35,21	175,565

Table C.4: Relative inertias of moderating factors and information demand codes of all conditions in the Key scenario.

	What	Why	Who	When	How	Status	Report	Usage	Prevent	Message	Trblshoot	Recomm	Visual	Audio	None	Total
HiCrit	0,008	0	0,001	0,02	0,004	0,009	0,001	0,004	0,022	0,001	0,003	0,011	0,002	0,001	0,021	0,109
LessCrit	0,006	0,001	0,007	0,029	0,005	0,012	0	0,003	0,01	0,007	0,01	0,01	0,001	0,005	0,017	0,122
TM	0,001	0	0,004	0,018	0,003	0,021	0	0	0,006	0,003	0,012	0,002	0,006	0,001	0,001	0,079
MI	0,008	0	0,013	0,031	0,003	0,011	0,001	0,004	0,02	0,001	0,003	0,009	0,002	0,001	0,022	0,129
Human	0,002	0,018	0,001	0,004	0,001	0,009	0,001	0,023	0,001	0,007	0,003	0,004	0	0	0,061	0,135
Design	0,004	0,008	0,017	0,008	0,03	0,011	0,001	0,017	0,002	0,009	0,055	0,005	0,001	0	0	0,167
Pos	0,001	0,013	0,01	0,011	0,002	0	0,005	0	0	0	0,007	0,005	0,002	0,002	0,003	0,06
Neg	0	0,005	0,005	0	0,003	0,003	0	0,005	0,002	0	0,001	0,004	0,002	0,002	0,026	0,059
Neut	0,004	0,003	0,003	0,013	0,002	0,001	0,002	0,003	0,005	0	0,009	0,013	0,011	0,021	0,05	0,139
Total	0,035	0,047	0,061	0,135	0,053	0,077	0,011	0,059	0,069	0,027	0,102	0,062	0,028	0,034	0,201	1

Table C.5: Contingency table of moderating factors and information demand codes of all conditions in the Crash scenario.

	What	Why	Who	When	How	Status	Report	Decision	Param	Usage	Prevent	Message	Trblshoot	Recomm	Visual	Audio	Before	None
HiCrit	21	21	1	2	3	10	8	2	3	1	5	22	3	3	3	4	18	2
LessCrit	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TM	14	14	0	0	0	10	7	2	3	0	2	15	0	1	3	4	16	1
MI	7	7	1	2	3	0	0	0	0	0	3	6	3	2	0	0	1	1
Human	8	8	0	1	1	5	5	0	3	1	3	5	2	1	2	2	5	1
Design	4	4	0	0	0	4	2	0	1	0	1	2	0	1	1	1	2	0
Threat	4	4	0	0	1	4	3	1	1	0	2	7	0	0	1	1	8	1
Pos	1	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
Neg	19	19	1	2	3	8	6	1	2	0	5	21	2	3	3	4	17	2
Neut	2	2	0	0	0	0	1	0	1	0	0	2	1	0	0	1	2	0

Table C.6: Chi-Square Distances of moderating factors and information demand codes of all conditions in the Crash scenario.

	What	Why	Who	When	How	Status	Report	Decision	Param	Usage	Prevent	Message	Trblshoot	Recomm	Visual	Audio	Before	None	Total
HiCrit	0,002	0,002	0,062	0,018	0,007	0,076	0,012	0,018	0,112	0,444	0,038	0,071	0,007	0,007	0,042	0,039	0	0,003	0,962
TM	0,016	0,016	0,543	1,268	1,992	0,753	0,25	0,423	0,085	0,362	0,855	0,018	1,992	0,494	0,177	0,276	0,983	0,139	10,644
MI	0,312	0,312	2,916	4,56	6,325	2,976	2,268	0,496	0,992	0,142	1,536	0,019	6,325	1,911	0,921	1,205	3,094	0,331	36,642
Human	0,014	0,014	0,313	0,1	0,019	0,087	0,827	0,73	1,622	3,001	0,299	1,342	0,633	0,019	0,306	0,029	0,672	0,033	10,059
Design	0,039	0,039	0,136	0,317	0,498	2,316	0,21	0,317	0,211	0,091	0,003	0,726	0,498	0,506	0,288	0,069	0,404	0,362	7,03
Threat	0,658	0,658	0,224	0,524	0,038	0,234	0,154	0,433	0,002	0,15	0,117	0,172	0,823	0,823	0,001	0,058	1,561	0,269	6,9
Pos	0,217	0,217	0,024	0,055	0,087	1,355	0,252	16,2	0,11	0,016	0,165	0,63	0,087	0,087	0,102	0,134	0,543	0,063	20,342
Neg	0,009	0,009	0,132	0,086	0,077	0,316	0,276	0,241	0,482	0,465	0,003	0,314	0,121	0,077	0	0,001	0,059	0,011	2,68
Neut	0,006	0,006	0,071	0,165	0,26	0,992	0,079	0,165	1,355	0,047	0,496	0,006	2,108	0,26	0,307	0,892	0,084	0,189	7,49
Total	1,276	1,276	4,422	7,092	9,303	9,106	4,327	19,02	4,972	4,717	3,512	3,299	12,59	4,184	2,144	2,702	7,401	1,4	**102,748**

Table C.7: Relative Inertias of moderating factors and information demand codes of all conditions in the Crash scenario.

	What	Why	Who	When	How	Status	Report	Decision	Param	Usage	Prevent	Message	Trblshoot	Recomm	Visual	Audio	Before	None	Total
HiCrit	0	0	0,001	0	0	0,001	0	0	0,001	0,004	0	0,001	0	0	0	0	0	0	0,009
TM	0	0	0,005	0,012	0,019	0,007	0,002	0,004	0,001	0,004	0,008	0	0,019	0,005	0,002	0,003	0,01	0,001	0,104
MI	0,003	0,003	0,028	0,044	0,062	0,029	0,022	0,005	0,01	0,001	0,015	0	0,062	0,019	0,009	0,012	0,03	0,003	0,357
Human	0	0	0,003	0,001	0	0,001	0,008	0,007	0,016	0,029	0,003	0,013	0,006	0	0,003	0	0,007	0	0,098
Design	0	0	0,001	0,003	0,005	0,023	0,002	0,003	0,002	0,001	0	0,007	0,005	0,005	0,003	0,001	0,004	0,004	0,068
Threat	0,006	0,006	0,002	0,005	0	0,002	0,001	0,004	0	0,001	0,001	0,002	0,008	0,008	0	0,001	0,015	0,003	0,067
Pos	0,002	0,002	0	0,001	0,001	0,013	0,002	0,158	0,001	0	0,002	0,006	0,001	0,001	0,001	0,001	0,005	0,001	0,198
Neg	0	0	0,001	0,001	0,001	0,003	0,003	0,002	0,005	0,005	0	0,003	0,001	0,001	0	0	0,001	0	0,026
Neut	0	0	0,001	0,002	0,003	0,01	0,001	0,002	0,013	0	0,005	0	0,021	0,003	0,003	0,009	0,001	0,002	0,073
Total	0,012	0,012	0,043	0,069	0,091	0,089	0,042	0,185	0,048	0,046	0,034	0,032	0,123	0,041	0,021	0,026	0,072	0,014	1

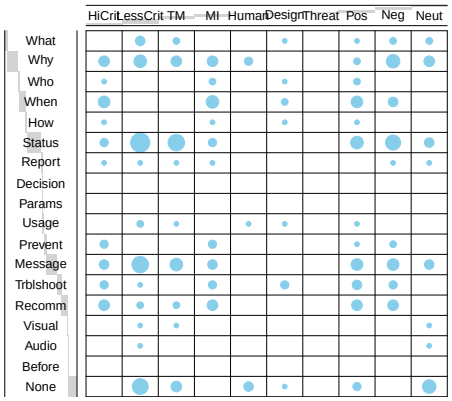


Figure C.1: *Key* scenario

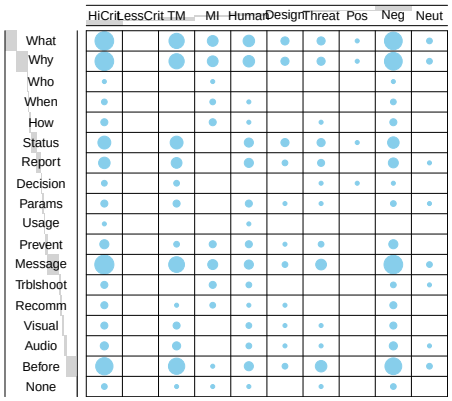


Figure C.2: *Crash* scenario

Figure C.3: Balloon plot representation of the contingency table table of *moderating factors* and *information demand* codes. Bigger dots indicate larger chi-square distances. Refer to tables C.3 and C.6 in the Appendix for exact results.

D

Security and Privacy of Content Creators in Pakistan

D.1 Screening Survey

We're recruiting content creators and influencers for a paid research study about your online experiences. The survey that follows will ask you a series of questions to evaluate if you qualify for our interview study. If you are eligible, we will reach out to you to schedule a 1 hour interview, where you will be compensated with a Rs 5000. We are a group of researchers from *INSTITUTE* in collaboration with the *INSTITUTE* and *INSTITUTE* in *LOCATION*. If you have logistical questions please email *EMAIL*. Your participation in our research is voluntary. You may decline further participation and abort this survey, at any time. However, you will forfeit the possibility to be invited to the interview by doing so. The collected data will be used for scientific purposes only.

Personal Information

1. Please enter your email address
2. What year were you born in?
3. Are you 18 years old or older?
4. Where do you live? [Pakistan, Outside Pakistan]
5. What is your gender? [Man, Woman, Non-Binary, Prefer to self-describe, Prefer to not disclose]
6. What province do you belong to? [Punjab, Sindh, KPK, Balochistan, Gilgit/Balistan, Federal Territory, Prefer not to say]
7. Do you have any disability? [Yes, No, Prefer not to respond]
8. What is the highest level of school you have completed or the highest degree you have received? [Less than a high school degree, High school graduate (or FA/Fsc or O/A level), Some college but no degree, Bachelor's degree, Advanced degree (Master's, doctorate), Prefer not to answer]

9. Which one of the following includes your total HOUSEHOLD monthly income for last year, before taxes? [Less than 20,000, 20,000-40,000, 40,000-50,000, 50,000-60,000, 60,000-80,000, 80,000-100,000, 100,000-150,000, 150,000-200,000, 200,000 or more, Prefer not to answer]

Platforms

- What Platforms do you create content (eg. Post) on? Select all that apply. [Instagram, Twitter, Youtube, Reddit, Facebook, Tiktok, Snapchat, Twitch, Other (free text)]
- Have you ever received compensation for the content that you create on any of these platforms? Check all that apply. [Yes, I have received payment from a brand or sponsor. | Yes, I have received comission or tip from a follower. | Yes, I have received payment from a platform. (e.g., payment from the TikTok Creator Fund or the Youtube Partnership program)) | Yes, I have received a different kind of payment. | No, I have never received compensation for the content I create.]
- Please describe what kind of content you create. Feel free to describe this overall, or for each platform that you selected above (free text)
- For each of the selected social media applications, how many times do you post a week? Please mention it in the format: AppName : NumberOfPostsPerWeek [free text]
- For each of the selected social media applications, estimate your follower count? Please mention it in the format: AppName : FollowerCount [free text]

Harassement

- Have you ever experienced bullying/harassment on any of your social media platforms (by your followers or fellow content creators)? [Yes I have, No I have not]

Call back

- Would you be willing to share your experience with us in an interview? To reiterate, all data will be anonymous and you will be compensated for your time [Yes, I'd be willing to share my experiences, I do not feel comfortable sharing those experiences]

D.2 Interview Guideline

Introduction Hello. Thank you for your interest in participating in our study! My name is $NAME$ and I am a research assistant at $INSTITUTION$. This project is in collaboration with $ANOTHERINSTITUTION$. We're trying to mitigate the privacy and security issues of content creators in Pakistan. To quickly go over how today's study is going to work, we will start with a brief verification of your status as a content creator. Then I'm going to ask you questions about your experiences with online content creation work. Our conversation should not take more than an hour. Feel free to let me know if you don't want to answer any questions, and we'll move on to the next question or we can stop the study, just let me know. If you have any questions about me or the study, we can talk about them first, then I'll ask your consent to start the recording and we'll start. [pause for questions] Are you ok to start the recording?

[Potential administrative questions]:

- What will the study be used for?
  The purpose of this study is to understand digital safety and security needs of Influencers. Our findings can be used to identify areas of improvement and to offer context-dependent security advice for Influencers.

