Privacy Preserving User Quantification with Smartphone Sensing

Dissertation

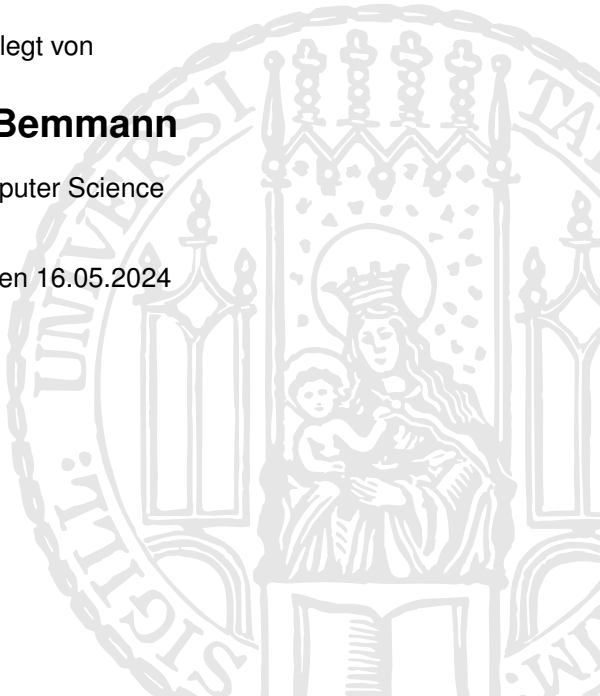
an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilians-Universität München

vorgelegt von

Florian Bemmann

M.Sc. Computer Science

München, den 16.05.2024



Erstgutachter:	Prof. Dr. Sven Mayer
Zweitgutachter:	Prof. Dr. Matt Jones
Drittgutachter:	Prof. Dr. Marc Langheinrich

Tag der mündlichen Prüfung: 10.07.2024

Zusammenfassung

Smartphones sind im Alltag der Menschen allgegenwärtig geworden, und damit auch die Erfassung und Nutzung mobiler Sensordaten. Wahrgenommene Datenschutzbedenken und -ängste führen jedoch zu Skepsis bei den Smartphone-Nutzenden und bremsen schließlich die Verbreitung und Nutzung von Apps, die auf der Verarbeitung mobiler Sensordaten basieren. Um die Auswirkungen der Sicherheits- und Datenschutzprobleme zu reduzieren, beschränkt das Betriebssystem den Zugriff für einige Datenarten auf ausgewählte Anwendungsfälle, um die Nutzentenden zu schützen. Dies behindert die Entwicklung neuartiger, anpassungsfähiger intelligenter Nutzerschnittstellen und die Erforschung besserer Technologien zur Verbesserung der Privatsphäre. In meiner Dissertation untersuchen wir, wie wir die Privatsphäre von Smartphone-Nutzenden verbessern und gleichzeitig die Daten für Anwendungszwecke nutzbar machen können. Zunächst begründe ich den Bedarf von mobilen Sensordaten, indem ich zeige, welche Informationen mit modernen mobilen Sensorsystemen extrahiert werden können, welche Anwendungsfälle sie ermöglichen und wie drei Interessengruppen, nämlich Nutzende, Forschende und die Gesellschaft, von ihrer Implementierung profitieren könnten. Danach hebe ich die aufkommenden Datenschutzprobleme hervor und zeige, wie sehr sie die Einführung solcher Apps behindern. Abschließend schlage ich nutzerzentrierte Ansätze vor, die diese Datenschutzprobleme entschärfen und gleichzeitig den Informationsgehalt beibehalten, der die Basis für die Anwendungsfälle der Daten bildet.

Wir stellen fest, dass aktuelle Benutzerschnittstellen zu wenig Nutzerorientierung aufweisen, und identifizieren Transparenz und Kontrollfunktionen als Schlüsselelemente für eine verbesserte Privatsphäre. Während das Angebot von Kontrollfunktionen die Bedenken hinsichtlich des Schutzes der Privatsphäre direkt verringert, verschlechtert Transparenz die Situation zunächst, wenn sie nicht umfassend gewährleistet wird. Wir bezeichnen diesen Effekt als “Tal der Transparenz” und erörteren auch, wie Transparenz und Kontrolle mit vorteilhaften Effekten integriert werden können. Wir diskutieren verschiedene Ansätze kontextbezogener Privatsphäre, die (Un-)Angemessenheit von Privatsphäre-Funktionen während Nutzende einer Aufgabe nachgehen, und das Problem der mangelnden Motivation der Nutzenden sich mit Privatsphäre zu beschäftigen. Unsere Arbeit trägt zu einem besseren Verständnis der Privatsphäre von Smartphone-Sensorik bei. Sie erleichtert die Nutzung mobiler Sensordaten für adaptive intelligente Anwendungen unter Wahrung der Privatsphäre der Nutzenden.

Abstract

With smartphones becoming omnipresent in people's everyday lives, mobile sensing data logging and usage have proliferated widely in the last decade. However, perceived privacy concerns and fears raise skepticism among smartphone users and finally throttle the spread and use of mobile sensing-based apps. To accommodate security and privacy issues, operating system developers impose general access restrictions on some kinds of data to protect the users. This obstructs the development of novel, adaptive intelligent interfaces and the study of better privacy-enhancing technologies. In this thesis, we study how we can improve the privacy of smartphone users while keeping the data usable for application purposes. We first motivate the demand for mobile sensing data by showing what information can be extracted with state-of-the-art mobile sensing systems, which use cases they fuel, and how three stakeholders, namely users, researchers, and society, could benefit from their implementation. After that, we highlight the emerging privacy issues and show how much they hinder app adoption. Finally, we propose user-centered approaches that mitigate these privacy issues while keeping the output that fuels the data's use cases. We found an overall lack of user-centeredness and identified transparency and control features as key elements towards an improved privacy perception. While offering control features directly reduces privacy concerns, transparency initially worsens the situation unless it is applied comprehensively. We outline how transparency and control can be integrated with beneficial effects, discuss different approaches to contextual privacy, the (in)appropriateness of

privacy interfaces in situ, and the issue of a lack of user motivation regarding privacy belongings. My thesis contributes to a better understanding of privacy in smartphone sensing. It facilitates using mobile sensing data for adaptive intelligent applications while preserving the users' privacy.

Acknowledgments

First of all, I want to thank **Heinrich Hußmann**, the grandfather of our lab and field of study, and person who gave me the opportunity to get started with all this! Allowing me to travel to my first conference during my masters caught my interest and motivated me a lot. Finally, his goodwill in support of the interdisciplinary PhoneStudy project, which was my main duty in the first years, allowed me to be where I am now. Unfortunately, he passed away too early, which was unexpected to all of us, leaving a gap at our lab that will never be completely filled. I really want to thank **Sven Mayer** for taking over my supervision - you are the most hands-on, engaged senior I've ever met, spending nights on making awesome figures for my papers and being always available, albeit you had to deal with a lot of stuff at our lab. Your engagement has shaped the lab in the last few years and made many of our projects a great success! You do a great job of caring about our lab and the people. I want to thank my examiners **Matt Jones** and **Marc Langheinrich**, for reviewing this huge project and making all the way to Munich to attend my defense event in-person.

A great thank you deserve my early mentors **Nađa Terzimehić** and **Daniel Buschek**. Having been involved in tutor jobs and practical projects since 2017, you were my first intensive contact at the media informatics lab. Your great spirit, kind, and calm manner contributed a lot to the great environment that we had an have. I will never forget the inspiring teatimes with Daniel and my shared love for good coffee and Max Beef's #25 with Nađa. Two essential components that also fueled this dissertation's

writing process. **Sarah Aragon-Hahner**, better known as Saragon, also played an essential role in my early Mimuc years. I collaborated on many of my early projects with you, and you were also an essential collaborator in our tea times!

A special thanks goes to the PhoneStudy team, with whom I have worked closely, especially in the first years of my PhD. I want to thank **Sarah Theres-Völkel**, with whom I started in the project as student assistant in 2018 and who considered myself worthy to become her successor as dev lead in the project. You have put a lot of effort into the project, albeit it was not the main essence of your research. I want to thank **Ramona Schödel** for having managed the PhoneStudy project during the most time of my PhD. I always admired your way of dealing with lots of (partially really annoying) things that needed to be done and dealt with. I could always consult you in any belonging, and really enjoy our meetings (the ones with the great homemade cake were best!). I want to thank all project members who I have been working with, especially **Clemens Stachl** who founded the project and sparked a lot of motivation for it in me, **Timo Koch** and **Larissa Sust** (I will never forget our business trip to Cambridge), **Fiona Kunz**, and **Markus Bühner**. Thank you **Sophia Sakel** and **Thomas Reiter** for joining the project in the last years and continuing the pleasant PhoneStudy spirit.

All our lab members deserve a great thank you! **Thomas Weber**, **Steeven Villa** and **Mochi** for being great roommates, **Florian Lang** for being my teaching-mate in most of my courses, **Francesco Chiossi** for many thoughtful discussions about dissertations and life, as well as **Maximiliane Windl** for joining me in the privacy game and working with me on many great collaborations! - To name just the ones with whom I have been working closest. A special thanks also goes to **Rainer Fink**, **Christa Feulner**, and **Franziska Schwamb**, who make our lab run and survive from behind the scenes. I can always count on your support and quick response to my concerns!

The biggest thanks go to my family and friends, who laid the basis for all this! My **mum**, **dad**, and **sister** who made me the person I am now and enabled me to pursue my journey. They have been supporting me from ever on and still support me in the things I do, and constitute a base I can always consult and refer to. I thank all my **old and young friends** for giving me the balance I need from work. Without the good times that we had in between and within my project phases and the energy, fun, and lovely vibes that you are spreading, I can hardly imagine that the last five years might have finally worked out well.