- How secure is my data / how will my data be anonymized / etc?
  The recording of this interview is uploaded to our secure INSTITUTION cloud. We are legally bound by research ethics and data privacy under the NAMEOFREGULATORY AUTHORITY Regulations. The anonymity and security of our participants is of the highest importance to us, but as mentioned in the consent form that you saw and agreed to in the screening survey, you have the right to stop the interview and withdraw your data at any time if you change your mind. The interviews and emails to you regarding it will be deleted within 12 months and only the anonymised data will be retained (apart from your email address if you choose to be on the list to be kept informed of publications). If we come across anything that compromises your anonymity, that piece of information will be redacted and replaced with a generalized version of what you say.

(1) Verification Section Before we begin, we'd like to start by confirming your role as a content creator to make sure that you match our criteria for the project.

- Please list all the platforms you use to post and create content.
- For each platform, how many posts do you create in a week?
- How many sponsored posts do you create for each platform in a week?
- For each platform, give an estimate of your follower/subscriber count.
- What ways do you employ for each platform to keep your reach desirable?
- When did you start receiving compensation for your social media content? And are you still actively receiving compensation for your social media content?
- How much of your annual income would you estimate is earned through the content you post on social media?
- Do you have multiple accounts on the same platform that generate your revenue?
- Could you give us an aggregate percentage of how much income you generate from all your accounts combined?
- Do you have any civilian jobs? How much do they contribute to your annual income?

(2) Defining their Work We would like to begin with a few baseline questions about your experience with content creation.

- How would you define your job title? Influencer, content creator, etc? *(this is how we determine what to call them moving forward)*
- You mentioned that you are active on [platforms]:
- What kind of influencing content do you create on [platforms]?
- Can you describe or characterize the audience or niche your influencing content caters to?
- Can you tell us what age demographic your followers fall in?
- Could you describe a typical day/week for you in terms of your influencing/content creation work?
- Does anyone help manage your online presence on any of the platforms? Who?
  - What aspects of your online presence do they help with?
  - Do they have access to any of your creator accounts? If so, which and why?
  - Who has access to your personal gadgets (mobile tab etc)?
  - How long have they been helping you manage your online presence?
  - Why did you primarily hire them?
  - Have they worked for more than what they were initially hired for? [If yes] Why is that?
- Do you work with others in order to create content or do your influencing work? [prompts: someone who takes pictures, someone who replies to messages on your behalf, editor]
  - (if they talk about collaboration) What type of content do you and the person(s) you mentioned collaborate to create together?

  - How did you find someone to help you / evaluate whether you wanted this particular person? Can you tell me more about your experience with this?
  - Do you share your credentials with this person / how do you collaborate?
- How does being a <platform> creator/influencer impact your life? Both positively and negatively. [Prompts: sleeping routines, relationships, civilian jobs, family life, finances, education, mental health]
- Are there certain people in your life who you don't want to know about your work? (If yes) How do you try to prevent them from finding out?
- Are there any specific topics you avoid talking about on your platforms, If yes then what and why?
- What are the main challenges you would say you face as an influencer?

(3) Safety and Negative Experiences We would now like to discuss safety while influencing and some of the more negative experiences you've had doing influencing work if any.

- What does safe or being/feeling safe mean to you as an influencer content creator?
- What are the things that make you feel vulnerable or unsafe (both in the digital and real world)?
- What does privacy mean to you as an influencer content creator? What would be private for you and would be ok to make public? (prompt only if the answer to the firs question is vague)
- As a content creator/influencer, would you say your work affects you more physically or mentally?
- As a content creator, what are your biggest digital and physical concerns? (probe if need be: including hate and harassment)
  - What are you worried will happen as a result of [concerns]?
  - Who are you most worried might do these things to you?
  - What have you done to address this? How effective has it been?
  - Did you reach out to anyone? How did they help you?
- Do you do anything to maintain your safety online as an influencer that you didn't do before you started content creation? How did you learn these strategies?
- Have you ever had negative experiences doing influencing work using [platform(s)]? What about any negative experiences offline, "in real life"?
  [digital and physical safety prompts only if needed: Privacy and Security, Deep or cheap fakes, Misinformation or conspiracy theories, Hate/ harassment,Account takeover - Hijacking, Monetization redirected, Impersonation, Stalking, Bullying, trolling, Personal content being leaked, Excessive negative reviews]
  [Follow up if not] Do you know anyone who has ever experienced anything negative, online or in real life, while doing influencer work?
  - If yes, please describe this/these incident(s). Remember that you do not have to mention any personal or sensitive information.
  - Which platform(s)? Was this experience happening across platforms or just once?
  - Can you tell us who did this to you? Do you think you know the reason they did this with you?
  - Can you tell us how you responded to it? What happened as a result of what you did?
  - Would you say you had to go through these experiences because of your personal traits? Such as your beliefs, faith, or characteristics? Why?
  - Did these experiences impact the people around you? Your family and friends for example? How often do you go to them for counsel on your online life?
  - How have these experiences informed your impressions of these platforms?
  - Do you feel you will ever fall in the same situation again? Why is that? Will the wisdom you have gained as a result of previous experiences help you mitigate the

problem as fast as possible?

- Considering your experiences as a creator across platforms, can you think of any experiences with digital or physical threats that you chose to ignore? What happened?
- As a creator, do you think there is anything unique when it comes to hate and harassment (digital or physical) one might face? Does this happen across platforms?
- Do you think digital and physical safety concerns or experiences are an inevitable part of being a creator, including hate and harassment? Why/why not?
- Can you imagine a world as a creator where you would NOT experience those concerns? What would that be like?
  - (if they can) What aspects of that world would you cherish the most?
  - (if they can't) Why does that feel impossible to imagine?

(4) Defense and Mitigation Strategies

- Earlier, we talked about experiences that cross a line. What is that line for you?
  - Do you have different expectations for the other platforms you're on? Why/why not?
  - Which platforms do you feel are not doing a good job of addressing these concerns?

You mentioned some specific examples above:

- Have you generally changed your behavior/activities/thinking as a creator due to your digital or physical safety concerns or experiences - including hate and harassment?
- How did you learn these strategies to increase your safety? When did you start adopting them?

[Probe 1: Offline]

- Have you changed any of your behaviors in real life, in the physical world because of these concerns or experiences?
- Can you tell me about a time that happened? Why did you decide to change this?
- Was this a result of one particular event or multiple Experiences?

[Probe 2: Online]

- What about online, have you changed your behaviors on social media platforms?
- Can you tell me about a time that happened? Why did you decide to change this?
- Was this a result of one particular event or multiple experiences?

[Probe 3: Content Change]

- Did you change anything about the content you make?
- How do you feel like your content has changed [If it hasn't changed] Have you ever considered changing the things you post, or the information in them?
- Do you wish you didn't have to change your content because of these concerns?
- How important a role does safety play in determining how you plan future content (vs other concerns, what gets views, the direction you want to develop your brand, etc.)

Mindsets

- Have you ever left or considered leaving a platform that you created on due to these concerns or experiences? Can you tell us more about the thought process that led you to take those steps? [Probe] What was the final straw?
- Was there anything else you considered doing before leaving the platform?
- When you left this platform, did you consider leaving any others? Why/why not? [if they stayed on other platforms] why did you decide to stay on some platforms but not others? Was there anything special about the platforms you stayed on?
- How would leaving platforms impact your income? Were finances a concern when thinking about leaving?
- How has leaving the platform affected how safe you feel, if at all?
- What is the situation now? Did you return to the platform? If so, why?

Resources

- Are there any platforms you feel do a good job of protecting you from digital or physical safety threats or hate and harassment? Why?
- Are there any platforms that you feel don't do a good job of protecting you from these threats?
- What are some things you do to stay safe online? Where did you learn how to do that? How effective do you think it is?
- What are the best things creators can do to protect themselves from digital safety threats?
- What do you do when you have a safety concern? Who or what do you turn to? Are there any resources you use when you have safety concerns?

  [can probe on]
  - 2FA (if they use, probe) What type - (e.g., text codes, app generator like Authenticator, hardware security key), On what accounts
  - Password manager (if they use, probe) What type - Browser's built-in password manager (e.g. Chrome, Firefox, ...), Standalone app (e.g. 1Password, LastPass, ...)
  - Security checkup
  - Different password across creator accounts
  - Moderation tools (if they use, probe) What type - Word filter lists, Blocking, Use of moderators
  - Platform guidelines
  - Lawyers or experts to help navigate experiences
  - Documentation of threats online - record keeping
  - Advice guides (e.g., Security Planner or EFF's Surveillance Self Defense)
- Do you feel you're well equipped to handle these types of threats yourself?
- Have you had experiences with digital or physical threats including hate and harassment that caused you to seek support or resources? If so, what happened? What made you decide to seek support? How did you decide to do that?
- Have you ever come across a situation where you felt it was necessary to get the authorities involved? Did you involve them? Why/why not? [if yes] how helpful were they? Were they well equipped to help you?
- Are there any communities or groups of people you turn to when you have safety concerns?
  - Can you describe the group?
  - How did you come across them?
  - How helpful do you feel like they are?
  - Are there any rules associated with being in the group? Who is allowed to join?
  - Is there anything that could help make it easier for you to find support groups?
- Would you expect a platform to help you find other people to talk to?

Hypotheticals

- Are there any resources or support that you think could help you feel safe and supported moving forward?
- Are there any suggestions or advice people gave you that you didn't take but wish you did? (precautions, digital steps)
- Thinking about what we discussed today, is there any information that you wish you knew when you first started creating? How would you have wanted to learn about this information when you were first starting out? What would you have paid attention to?
- Imagine a world where there's a resource for creators on digital and physical safety. If that existed, what would you want it to cover?
- Ignoring technical constraints, what would wish for [Magic wand question]?

Debriefing Is there anything else you would like to share with us that you have not gotten a chance to mention yet? Thank you very much for speaking with me today! Two final questions for you: Do you know someone who might be eligible and willing to take part in the study? Are you interested in being notified of the results of the study + any publications or presentations made about this study? Thanks again. I'm going to turn off the recording and then I'll check your payment method with you.

Bibliography

Author's Papers for this Thesis

[P1] **Gröber, L.**, Mrowczynski, R., Vijay, N., Muller, D. A., Dabrowski, A., and Krombholz, K. To Cloud or not to Cloud: A Qualitative Study on Self-Hosters' Motivation, Operation, and Security Mindset. In: *32nd USENIX Security Symposium (USENIX Security 23)*. 2023, 2491–2508.

[P2] **Gröber, L.**, Lenau, S., Weil, R., Groben, E., Schilling, M., and Krombholz, K. Towards Privacy and Security in Private Clouds: A Representative Survey on the Prevalence of Private Hosting and Administrator Characteristics. In: *33rd USENIX Security Symposium (USENIX Security 24)*. 2024, 6057–6074.

[P3] **Gröber, L.**, Fassl, M., Gupta, A., and Krombholz, K. Investigating Car Drivers' Information Demand after Safety and Security Critical Incidents. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI. ACM, Yokohama, Japan, May 2021, 1–17.

[P4] **Gröber, L.**, Arshad, W., Shanza, Goetzen, A., Redmiles, E. M., Mustafa, M., Krombholz, K., et al. "I chose to fight, be brave, and to deal with it": Threat Experiences and Security Practices of Pakistani Content Creators. In: *33rd USENIX Security Symposium (USENIX Security 24)*. 2024.

Other Papers of the Author

[S1] Anell, S., **Gröber, L.**, and Krombholz, K. End User and Expert Perceptions of Threats and Potential Countermeasures. In: *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW 20)*. IEEE. 2020, 230–239.

[S2] Fassl, M., **Gröber, L.**, and Krombholz, K. Exploring User-Centered Security Design for Usable Authentication Ceremonies. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI 21)*. 2021, 1–15.

[S3] Fassl, M., **Gröber, L.**, and Krombholz, K. Stop the Consent Theater. In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (alt.CHI 21)*. 2021, 1–7.

[S4] Hashmi, S., Sarfaraz, R., **Gröber, L.**, Javed, M., and Krombholz, K. Understanding the Security Advice Mechanisms of Low Socioeconomic Pakistanis. In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI 25)*. 2025.

[S5] Roth, S., **Gröber, L.**, Backes, M., Krombholz, K., and Stock, B. 12 Angry Developers–A Qualitative Study on Developers' Struggles with CSP. In: *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security (CCS 21)*. 2021, 3085–3103.

[S6] Roth, S., **Gröber, L.**, Baus, P., Krombholz, K., and Stock, B. Trust Me If You Can–How Usable is Trusted Types in Practice? In: *33rd USENIX Security Symposium (USENIX Security 24)*. 2024, 6003–6020.

Other references

[1] (pseudonym), awesome-selfhosted. *"awesome-selfhosted"*. Github, `https://github.com/awesome-selfhosted/awesome-selfhosted`, accessed 2023-02-07.

[2] (pseudonym), u/muchTasty. *r/selfhosted: Beginner guide: How to secure your self-hosted services*. Reddit, `https://www.reddit.com/r/selfhosted/comments/pufhs0/beginner_guide_how_to_secure_your_selfhosted/`, accessed 2023-02-07.