Table of Contents

1

Introduction

Ubiquitous mobile devices, especially smartphones, have proliferated within our society in the last decade. While mobile phones around the year 2000 were limited in their abilities, i.e., making calls and writing SMS, they did not process much user data. With the introduction of the smartphone, the amount of tracked user data has increased. Nowadays, smartphones possess a plethora of data about their users and surroundings. Data can be passively obtained through device sensors and application programming interfaces (APIs) or actively entered by the user in apps and device interfaces. Increasing computational power further enlarges the space for data processing and inference opportunities. Analyses regard multiple data items in relation to each other, and can cover longer time spans. Low-level data unveil high-level insights. Therefore, our devices can quantify their users and gain context awareness. Context-awareness means understanding the user. Through being context-aware, the device understands the user's current situation and what they are doing. Understanding the user is essential for ubiquitous technology to well-integrate into the user's life. Without context-awareness, even the most sophisticated prediction system cannot deliver relevant interventions that integrate with the users real-world life. Especially for proactive artificial intelligence (AIs), such as the novel interaction paradigm of personal large language model (LLM) agents [254], a detailed understanding of the user's situation and context is essen-

tial. The output of every intelligence can only be as good as its input data has been. Two aspects especially distinguish smartphones from previous technologies: They are omnipresent and always connected. Due to their pocket-size form factor and the ubiquity of internet access through mobile networks and WiFi these devices are carried by many people 24/7, and their services are always available. Thereby, yet unseen challenges for user privacy arise, which demand novel privacy-enhancing technologies [103]. Especially the two factors *omnipresence* and *always-connectedness* have not been implemented with previous technologies and require new solutions. Therefore, existing privacy-enhancing technologies are not designed and pose limitations in that context. In the literature and our work, we found that users are often critical about collecting and using their data. They face uncertainty of what data is collected [385] and which actions are performed with the collected information [347, 388]. A general lack of information and understanding makes people skeptical privacy-wise [221, 338]. Furthermore, people feel they need to be more in control of their data as current control mechanisms are not satisfactorily usable with omnipresent data collection. Besides making users uncomfortable, privacy concerns also lead to users rejecting services and using devices and applications less. Also, in mobile sensing research studies, where a smartphone app is deployed in the wild to collect smartphone sensing data passively, such concerns led to much lower consent rates than in traditional studies [221, 234, 338]. Beyond affecting users' perception of an app and behavior with the device, the effects of privacy issues also reach app and OS developers. Reduced usage and users avoiding certain features throttle business models. Furthermore, security issues that can follow as a consequence of privacy issues pose a danger to their system and users. Operating system developers, who are in charge of providing a safe and usable system to their users, counter privacy and security issues with access restrictions to data. Detailed, contentful data is made available for specific use cases only, such as screen and textual contents only being available to accessibility-fostering applications. Due to a lack of satisfying privacy-enhancing interface concepts, restricting access to potentially sensitive data sources is the only approach currently protecting users. However, broader access to detailed, contentful data could enable implementing more context-aware systems. Through access restrictions, the aforementioned use cases are no longer possible - privacy barriers obstruct the potential benefits of mobile sensing technology.

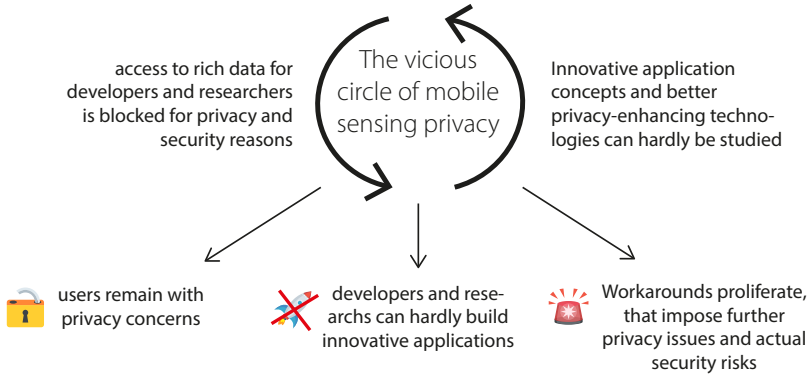


Figure 1.1 : Mobile sensing privacy is stuck in a vicious circle. The operating system cannot provide access to rich data for privacy and security reasons, so developers and researchers cannot study innovative application concepts and better privacy-enhancing technologies.

Data access restrictions also impede research on novel privacy solutions. Thus, mobile sensing is stuck in a vicious circle of restricted availability, opposing the need to study privacy-enhancing technologies more in-depth (see Figure 1.1). Currently, intelligence opposes privacy: Intelligent applications make compromises privacy-wise to obtain the information their intelligence is based on. In contrast, applications that provide real privacy dispense with intelligent features. In this thesis, we investigate how we can improve user quantification through mobile devices while preserving privacy. We aim to improve users' privacy while keeping the data usable for the users' good. While most current approaches instead improve privacy by reducing data usage, we aim for privacy going hand-in-hand with an application's use case.

Mission Statement

With this thesis, we improve smartphone users' privacy while facilitating the usage of mobile sensing data for innovative application use cases.

It is important for me to find solutions to how data can be used for good in a privacy-friendly manner. Velykoivanenko et al. [405], who study the privacy of fitness trackers, also conclude that opportunities to increase privacy without reducing the technology's utility are understudied.

1.1 Motivation: Mobile Sensing Data Is in the Interest Of Users, Researchers, and Society

We motivate our research by showing what mobile sensing technology is capable of now and what it could be in the future. To realize their purpose, applications require information about the user's situation and context. Without sharing information and data with the smartphone, its user experience would be way worse. Information-retrieval tasks would yield less accurate results, i.e., they would require more detailed search query specifications from the user. Monitoring activities in the background would become practically impossible. For example, fitness and health trackers could not passively keep track of one's activity, and sleep-tracking applications would have to rely on manual user input.

With an increasing amount of data, more preprocessing becomes necessary; furthermore, the increase in computational capabilities enables more pattern discovery and knowledge extraction processes. Combining both, mobile sensing often allows for the application of data mining and machine learning techniques. Models can be built offline with sensing-collected data (e.g., [261, 416]) or trained and optimized online with continuous user data (e.g., [260]). Multiple devices can also work together in a multi-agent architecture, yielding ambient intelligence [22], i.e., an intelligent environment aware of surroundings and people. For the end user this leads to more **context-aware applications**, **pro-active application behavior**, more **precise recommender systems and personal daily support** . These benefits mainly target the end-users, although app developers also indirectly benefit from an improved app experience. Due to the dissemination of smartphones in our society, sensing can be deployed at scale efficiently from a technical perspective. In contrast to the aforementioned user-sided benefits, crowd-sensing applications usually fulfill a purpose that is in the interest of the app-publishing company or organization, for example, a governmental organization

tracking parameters in cities (e.g., [126]) or a university that deploys a sensing app to collect data for a study (e.g., [384]). Especially as **data source for interdisciplinary research**, for example, in the domains of health and well-being [98, 280], psychology [180, 384], and HCI [399], it is valuable. Furthermore, the *society* can be regarded as a third stakeholder. Data on environmentally-relevant behaviors, such as mobility or consumption, can be used by applications to track their progress over time, or **support behavior change** [386].

Mobile sensing technology supports these application concepts by providing ubiquitous behavioral data, i.e., the ability to unobtrusively access data on the user's behavior, context, and situation in the background [181, 182]. Common behavioral data encompasses but is not limited to device usage and mobility behavior, including the choice of means of transport (e.g., via Google's Awareness API[1]), and mobile language use. Information on behaviors that cannot be directly sensed by the smartphone, such as consumption and nutrition behaviors, can either be gathered with journaling methods [401] (e.g., asking the user daily for their consumed amount of meat) via third-party devices or services (e.g., financial APIs that have access to purchases), or a semi-automatic approach combining both (e.g., taking a photo of each meal that is processed by image recognition) [41]. Most data is available immediately in the situation (in situ), allowing the user to follow their progress live. More rich, contentful data that provides detailed information on user state and behavior is especially promising. However, potential research and app design directions are yet unknown, as such data is hardly made available by the operating system for privacy and security reasons.

Research Gap

We are in a vicious circle of (1) privacy and security issues hindering OS developers from making rich, contentful data available to developers and researchers, whereby (2) innovative application use cases and better privacy-enhancing technologies cannot be developed and studied. Future directions for mobile sensing-based applications need to be pointed out, and requirements on data need to be defined.

[1] https://developers.google.com/awareness/overview, last accessed 2024-11-22

1.2 Current Barriers: Privacy Concerns and App Adoption Refusal

There is a general lack of user perspective on privacy. Existing research studied barriers to the adoption of smartphone apps that include passive sensing, e.g., [83, 348] and mobile sensing research apps [221, 338]. Moreover, Christin et al. [89] provide an overview of technical privacy issues and measures. Many papers propose technical concepts to reduce security issues (e.g., [40, 256]). Today's literature concludes that privacy is the most significant social barrier to adoption [221, 266]. Still, the underlying mechanisms of such effects are unclear: To tackle user privacy concerns, a deeper understanding of what users actually are afraid of, which outcomes they fear, and which solutions they desire are necessary. Making an app technically safe and privacy-friendly does not go far enough to ensure user acceptance.

It is essential to understand these aspects to build privacy-enhancing technologies that match the user's desires. Designers of future data-using smartphone apps and mobile sensing systems need to know how to foster user trust and lower privacy concerns to reach satisfactory adoption rates. Furthermore, we do not know in detail which aspects contribute how much to app adoption decisions. In detail, research does not understand the effects of individual data types, privacy-enhancing technologies, the aforementioned privacy concerns, and other app characteristics. However, that would be important to understand, to build privacy-enhancing technologies that match the users' desires. People are reluctant to install a sensing app, especially if they cannot expect a personal benefit thereof [74].

Research Gap

In-depth privacy concerns, their underlying reasons, and the feared consequences of mobile sensing applications are yet unclear. To design better privacy-enhancing technologies, we need to understand the user perspective on smartphone privacy and the factors and influences that underlay app adoption decisions.

1.3 Solutions: Improving Privacy While Preserving Data Usability

To tackle the above-mentioned privacy challenges, prior research proposed the concept of "consent as a process," including making data logging processes transparent and giving the users control over their data [179]. However, more information is needed about the effect transparency and control have on smartphone app adoption. In the context of privacy dashboards [187], studies have been conducted with diverging results. Privacy dashboards are a privacy design pattern that makes users aware of the data services have collected about them. They should provide successive summaries of the collected data and give an easily understandable overview [121, 457]. While some studies indicate positive results, such as Tsai et al. [397], other studies revealed none (e.g., [213]) or even contradictory effects (e.g., [203]). A promising direction is supplementing transparency with control, which has already been shown to mitigate the adverse side effects of transparency [141, 187]. For example, increased transparency decreased trust and willingness to share data [206, 343]. However, those studies were either conducted in the domain of actively donated data [187], conducted as vignette studies [213, 221] or did not evaluate transparency and control independently [213, 397]. Thus, developers of mobile sensing applications in industry and research cannot build on insights into the effects of transparency and control features.

Transparency is currently limited, as Shen et al. [368] argue that current mobile systems hardly convey to users what happens with their data and which specific data is used. Especially when it comes to machine learning procedures and data inference, users have a hard time understanding the procedures and estimating the inherited capabilities. **Control** is also not given to a satisfactory extent, as current permission systems do not allow fine-grain control but only some toggle switches for groups of data access [248]. Only a few studies investigated the users' perspective of control features (e.g., [312] rated usability themselves). Finer-grained permissions (e.g., [363]) were rarely studied and, if so, emphasized on technical aspects rather than on the user. Research needs to find solutions for better mobile sensing privacy while keeping data usable for the good. Transparency and control, thereby, are the key aspects underlying most privacy-enhancing interface concepts.

Research Gap

There is a lack of privacy-enhancing technologies appropriate in light of mobile sensing systems' specific characteristics. Current interfaces do not provide satisfactory transparency and control, do not keep the user in the loop, and hinder rather than facilitate the use of rich, contentful data for innovative application purposes.

CONTRIBUTION

With this thesis, we contribute to two areas: With the studies on users' perspective on privacy and app adoption decisions, which we present in Chapter 4, we help system designers understand how users perceive current systems privacy-wise. Based on these, we propose means to improve users' privacy while keeping the data usable for innovative application use cases, based on the studies presented in Chapter 5.

We contribute insights on the relevance of app and publisher characteristics for app adoption decisions, focusing on mobile sensing particularities. Direct user benefits are essential, as the user's decision process is mainly based on an evaluation of the ratio between service value and privacy cost. We furthermore give insights on what users are actually concerned about, i.e., report on potential privacy invasions that users judge to be dangerous, real-world consequences that they are afraid of, and which mitigation measures they envision and desire.

The insights from our second main block help app designers, developers, and the HCI community to create more privacy-friendly applications. As a basis for nearly all privacy-enhancing technologies, we provide general insights into the effect of transparency and control. We quantified the initially averse effect of transparency and showed how control can surpass these. We frame the phenomenon *valley of transparency*, where we argue based on prior work and our studies that current transparency measures do not go far enough,

to surpass the initial negative effects of users becoming aware of a system's data practices. We conclude on critical aspects that should be considered for smartphone privacy-enhancing technologies and propose three types of practical approaches. We point out open points for discussion and future research directions that we would like to motivate the HCI community to study further. These evolve around both interface design problems, such as control interfaces and the implementation of holistic transparency, and psychological aspects, like achieving user motivation to spend effort on privacy belongings and yielding contextual understanding.

In the big picture, our research has relevant implications for three stakeholders: Users can benefit from better mobile applications that invade their privacy less. Context-aware systems can fulfill tasks faster and proactively support their users. Research projects benefit from better data than was previously available with manual data collection methods. Finally, due to its scalability and the omnipresence of mobile technologies, mobile sensing-based applications have the potential to support our society in dealing with its challenges. We envision creating a better-educated user base that can make informed decisions. Technology skepticism could be dismantled when users are kept in the loop more effectively. Confronting people with things they do not understand rather fears them and raises technology skepticism rather than fostering a literate relationship with daily technology.

1.4 Thesis Structure and Research Questions

After presenting definitions and related work relevant to the remainder of my thesis, it is structured into three chapters. Each is devoted to one overarching research question (RQ 1-3). Further specific research questions are then presented at the beginning of each chapter. In Figure 1.2, we show how the projects contribute to the chapters.

In Chapter 3, we motivate the demand for and benefit of mobile sensing. We present application use cases that demonstrate that mobile sensing data interests

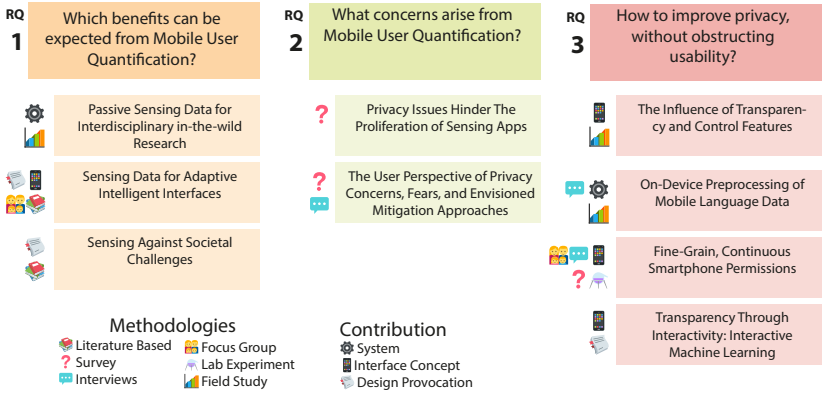


Figure 1.2 : This thesis is structured into three blocks. Within each block are two to four projects. Icons indicate the kind of methodologies that we applied in each block and the kind of artifact we contribute.

three stakeholders: Researchers, users, and society. We point out which data these applications require or might require for future advancements. It follows the overarching research question:

RQ1 *Which benefits can be expected from Mobile User Quantification?*

After that, in Chapter 4, we report on privacy and adoption issues that come with mobile sensing. Referring to the data demands we have shown in Chapter 3, we show that these demands can currently not be satisfied. With two online surveys and one interview study, we show that privacy concerns and security issues throttle app adoption behavior from the user side and also hinder data provisioning by the operating system developer. Its research question is:

RQ2 *What concerns arise from Mobile User Quantification?*

In Chapter 5, we then propose three concepts to improve privacy without obstructing the data's usability. This means that we emphasized not to simply limit the amount of

available data. Instead, my proposed concepts aim to bring humans back into the loop by realizing comprehensive transparency and control mechanisms. This follows the research question:

RQ3 *How to improve privacy without obstructing usability?*

We finally reflect on our findings in Chapter 6 and point out the key takeaways and learnings from our studies. Bringing findings from prior work back in, we discuss the contradictory results regarding the effects of transparency and control. We propose a model concerning the relationship between transparency, privacy concerns, and app adoption behaviors that informs the design of future privacy-enhancing technologies on overcoming current adverse effects.

1.5 Research Context

The research that leads to this thesis was conducted at the LMU Munich in the years between 2019 and 2024 in the the Media Informatics Group. I give an overview of all incorporated publications with clarification of my contribution in Table 8.1 on page 328.

The PhoneStudy Project - Department of Psychology, LMU Munich My research was mainly conducted with relation to the *PhoneStudy* project[1], an interdisciplinary research project with the aim of leveraging mobile sensing technologies to answer psychological research questions. As lead developer I was technically responsible for developing and running the sensing and analysis infrastructure in multiple field studies (e.g., [55, 170, 232, 341, 355, 359]), that were conducted at LMU Munich, in cooperation with national partners such as the university of Heidelberg and the German Institute for Economic Research (DIW)[2], and incorporated into panel studies such as the Socio-Economic Panel (SOEP)[3]. The following publications, that are included in this thesis, were conducted in relation with the PhoneStudy project: [42, 47].

[1] https://phonestudy.org, last accessed 2024-11-22

[2] https://www.diw.de/en, last accessed 2024-11-22

[3] https://www.diw.de/en/diw_01.c.615551.en/research_infrastructure__socio-economic_panel__soep.html, last accessed 2024-11-22

Media Informatics Group My early research at the Media Informatics Group of LMU Munich was mainly inspired by and conducted in collaboration with Daniel Buschek [16, 38–40, 49, 351, 392]. I initially studied research questions regarding the topic of self-reflection and behavior change [16, 41, 391, 392], mainly in collaboration with Nađa Terzimehić, and conducted research in the domain of environmental sustainability [41, 43, 45]. My latest project evolve especially around the topic of privacy in the context of smartphones and mobile sensing. I conducted that research mainly under the guidance of Sven Mayer, leading to the following publications that are incorporated in this thesis: [44, 46, 48].

2

Related Work

In this chapter, we outline the related work that constitutes the basis for our work. We, therefore, start with a definition of **mobile sensing** and describe the scope that this thesis evolves. Afterward, **privacy** is introduced, starting with general definitions and becoming more specific to mobile sensing-related privacy throughout the remainder of this chapter. Finally, we describe the state of research regarding **privacy-enhancing technologies**, outline current directions research is facing, and point out limitations.

2.1 Mobile Sensing

Passive smartphone sensing enables access to data about its user and surroundings, with no extra effort for the user [98]. Their capabilities range from acceleration and position of the device, over user location via GPS, to environmental data, e.g., via the microphone [111, 242]. An overview of the available passive data sources is given e.g., by Delgado-Santos et al. [111], or Cornet and Holden [98] in the health context. The passive data sources can be complemented by actively entered data, for example through Experience Sampling [401], journaling [135], or data donations [167, 322, 374].

Mobile sensing based systems can be defined by having three essential characteristics [242]: (1) *Sense*, i.e. collecting data through a mobile device, (2) *Learn*, i.e. deriving some higher level information or insight from that data, and (3) *Inform-Share-Persuade*, i.e. fueling a use case with that information. In the literature does not exist a common and precise definition for the term *mobile sensing* or *smartphone sensing*. Its scope varies, depending on the context and application domain. In the scope of my thesis I regard *mobile sensing* as:

- Smartphone infrastructure as outlined by Lane et al. [242], where data is collected with a mobile device to derive information which serves a use case
- I restrict my research to smartphones, and deliberately do not regard other ubiquitous mobile devices such as wearables, IoT devices, vehicles, or custom dedicated sensing hardware.
- Data can originate from passive or participatory sensing (cf. [242])
- I do not apply any restriction to datatypes, i.e. generalize for all kinds of people-centric and environmental-centric data (cf. [244]).

To bring structure into the space of possible sensing, research has come up with multiple taxonomies. Taxonomies classify mobile sensing apps regarding user *involvement* (participatory sensing vs. passive sensing) [242], and the data subject by the two *approaches* people-centric and environmental-centric [244]. Passive sensing is an essential part of context-aware and data-assessing applications, as it lowers the burden for the user [381], resulting in a higher data frequency and resulting data has higher quality due to avoiding the self-report bias [107, 383]. Khan et al. [223] on another dimension describe by how the data of a sensing system is used. They distinguish between *personal sensing*, *social sensing* and *public sensing*.

2.2 Mobile Sensing as Research Method

Mobile sensing has proliferated as research methodology, to collect data on human behavior and their environment. With smartphones having become a constant companion in all daily situations for most people [220, 222], they offer themselves as tool for observational studies. The yet prevalent self-report methodologies (e.g., [245]), i.e. people

are asked to report their behavior manually, are known to differ from actual behavior [33, 169]. They impose several biases, for example, *self-report bias* also known as *affirmation bias*, that is driven by social desirability and leads to people report behaviors that they think their observer expects from them [104, 119, 403], or (non-)compliance biases [459]. Furthermore, data collection manually involving people's cooperation is limited in its extensiveness and sampling frequency [132]. Besides being used in human-computer interaction [399, 447], it finds application in the social sciences [328], psychology [180], especially personality sciences [36, 384], and medical and mental health contexts [98, 285].

The objective of mobile sensing in research applications is mostly to gather data about people in the wild. That can encompass their (a) behavior, (b) state and context, and (c) aspects on their environment. Data is used to collect large datasets from which, through statistical analysis or machine learning methods, insights are generated (e.g., [358, 365, 384]). Besides classical knowledge discovery methods, research increasingly aims to build prediction models (e.g.[381]). Here mobile sensing is studied as a tool to detect some user state, which is envisioned to fuel some application scenario (e.g., [399]). As measurement added on a study that deploys an artifact or prototype, mobile sensing data is used to study the influence of the artifact on the user [65]. Accordingly, also the effects of existing features of devices are studied with mobile sensing (e.g., [170]). Crowd sensing leverages the ubiquity of mobile phones, to collect data on an area, for example a city (e.g., [126, 448]).