[3] 24 News HD. *Jannat Mirza returns to Pakistan after long stay in Tokyo*. 24 NEWS HD, `https://www.24newshd.tv/10-Nov-2020/jannat-mirza-returns-to-pakistan-after-long-stay-in-tokyo`, last accessed: 05.04.2023. 2020.

[4] Abbas, S. K., Hassan, H. A., Asif, J., and Zainab, F. How income level distribution responds to poverty: Empirical evidence from Pakistan. *Global Scientific Journals* 6, 3 (2018), 131–142.

[5] Abu-Salma, R., Sasse, M. A., Bonneau, J., Danilova, A., Naiakshina, A., and Smith, M. Obstacles to the adoption of secure communication tools. In: *2017 IEEE Symposium on Security and Privacy (SP)*. IEEE. 2017, 137–153.

[6] Acar, Y., Fahl, S., and Mazurek, M. L. You are not your developer, either: A research agenda for usable security and privacy research beyond end users. *2016 IEEE Cybersecurity Development (SecDev)* (2016), 3–8.

[7] Acquisti, A. and Gross, R. Imagined communities: Awareness, information sharing, and privacy on the Facebook. In: *International workshop on privacy enhancing technologies*. Springer. 2006, 36–58.

[8] Adams, A. and Sasse, M. A. Users are not the enemy. *Commun. ACM* 42, 12 (Dec. 1999), 40–46.

[9] Addae, J. H., Brown, M., Sun, X., Towey, D., and Radenkovic, M. Measuring attitude towards personal data for adaptive cybersecurity. *Information & Computer Security* 25, 5 (2017), 560–579.

[10] Adkinson-Orellana, L., Rodríguez-Silva, D. A., Gil-Castiñeira, F., and Burguillo-Rial, J. C. Privacy for google docs: Implementing a transparent encryption layer. In: *International Conference on Cloud Computing*. 2010.

[11] *Advanced Data Protection Control (ADPC)*. https://www.dataprotectioncontrol.org/, last accessed: 01.12.2024. 2022.

[12] *advertisingIdentifier*. https://developer.apple.com/documentation/adsupport/asidentifiermanager/advertisingidentifier, last accessed: 11.18.2024.

[13] Agarwal, R. and Prasad, J. A conceptual and operational definition of personal innovativeness in the domain of information technology. *Information systems research* 9, 2 (1998), 204–215.

[14] Aharony, N. An exploratory study on factors affecting the adoption of cloud computing by information professionals. *The Electronic Library* (2015).

[15] Ahmed, S. I., Haque, M. R., Chen, J., and Dell, N. Digital privacy challenges with shared mobile phone use in Bangladesh. *Proceedings of the ACM on Human-computer Interaction* 1, CSCW (2017), 1–20.

[16] Ahmed, S. I., Haque, M. R., Chen, J., and Dell, N. Digital Privacy Challenges with Shared Mobile Phone Use in Bangladesh. *Proceedings of the ACM Conference on Human-Computer Interaction* 1, CSCW (Dec. 2017).

[17] Ahmed, T., Hoyle, R., Connelly, K., Crandall, D., and Kapadia, A. Privacy concerns and behaviors of people with visual impairments. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2015, 3523–3532.

[18] Akaike, H. Information Theory and an Extension of the Maximum Likelihood Principle. In: *Proceedings of the 2nd International Symposium on Information Theory*. Akadémiai Kiado, Budapest, 1973, 267–281.

[19] Akhawe, D. and Felt, A. P. Alice in warningland: A large-scale field study of browser security warning effectiveness. In: *Presented as part of the 22nd USENIX Security Symposium (USENIX Security 13)*. USENIX Association, Berkeley, 2013, 257–272.

[20] Akter, M., Alghamdi, L., Kropczynski, J., Lipford, H. R., and Wisniewski, P. J. It takes a village: A case for including extended family members in the joint oversight of family-based privacy and security for mobile smartphones. In: *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 2023, 1–7.

[21] Akter, M., Godfrey, A. J., Kropczynski, J., Lipford, H. R., and Wisniewski, P. J. From parental control to joint family oversight: Can parents and teens manage mobile online safety and privacy as equals? *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–28.

[22] Akter, M., Tabassum, M., Miazi, N. S., Alghamdi, L., Kropczynski, J., Wisniewski, P. J., and Lipford, H. Evaluating the impact of community oversight for managing mobile privacy and security. In: *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*. 2023, 437–456.

[23] Alaoui, L. and Fons-Rosen, C. Know when to fold'em: The flip side of grit. *European Economic Review* 136 (2021), 103736.

[24] Ali, F. Z. *Gender Discourse and Transphobia in Pakistan's Digital Sphere*. https://blogs.lse.ac.uk/southasia/2022/12/19/gender-discourse-and-transphobia-in-pakistans-digital-sphere/. 2022.

[25] Allison-Hope, D. *Human Rights Due Diligence of Meta's Impacts in Israel and Palestine in May 2021*. BSR, https://www.bsr.org/en/blog/human-rights-due-diligence-of-meta-impacts-in-israel-and-palestine-may-2021, last accessed: 31.05.2023. 2022.

[26] Ammari, T., Nofal, M., Naseem, M., and Mustafa, M. Moderation as Empowerment: Creating and Managing Women-Only Digital Safe Spaces. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–36.

[27] Andre, P., Chetwani, S., Dornadula, K., and Teckchandani, A. *Automotive security and monitoring system*. US Patent App. 10/073,725. Aug. 2003.

[28] Are, C. The Shadowban Cycle: an autoethnography of pole dancing, nudity and censorship on Instagram. *Feminist Media Studies* 22, 8 (2022), 2002–2019.

[29] Are, C. An autoethnography of automated powerlessness: lacking platform affordances in Instagram and TikTok account deletions. *Media, Culture & Society* 45, 4 (2023), 822–840.

[30] Are, C. and Briggs, P. The emotional and financial impact of de-platforming on creators at the margins. *Social Media+ Society* 9, 1 (2023), 20563051231155103.

[31] Asgharpour, F., Liu, D., and Camp, L. J. Mental models of security risks. In: *Financial Cryptography and Data Security: 11th International Conference, FC 2007, and 1st International Workshop on Usable Security, USEC 2007, Scarborough, Trinidad and Tobago, February 12-16, 2007. Revised Selected Papers 11*. Springer. 2007, 367–377.

[32] Ashraf, A., König, C. J., Javed, M., Mustafa, M., et al. " Stalking is immoral but not illegal": Understanding Security, Cyber Crimes and Threats in Pakistan. In: *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*. 2023, 37–56.

[33] Avila Pinto, R. Digital sovereignty or digital colonialsim. *SUR-Int'l J. on Hum Rts.* 15 (2018), 15.

[34] Axtell, B. and Munteanu, C. Back to Real Pictures: A Cross-generational Understanding of Users' Mental Models of Photo Cloud Storage. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* (2019).

[35] Banyard, P., Grayson, A., and Orne, M. Demand characteristics. *Introducing psychological research: Sixty studies that shape psychology* (1996), 395–401.

[36] Barakat, H. and Redmiles, E. M. Community under surveillance: Impacts of marginalization on an online labor forum. In: *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 16. 2022, 12–21.

[37] Barbrook, R. and Cameron, A. The californian ideology. *Science as culture* 6, 1 (1996), 44–72.

[38] Barlow, J. P. *A Declaration of the Independence of Cyberspace*. `https://www.eff.org/cyberspace-independence`, last accessed: 01.12.2024. 1996.

[39] Barnes, S. B. A privacy paradox: Social networking in the United States. *First Monday* (2006).

[40] Barroso Domínguez, A. et al. *¿ Redes de apoyo, comunicación y sororidad entre las mujeres creadoras de contenido erótico en OnlyFans?* `http://hdl.handle.net/10651/64197`. 2022.

[41] Barwulor, C., McDonald, A., Hargittai, E., and Redmiles, E. M. "Disadvantaged in the American-dominated internet": Sex, Work, and Technology. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021, 1–16.

[42] Begum Sadaquat, M. and Sheikh, Q.-t.-a. A. Employment situation of women in Pakistan. *International Journal of Social Economics* 38, 2 (2011), 98–113.

[43] Bellini, R., Tseng, E., Warford, N., Daffalla, A., Matthews, T., Consolvo, S., Woelfer, J. P., Kelley, P. G., Mazurek, M. L., Cuomo, D., et al. Sok: Safer digital-safety research involving at-risk users. In: *2024 IEEE Symposium on Security and Privacy (SP)*. IEEE. 2024, 635–654.

[44] Bellotti, V. and Edwards, K. Intelligibility and accountability: human considerations in context-aware systems. *Human–Computer Interaction* 16, 2-4 (2001), 193–212.

[45] Beres, N. A., Frommel, J., Reid, E., Mandryk, R. L., and Klarkowski, M. Don't you know that you're toxic: Normalization of toxicity in online gaming. In: *Proceedings of the 2021 CHI conference on human factors in computing systems*. 2021.

[46] Berners-Lee, T., Cailliau, R., Luotonen, A., Nielsen, H. F., and Secret, A. The world-wide web. *Communications of the ACM* 37, 8 (1994), 76–82.

[47] Binkhorst, V., Fiebig, T., Krombholz, K., Pieters, W., and Labunets, K. 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*. 2022.

[48] Birks, M. and Mills, J. *Grounded theory: A practical guide*. ISBN 9781446295786. Sage, 2015.

[49] Blackwell, L., Hardy, J., Ammari, T., Veinot, T., Lampe, C., and Schoenebeck, S. LGBT parents and social media: Advocacy, privacy, and disclosure during shifting social movements. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2016, 610–622.

[50] Bly, S., Schilit, B., McDonald, D. W., Rosario, B., and Saint-Hilaire, Y. Broken expectations in the digital home. In: *CHI'06 extended abstracts on Human factors in computing systems*. 2006.

[51] Böhme, R. and Köpsell, S. Trained to accept? A field experiment on consent dialogs. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 2010, 2403–2406.

[52] Bohnsack, R., Pfaff, N., and Weller, W., eds. *Qualitative analysis and documentary method in international educational research*. 2010, 369.

[53] BOL News. *TikTok star Jannat Mirza Does Not Want to Hear Criticism From Bushra Ansari*. BOL NEWS, https://www.bolnews.com/entertainment/2021/06/tiktok-star-jannat-mirza-does-not-want-to-hear-criticism-from-bushra-ansari, last accessed: 05.04.2023. 2021.

[54] Borges, H., Hora, A., and Valente, M. T. Understanding the factors that impact the popularity of GitHub repositories. In: *2016 IEEE international conference on software maintenance and evolution (ICSME)*. IEEE. 2016, 334–344.

[55] Brackenbury, W., McNutt, A., Chard, K., Elmore, A., and Ur, B. KondoCloud: Improving information management in cloud storage via recommendations based on file similarity. In: *The 34th Annual ACM Symposium on User Interface Software and Technology*. 2021.

[56] Braun, V. and Clarke, V. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[57] Brown, B. Studying the internet experience. *HP laboratories technical report HPL* 49 (2001).

[58] Brush, A. IT@ home: Often best left to professionals. In: *Position paper for the CHI 2006 Workshop on IT@ Home*. 2006.

[59] Büchi, M., Just, N., and Latzer, M. Caring is not enough: the importance of Internet skills for online privacy protection. *Information, Communication & Society* 20, 8 (2017), 1261–1278.

[60] Buf, D.-M. and Ștefăniță, O. Uses and gratifications of YouTube: A comparative analysis of users and content creators. *Romanian Journal of Communication and Public Relations* 22, 2 (2020), 75–89.

[61] Bund, C. *Digitale Souveränität*. https://www.cio.bund.de/Webs/CIO/DE/digitale-loesungen/digitale-souveraenitaet/digitale-souveraenitaet-node.html, last accessed: 11.05.2024.

[62] Burkert, H. et al. Privacy-enhancing technologies: Typology, critique, vision. *Technology and privacy: The new landscape* 125 (1997).

[63] Carroll, J. M., Rosson, M. B., and Zhou, J. Collective efficacy as a measure of community. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 2005, 1–10.

[64] Castells, M. Informationalism, networks, and the network society: a theoretical blueprint. In: *The Network Society*. Ed. by Castells, M. 2004.

[65] Chang, J. B., Lusk, J. L., and Norwood, F. B. How closely do hypothetical surveys and laboratory experiments predict field behavior? *American Journal of Agricultural Economics* 91, 2 (2009), 518–534.

[66] Charmaz, K. C. *Constructing Grounded Theory*. ISBN 9780857029140. Sage, 2014.