2.3 Foundations of Privacy

Privacy can roughly be described as "The ability of an individual to control the terms under which their personal information is acquired and used." [426, p. 42]. Legal privacy thereby usually differs from what people perceive as their privacy. A plethora of literature studies the many aspects of privacy [198], such as Westin [426] defining four states of privacy (solitude, intimacy, anonymity, and reserve). In the context of HCI privacy encompasses the aspects of controlling information flow, security and concerns risks, and social aspects raising ethical questions [4]. However, comprehensive a definition can hardly be made [262]. Privacy is closely related to human freedom, dignity and independence, has existed ever since, and is thus not a novel phenomenon

related to digital technologies. However, information technology increasingly challenges our privacy [5, 445]. Sophisticated knowledge discovery, the omnipresence of digital footprints, and the blurring line between private and public data make it challenging for people to obtain their privacy [274]. Helen Nissenbaum's privacy as contextual integrity [291] describes privacy rather as a matter of context.

Various theories explain how and why users behave privacy-wise. The Privacy Calculus Theory [314] states that users weigh the risks and benefits of disclosing their data to come to a decision. For example, the perceived social benefits outweigh the risk of data privacy issues for social media apps.

The construct of *Service - Privacy Fit* is an antecedent to the Privacy Calculus. It describes whether the service of an app matches its requests Hsieh and Li, Hurwitz [192, 197], and is thereby part of the user's risk assessment. Its mediating effects are mainly benefit expectancy and perceived privacy concern about whose trade-off users decide. It stems from the older construct of *task technology fit* [168]. A model of factors yielding user's perceived information privacy perception is described by Dinev et al. [115]. The particularly relevant correlates to information privacy are anonymity, secrecy, confidentiality, and control.

The relation between willingness to be profiled and the desire for transparency features is described by the Personalization-Privacy Paradox [20]: Users who value transparency features are less willing to be tracked and profiled. Karwatzki et al. [213] justify this with those people being "privacy fundamentalists," which means that they are careful with their data in general and value privacy more. Moreover, the Privacy Paradox states that users are generally concerned about their privacy; however, this is not reflected in their behavior [5, 292].

When deciding for or against installing an app, users weigh costs and benefits [30]. While the most relevant factors in this decision model are an app's costs, design, and functionality, privacy is reported to play only a minor role [30]. However, as most users are initially unaware of their concerns [157], it is wrong to conclude that privacy has only minor importance.

Research has developed multiple mental models that describe and explain the thought and decision processes of people regarding privacy. Coopamootoo and Groß [97] compile definitions of mental models from privacy-independent research

of Johnson-Laird [207] and Craik [101], as “internalized, mental representations of a device or idea that facilitates reasoning. They [mental models] are simplistic and small-scale representations of reality."

The model of *Privacy As Expectations* by Lin et al. [255] explains privacy issues as a mismatch of expectations and reality: People have a simplified model of an app in their mind, describing how it works. Privacy problems arise if what the app actually does deviates from the user’s mental model. Bonné et al. [59] studied smartphone permission decisions, and found that a main decision reason is the user’s expectation about whether an app should need a permission.

We can also look at this from a *Cognitive Behavioral Theory* perspective. Here, the process of informing one’s behavior starts from a privacy attitude that has been learned and developed over the course of life [97]. From this attitude arise privacy concerns, which then lead to behavioral intentions. Intentions are then followed by behavior but may be intercepted by the attitude-behavior gap (also see *Theory of Planned Behavior* [9]). Users’ understanding of how a system is working is studied by Wash [421] and compiled into a set of mental folk models on security threats.

The *Privacy Calculus* states that users outweigh anticipated risks and potential benefits to decide for or against disclosing personal data [105]. Constituting the central behavioral theory that explains users’ decision process regarding privacy, Kehr et al. [215] extended it also to incorporate user’s dispositions and attitudes. Users mostly accept the cost of data being collected if they want to use a service (cf. Price of Convenience) Ketelaar and Van Balen [219].

2.3.1 Defining Privacy Concerns

Coopamootoo and Groß [97] view privacy concerns as an instantiation of one’s attitude, referring to the user’s disposition. Colnago et al. [94], who interviewed experts in order to refine definitions of common privacy constructs, define privacy concern as “an expression of worry towards a specific privacy-related situation.” We thereby conclude that privacy concerns have two underlying components: (1) users’ predisposition, which is a result of past experiences (e.g., experienced incidences or things that happened to friends) and learned values and standards, e.g., developed by their education and social bubble. The second component is situations, e.g., when the smartphone requests sensitive data or when filling out a form. These two aspects are processed through the

user's mental model (cf. *Privacy as Expectation*, [255]) of how a system is working, possibly leading to feared consequences that might happen (= *privacy concerns*). To decide on a consequence (e.g., rejecting to provide data to an app) users apply the Privacy Calculus, outweighing situational perceived risk, and potential benefits.

2.3.2 Concern Taxonomies: Situational

Literature has proposed taxonomies for these arising privacy concerns. Merging models from Smith et al. [377] (dimensions of individual's concerns about information privacy practices) and [379] (structure their taxonomy by activities that invade privacy) privacy concerns can be described by four dimensions: 1) Data collection, 2) Data processing, 3) Data dissemination / improper access (i.e., how does data get into the hands of others?), and 4) Invasion and secondary use (i.e., what is the gained data used for, which implications on the user arise).

2.4 Privacy in Mobile Sensing Applications

Research on privacy and security in mobile sensing apps find that existing privacy-enhancing systems lack clarifying privacy implications, and users behave inconsistently with their concerns [89]. Klasnja et al. [227] interviewed participants of a personal context-sensing study on their privacy concerns. They found that people were primarily concerned about higher-level features derived from data than the raw data itself. For example, the ability to derive one's home address was rated more critical than the continuous GPS data collection. As a remedy, they suggest explanations of what an app is doing in the background, alongside features offering transparency and control on what data is recorded. Wang et al. [418] introduced a threat model and a taxonomy for privacy issues to bring some structure to the space of potential attacks. They distinguish between *task privacy*, *identity privacy*, *attribute privacy*, and *data privacy* to further propose privacy protection schemes for each privacy issue.

To manage privacy on smartphones, all major mobile operating systems implement a permission system, where users must grant data access for specific datatypes to apps individually [312]. However, the implemented smartphone privacy concepts face limitations and rarely introduce real privacy from a user perspective [140]. Christin et al. [89] found that existing privacy-enhancing systems lack clarifying privacy implications,

and users behave inconsistently with their concerns. Balebako et al. [25] found that users are not careless on that topic, but instead have misconceptions about data sharing that happens through smartphone apps and lack sufficient information. On the other hand, users outweigh anticipated costs and potential benefits, referred to as *privacy calculus* [105]. Here, users accept data being collected in exchange for being able to use a service, cf. *Price of Convenience* [219]. Additionally, today we see *digital resignation* [123, 364] or *privacy fatigue* [86] as an overload or lack of control leads to resignation, i.e., users giving up dealing with privacy decisions. As such, users face a challenge with the fundamental concepts of permissions and the associated privacy.

An increasing challenge in the privacy domain is data inference. Through machine learning and knowledge discovery processes further information can be derived from on the first sight uncritical data. Users hardly understand these processes [159], underestimate the underlying potential of inference threats [352, 405].

We also see limitations in the user interface itself. For instance, during permission requests, users show low comprehension [144], leading to a *lack of transparency*. Also, apps often do not convey understandably which information leaves the phone, making it hard for users to understand potential privacy leakages [25]. The wording in the permission UI was also found to be hardly understandable, and it was hard for users to grasp the implications [217]. A general *lack of control* is a crucial cause of privacy concerns [268], which has been shown in the online shopping and social media context [415]. Keusch et al. [221] raised concerns about the lack of control. In some cases, users do not even have privacy in their own hands, e.g., if one user leaks a contact list to a service, the other users (who are contained in that contact list) can not do anything against it [312]. The aforementioned two aspects, (1) transparency and (2) control, are identified as the two main pillars of information privacy [47, 179], also coined as the principles of *notice and choice* [346, 438]. A privacy issue introduced by app developers is *permission overclaiming*, also called *permission overdeclaration* [19]. By setting too coarse permissions, developers may claim less data access than they technically have permission-wise. An inappropriately huge amount of permissions also reduces user trust in an application [390]. Fang et al. [140] studied permission overclaiming on the example of the internet permission. This permission poses insufficient expressiveness to enforce control over internet access (i.e., access could be restricted) [28]. They found that many applications would tolerate stricter permission

here. 62% request internet permissions, but 36% make requests to specific domains only. Furthermore, third-party libraries that request permission for their purpose lead to that permission being claimed to the full application [308]. Finally, laziness among developers can lead to permission overclaiming due to confusion about the scope of individual permissions [390] and the aim to “just make it work” [28].

2.4.1 Users’ Privacy Concerns

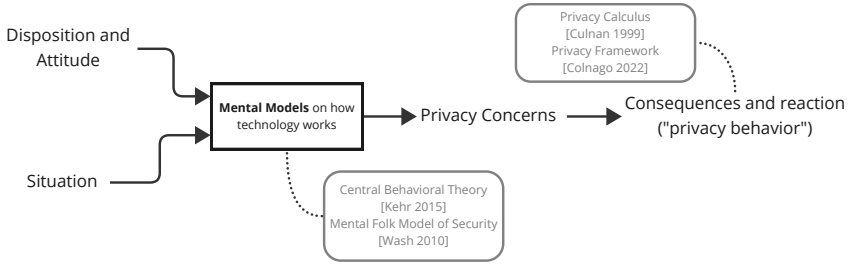


Figure 2.1 : A model compiled from related work that visualizes how various constructs in the privacy domain interplay. Privacy concerns base on users’ disposition and situation, depending on how a user assumes that a system is working. Thereon, users decide for consequences, i.e., mitigation behaviors, as described by decision theories such as the privacy calculus theory.

Privacy literature from other domains, such as smart homes [449], distinguishes between *assets*, *adversaries*, *vulnerabilities*, and *threats.* Building on these models from the literature, we group smartphone privacy findings into four factors by Zeng et al. [449]. We compile an overview of how users’ disposition, situation, and mitigation behavior relate to each other in Figure 2.1.

Differences Between Datatypes (cf. *Assets*) Summarizing findings from various studies, users are most concerned about login credentials [83, 152, 155, 324], which might lead to financial loss or identity theft. It is followed by contextual data, especially text messages [155] and address book/contact information [129]. Next comes personal

high-level behavioral data like GPS [152, 155, 227]. Behavioral sensor data like accelerometers were judged less concerning [227], likely due to the missing direct relation to personal high-level behaviors.