[67] Chatterjee, R., Doerfler, P., Orgad, H., Havron, S., Palmer, J., Freed, D., Levy, K., Dell, N., McCoy, D., and Ristenpart, T. The spyware used in intimate partner violence. In: *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE. 2018, 441–458.

[68] Chaudhary, S., Zhao, Y., Berki, E., Valtanen, J., Li, L., Helenius, M., and Mystakidis, S. A Cross-Cultural and Gender-Based Perspective for Online Security: Exploring Knowledge, Skills and Attitudes of Higher Education Students. *IADIS International Journal on WWW/Internet* 13, 1 (2015).

[69] Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Van der Kaap-Deeder, J., Duriez, B., Lens, W., Matos, L., Mouratidis, A., et al. Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and emotion* 39 (2015), 216–236.

[70] Chen, J., Sitter, R., and Wu, C. Using Empirical Likelihood Methods to Obtain Range Restricted Weights in Regression Estimators for Surveys. *Biometrika* 89, 1 (2002), 230–237.

[71] Cheng, J., Danescu-Niculescu-Mizil, C., and Leskovec, J. Antisocial behavior in online discussion communities. In: *International AAAI conference on web and social media*. 2015.

[72] Choi, D., Lee, U., and Hong, H. "It's not wrong, but I'm quite disappointed": Toward an Inclusive Algorithmic Experience for Content Creators with Disabilities. In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 2022, 1–19.

[73] Choi, H., Lee, W.-C., Aafer, Y., Fei, F., Tu, Z., Zhang, X., Xu, D., and Deng, X. Detecting attacks against robotic vehicles: A control invariant approach. In: *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*. ACM, New York, 2018, 801–816.

[74] Chopra, R. *Dissenting Statement of Commissioner Rohit Chopra In re Facebook, Inc.* `https://www.ftc.gov/system/files/documents/public_statements/1536911/chopra_dissenting_statement_on_facebook_7-24-19.pdf`, last accessed: 11.18.2024.

[75] Chouhan, C., LaPerriere, C. M., Aljallad, Z., Kropczynski, J., Lipford, H., and Wisniewski, P. J. Co-designing for community oversight: Helping people make privacy and security decisions together. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–31.

[76] Christofides, E., Muise, A., and Desmarais, S. Information disclosure and control on Facebook: are they two sides of the same coin or two different processes? *Cyberpsychology & behavior* 12, 3 (2009), 341–345.

[77] Clark, J. W., Snyder, P., McCoy, D., and Kanich, C. " I Saw Images I Didn't Even Know I Had" Understanding User Perceptions of Cloud Storage Privacy. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2015.

[78] Cloud, G. *UNC5537 Targets Snowflake Customer Instances for Data Theft and Extortion*. 2024.

[79] Coles-Kemp, L., Jensen, R. B., and Heath, C. P. Too much information: Questioning security in a post-digital society. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020, 1–14.

[80] Collier, A. F. and Wayment, H. A. Psychological benefits of the "maker" or do-it-yourself movement in young adults: A pathway towards subjective well-being. *Journal of Happiness Studies* 19 (2018), 1217–1239.

[81] Consolvo, S., Kelley, P. G., Matthews, T., Thomas, K., Dunn, L. C., and Bursztein, E. "Why wouldn't someone think of democracy as a target?": Security practices & challenges of people involved with US political campaigns. In: *Proceedings of the 30th USENIX Security Symposium*. USENIX. 2021, 1181–1198.

[82] Cox, J. *Inside the U.S. Government-Bought Tool That Can Track Phones at Abortion Clinics*. https://www.404media.co/inside-the-u-s-government-bought-tool-that-can-track-phones-at-abortion-clinics/, last accessed: 11.18.2024.

[83] Cranor, L. F. A framework for reasoning about the human in the loop (2008).

[84] Cranor, L. F. *Security and usability: designing secure systems that people can use*. " O'Reilly Media, Inc.", 2005.

[85] d'Angelo, G., Vitali, F., and Zacchiroli, S. Content cloaking: preserving privacy with google docs and other web applications. In: *Proceedings of the 2010 ACM symposium on applied computing*. 2010.

[86] Daffalla, A., Simko, L., Kohno, T., and Bardas, A. G. Defensive technology use by political activists during the Sudanese revolution. In: *2021 IEEE Symposium on Security and Privacy (SP)*. IEEE. 2021, 372–390.

[87] Daily Pakistan. *TikToker assaulted by 400 men at Minar-e-Pakistan shares her ordeal (DP Exclusive)*. Daily Pakistan, https://en.dailypakistan.com.pk/18-Aug-2021/tiktoker-assaulted-by-400-men-at-minar-e-pakistan-shares-her-ordeal-dp-exclusive, last accessed: 05.04.2023. 2021.

[88] *Data Protection and Privacy Legislation Worldwide*. https://unctad.org/page/data-protection-and-privacy-legislation-worldwide, last accessed: 07.14.2024.

[89] De Luca, A., Das, S., Ortlieb, M., Ion, I., and Laurie, B. Expert and non-expert attitudes towards (secure) instant messaging (2016).

[90] Degeling, M., Utz, C., Lentzsch, C., Hosseini, H., Schaub, F., and Holz, T. We value your privacy... now take some cookies: Measuring the GDPR's impact on web privacy. *arXiv preprint arXiv:1808.05096* (2018).

[91] Deibert, R. J. and Crete-Nishihata, M. Global governance and the spread of cyberspace controls. *Global Governance* 18 (2012), 339.

[92] DeNardis, L. Hidden levers of Internet control: An infrastructure-based theory of Internet governance. *Information, Communication & Society* 15, 5 (2012), 720–738.

[93] Dev, J., Moriano Salazar, P., and Camp, J. Lessons learnt from comparing WhatsApp privacy concerns across Saudi and Indian populations. In: *Proceedings of the Sixteenth Symposium on Usable Privacy and Security (SOUPS)*. USENIX Association. 2020, 81–97.

[94] Deville, J.-C. and Särndal, C.-E. Calibration estimators in survey sampling. *Journal of the American statistical Association* 87, 418 (1992), 376–382.

[95] DeVito, M. A. How transfeminine TikTok creators navigate the algorithmic trap of visibility via folk theorization. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–31.

[96] Dienlin, T. and Trepte, S. Is the privacy paradox a relic of the past? An in-depth analysis of privacy attitudes and privacy behaviors. *European journal of social psychology* 45, 3 (2015), 285–297.

[97] Dietrich, C., Krombholz, K., Borgolte, K., and Fiebig, T. Investigating system operators' perspective on security misconfigurations. In: *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*. 2018, 1272–1289.

[98] Dimmroth, K., Steiger, S., and Schünemann, W. J. Outrage without consequences? Post-Snowden discourses and governmental practice in Germany. *Media and Communication* 5, 1 (2017), 7–16.

[99] Dinev, T. and Hart, P. An extended privacy calculus model for e-commerce transactions. *Information systems research* 17, 1 (2006), 61–80.

[100] Distler, V., Zollinger, M.-L., Lallemand, C., Roenne, P. B., Ryan, P. Y., and Koenig, V. Security-visible, yet unseen? In: *Proceedings of the 2019 CHI conference on human factors in computing systems*. 2019, 1–13.

[101] Doshi-Velez, F., Kortz, M., Budish, R., Bavitz, C., Gershman, S., O'Brien, D., Schieber, S., Waldo, J., Weinberger, D., and Wood, A. Accountability of AI under the law: The role of explanation. *arXiv* (2017). eprint: `arXiv:1711.01134`.

[102] Douglas, B. D., Ewell, P. J., and Brauer, M. Data quality in online human-subjects research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. *Plos one* 18, 3 (2023), e0279720.

[103] Duckworth, A. L., Peterson, C., Matthews, M. D., and Kelly, D. R. Grit: perseverance and passion for long-term goals. *Journal of personality and social psychology* 92, 6 (2007), 1087.

[104] Duckworth, A. L. and Quinn, P. D. Development and validation of the Short Grit Scale (GRIT–S). *Journal of personality assessment* 91, 2 (2009), 166–174.

[105] Dutta, R. G., Yu, F., Zhang, T., Hu, Y., and Jin, Y. Security for safety: a path toward building trusted autonomous vehicles. In: *Proceedings of the International Conference on Computer-Aided Design*. ACM, New York, 2018, 1–6.

[106] Dwyer, C., Hiltz, S., and Passerini, K. Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace. *AMCIS 2007 proceedings* (2007), 339.

[107] Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., and Beck, H. P. The role of trust in automation reliance. *International journal of human-computer studies* 58, 6 (2003), 697–718.

[108] Ebersole, C. Online sex work in the time of Covid-19. *Illinois State University Research Symposium* 403 (2022).

[109] Edwards, W. K. and Grinter, R. E. At home with ubiquitous computing: Seven challenges. In: *Ubicomp: Ubiquitous Computing: International Conference*. 2001.

[110] Edwards, W. K., Poole, E. S., and Stoll, J. Security automation considered harmful? In: *Proceedings of the 2007 Workshop on New Security Paradigms*. 2008, 33–42.

[111] Elbitar, Y., Schilling, M., Nguyen, T. T., Backes, M., and Bugiel, S. Explanation beats context: The effect of timing & rationales on users' runtime permission decisions. In: *30th USENIX Security Symposium (USENIX Security 21)*. 2021, 785–802.

[112] ElSherief, M., Nilizadeh, S., Nguyen, D., Vigna, G., and Belding, E. Peer to peer hate: Hate speech instigators and their targets. In: *International AAAI Conference on Web and Social Media*. 2018.

[113] European Parliament and Council of the European Union. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation)*. Official Journal of the European Union, L 119, 4 May 2016. 2016.

[114] Ewing, J. and Boudette, N. E. *A Tiny Part's Big Ripple: Global Chip Shortage Hobbles the Auto Industry*. `https://www.nytimes.com/2021/04/23/business/auto-semiconductors-general-motors-mercedes.html`, last accessed: 11.11.2024.

[115] Fahl, S., Acar, Y., Perl, H., and Smith, M. Why Eve and Mallory (also) love webmasters: A study on the root causes of SSL misconfigurations. In: *Proceedings of the 9th ACM symposium on Information, computer and communications security (CCS)*. 2014, 507–512.

[116] Fiebig, T., Gürses, S., Gañán, C. H., Kotkamp, E., Kuipers, F., Lindorfer, M., Prisse, M., and Sari, T. Heads in the clouds: Measuring the implications of universities migrating to public clouds. *The 23rd Privacy Enhancing Technologies Symposium* (2023).

[117] Floridi, L. The Fight for Digital Sovereignty: What It Is, and Why It Matters, Especially for the EU. *Philosophy & Technology* 33, 3 (Sept. 2020), 369–378.

[118] Floridi, L. The fight for digital sovereignty: What it is, and why it matters, especially for the EU. *Philosophy & technology* 33 (2020), 369–378.

[119] Forum, W. E. *Global Gender Gap Report 2022*. https://www3.weforum.org/docs/WEF_GGGR_2022.pdf, last accessed: 31.05.2023. 2022.

[120] Foundation, R. P. *Raspberry Pi Web Site*. https://www.raspberrypi.org/, accessed 2023-07-27.

[121] Franke, T., Attig, C., and Wessel, D. A personal resource for technology interaction: development and validation of the affinity for technology interaction (ATI) scale. *International Journal of Human–Computer Interaction* 35, 6 (2019), 456–467.

[122] Freed, D., Palmer, J., Minchala, D., Levy, K., Ristenpart, T., and Dell, N. "A Stalker's Paradise": How Intimate Partner Abusers Exploit Technology. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018, 1–13.

[123] Garcia, F. D., Oswald, D., Kasper, T., and Pavlidès, P. Lock it and still lose it—on the (in) security of automotive remote keyless entry systems. In: *25th USENIX Security Symposium (USENIX Security 16)*. USENIX Association, Berkeley, 2016.

[124] Garrison, G., Rebman Jr, C. M., and Kim, S. H. An identification of factors motivating individuals' use of cloud-based services. *Journal of Computer Information Systems* 58, 1 (2018).

[125] *Get a user-resettable advertising ID*. https://developer.android.com/identity/ad-id, last accessed: 11.18.2024.

[126] Giraldo, J., Kafash, S. H., Ruths, J., and Cardenas, A. A. DARIA: Designing Actuators to Resist Arbitrary Attacks Against Cyber-Physical Systems. In: *2020 IEEE European Symposium on Security and Privacy (EuroS&P)*. IEEE, New York, 2020, 339–353.

[127] Glass, A., McGuinness, D. L., and Wolverton, M. Toward establishing trust in adaptive agents. In: *Proceedings of the 13th international conference on Intelligent user interfaces*. ACM, New York, 2008, 227–236.

[128] Goldacker, G. *Digitale Souveränität*. Kompetenzzentrum Öffentliche IT, Fraunhofer-Institut für Offene ..., 2017.

[129] Goode, S. i. Keeping the user in the cloud: a cognitive social capital antecedent to use continuance and trust-commitment in personal cloud storage services. *Behaviour & Information Technology* 38, 7 (2019).

[130] Gorski, P. L., Iacono, L. L., Wermke, D., Stransky, C., Möller, S., Acar, Y., and Fahl, S. Developers deserve security warnings, too: On the effect of integrated security advice on cryptographic {API} misuse. In: *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*. 2018, 265–281.

[131] Goscinny, R. and Uderzo, A. *Asterix Comics*. https://asterix.com/, accessed 2023-02-07.

[132] Granovetter, M. S. The Strength of Weak Ties. *The American Journal of Sociology* 78, 6 (1973), 1360–1380.

[133] Granovetter, M. S. Economic Action and Social Structure: The Problem of Embeddedness. *American Journal of Sociology* 91, 3 (1985), 481–510.

[134] Green, M. and Smith, M. Developers are not the enemy!: The need for usable security apis. *IEEE Security & Privacy* 14, 5 (2016), 40–46.

[135] Greenacre, M. *Correspondence analysis in practice*. CRC press, Boca Raton, 2017.

[136] Greenberg, A. *Forbes, "Hackers Reveal Nasty New Car Attacks–With Me Behind The Wheel (Video)"*. English. https://www.forbes.com/sites/andygreenberg/2013/07/24/hackers-reveal-nasty-new-car-attacks-with-me-behind-the-wheel-video/. Accessed: 2021-01-11. 2013.

[137] Greenberg, A. *Wired, "Hackers Remotely Kill a Jeep on the Highway—With Me in It"*. English. https://www.wired.com/2015/07/hackers-remotely-kill-jeep-highway/. Accessed: 2021-01-11. 2015.

[138] Greenberg, A. *Wired, A New Wireless Hack Can Unlock 100 Million Volkswagens*. English. https://www.wired.com/2016/08/oh-good-new-hack-can-unlock-100-million-volkswagens/. Accessed: 2019-15-11. 2016.

[139] Greenberg, A. *Wired, "Just a Pair of These $11 Radio Gadgets Can Steal a Car"*. English. https://www.wired.com/2017/04/just-pair-11-radio-gadgets-can-steal-car/. Accessed: 2021-01-11. 2017.

[140] Greenberg, A. *Wired, "Hackers Can Steal a Tesla Model S in Seconds by Cloning Its Key Fob"*. English. https://www.wired.com/story/hackers-steal-tesla-model-s-seconds-key-fob/. Accessed: 2021-01-11. 2018.

[141] Greenberg, A. *Wired, "Hackers Could Steal a Tesla Model S by Cloning Its Key Fob—Again"*. English. https://www.wired.com/story/hackers-steal-tesla-model-s-key-fob-encryption/. Accessed: 2021-01-11. 2019.

[142] Greenberg, A. *Wired, "Split-Second 'Phantom' Images Can Fool Tesla's Autopilot"*. English. https://www.wired.com/story/tesla-model-x-autopilot-phantom-images/. Accessed: 2021-01-11. 2020.

[143] Greenberg, A. *Wired, "This Bluetooth Attack Can Steal a Tesla Model X in Minutes"*. English. https://www.wired.com/story/tesla-model-x-hack-bluetooth/. Accessed: 2021-01-11. 2020.

[144] Gregor, S. and Benbasat, I. Explanations from intelligent systems: Theoretical foundations and implications for practice. *MIS quarterly* (1999), 497–530.

[145] Grimm, D. *Sovereignty: The origin and future of a political and legal concept*. Columbia University Press, 2015.

[146] Grinter, R. E., Edwards, W. K., Newman, M. W., and Ducheneaut, N. The work to make a home network work. In: *ECSCW: European Conference on Computer-Supported Cooperative Work*. 2005.

[147] Group, G. P. C. *Global Privacy Control (GPC)*. https://w3c.github.io/gpc/, last accessed: 01.12.2024. 2022.

[148] Hadi, A. Patriarchy and gender-based violence in Pakistan. *European Journal of Social Science Education and Research* 4, 4 (2017), 289–296.

[149] Hamilton, V., Barakat, H., and Redmiles, E. M. Risk, resilience and reward: Impacts of shifting to digital sex work. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–37.

[150] Hamilton, V., Soneji, A., Mcdonald, A., and Redmiles, E. M. "Nudes? Shouldn't I charge for these?": Motivations of New Sexual Content Creators on OnlyFans. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 2023, 1–14.

[151] Han, J., Deng, S., Xia, X., Wang, D., and Yin, J. Characterization and prediction of popular projects on github. In: *2019 IEEE 43rd annual computer software and applications conference (COMPSAC)*. Vol. 1. IEEE. 2019, 21–26.

[152] Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J., and Parasuraman, R. A meta-analysis of factors affecting trust in human-robot interaction. *Human factors* 53, 5 (2011), 517–527.

[153] Haque, R. *Purdah of the heart and eyes: An examination of purdah as an institution in Pakistan*. University of New South Wales, 2003.

[154] Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., and Khan, S. U. The rise of "big data" on cloud computing: Review and open research issues. *Information systems* 47 (2015), 98–115.

[155] Hastie, H., Chiyah Garcia, F. J., Robb, D. A., Laskov, A., and Patron, P. MIRIAM: A multimodal interface for explaining the reasoning behind actions of remote autonomous systems. In: *Proceedings of the 2018 on International Conference on Multimodal Interaction*. ACM, New York, 2018, 557–558.

[156] Havron, S., Freed, D., Chatterjee, R., McCoy, D., Dell, N., and Ristenpart, T. Clinical Computer Security for Victims of Intimate Partner Violence. In: *USENIX Security Symposium*. 2019, 105–122.

[157] Heckman, J. J. Selection bias and self-selection. In: *Econometrics*. Springer, 1990, 201–224.

[158] Held, D. *Democracy and the global order: From the modern state to cosmopolitan governance*. Stanford University Press, 1995.

[159] Helsley, R. W. and Strange, W. C. Knowledge barter in cities. *Journal of Urban Economics* (2004).

[160] Hennink, M. and Kaiser, B. N. Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social science & medicine* 292 (2022), 114523.

[161] Hennink, M. M., Kaiser, B. N., and Marconi, V. C. Code saturation versus meaning saturation: how many interviews are enough? *Qualitative health research* 27, 4 (2017), 591–608.

[162] Henrich, J., Heine, S. J., and Norenzayan, A. The weirdest people in the world? *Behavioral and brain sciences* 33, 2-3 (2010), 61–83.

[163] Herbert, F., Becker, S., Schaewitz, L., Hielscher, J., Kowalewski, M., Sasse, A., Acar, Y., and Dürmuth, M. A World Full of Privacy and Security (Mis) conceptions? Findings of a Representative Survey in 12 Countries. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 2023, 1–23.

[164] Hindman, M. *The Internet trap: How the digital economy builds monopolies and undermines democracy*. Princeton University Press, 2018.

[165] Ho, A., Maiga, A., and Aïmeur, E. Privacy protection issues in social networking sites. In: *2009 IEEE/ACS International Conference on Computer Systems and Applications*. IEEE. 2009, 271–278.

[166] Hoff, K. A. and Bashir, M. Trust in automation: Integrating empirical evidence on factors that influence trust. *Human factors* 57, 3 (2015), 407–434.

[167] Hoffmann, C. P., Lutz, C., and Ranzini, G. Privacy cynicism: A new approach to the privacy paradox. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 10, 4 (2016).

[168] Holzinger, A. From machine learning to explainable AI. In: *2018 World Symposium on Digital Intelligence for Systems and Machines (DISA)*. IEEE, New York, 2018, 55–66.

[169] Hope, A. *Toyota Connected Service Decade-Long Data Leak Exposed 2.15 Million Customers*. CPO Magazine, `https://www.cpomagazine.com/cyber-security/toyota-connected-service-decade-long-data-leak-exposed-2-15-million-customers/`, last accessed: 2/9/2023. 2023.

[170] Hu, W., Yang, T., and Matthews, J. N. The good, the bad and the ugly of consumer cloud storage. *ACM SIGOPS Operating Systems Review* 44, 3 (2010).

[171] Hughes-Roberts, T. Privacy and social networks: Is concern a valid indicator of intention and behaviour? In: *2013 International Conference on Social Computing*. IEEE. 2013, 909–912.

[172] Information, C. L. *California Online Privacy Protection Act*. 2003.

[173] Ion, I., Reeder, R., and Consolvo, S. {"... No} one Can Hack My {Mind"}: Comparing Expert and {Non-Expert} Security Practices. In: *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*. 2015, 327–346.

[174] Irfan, A. *Gruesome Murder Prompts Irfan Junejo to Break His Non-Political Stance*. Lens, `https://propakistani.pk/lens/gruesome-murder-prompts-irfan-junejo-to-break-his-non-political-stance/`, last accessed: 05.04.2023. 2020.

[175] Jack, M. C., Sovannaroth, P., and Dell, N. "Privacy is not a concept, but a way of dealing with life": Localization of Transnational Technology Platforms and Liminal Privacy Practices in Cambodia. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–19.

[176] Jakobi, T., Stevens, G., Castelli, N., Ogonowski, C., Schaub, F., Vindice, N., Randall, D., Tolmie, P., and Wulf, V. Evolving needs in iot control and accountability: A longitudinal study on smart home intelligibility. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 4 (2018), 1–28.

[177] Jamal, S. *Social media users demand justice for Pakistani blogger stabbed to death.* Gulf News, https://gulfnews.com/world/asia/pakistan/social-media-users-demand-justice-for-pakistani-blogger-stabbed-to-death-1.64684353, last accessed: 05.04.2023. 2019.

[178] Jen Caltrider Misha Rykov, Z. M. *It's Official: Cars Are the Worst Product Category We Have Ever Reviewed for Privacy.* https://foundation.mozilla.org/en/privacynotincluded/articles/its-official-cars-are-the-worst-product-category-we-have-ever-reviewed-for-privacy/, last accessed: 11.18.2024.

[179] Jian, J.-Y., Bisantz, A. M., and Drury, C. G. Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics* 4, 1 (2000), 53–71.

[180] Johnson, D. R. and Post, D. Law and borders: The rise of law in cyberspace. *stanford law review* (1996), 1367–1402.

[181] Johnson, W. L. Agents that Learn to Explain Themselves. In: *AAAI-94 Proceedings.* AAAI, Palo Alto, 1994, 1257–1263.

[182] Kalkbrenner, M. T. Alpha, omega, and H internal consistency reliability estimates: Reviewing these options and when to use them. *Counseling Outcome Research and Evaluation* 14, 1 (2023), 77–88.

[183] Kamkar, S. Drive it like you hacked it: New attacks and tools to wirelessly steal cars. *Presentation at DEFCON* 23 (2015).

[184] Kannen, A. *Unpacking TikTok Surveillance: Understanding Privacy Concerns and Implications.* https://aims.amnesty.nl/2024/01/09/unpacking-tiktok-surveillance-understanding-privacy-concerns-and-implications/, last accessed: 11.18.2024.

[185] Katz, J. *Birth of a Digital Nation.* https://www.wired.com/1997/04/netizen-3/, last accessed: 01.12.2024. 1997.

[186] Kaur, M., Sri Ramulu, H., Acar, Y., and Fiebig, T. "Oh yes! over-preparing for meetings is my jam:)": The Gendered Experiences of System Administrators. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–38.

[187] Kehayias, J. *Taking Back the Internet One Server at a Time.* Vice, https://www.vice.com/en/article/pkb4ng/meet-the-self-hosters, accessed 2023-02-07. Sept. 2021.

[188] Keller, D. C. I. Exception and Harmonization: Three Theoretical Debates on Internet Regulation (2019).

[189] Khan, F. A. Khwaja sira: Culture, identity politics, and" transgender" activism in Pakistan. PhD thesis. Syracuse University, 2014.

[190] Khan, M. T., Hyun, M., Kanich, C., and Ur, B. Forgotten but not gone: Identifying the need for longitudinal data management in cloud storage. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018.

[191] Khan, M. T., Tran, C., Singh, S., Vasilkov, D., Kanich, C., Ur, B., and Zheleva, E. Helping Users Automatically Find and Manage Sensitive, Expendable Files in Cloud Storage. In: *USENIX Security Symposium*. 2021.

[192] Khandelwal, S. *Hackernews, "Car Thieves Can Unlock 100 Million Volkswagens With A Simple Wireless Hack"*. English. `https://thehackernews.com/2016/08/hack-unlock-car-door.html`. Accessed: 2021-01-11. 2016.

[193] Kingsley, S., Sinha, P., Wang, C., Eslami, M., and Hong, J. I. " Give Everybody [..] a Little Bit More Equity": Content Creator Perspectives and Responses to the Algorithmic Demonetization of Content Associated with Disadvantaged Groups. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–37.