Differences Between Whom One is Sharing With (cf. *Adversaries*) Egelman et al. [129] reported that giving (un)authorized access to personal data is the largest factor in sharing decisions and an essential aspect of users' privacy concerns [127, 129]. Data sharing with can be categorized as *second-party* (sharing with the device or OS developing company itself [219]) or *third-party* (advertising companies or data brokers) data sharing. Users' attitudes and opinions towards third parties have been extensively studied= [219]. Yet, it is unclear whether third- or second-party sharing induces more concerns, cf. [152, 202, 373]. Nevertheless, Balebako et al. [25] reported that generally aware users, users are still unaware of the actual sharing scope (frequency, target, apps). Moreover, this varies depending on the surveyed population [6].

Underlying Events (cf. *Threats*) Users are generally unaware of which real-world implications they are afraid of and, as such, can not specify the purpose and reasons [152]. Among the few concrete reasons, financial and physical loss is most prominent [152, 155], followed by concerns about location data, which can lead to fears of physical threats [227]. However, users are more precise regarding specific domains and situations. For instance, Afnan et al. [6] reported that participants (Muslim women in the U.S.) feared being disproportionately subjected to security checks. Efstratiou et al. [128] found specific fears among their participants when sharing behavioral data in the workplace. Regarding identity theft, literature points to financial loss and destruction of personal reputation as concerning [295]. Users become especially concerned when they lose awareness and control of what happens to their data [373].

Underlying Reasons (cf. *Vulnerabilities*) Unauthorized remote access, like hacking, malware, data breaches, or compromised passwords, is mentioned frequently [152, 449]. Wifi and mobile networks are also often perceived as unsafe [83, 449]. Companies deliberately transmitting/selling data to others is also a frequently mentioned issue [6, 152, 202, 373]. Some people also fear the physical loss of their device, fearing that someone finding it could access their data [83]. Compiling a ranking of the issues

is difficult, as holistic, user-centered literature is lacking. Many studies are from a technical point of view instead of considering the user perspective or surveying only specific aspects.

2.4.2 Privacy Behaviors: The User Perspective on How to Mitigate Privacy Issues

Users accept to a huge extent that data is logged when they want to use technology [327]. However, they do not appreciate if practices are veiled and happening behind the scenes, as they always want to be aware and in control of the data flow [327], also including inferred data [89]. In our literature-derived model of privacy concerns (see Figure 2.1), users finally decide on *privacy behaviors* to mitigate their concerns. Colnago et al. [94] define these as *"What an individual actually does or has done in an attempt to achieve the level of privacy that they prefer."* The first group of privacy behaviors we found in the literature is about *improving a device's actual security*. Within the limited room for measures from a user perspective, studies primarily report actions on authentication management [152]. Such measures include choosing strong passwords, using 2FA, and password managers [155]. Further security strategies include clearing history data where possible [445, 449] and using security software [152]. As the second step of privacy behaviors, when users still do not have sufficient trust in a technology's security, literature distinguishes between measures that aim to *avoid behaviors* and *control data collection*. In the first case, users apply behavioral changes, leading to less data being provided to a device [152]. For example, making voice calls only in specific environments [127] or avoiding certain things in rooms with a smart home device [449]. In the online context, measures include avoiding behaviors by not doing certain tasks via mobile devices, such as opening attachments [152]. The latter measures to control the data collection include avoiding specific apps, websites, or networks [152].

2.4.3 Implications of Privacy Issues

Related work comes to the conclusion that privacy is the most important barrier to mobile app adoption [56, 80, 128]. Studies on adoption rates and reasons against the usage of mobile sensing apps identified a wide range of privacy concerns: General privacy and data security [80, 221, 266, 338], poor personalization [80], lack of

usefulness/trust in provided information [80], and general trust [338]. Over half of a smartphone app's users no longer want to install an app when they discover how much personal data is collected [62, 445], and about a third uninstalls applications when learning about collected information [62].

2.4.4 Role of Transparency and Control

According to Harari et al. [179], privacy should be incorporated into the entire process of self-tracking systems. Hence, transparency must be facilitated during each stage, opt-in should be adopted as the default setting, and control should be provided throughout each stage. Harari et al. [179] thereby distinguish between the two general privacy concepts of offering (1) transparency and (2) control. The demand for control resulted from a vignette study on the willingness to use passive mobile data collection technologies by Keusch et al. [221], where participants commented on a lack of control over their data. Literature provides design guidelines for features incorporating both transparency and control, for example a design space for privacy notices by Schaub et al. [347], and a design space on privacy control by Feng et al. [145].

2.5 Privacy Enhancing Technologies with Smartphone Sensing

For the implementation of various technical privacy protection methods, a large body of research exists, e.g., differential privacy [111, 258], on-device preprocessing [40, 111, 440], early aggregation of data [253], anonymous assessment [89], and cryptographic approaches such as self-destructive data [163]. From the literature, we found two major lines of advancements on privacy enhancing technology [60]. Harari [179] call to treat informed consent as a process rather than a one-time thing, to make data practices transparent. Their key aspects are what data is being collected, and how it is used. Reducing the amount of used data is an, in theory, easy step to improve privacy. However in practice it is often not possible to reduce logging and processing activities without obstructing a system's use case. A viable way to reduce the amount of that leaves the device, and thereby the scope of control of the user, is to apply more pre-processing on the client device. On-device preprocessing has shown beneficial in

other contexts, e.g., smarthome [334]. Recent research (e.g., [199]) motivate on-device preprocessing with mobile sensing applications. Overall, recent research emphasizes the necessity of a user-centered design, as non-technical measures show a high importance for users [96]. Thus, both consent mechanisms and data minimization approaches need to be designed for the user, in order to make an actual improvement in perceived privacy.

2.5.1 Data Minimization

Data minimization is an essential approach towards fulfilling privacy. It stands for using only the least possible amount of data that is necessary to realize an application's purpose [76, 243]. Unnecessary information, that may be included when raw data is initially collected for later extraction of higher-level features, should be discarded as early as possible. The principle of data minimization is also manifested in the General Data Protection Regulation (GDPR) stating "Personal data shall be adequate, relevant, and limited to what is necessary in relation to the purposes for which they are processed."[1] Although the principle of data minimization is relatively straight forward to operationalize and effective, it is in practice difficult to incorporate into apps from both design and technical perspective [205]. To operationalize data minimization, literature proposes various concepts. They can be structured by terminologies, such as by Pfitzmann and Hansen [316] who distinguish different applied strategies. Data minimization is often violated by the permission system [452], which does not pose sufficient granularity, requiring post-hoc data cleanup procedures. Trusted third-party runtimes are studied for example by Jin [205], where developers specify which fine-grain data they need, what is then enforced by the runtime. Data preprocessing on-device is also promising and studied in many projects [405], however mostly application-specific. Velykoivanenko et al. [405] thereby name and distinguish between preprocessing data on-device, and reducing temporal granularity, as most promising data minimization approaches.

Preprocessing approaches need to be adapted to the application's specific purpose, thus it is difficult to come up with generalizable approaches (e.g., [85, 172, 371]). For machine learning-based applications the amount of data that has to leave the device can be minimized through model training and improvement that happens solely on-device, namely model adaptation and federated learning approaches [58, 184]. For example,

[1] GDPR, https://eur-lex.europa.eu/eli/reg/2016/679/oj, Article 5 (1) (c), last accessed 2024-11-22

Xu et al. [443] leverage that to improve a smartphone keyboard's word suggestion model without one's typing data leaving the device. Challenging again is implementing such procedures understandably, i.e. so that users understand what happens and trust it, in order to actually yield an improvement in privacy perception [405].

2.5.2 Information and Consent Mechanisms

Research has investigated various information and consent mechanisms. In the following, we give an overview of different popular approaches.

2.5.2.1 Privacy Dashboards

The privacy dashboard is a common privacy design pattern [121]. Other privacy design patterns are, for example, the *Personal Data Table* and *Privacy Policy Icons* [372]. Privacy dashboards make users aware of the data which services have collected about them. They should provide successive summaries of the collected data and give an easily understandable overview [121, 457]. For this, they can use demonstrative examples, predictive models, visualizations, or statistics. Additionally, a privacy dashboard can provide control options and privacy settings to empower users to control the processing and collection of future data, cf. [121, 457]. Especially actions like deletion and correction of data are highlighted. Finally, privacy dashboards should give an overview rather than presenting every detail of possibly thousands of data items [372].

Privacy dashboards are spreading in practice, for example, the Google Privacy Dashboard[1], and have become subject to research. They were examined as a GDPR compliant alternative to consent forms [54]. They were studied as tools to give users a sense of what data is collected and inform the user instead of listing every detail [20, 213]. Raschke et al. [329] implemented a comprehensive privacy dashboard that adapts the newsfeed concept from social media. The dashboard incorporates transparency and control features so that users can view the collected data and learn about the purpose by obtaining information about involved processors, requesting rectification or erasure of each data item in the timeline, or reviewing and withdrawing the consent for each individual data type.

[1] https://myaccount.google.com/dashboard, last accessed 2024-11-22

Privacy dashboards are a tool to implement the principles of *notice and choice* [346, 438], often through features that provide transparency and control [353]. Transparency and control have long been studied in the context of mobile systems. Permission popups force mobile apps to provide transparency and control about what data an app can access [144]. However, the context is limited [410]. Permission popups lack appropriate information and contain hardly understandable terms, making it hard for users to grasp the implications of granting permission [217]. Also, the amount of information conveyed to the user leaves space for improvement [53, 144]. In contrast, interfaces that provide more detailed information on what happens with the data increase user confidence [402]. In the web context, similar issues have proliferated. Privacy policies are long and hard to understand and, thus, often ignored [294]. In addition, they fail to provide sufficient transparency to the user [53]. Here, consent popups may even be designed to nudge users towards illegal configurations [293].

Permission popups offer transparency and control before the data logging happens. In contrast, privacy dashboards take effect afterward. The retrospective approach has the advantage that the user can be informed about what has actually been logged. Transparency and control features, incorporated through privacy dashboards, have shown positive effects: The Google privacy dashboard [141] and a dashboard for online shopping [187] increased user trust. However, this is only valid for raw data: In the study of Herder and van Maaren [187], showing derived data increased perceived privacy risk and reduced user trust. Perceived risks and trust may lead to fewer people using a service, not sharing required data, or dropping out early. For example, in vignette studies, participants indicated to prefer using a service that provides transparency over the logged data [20, 221, 397] or an option to switch off the data collection [221]. However, while control features show a positive effect, they are only seldomly used. In the studies by Farke and Elevelt only a quarter of the participants involved had already used or indicated a willingness to use such features in the future [134, 141].

While the previously reported studies in the contexts of webshops, surveys, and personalization of online services agree that the provision of decision-relevant information is positive [456], the literature is contradictory in the context of sensing data [213]. Here, transparency increases privacy concerns resulting in less data disclosure [206].

This inverse impact of transparency features can also be found in studies on personalized advertisements [422] and inferred user interests [353]. Also, the reaction to transparency features depends on the user's privacy predisposition [353].