[194] Kokolakis, S. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & security* 64 (2017), 122–134.

[195] Komok, H. *Does the Pay Gap Exist on Instagram? Remuneration of Male vs Female Creators*. HypeAuditor, `https://hypeauditor.com/blog/does-the-pay-gap-exist-on-instagram-remuneration-of-male-vs-female-creators/`, last accessed: 19.04.2023. 2020.

[196] Kow, Y. M., Kou, Y., Semaan, B., and Cheng, W. Mediating the undercurrents: Using social media to sustain a social movement. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2016, 3883–3894.

[197] Kraemer, S. and Carayon, P. Human errors and violations in computer and information security: The viewpoint of network administrators and security specialists. *Applied ergonomics* 38, 2 (2007), 143–154.

[198] Krippendorff, K. Content analysis: An introduction to its methodology. In: ISBN 9780761915454. Sage, 2004. Chap. Reliability, chap. 11.

[199] Krippendorff, K. *Content analysis: An introduction to its methodology*. ISBN 9780761915454. Sage, 2004, chap. 11.

[200] Krombholz, K., Busse, K., Pfeffer, K., Smith, M., and Von Zezschwitz, E. " 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 (SP)*. 2019.

[201] Krombholz, K., Mayer, W., Schmiedecker, M., and Weippl, E. " I Have No Idea What I'm Doing"-On the Usability of Deploying {HTTPS}. In: *26th USENIX Security Symposium (USENIX Security 17)*. 2017, 1339–1356.

[202] Kropczynski, J., Aljallad, Z., Elrod, N. J., Lipford, H., and Wisniewski, P. J. Towards building community collective efficacy for managing digital privacy and security within older adult communities. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (2021), 1–27.

[203] Kropczynski, J., Ghaiumy Anaraky, R., Akter, M., Godfrey, A. J., Lipford, H., and Wisniewski, P. J. Examining collaborative support for privacy and security in the broader context of tech caregiving. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–23.

[204] Kuckartz, U. Qualitative Content Analysis: From Kracauer's Beginnings to Today's Challenges. *Forum: Qualitative Sozialforschung / Forum: Qualitative Social Research* 20, 3 (2019), Art. 12.

[205] Kundinger, T., Wintersberger, P., and Riener, A. (Over) Trust in Automated Driving: The Sleeping Pill of Tomorrow? In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, New York, 2019, 1–6.

[206] Kwet, M. Digital colonialism: US empire and the new imperialism in the Global South. *Race & Class* 60, 4 (2019), 3–26.

[207] Lake, M. *The New York Times, "HOW IT WORKS; Remote Keyless Entry: Staying a Step Ahead of Car Thieves"*. English. https://www.nytimes.com/2001/06/07/technology/how-it-works-remote-keyless-entry-staying-a-step-ahead-of-car-thieves.html. Accessed: 2021-01-11. 2001.

[208] Lambach, D. and Oppermann, K. Narratives of digital sovereignty in German political discourse. *Governance* 36, 3 (2023), 693–709. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/gove.12690.

[209] Langley, P., Meadows, B., Sridharan, M., and Choi, D. Explainable agency for intelligent autonomous systems. In: *Twenty-Ninth AAAI Conference*. AAAI, Palo Alto, 2017.

[210] Lastovicka, J. L., Bettencourt, L. A., Hughner, R. S., and Kuntze, R. J. Lifestyle of the tight and frugal: Theory and measurement. *Journal of consumer research* 26, 1 (1999), 85–98.

[211] Lee, H., Park, H., and Kim, J. Why do people share their context information on Social Network Services? A qualitative study and an experimental study on users' behavior of balancing perceived benefit and risk. *International Journal of Human-Computer Studies* 71, 9 (2013). Social Networks and Ubiquitous Interactions, 862–877.

[212] Lee, J. D. and See, K. A. Trust in automation: Designing for appropriate reliance. *Human factors* 46, 1 (2004), 50–80.

[213] Lee, S. and Valliant, R. Estimation for Volunteer Panel Web Surveys Using Propensity Score Adjustment and Calibration Adjustment. *Sociological Methods & Research* 37, 3 (2009), 319–343.

[214] Lerner, A., He, H. Y., Kawakami, A., Zeamer, S. C., and Hoyle, R. Privacy and activism in the transgender community. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020, 1–13.

[215] Leung, L. User-generated content on the internet: An examination of gratifications, civic engagement and psychological empowerment. *New Media & Society* 11, 8 (2009), 1327–1347.

[216] Li, F., Rogers, L., Mathur, A., Malkin, N., and Chetty, M. Keepers of the machines: Examining how system administrators manage software updates for multiple machines. In: *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. 2019, 273–288.

[217] Lilienfeld, S. O. Correlation still isn't causation. *APS Observer* 19 (2006).

[218] Lim, B. Y. and Dey, A. K. Assessing demand for intelligibility in context-aware applications. In: *Proceedings of the 11th international conference on Ubiquitous computing*. ACM, New York, 2009, 195–204.

[219] Lim, B. Y. and Dey, A. K. Toolkit to support intelligibility in context-aware applications. In: *Proceedings of the 12th ACM international conference on Ubiquitous computing*. ACM, New York, 2010, 13–22.

[220] Lim, B. Y., Dey, A. K., and Avrahami, D. Why and why not explanations improve the intelligibility of context-aware intelligent systems. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, 2009, 2119–2128.

[221] Lin, N. Social Networks and Status Attainment. *Annual Review of Sociology* 25 (1999), 467–487.

[222] Lohr, S. *He Created the Web. Now He's Out to Remake the Digital World*. New York Times, `https://www.nytimes.com/2021/01/10/technology/tim-berners-lee-privacy-internet.html`, accessed 2023-02-08. Jan. 2010.

[223] Lucas, G. M., Gratch, J., Cheng, L., and Marsella, S. When the going gets tough: Grit predicts costly perseverance. *Journal of Research in Personality* 59 (2015), 15–22.

[224] M., C. J. and Strauss, A. L. Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology* 19, 6 (1990), 418–427.

[225] MacCormick, N. *Questioning Sovereignty: Law, State, and Nation in the European Commonwealth*. 1999.

[226] Mackey, A. *EFF Lawsuit Discloses Documents Detailing Government's Social Media Surveillance of Immigrants*. `https://www.eff.org/deeplinks/2024/11/eff-lawsuit-discloses-documents-detailing-governments-social-media-surveillance`, last accessed: 11.18.2024.

[227] Madhavan, P. and Wiegmann, D. A. Similarities and differences between human–human and human–automation trust: an integrative review. *Theoretical Issues in Ergonomics Science* 8, 4 (2007), 277–301.

[228] Maher, S. *Viewpoint: Qandeel Baloch was killed for making lives 'difficult'*. The British Broadcasting Corporation (BBC), https://www.bbc.com/news/world-asia-49874994, last accessed: 05.04.2023. 2019.

[229] Majid, H. and Mustafa, M. Transformative digital spaces? Investigating women's digital mobilities in Pakistan. *Gender & Development* 30, 3 (2022), 497–516.

[230] Malcolm, J. *Multi-stakeholder governance and the Internet Governance Forum*. Terminus Press, 2008.

[231] marcelklehr (pseudonym). *Re: Creating a community survey*. Nextcloud Forum, https://help.nextcloud.com/t/123092/1, accessed: 2023-01-28.

[232] marcelklehr (pseudonym). *Who are we? - Take part in the Nextcloud Community Survey*. Nextcloud Forum, https://help.nextcloud.com/t/124056/18, accessed: 2023-01-28.

[233] Marczak, W. R. and Paxson, V. Social Engineering Attacks on Government Opponents: Target Perspectives. In: *Proceedings on Privacy Enhancing Technologies*. 2017, 152–164.

[234] Marne, S. T., Al-Ameen, M. N., and Wright, M. K. Learning System-assigned Passwords: A Preliminary Study on the People with Learning Disabilities. In: *Proceedings of the Thirteenth Symposium on Usable Privacy and Security (SOUPS)*. 2017.

[235] Martin, M. *Computer and Internet Use in the United States: 2018*. American Community Survey Reports – ACS-49. 2021.

[236] Masood, S. *Pakistan Strengthens Already Harsh Laws Against Blasphemy*. New York Times, https://www.nytimes.com/2023/01/21/world/asia/pakistan-blasphemy-laws.html, last accessed: 19.04.2023. 2023.

[237] *Mastodon*. https://mastodon.social, last accessed: 11.18.2024.

[238] Matthews, T., O'Leary, K., Turner, A., Sleeper, M., Woelfer, J. P., Shelton, M., Manthorne, C., Churchill, E. F., and Consolvo, S. Stories from survivors: Privacy & security practices when coping with intimate partner abuse. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2017, 2189–2201.

[239] Mayring, P. Qualitative Content Analysis. *Forum: Qualitative Sozialforschung / Forum: Qualitative Social Research* 1, 2 (2000), Art. 20.

[240] McClain, C., Faverio, M., Anderson, M., and Park, E. How Americans view data privacy. *Pew Research Center* (2023).

[241] McDonald, A., Barwulor, C., Mazurek, M. L., Schaub, F., and Redmiles, E. M. "It's stressful having all these phones": Investigating Sex Workers' Safety Goals, Risks, and Practices Online. In: *30th USENIX Security Symposium*. USENIX. 2021, 375–392.

[242] McDonald, N., Badillo-Urquiola, K., Ames, M. G., Dell, N., Keneski, E., Sleeper, M., and Wisniewski, P. J. Privacy and power: Acknowledging the importance of privacy research and design for vulnerable populations. In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020, 1–8.

[243] McGregor, S. E., Charters, P., Holliday, T., and Roesner, F. Investigating the computer security practices and needs of journalists. In: *24th USENIX Security Symposium (USENIX Security 15)*. 2015, 399–414.

[244] McGregor, S. E., Roesner, F., and Caine, K. Individual versus Organizational Computer Security and Privacy Concerns in Journalism. *Proceedings on Privacy Enhancing Technologies* 2016, 4 (2016), 1–18.

[245] McGregor, S. E., Watkins, E. A., Al-Ameen, M. N., Caine, K., and Roesner, F. When the Weakest Link is Strong: Secure Collaboration in the Case of the Panama Papers. In: *USENIX Security Symposium*. 2017, 505–522.

[246] McGuinness, D. L., Glass, A., Wolverton, M., and Da Silva, P. P. A Categorization of Explanation Questions for Task Processing Systems. In: *ExaCt*. AAAI, Palo Alto, 2007, 42–48.

[247] Mehmood, H., Ahmad, T., Razaq, L., Mare, S., Usmani, M. Z., Anderson, R., and Raza, A. A. Towards digitization of collaborative savings among low-income groups. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–30.

[248] Mell, P. and Grance, T. The NIST definition of cloud computing. *NIST Special Publication 800-145* (2011).

[249] Merritt, S. M. and Ilgen, D. R. Not all trust is created equal: Dispositional and history-based trust in human-automation interactions. *Human Factors* 50, 2 (2008), 194–210.

[250] Miltgen, C. L. and Peyrat-Guillard, D. Cultural and generational influences on privacy concerns: a qualitative study in seven European countries. *European journal of information systems* 23, 2 (2014), 103–125.

[251] Mitchell, J. Hierarchies. Introduction. In: *Markets, Hierarchies and Networks*. Ed. by Thompson, G., Frances, J., Levačić, R., and Mitchell, J. Sage Publications, 1998.

[252] Molnar, D. and Schechter, S. E. Self Hosting vs. Cloud Hosting: Accounting for the Security Impact of Hosting in the Cloud. In: *WEIS*. 2010, 1–18.

[253] Moura, G. C., Castro, S., Hardaker, W., Wullink, M., and Hesselman, C. Clouding up the internet: How centralized is dns traffic becoming? In: *Proceedings of the ACM Internet Measurement Conference*. 2020, 42–49.

[254] Mueller, M. L. *Networks and States: The Global Politics of Internet Governance*. MIT Press, 2010.

[255] Mumtaz, Z. and Salway, S. 'I never go anywhere': Extricating the links between women's mobility and uptake of reproductive health services in Pakistan. *Social Science & Medicine* 60, 8 (2005), 1751–1765.

[256] Mun, H., Han, K., and Lee, D. H. Ensuring Safety and Security in CAN-based Automotive Embedded Systems: A Combination of Design Optimization and Secure Communication. *IEEE Transactions on Vehicular Technology* (2020).

[257] Murthy, S., Bhat, K. S., Das, S., and Kumar, N. Individually vulnerable, collectively safe: The security and privacy practices of households with older adults. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–24.

[258] Mustafa, M. Negotiating borders through feminisms. *Interactions* 27, 6 (2020), 49–51.

[259] Naiakshina, A., Danilova, A., Gerlitz, E., Von Zezschwitz, E., and Smith, M. " If you want, I can store the encrypted password" A Password-Storage Field Study with Freelance Developers. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019, 1–12.

[260] Naiakshina, A., Danilova, A., Tiefenau, C., Herzog, M., Dechand, S., and Smith, M. Why do developers get password storage wrong? A qualitative usability study. In: *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*. 2017, 311–328.

[261] Naveed, S., Naveed, H., Javed, M., and Mustafa, M. "Ask this from the person who has private stuff": Privacy Perceptions, Behaviours and Beliefs Beyond WEIRD. In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 2022, 1–17.

[262] Nee, V. Norms and networks in economic and organizational performance. *American Economic Review* 88, 2 (1998), 85–89.

[263] Nextcloud GmbH. *Nextcloud About Page*. `https://nextcloud.com/about/`, accessed: 2023-01-28.

[264] Nextcloud GmbH. *Nextcloud Online Collaboration Plattform*. `https://nextcloud.com/`, accessed 2023-01-30.

[265] Nextcloud GmbH. *Threat model & accepted risks*. `https://nextcloud.com/security/threat-model/`, accessed: 2023-01-28.

[266] Nisar, M. A. (Un)becoming a man: Legal consciousness of the third gender category in Pakistan. *Gender & Society* 32, 1 (2018), 59–81.

[267] NIST. *Cybersecurity Framework*. English. `https://www.nist.gov/cyberframework`. Accessed: 2021-01-11. 2021.

[268] North, D. C. *Institutions, Institutional Change and Economic Performance*. ISBN 9780511808678. 1990.

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

[270] Nthala, N. and Flechais, I. Informal support networks: an investigation into home data security practices. In: *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*. 2018, 63–82.

[271] Nürnberger, S. and Rossow, C. –vatiCAN–Vetted, Authenticated CAN Bus. In: *International Conference on Cryptographic Hardware and Embedded Systems*. Springer, Cham, 2016, 106–124.

[272] Oliver, J. *36C3 - Server Infrastructure for Global Rebellion*. Youtube, https://www.youtube.com/watch?v=I_O3zj3p52A, last accessed: 2/9/2023. 2019.

[273] OPM.GOV. *EDUCATION LEVEL*. English. https://dw.opm.gov/datastandards/referenceData/1435/current?index=E. Accessed: 2021-01-11. 2021.

[274] Oppenheimer, D. M., Meyvis, T., and Davidenko, N. Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of experimental social psychology* 45, 4 (2009), 867–872.

[275] Page, C. *Microsoft AI researchers accidentally exposed terabytes of internal sensitive data*. TechCrunch, https://techcrunch.com/2023/09/18/microsoft-ai-researchers-accidentally-exposed-terabytes-of-internal-sensitive-data/, last accessed: 2/9/2023. 2023.

[276] *Pakistan Labour Force Survey 2020-21*. https://www.pbs.gov.pk/sites/default/files/labour_force/publications/lfs2020_21/LFS_2020-21_Report.pdf, last accessed: 20.09.2023.

[277] *Pakistan's third gender recognized but still discriminated*. DW, https://www.dw.com/en/pakistans-third-gender-recognized-but-still-discriminated/audio-16958385, last accessed: 12.04.2023.

[278] Park, C. Y., Faklaris, C., Zhao, S., Sciuto, A., Dabbish, L., and Hong, J. Share and share alike? An exploration of secure behaviors in romantic relationships. In: *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*. 2018, 83–102.

[279] Pasquale, F. Two narratives of platform capitalism. *Yale L. & Pol'y Rev.* 35 (2016), 309.

[280] Paterson, N. Walled gardens: the new shape of the public Internet. In: *Proceedings of the 2012 iConference*. 2012, 97–104.

[281] Peer, E., Rothschild, D., Gordon, A., Evernden, Z., and Damer, E. Data quality of platforms and panels for online behavioral research. *Behavior Research Methods* (2022), 1.

[282] Petrosyan, A. *Number of internet and social media users worldwide as of October 2024*. https://www.statista.com/statistics/617136/digital-population-worldwide/, last accessed: 11.18.2024.

[283] Pfeffermann, D. Methodological Issues and Challenges in the Production of Official Statistics: 24th Annual Morris Hansen Lecture. *Journal of Survey Statistics and Methodology* 3, 4 (2015), 425–483.

[284] Pfitzmann, A. and Hansen, M. *A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management*. 2010.

[285] Philipps, A. and Mrowczynski, R. Getting more out of interviews. Understanding interviewees' accounts in relation to their frames of orientation. *Qualitative Research* 21, 1 (2021). eprint: https://doi.org/10.1177/1468794119867548.

[286] Pinto, R. Á. DIGITAL SOVEREIGNTY OR DIGITAL COLONIALISM? (2018).

[287] Pohle, J. and Thiel, T. Digital sovereignty. *Internet Policy Review* 9, 4 (Dec. 17, 2020).

[288] Poole, E. S., Chetty, M., Grinter, R. E., and Edwards, W. K. More than meets the eye: transforming the user experience of home network management. In: *ACM conference on Designing interactive systems*. 2008.

[289] Portes, A. and Sensenbrenner, J. Embeddedness and Immigration: Notes on the Social Determinants of Economic Action. *American Journal of Sociology* 98, 6 (1993), 1320–1350.

[290] Post, D. G. Governing Cyberspace: Law. *Santa Clara Computer & High Tech. LJ* 24 (2007), 883.

[291] Powell, W. W. Neither Market nor Hierarchy: Network Forms of Organization. *Research in Organizational Behavior* 12 (1990), 295–336.

[292] Prendergast, C. and Stole, L. Barter relationships. In: *The Vanishing Rouble*. Ed. by Seabright, P. Cambridge University Press, 2000.

[293] prolific.com. *How does Prolific create the demographic subgroups used for representative samples?* https://researcher-help.prolific.com/hc/en-gb/articles/360019238413-Representative-samples-FAQ#heading-2. [Accessed: 2024/02/06]. 2022.

[294] Rajivan, P., Moriano, P., Kelley, T., and Camp, L. J. Factors in an end user security expertise instrument. *Information & Computer Security* 25, 2 (2017), 190–205.

[295] Redmiles, E. M. " Should I Worry?" A Cross-Cultural Examination of Account Security Incident Response. In: *2019 IEEE Symposium on Security and Privacy (SP)*. IEEE. 2019, 920–934.

[296] Redmiles, E. M., Kross, S., and Mazurek, M. L. How I learned to be secure: A census-representative survey of security advice sources and behavior. In: *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*. 2016, 666–677.

[297] Redmiles, E. M., Kross, S., and Mazurek, M. L. Where is the digital divide? A survey of security, privacy, and socioeconomics. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2017, 931–936.

[298] Redmiles, E. M., Kross, S., and Mazurek, M. L. How well do my results generalize? comparing security and privacy survey results from mturk, web, and telephone samples. In: *2019 2019 IEEE Symposium on Security and Privacy (SP), Vol. 00*. IEEE, New York, 2019, 227–244.

[299] Redmiles, E. M., Liu, E., and Mazurek, M. L. You Want Me To Do What? A Design Study of Two-Factor Authentication Messages. In: *SOUPS*. USENIX Association, Berkley, 2017.

[300] Redmiles, E. M., Malone, A. R., and Mazurek, M. L. I Think They're Trying to Tell Me Something: Advice Sources and Selection for Digital Security. In: *Security and Privacy (SP), 2016 IEEE Symposium on*. IEEE, New York, 2016, 272–288.

[301] Redmiles, E. M., Warford, N., Jayanti, A., Koneru, A., Kross, S., Morales, M., Stevens, R., and Mazurek, M. L. A comprehensive quality evaluation of security and privacy advice on the web. In: *29th USENIX Security Symposium (USENIX Security 20)*. 2020, 89–108.

[302] Reeder, R. W., Felt, A. P., Consolvo, S., Malkin, N., Thompson, C., and Egelman, S. An experience sampling study of user reactions to browser warnings in the field. In: *Proceedings of the 2018 CHI conference on human factors in computing systems*. ACM, New York, 2018, 1–13.

[303] Reichel, J., Peck, F., Inaba, M., Moges, B., Chawla, B. S., and Chetty, M. 'I have too much respect for my elders': Understanding South African mobile users' perceptions of privacy and current behaviors on Facebook and WhatsApp. In: *Proceedings of the 29th USENIX Conference on Security Symposium*. 2020, 1949–1966.

[304] Reist, B. M. and Valliant, R. Model-assisted estimators for time-to-event data from complex surveys. *Statistics in Medicine* 39, 29 (), 4351–4371.

[305] Rempel, J. K., Holmes, J. G., and Zanna, M. P. Trust in close relationships. *Journal of personality and social psychology* 49, 1 (1985), 95.

[306] Reynolds, B., Venkatanathan, J., Gonçalves, J., and Kostakos, V. Sharing ephemeral information in online social networks: privacy perceptions and behaviours. In: *Human-Computer Interaction–INTERACT 2011: 13th IFIP TC 13 International Conference, Lisbon, Portugal, September 5-9, 2011, Proceedings, Part III 13*. Springer. 2011, 204–215.

[307] Riebe, T., Biselli, T., Kaufhold, M.-A., and Reuter, C. Privacy Concerns and Acceptance Factors of OSINT for Cybersecurity: A Representative Survey. *Proceedings on Privacy Enhancing Technologies* 1 (2023), 477–493.

[308] Roller, M. R. A Quality Approach to Qualitative Content Analysis: Similarities and Differences Compared to Other Qualitative Methods. *Forum: Qualitative Sozialforschung / Forum: Qualitative Social Research* 20, 3 (2019), Art. 31.

[309] Rosenblum, D. What anyone can know: The privacy risks of social networking sites. *IEEE Security & Privacy* 5, 3 (2007), 40–49.

[310] Rouf, I., Miller, R. D., Mustafa, H. A., Taylor, T., Oh, S., Xu, W., Gruteser, M., Trappe, W., and Seskar, I. Security and Privacy Vulnerabilities of In-Car Wireless Networks: A Tire Pressure Monitoring System Case Study. In: *USENIX Security Symposium*. Vol. 10. USENIX Association, Berkley, 2010.

[311] Ryan, C. Computer and Internet Use in the United States: 2016 (2018).

[312] Saltzer, J. H., Reed, D. P., and Clark, D. D. End-to-end arguments in system design. *ACM Transactions on Computer Systems (TOCS)* 2, 4 (1984), 277–288.

[313] Sambasivan, N., Ahmed, N., Batool, A., Bursztein, E., Churchill, E., Gaytan-Lugo, L. S., Matthews, T., Nemar, D., Thomas, K., and Consolvo, S. Toward gender-equitable privacy and security in South Asia. *IEEE Security & Privacy* 17, 4 (2019), 71–77.

[314] Sambasivan, N., Batool, A., Ahmed, N., Matthews, T., Thomas, K., Gaytán-Lugo, L. S., Nemer, D., Bursztein, E., Churchill, E., and Consolvo, S. " They Don't Leave Us Alone Anywhere We Go" Gender and Digital Abuse in South Asia. In: *proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019, 1–14.

[315] Sambasivan, N., Checkley, G., Batool, A., Ahmed, N., Nemer, D., Gaytán-Lugo, L. S., Matthews, T., Consolvo, S., and Churchill, E. " Privacy is not for me, it's for those rich women": Performative Privacy Practices on Mobile Phones by Women in South Asia. In: *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*. 2018, 127–142.

[316] Sambasivan, N., Cutrell, E., Toyama, K., and Nardi, B. Intermediated technology use in developing communities. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2010, 2583–2592.

[317] Samek, W. *Explainable AI: Interpreting, explaining and visualizing deep learning*. Springer Nature, Cham, 2019.

[318] Samermit, P., Turner, A., Kelley, P. G., Matthews, T., Wu, V., Consolvo, S., and Thomas, K. {"Millions} of people are watching {you"}: Understanding the {Digital-Safety} Needs and Practices of Creators. In: *32nd USENIX Security Symposium (USENIX Security 23)*. 2023, 5629–5645.

[319] Sangaroonsilp, P., Dam, H. K., and Ghose, A. On Privacy Weaknesses and Vulnerabilities in Software Systems. In: *2023 IEEE/ACM 45th International Conference on Software Engineering (ICSE)*. IEEE. 2023, 1071–1083.

[320] Sarah Friedrich Frank Konietschke, M. P. *MANOVA.RM Documentation*. English. `https://cran.r-project.org/web/packages/MANOVA.RM/MANOVA.RM.pdf`. Accessed: 2021-01-11. 2020.

[321] Saravanakumar, K., Deepa, K., et al. On privacy and security in social media–a comprehensive study. *Procedia computer science* 78 (2016), 114–119.

[322] Sassen, S. *Losing control? Sovereignty in an age of globalization*. Vol. 5. Columbia University Press, 1996.