2.5.2.2 Alternative Smartphone Permission Systems

Many technical solutions have been proposed to serve as a middleware between the app and the user as privacy-enhancing technologies (PETs) for smartphones (e.g., [23, 136, 363], for an in-depth survey see [370]). Moreover, Pennekamp et al. [312] reviewed privacy enforcement strategies on smartphones. On the level of user manipulation, they structure concepts regarding privacy mechanisms into three categories: 1) **reporting** (e.g., omnipresent install prompts, permission visualization, and ways to allow tracking the flow of private information), 2) **fine-grained tuning** (such as user-based configurations), and 3) **fencing** information (e.g., Mockdroid [52], TISSA [455], SHAMDROID [68]).

With the evolution of mobile operating systems, fine-grained control has proliferated in slow steps. Permission popups allowing to choose "only one time" access mitigate the issue of permanent access [283]. Hong et al. [190] already proposed sliders as interface elements as an extension for the one-time-only feature with three options options: allow, ask, and deny. Research proposed various approaches to give users finer control of their data. Jeon et al. [204] categorize permissions into four classes (e.g., outside resources, sensors), each of which common strategies for permission subdivision can be applied to. Zhou et al. [455] enables users to bypass the compulsion to grant a permission to use an application by giving the option to pass falsified data such as empty data, anonymized data, or bogus. Other approaches involve restrictions on how many times a critical resource may be accessed [289], and context-dependent privacy policy configuration [95].

More drastically, Scoccia et al. [363] restructure the Android permission interface by allowing users to (1) make permissions on a feature level and (2) grant finer-grained permissions by introducing permission levels, i.e., a granularity at which data or a resource can be accessed. They found that users appreciated the greater choice, felt more control, and had higher trust. The traditional Android permissions, in contrast, were described as misleading, and the enforced binary choice was not preferred. However, their study is rather proof of the general concept of more granularity in

smartphone permissions, emphasizing the realizability in Android. The major part of their contribution is an implemented app instrumenter and its evaluation. The design process of their so-called permission levels has come up rather short.

3

Extracting Information With Mobile Sensing

In this chapter, we show which application use cases smartphone-based mobile sensing enables. We show that **mobile sensing can yield benefits for users, researchers and the society**. Through the presented studies, we point out the data demands that such use cases have in order to work properly, and conclude on current **limitations in data availability and usability** for privacy and security reasons. These findings **motivate the subsequent chapters** of this thesis: To tackle the privacy issues and data access restrictions that current mobile sensing apps face, we need deeper insights into the users' privacy perceptions of such apps. Based on these, we can then propose concepts to mitigate these issues.

To realize their purpose, applications require information about the user's situation and context. Without sharing information and data with our smartphone, its user experience was way worse. Information-retrieval tasks would yield less accurate results, i.e. require more detailed search query specifications of the user. Monitoring activities

in the background would become practically impossible. In mobile sensing systems, we can distinguish between two types of stakeholders: The *participant* who creates the data, and the *consumer* who makes use of it [266]. These may be different or the same entities. The application of mobile sensing technologies yields benefits for different types of consumers (most prominently the user themselves, the app developing company, or a data-collecting organization such as a university) and has proliferated in various domains (see for example [244] for an overview). With mobile sensing, it becomes possible to **generate patterns and extract knowledge** from the data. The, in comparison to other manual data sources, rather large amount of available data through passive sensing enables the application of data mining and machine learning techniques. Models can be built offline with sensing-collected data (e.g., [416] or [261]), or trained and optimized online with continuous user data (e.g., [260]). Multiple devices can also work together in a multi-agent architecture, yielding ambient intelligence [22], i.e. an intelligent environment that is aware of surroundings and people. For the end user this leads to more context-aware applications, pro-active application behavior, more precise recommender systems and personal daily support. These benefits mainly target the end-users, although app developers also indirectly benefit by an improved app experience.

Crowd Sensing Allows Data Collection at Scale. Due to the dissemination of smartphones in our society, sensing can be deployed at scale easily from a technical perspective. However, in practice individuals are often reluctant to install a sensing app due to privacy concerns and other possible disadvantages such as reduced battery endurance, especially if they cannot expect a personal benefit thereof [74]. In contrast to the aforementioned kinds of benefits, crowd-sensing applications usually fulfill a purpose that is in the interest of the app-publishing company or organization, for example a governmental organization tracking parameters in cities (e.g., [126]) or a university that deploys a sensing app to collect data for a study (e.g., [384]).

A major barrier of mobile sensing for its adoption are privacy concerns of the smartphone users [221, 266]. They weigh perceived privacy risks and expected benefits (cf. [105]) to make a decision for or against adopting an app. This inherently restricts the space of implemented benefits, as OS and app developers have to mind these privacy aspects. Disregarding privacy-induced limitations, more would be possible today. In the vicious cycle where privacy restricts possible implementations and studies,

it is hardly possible to gain knowledge on future directions for improvement. However, insights on what kind of data would be of interest for different stakeholders would be important to motivate improved privacy-enhancing technologies in smartphones, and steer which key characteristics they should implement.

In this chapter, we investigate which kinds of benefits can be expected from mobile user quantification. We point out which requirements regarding privacy these use cases bring, what informs the design of appropriate information and consent approaches. Our work proposes mobile sensing use cases that hardly exist yet due to privacy reasons, while outlining a perspective towards making them possible with further research.

We show that both stakeholder types *user* and *app developing organization* would benefit from more in-depth user quantification, i.e. that the usage of more detailed mobile sensing data would be in their interest. Furthermore, we introduce the *society* as third stakeholder. Technology has shown to have strong effects on our societies, e.g., through influencing voting behavior [122]. We show that mobile sensing data can not only yield a benefit for individual users or organizations, but also for the society as a whole.

Our work motivates further research towards a privacy-paradigm in mobile sensing, that goes beyond solely limiting data access. The collection and processing of mobile sensing data is in the interest of the data-generating user. However, this is not valid for all use cases - that may also be used for purposes that are against the users will. As a next step, it is therefore important to, first, study privacy-issues that arise from more in-depth mobile sensing data usage, what users are afraid of, and what are important factors for them. Second, these insights can inform the design of novel privacy-enhancing technology concept, that protect the users privacy and helps them applying their will, while keeping the data usable so that all stakeholders can benefit.

In this section, we present three benefits that mobile sensing can generate. Each is devoted to a different stakeholder: First, we describe a logging and context-enrichment approach for mobile language data, which is to the benefit of researchers. Mobile language data is a powerful data source fueling research questions in various fields and interdisciplinary projects. Second, we investigate how the joined usage of deep smartphone usage data and contextual sensing data could support the user, by fastening their interaction with their device. By predicting the next action performed, the smartphone could anticipate desired results before the user requested them. Lastly, we

	Research Question	Paper	Section
RQ1	Which benefits can be expected from Mobile User Quantification?		Chapter 3
RQ1a	Can the evaluation of input prompt text meta data create a benefit for interdisciplinary mobile sensing research?	[42]	Section 3.1
RQ1b	How can contextual information improve mobile sensing based research data?	[42]	Section 3.1
RQ1c	How can mobile sensing enable novel adaptive and intelligent smartphone app use cases that are to the users' benefit?	[393]	Section 3.2
RQ1d	How does mobile sensing at scale interplay with our society and its challenges?	[43, 45]	Section 3.3

Table 3.1 : Overview of the studied sub research questions of RQ1.

show that also the society as a whole can benefit from mobile sensing technology. We propose concepts how mobile sensing technology could be applied to target societal challenges, such as distorted public opinion making and climate change.

All three have in common that they are nowadays hardly possible due to privacy issues. Smartphones are lacking means to provide the used data in a privacy-friendly manner. Thus, users either show a low willingness to adopt the envisioned application scenarios, or the OS developers deny access to such data at all.

To motivate work on improved privacy concepts, the studies in this chapter follow the overall research question:

RQ1 *Which benefits can be expected from mobile user quantification?*

An overview of the research questions of the contained subsections is given in Table 3.1.

3.1 Passive-Sensing Data for Interdisciplinary In-the-Wild Research

This section is based on the following publication:

Publication planned: Florian Bemmann, Timo Koch, Maximilian Bergmann, Clemens Stachl, Daniel Buschek, Ramona Schoedel, and Sven Mayer. “Putting Language into Context Using Smartphone-Based Keyboard Logging.” In: *arXiv preprint arXiv:2403.05180* (2024)

Mobile sensing methods allow researchers to collect information about people’s behavior unobtrusively in the background over longer periods of time [182, 384]. In various research disciplines, such data collection methods have shown to outperform classical methods, such as observations and self-reports.

Of special interest for interdisciplinary research thereby is language data: Language is one of the most effective ways to gain insights into people’s minds [87, 309]. Designing personalized interfaces with great user experience in technical systems requires a profound understanding of user‘s inner feelings and thoughts in relation to their interaction behaviors in specific contexts. Digital footprints are left behind in people’s everyday language, such as Facebook posts [362] or WhatsApp instant messages [228, 407] have been shown to provide useful information on psychologically relevant traits such as personality traits or depression [131, 362]. Hence, understanding the mental models, attitudes, psychological states and dispositions of users is a core goal of research in human computer interaction and in neighboring fields (e.g., psychology, behavioral science). User modeling is often used in HCI to quantify individual differences for system adaption in a technical systems.

Regarding language in context is thereby very important, as the meaning of textual contents highly varies with the situation in that it is expressed. To regard text inputs into context when logging smartphone typing behavior, usually, the target app is used as a proxy, for example, by filtering for a defined set of communication apps (i.e., WhatsApp, Signal, Telegram), or using an app categorization mapping (cf. [357, 384]). The target-app name is a data feature that is available on most operating systems easily and common to be evaluated.

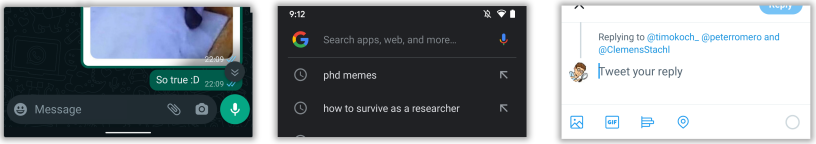


Figure 3.1 : Screenshots of three text fields of the three Android apps: Google WhatsApp (left), Google Search (middle), and Twitter (right). All three text fields have input prompt texts, that give the user a hint about what the text field is intended to be used for.

However, the target app name only roughly describes the context in which a text input happened. It has two major drawbacks: 1) Defining a list of apps that contain the desired behavior (i.e., messaging or posting on social media) is difficult. A wide variety of apps exist and their relevance changes rapidly. Thus, categorization approaches require frequent adaptation. 2) An app may be used for various purposes and thus, cannot be assigned to one category. For example, the very popular app Instagram is used frequently to post public content and direct messaging. However, it is regarded as a social media app only in most app categorizations. Also, search fields that exist in nearly every app contaminate the resulting text data.

In this study, we demonstrate that the evaluation of more detailed, contextual data can bring a benefit. User interface properties, such as properties of the typed-in text field and the surrounding UI can give the text input more context. Selected as an exemplary property, I evaluate a input field's input prompt text as contextual property. The input prompt text gives the user a hint about what kind of content the app is expecting, which researchers can leverage to understand the input motive of the user. With that, we propose researchers who study mobile language to filter for *WHAT kind of content* users type instead of *WHERE the content was typed*. We show some examples of common input prompt texts in Figure 3.1.

In this study, we deploy a holistic logging and preprocessing approach for smartphone typing data. On the example of the input prompt text, we investigate how input prompt UI metadata can be used to contextualize language data and, in this case, allows us to distinguish language contents by their input motive (i.e., posts, comments, messaging, search inputs).

We base the categorization on the dataset from a six-month representative study ($N = 624$) that is the first of its kind using this approach. With this context-enriched keyboard logging and analysis method, we show the untapped potential for research in sensing language data. We publish our input prompt motive mapping, alongside an Android library that allows other researchers to reuse our system in the wild. We critically discuss the threats and dangers that are posed by such technology. While our logging approach already improves user privacy, we point out space for further improvements and motivate future work in that area.

Building on our study of the example of the input prompt text, our insights motivate further efforts to regard more properties of the UI elements that surround a text input field. We discuss the unraveled opportunities for interdisciplinary researchers who deploy mobile sensing to collect study data in my discussion.

3.1.1 Research Gap: Leveraging Typing Meta Data

Implementations that access the input events by replacing the keyboard with a custom implementation allow access to typing metadata (cf. [40, 412]). Only a few works have been conducted that make use of mobile typing metadata (e.g., [64, 165]), and the derivation of higher-level contextual features that are relevant to interdisciplinary research is rare. Nowadays, when analyzing language use, researchers mostly select and contextualize their data by filtering by the used app, which is only a proxy for the type of content and the user's intention, but in practice not accurate. For example, social media apps can be used for both crafting public content and sending private messages.

We fill this gap and motivate further research by proposing a context-enriched keyboard logging approach. Choosing a language logging approach that exchanges the keyboard with a custom implementation or leverages device APIs, meta information on where a user did type can be accessed. We specifically regard the input target app and the input prompt text, i.e., the text that is visible as a background or placeholder inside a text field as long as the user did not type any content. The usage of input prompt texts has not yet been studied as a data source for psychological research. Studied application scenarios in HCI encompass UI automation and privacy-enhancing tools, e.g., by Pereira Borges Junior [315] to access a specific text field for automated UI testing or by Wanwarang et al. [420] to help a system understand the meaning and

input concept of a UI element, respectively the latter by Andow et al. [14] and Huang et al. [194] who use input prompt text to estimate whether a text field is intended to contain potentially sensitive contents.Thus we put up RQ1a of this thesis:

RQ1a *Can the evaluation of input prompt text meta data create a benefit for interdisciplinary mobile sensing research?*

Building on my insights obtained on the example of input prompt texts with mobile language input, we generalize our results and answer RQ1b:

RQ1b *How can contextual information improve mobile sensing based research data?*

3.1.2 In-the-Wild Logging of Text Input and UI Contextual Information

To tackle the aforementioned research gap, we study the potential for language use research that lies in UI metadata. We propose to use device APIs to log mobile typing behavior and create an approach that makes provided UI metadata usable to create contextual variables. We develop and provide (a) logging libraries for the Android operating systems to track language input data, and (b) a categorization approach for UI metadata that extracts a variable on the user motive of text input from the UI property input prompt text.

Our method puts an emphasis on:

- *Comprehensiveness.* Being able to regard text inputs across all apps and input methods.
- *Context-sensitivity.* It should be possible to regard text inputs in the user's context.
- *Reducing Observer Biases.* The logging method should have the least possible observer bias on the participant.
- *Privacy.* Text input data contains privacy-invasive information, and should thus not simply be recorded as it is.
- *Replicability and Adaptability.* The presented approach should be applicable to many different kinds of research from interdisciplinary fields. Therefore we report our full research pipeline so that it is reproducible, and can be adapted to other research's demands.

3.1.2.1 Input UI Metadata

We logged the input prompt text alongside a text input through an Android API. This gives us access to all typing events at full granularity, and allows to retrieve UI metadata. We abstract each typed word into psychological categories on-device, to ensure that no raw text contents are logged. The available text input metadata encompasses the input prompt text, also a title, label, and styling/sizing properties. Label and input prompt texts tell users what they are supposed to enter and which type of input the app is expecting. Input prompt texts are visualized as placeholders, that indicate what the user is supposed to type, e.g., “Message” (WhatsApp) or “Tweet your reply” (Twitter), see Figure 3.1.

3.1.2.2 Deriving Contextual Feature: Input Motive

In our approach, we solely rely on the input prompt texts. We create a categorization, that maps input prompt texts into one of 7 categories, i.e. *input motives*. An *input motive* describes the purpose and kind of content that a user is supposed to enter into a text field, such as search inputs, direct messaging, or public social media posts. We visualize this concept on the example of the app Instagram in Figure 3.2, list the input motives in Table 3.2, and describe the categorization procedure in detail in Section 3.1.4.

3.1.3 Mobile Sensing Field Study

In this section, we present our field study, which to the best of our knowledge, is the first large-scale deployment ($N = 624$, 3 to 6 months) of smartphone-typed language logging in the wild that applies privacy-respectful on-device preprocessing. Its purpose for this paper mainly is to retrieve a body input prompt texts from real smartphone usage to create an input prompt categorization that maps input prompt texts to motives. This can be used as basis for future projects, where it allows to categorize data directly on-device. Furthermore we describe the characteristics of the obtained data, to give researchers an overview of what data to expect from such a study.

3.1.3.1 Implementation

To collect a solid data basis for our categorization in the wild, we developed an Android app that participants had to install on their smartphones. To access typed language data, it subscribes to content change events of input fields[1]. Thereby it is notified on each change, i.e., each typed or removed character. For the preprocessing procedure we rely on the provided code of Bemmann and Buschek [40], however our logging is implemented differently, as we did not adopt their approach of replacing the device's keyboard with a custom implementation. With the receiving event comes a reference to the target input field[2] from which we retrieve additional metadata, namely the app name that the input field belongs to, and its input prompt text.

3.1.3.2 Data Collection: Representative Smartphone Sensing Panel Study

The study was not conducted exclusively for and tailored toward the specific needs of this paper. Data collection was conducted as part of a cooperation project between the LMU Munich and Leibniz-Institut für Psychologie (ZPID). Its purpose is to collect data to answer various research questions from psychology, sociology, and human-computer interaction.Mobile sensing, experience sampling, and survey data were collected during an individual study period of up to six months (from May until November 2020). A detailed description of data collection procedures is supplied in the pre-registration of the study protocol [356]. In the following description, we focus on the parts of the study that are relevant for this project. This research was ethically approved and carefully aligned with EU GDPR guidelines.

By means of a provider for non-probability-based online panels, a starting sample of 851 participants was recruited according to a pre-specified quota (gender, age, education, income, confession, and relationship status) representative of the German population in the age group 18-65. Participants were required to own an Android smartphone, on which they were asked to install our research app *PhoneStudy*: The app had access to the users' text inputs via the Android Accessibility services[3].The

[1] Android Accessibility Event type VIEW_TEXT_CHANGED https://developer.android.com/reference/android/view/accessibility/AccessibilityEvent#TYPE_VIEW_TEXT_CHANGED, last accessed 2024-11-28

[2] EditText object https://developer.android.com/reference/android/widget/EditText, last accessed 2024-11-28

[3] https://developer.android.com/guide/topics/ui/accessibility/service, last accessed 2024-11-28

research app ran continuously in the background of participants' smartphones during the study period of three to six months. Language data thereby were pre-processed with the Language Logger abstraction module of Bemmann and Buschek [40], using the LIWC dictionary[1] [310] for word categorization and a German dictionary [230] as a whitelist for word frequency counting. The full body of measures is reported in detail in the study preregistration [356].Participants were compensated in stages, i.e., depending on how long and in which parts of the study they participated. They were excluded from data collection after several reminders if they revoked permissions for mobile sensing data collection several times for longer than seven consecutive study days and failed to complete two out of three monthly online surveys per study half. Participants could withdraw their consent at any time and ask for their collected data to be deleted. All log data was deleted from the client devices when the study finished, i.e. immediately before prompting the participant to choose their compensation.

3.1.3.3 Sample

The initial sample consisted of 851 users. We excluded participants who did not grant all necessary permissions or experienced technical issues (42), switched their primary smartphone during the study (44), or participated less than two weeks (141).This leads to a final sample of $N = 624$ available for further analyses. The mean age in our sample was 42.65 years (SD = 12.60). The age distribution follows the population in Germany, having one peak around 40 years and another between 50 and 60. Gender groups were balanced with slightly more male participants (our sample: 55.34% male, 44.49% female, 0.18% other; national statistics: 49.34% male, 50.66% female, divers is not captured by the statistic)[2]. Regarding their highest completed education levels, our participants are rather well-educated in comparison to national averages. We report a higher amount of participants holding a school degree (80.00% of the participants hold a school degree as their highest education (national statistic: 52.58%[2]), while nearly none of our participants does not hold any degree (0.81%, national sample: 21.27%). 19.12% hold a university degree (national statistic: 26.15%[2]). Only native Android

[1] LIWC is proprietary software. Its usage must be cleared with the authors.

users were considered as participants, i.e., people who use an Android device as their primary phone. No devices were handed out for the purpose of participating in this study.

3.1.3.4 Data Analysis and Offline Preprocessing

The data analysis was conducted on a central server with the statistics software R[1]. No raw data was downloaded to local computers. The data analysis was divided into two steps: In the first step, we imported the data, did general preprocessing (e.g., parse timestamps, select the final sample), and grouped keyboard events by their text input. Word category occurrences were encoded in the many-hot format. The resulting data was in a table-like format, each row representing a text input. If not described otherwise, we analyzed events of type *added* and *changed* and discarded *removed* events. The data were grouped per text input. In the second step, we categorized apps and input prompt texts.

3.1.4 Input Prompt Text Categorization: Distinguishing Language Contents by Their Input Motive

In this section, we describe our input prompt text categorization and show the process of how we created it. We attach our input prompt text category mapping to this paper. However, we think it is important to enable researchers to create their own category mapping, to meet the specific characteristics of their dataset.