[323] Scherschel, F. A. *Heise, "Schlüssel-Hack: Autos von Tesla lassen sich in Sekunden öffnen"*. English. https://www.heise.de/security/meldung/Schluessel-Hack-Autos-von-Tesla-lassen-sich-in-Sekunden-oeffnen-4161136.html. Accessed: 2021-01-11. 2018.

[324] Scheuerman, M. K., Branham, S. M., and Hamidi, F. Safe spaces and safe places: Unpacking technology-mediated experiences of safety and harm with transgender people. In: *Proceedings of the ACM on Human-Computer Interaction*. Vol. 2. ACM New York, NY, USA, 2018, Article 155.

[325] Schreier, M. *Qualitative Content Analysis in Practice*. Sage, 2012.

[326] Semiconductor, P. *How Many Semiconductor Chips Are in a Car? [Infographic]*. https://polarsemi.com/blog/blog-semiconductor-chips-in-a-car/, last accessed: 11.11.2024.

[327] Shah, Z. *Muskan Sheikh and Rehan Shah: The TikTok stars gunned down in Karachi*. Geo News, https://www.geo.tv/latest/332999-muskan-sheikh-and-rehan-shah-the-tiktok-stars-gunned-down-in-karachi, last accessed: 05.04.2023. 2021.

[328] Smith, M., Strohmeier, M., Harman, J., Lenders, V., and Martinovic, I. A view from the cockpit: exploring pilot reactions to attacks on avionic systems. *NDSS* (2020).

[329] Smith, M., Szongott, C., Henne, B., and Von Voigt, G. Big data privacy issues in public social media. In: *2012 6th IEEE international conference on digital ecosystems and technologies (DEST)*. IEEE. 2012, 1–6.

[330] Spiekermann, S., Grossklags, J., and Berendt, B. E-privacy in 2nd generation E-commerce: privacy preferences versus actual behavior. In: *Proceedings of the 3rd ACM conference on Electronic Commerce*. 2001, 38–47.

[331] Srnicek, N. The challenges of platform capitalism: Understanding the logic of a new business model. *Juncture* 23, 4 (2017), 254–257.

[332] Strauss, A. L. and Corbin, J. M. *Grounded theory in practice*. Sage, 1997, 288.

[333] Strohmayer, A., Clamen, J., and Laing, M. Technologies for social justice: Lessons from sex workers on the front lines. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019, 1–14.

[334] Stuart, E. A. Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics* 25, 1 (2010), 1.

[335] Surbiryala, J. and Rong, C. Cloud computing: History and overview. In: *2019 IEEE Cloud Summit*. IEEE. 2019.

[336] Svantesson, D. and Clarke, R. Privacy and consumer risks in cloud computing. *Computer law & security review* 26, 4 (2010), 391–397.

[337] Syynimaa, N. and Viitanen, T. Is My Office 365 GDPR Compliant?: Security Issues in Authentication and Administration. In: *International Conference on Enterprise Information Systems*. 2018.

[338] Tabassum, M., Kosinski, T., and Lipford, H. R. I don't own the data: End User Perceptions of Smart Home Device Data Practices and Risks. In: *SOUPS, Symposium on Usable Privacy and Security*. 2019.

[339] Taddicken, M. The 'privacy paradox'in the social web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal of computer-mediated communication* 19, 2 (2014), 248–273.

[340] Tadic, B., Rohde, M., Wulf, V., and Randall, D. ICT use by prominent activists in Republika Srpska. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM. 2016, 3364–3377.

[341] Tang, J., Birrell, E., and Lerner, A. Replication: How well do my results generalize now? The external validity of online privacy and security surveys. In: *Eighteenth symposium on usable privacy and security (SOUPS 2022)*. 2022, 367–385.

[342] Tavani, H. T. and Moor, J. H. Privacy protection, control of information, and privacy-enhancing technologies. *ACM Sigcas Computers and Society* 31, 1 (2001), 6–11.

[343] Tesla. *Tesla's Autopilot*. English. `https://www.tesla.com/autopilot`. Accessed: 2019-07-10. 2020.

[344] Thomas, K., Akhawe, D., Bailey, M., Boneh, D., Bursztein, E., Consolvo, S., Dell, N., Durumeric, Z., Kelley, P. G., Kumar, D., et al. Sok: Hate, harassment, and the changing landscape of online abuse. In: *2021 IEEE Symposium on Security and Privacy (SP)*. IEEE. 2021, 247–267.

[345] Thomas, K., Kelley, P. G., Consolvo, S., Samermit, P., and Bursztein, E. "It's common and a part of being a content creator": Understanding How Creators Experience and Cope with Hate and Harassment Online. In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 2022, 1–15.

[346] Thomas Franke Christiane Attig, D. W. *ATI Scale*. English. `https://ati-scale.org/`. Accessed: 2021-01-11. 2021.

[347] Tolmie, P., Crabtree, A., Rodden, T., Greenhalgh, C., and Benford, S. Making the home network at home: Digital housekeeping. In: *European Conference on Computer-Supported Cooperative Work*. 2007.

[348] Tréguer, F. Intelligence reform and the Snowden paradox: The case of France. *Media and Communication* 5, 1 (2017), 17–28.

[349] Tréguer, F. US technology companies and state surveillance in the post-Snowden context: Between cooperation and resistance. PhD thesis. CERI, 2018.

[350] Trepte, S. The social media privacy model: Privacy and communication in the light of social media affordances. *Communication Theory* 31, 4 (2021), 549–570.

[351] Turk, A. M. *Amazon Mechanical Turk*. English. `https://www.mturk.com/`. Accessed: 2021-01-11. 2021.

[352] Turk, A. M. *FAQs - About Amazon Mechanical Turk*. English. `https://www.mturk.com/worker/help`. Accessed: 2021-01-11. 2021.

[353] Turner, F. *From counterculture to cyberculture: Stewart Brand, the Whole Earth Network, and the rise of digital utopianism*. University of Chicago Press, 2010.

[354] U.S. Census Bureau. *Decennial Census 2010*. Oct. 2022.

[355] UNSECO, D. R. F. *Policy Brief: Centering Pakistani Digital Content and Creators*. https://digitalrightsfoundation.pk/wp-content/uploads/2022/12/UNESCO-Policy-Paper_FINAL-December-2022.pdf, last accessed: 31.05.2023. 2022.

[356] Urquhart, C. *Grounded theory for qualitative research: A practical guide*. Sage, 2012.

[357] Uttarapong, J., Bonifacio, R., Jereza, R., and Wohn, D. Y. Social support in digital patronage: OnlyFans adult content creators as an online community. In: *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 2022, 1–7.

[358] Utz, C., Degeling, M., Fahl, S., Schaub, F., and Holz, T. (Un) informed consent: Studying GDPR consent notices in the field. In: *Proceedings of the 2019 acm sigsac conference on computer and communications security*. 2019, 973–990.

[359] Uzzi, B. and Lancaster, R. Embeddedness and Price Formation in the Corporate Law Market. *American Sociological Review* 69, 3 (2004), 319–344.

[360] Vaismoradi, M. and Snelgrove, S. Theme in Qualitative Content Analysis and Thematic Analysis. *Forum: Qualitative Sozialforschung / Forum: Qualitative Social Research* 20, 3 (2019), Art. 23.

[361] Van Blarkom, G., Borking, J. J., and Olk, J. E. Handbook of privacy and privacy-enhancing technologies. *Privacy Incorporated Software Agent (PISA) Consortium, The Hague* 198 (2003), 14.

[362] Van Dijk, M. and Juels, A. On the impossibility of cryptography alone for privacy-preserving cloud computing. *HotSec* 10, 1 (2010), 8.

[363] Vashistha, A., Anderson, R., and Mare, S. Examining security and privacy research in developing regions. In: *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies (COMPASS '18)*. 2018, 1–14.

[364] Vashistha, A., Garg, A., Anderson, R., and Raza, A. A. Threats, abuses, flirting, and blackmail: Gender inequity in social media voice forums. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019, 1–13.

[365] Visinescu, L. L., Azogu, O., Ryan, S. D., Wu, Y. ", and Kim, D. J. Better safe than sorry: A study of investigating individuals' protection of privacy in the use of storage as a cloud computing service. *International Journal of Human–Computer Interaction* 32, 11 (2016).

[366] Vitak, J., Liao, Y., Subramaniam, M., and Kumar, P. "I Knew It Was Too Good to Be True": The Challenges Economically Disadvantaged Internet Users Face in Assessing Trustworthiness, Avoiding Scams, and Developing Self-Efficacy Online. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–25.

[367] W3C. *Platform for Privacy Preferences (P3P) Project - Platform for Privacy Preferences (P3P) Project*. https://www.w3.org/P3P/, last accessed: 01.12.2024. 2002.

[368] W3C. *Tracking Preference Expression (DNT)*. https://www.w3.org/TR/tracking-dnt/, last accessed: 01.12.2024. 2019.

[369] Wang, Y. The third wave? Inclusive privacy and security. In: *Proceedings of the 2017 new security paradigms workshop*. 2017, 122–130.

[370] Warford, N., Matthews, T., Yang, K., Akgul, O., Consolvo, S., Kelley, P. G., Malkin, N., Mazurek, M. L., Sleeper, M., and Thomas, K. Sok: A framework for unifying at-risk user research. In: *2022 IEEE Symposium on Security and Privacy (SP)*. IEEE. 2022, 2344–2360.

[371] Watson, H., Moju-Igbene, E., Kumari, A., and Das, S. " We Hold Each Other Accountable": Unpacking How Social Groups Approach Cybersecurity and Privacy Together. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020, 1–12.

[372] Wei, M., Consolvo, S., Kelley, P. G., Kohno, T., Roesner, F., and Thomas, K. "There's so much responsibility on users right now": Expert Advice for Staying Safer From Hate and Harassment. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 2023, 1–17.

[373] Wei, M., Emami-Naeini, P., Roesner, F., and Kohno, T. Skilled or Gulliblef Gender Stereotypes Related to Computer Security and Privacy. In: *2023 IEEE Symposium on Security and Privacy (SP)*. IEEE. 2023, 2050–2067.

[374] Weigold, A. and Weigold, I. K. Measuring confidence engaging in computer activities at different skill levels: Development and validation of the Brief Inventory of Technology Self-Efficacy (BITS). *Computers & Education* 169 (2021), 104210.

[375] Wendt, N., Jensen, R. B., and Coles-Kemp, L. Civic empowerment through digitalisation: The case of Greenlandic women. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020, 1–13.

[376] Wermke, D., Stransky, C., Huaman, N., Busch, N., Acar, Y., and Fahl, S. Cloudy with a chance of misconceptions: exploring users' perceptions and expectations of security and privacy in cloud office suites. In: *SOUPS, Symposium on Usable Privacy and Security*. 2020.

[377] Whitten, A. and Tygar, J. D. Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0. In: *USENIX security symposium*. Vol. 348. 1999, 169–184.

[378] Wolf, M., Weimerskirch, A., and Paar, C. Security in automotive bus systems. In: *Workshop on Embedded Security in Cars*. isits AG, Bochum, 2004.

[379] Wu, J. and Zappala, D. When is a tree really a truck? exploring mental models of encryption. In: *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*. 2018, 395–409.

[380] Yağdereli, E., Gemci, C., and Aktaş, A. Z. A study on cyber-security of autonomous and unmanned vehicles. *The Journal of Defense Modeling and Simulation* 12, 4 (2015), 369–381.

[381] Yan, C., Xu, W., and Liu, J. Can you trust autonomous vehicles: Contactless attacks against sensors of self-driving vehicle. *DEF CON* 24 (2016).

[382] Younas, F., Naseem, M., and Mustafa, M. Patriarchy and social media: Women only Facebook groups as safe spaces for support seeking in Pakistan. In: *Proceedings of the 2020 International Conference on Information and Communication Technologies and Development*. 2020, 1–11.

[383] Young, A. L. and Quan-Haase, A. Privacy protection strategies on Facebook: The Internet privacy paradox revisited. *Information, Communication & Society* 16, 4 (2013), 479–500.

[384] YouTube. *"AReallyBadDay: Tesla Crash into Construction Barrels"*. English. `https://www.youtube.com/watch?v=i9r4nS5EjjQ`. Accessed: 2020-09-17. 2020.

[385] Zorz, Z. *A bug revealed ChatGPT users' chat history, personal and billing data.* Help Net Security, `https://www.helpnetsecurity.com/2023/03/27/chatgpt-data-leak/`, last accessed: 2/9/2023. 2023.

[386] Zuboff, S. The age of surveillance capitalism. In: *Social theory re-wired*. Routledge, 2023, 203–213.

[387] Zurko, M. E. and Simon, R. T. User-centered security. In: *Proceedings of the 1996 workshop on New security paradigms*. 1996, 27–33.