Prefiltering Input prompt texts can be dynamically created by the according smartphone app, and may thus contain private content. For example, some apps extend input prompts such as “Reply to message” with the corresponding user name, i.e. “Reply to John Smith.” As it is our premise not to have names of conversation partners in our dataset, we filtered such input prompt texts. We did this through the assumption, that an input prompt text that contains private information, is unlikely to occur for multiple

[2]Source: Federal Statistical Office of Germany https://www.destatis.de/EN/Home/_node.html, last accessed 2024-11-28

[1]https://www.r-project.org/, last accessed 2024-11-28

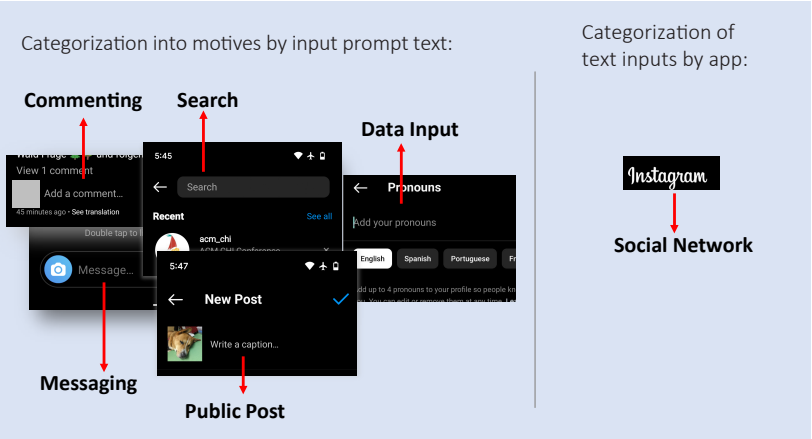


Figure 3.2 : Instead of categorizing collected in-the-wild text input data by the originating app, we propose to regard the originating text field's input prompt text. This figure shows on the example of the app Instagram, that text inputs into Instagram are not just social media contents such as posts and comments, but can also have other motives such as messaging, search, and data input.

distinct participants. Thus, we removed all input prompt texts that only occurred with one participant. To aggregate the input prompt texts into input motives for further analyses, we applied a semi-automatic, two-step categorization concept:

Identification of Major Motives We identified five major motives that users follow when composing texts on their smartphones. Those constitute our input motives: *Messaging*, *Posting*, *Commenting*, *Search*, and *Data Input*. Furthermore we have the two categories *Other* and *Ambiguous*: input prompt texts that cannot be assigned to one of these were labeled as *Other*. *Ambiguous* was used if the meaning of an input prompt text was unclear at all. We describe the input motives in Table 3.2. Our categorization regards all smartphone text entry interactions of all participants, regardless of app category or any other pre-selection.

Categorization Step 1: Automatic Keyword Stem Matching: To categorize input prompt text into input motives, we applied a semi-automatic approach. In the first stage,

Motive	Description	Example	keywords for automatic categorization
Messaging	A private message targeted to a defined person or group of people	"Type a message" (WhatsApp), "Enter your message here" (Facebook Messenger)	nachricht, message
Posting	(Semi-)public posts in social media applications. They are visible to either anyone (public posts) or limited to a group of people (e.g., friends of that user)	"Write a caption" (Instagram), "What are you doing?" (Facebook)	
Commenting	Content that is attached to an existing post, usually with the same visibility as the post	"Comment ..." (Facebook), "Tweet your reply" (Twitter)	komment, comment
Search	Content that constitutes a search query. e.g., inputs into search fields	"Search apps, web, and more..." (Google Quicksearch), "Search photos..." (Gallery app)	such, search
Data Input	Inputs that ask the user for some information, usually form fields	"email address" (on a login screen), "Stop, address, ..." (in a public transport service app), "Spanish translation" (in a language learning app)	
Other	The input prompt text cannot be assigned exactly one motive, or the purpose is clear but does not belong to one of the five main motives	e.g., experience sampling and questionnaire items, "write a note..."	
Ambiguous	The input prompt text's meaning and purpose is not understandable at all	"0," "???"	

Table 3.2 : We categorize text inputs on smartphones into *input motives*, using a text field's hint text.

we were looking for keywords that could be used to categorize input prompts automatically. We decided to classify all texts that contain German and respective English word snippets "such" / "search" as *Search*, "komment" / "comment" as *Commenting*, and those containing "nachricht" / "message" as *Messaging*. With this step we could categorize 3,671 distinct input prompt texts.

Categorization Step 2: Manual Coding: For further categories, we did not find a reliable matching scheme and thus, proceeded to categorize all remaining input prompt texts manually. We limited our manual categorization to only those input prompt texts which occur in more than 0.01% of all logged texts (465 input prompt texts). The categorization was done independently by three researchers (R1 to R3), in three iterations.

	Assigned Input Prompt Texts				
Input Motive	**Overall**	**Manually Coded**	**Disagreements**	**Cohen's Kappa**	**95% CI**
Messaging	568	31	4	0.88	[0.79;0.97]
Posting	20	20	4	0.75	[0.6;0.9]
Commenting	157	4	1	0.75	[0.41;1]
Search	2997	16	1	0.78	[0.63;0.93]
Data Input	248	248	12	0.85	[0.81;0.9]
Other	13	13	7	0.8	[0.57;1]
Ambiguous	105	105	16	0.82	[0.75;0.88]

Table 3.3 : Number of assigned motive categories, alongside disagreements and interrater agreement to each motive category of the manual coding process.

R1 in the previous step created the category definitions and the first coding iteration was done by two other researchers independently. We thereby intend to keep the coding free of biases that arise from the motive identification phase. R2 and R3 both coded 50 randomly selected input prompt texts independently. They afterward compared their coding and discussed the differences. Their notes and suggestions were given to R1, who then improved the motive definitions. Changes included refinement of the *Data Input* category, and splitting up and refining *Other* and *Ambiguous* (in R1's initial definition, they were joined). In the second coding iteration, R2 and R3 used the refined motive definition to code the remaining input prompt texts. They agreed in 89.7% cases, of 438 input prompts that were manually coded, R2 and R3 agreed on 393 cases, and for 45 cases their coding differed. To check the inter-coder reliability we evaluated Cohen's Kappa [304] resulting in a nearly perfect agreement, according to the classification of Landis and Koch [241] ($K = 0.83$, 95% CI: $[0.78; 0.88]$). The remaining discrepancies were resolved by R1 by coding the input prompts independent of the decisions of R2 and R3. Afterward the coding of R2 and R3 were additionally taken into account to overthink the coding of R1. The different coding decisions were discussed where necessary, and R1 made the final decision. This increased our final coverage to 88.4% of all text inputs, with 4,108 distinct input prompt texts being covered.

3.1.5 Descriptive Evaluation of Input Prompt Text Categorized Data

In this section, we evaluate descriptive statistics of our input prompt text categorization. We report the characteristics of our mobile typing language dataset with input motive as a new independent variable. We show the advantages that researchers gain when selecting and filtering their language data by input motive, instead of by app category.

3.1.5.1 The Dataset

In total, our analysis encompasses 1,868,416 text inputs across 624 users. The average smartphone user in our sample typed 23.40 text inputs per day ($SD = 28.23$). These made up 146 words ($SD = 155$) per user per day. Regarding all text inputs, each text input consists of on average 9.44 words ($SD = 30.7$). During a text input, the users added on average 8.13 words ($SD = 25.6$), changed 1.31 words ($SD = 6.35$), and removed 1.50 words ($SD = 6.86$) words.

3.1.5.2 Descriptives of the Input Prompt Categorisation

Of all 1,868,416 text inputs, 1,112,317 (59.5%) obtained an input prompt text. Our categorization labeled 983,281 (88.4%) of them, covering 52.63% of all text inputs.

The most frequent input motives were *Messaging* (44.0%), *Search* (33.8%) and *Data Input* (12.2%). The social media categories *Posting* (1.0%) and *Commenting* (3.2%) occurred less often. The occurrence of the remainders categories *Ambiguous* (4.9%) and *Other* (0.9%) is on a low-level, what shows that our category mapping covers the user's motives well in nearly 95% of all text inputs.

3.1.5.3 Input Prompt Motives vs. App Categories: Changes in Data Characteristics

We compare the number of words per text input, regarding their motive category. We found that text inputs that fulfill a functional purpose, such as *Search* ($M_{Search} = 2.30 words$, $SD_{Search} = 6.80$) and *Data Input* ($M_{DataInput} = 2.73$, $SD_{DataInput} = 9.29$), are significantly shorter than texts of motives whose content is targeted towards other people, such as *Messaging* ($M_{Messaging} = 12.43$, $SD_{Messaging} = 18.80$), *Posting* ($M_{Posting} = 12.84$, $SD_{Posting} = 19.00$), and *Commenting* ($M_{Commenting} = 12.65$,

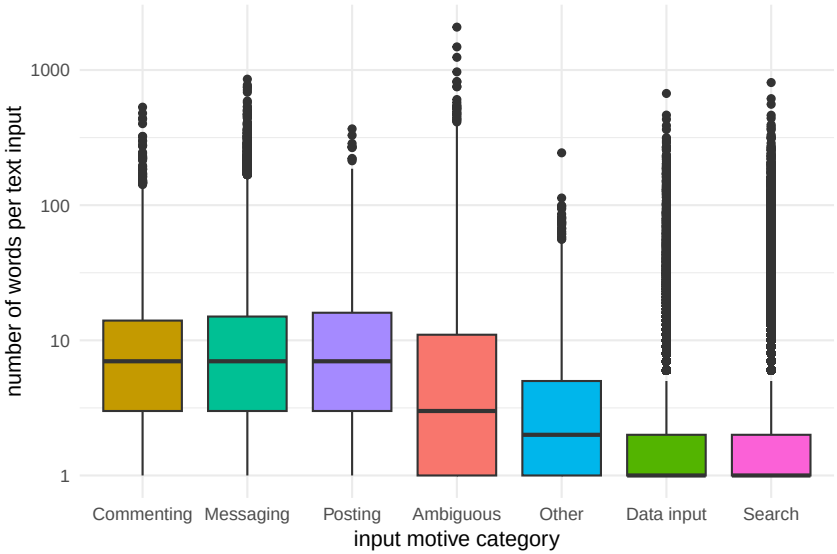


Figure 3.3 : Words typed per user per input motive. Search inputs are rather short (1 to 3 words), and Messaging inputs are rather long with 5 to 50 words. Social network contents like posts (Content Creation) and Comments range in between.

$SD_{Commenting} = 20.28$). *Other* text inputs range in between of both clusters ($M_{Other} = 5.32$, $SD_{Other} = 9.42$). A Kruskal Wallis rank sum test with a consecutive Dunn's test (p-values adjusted with Bonferroni method) revealed significant differences between the three groups *Search* and *Data Input* vs. *Messaging*, *Posting* and *Commenting* vs. *Other* for all pairwise comparisons with $\alpha < .01$.

Furthermore, we regard how much of the logged text contents could be matched with the applied dictionary. We analyze the different matching rates of text input (1) selected via our motive categories derived from the input prompt texts, and (2) filtered by the app categories of Schoedel et al. [357]. Therefore, we compare the matching rates of the extracted LIWC categories for the input motive *Messaging* with the app category *Messages*, for the input motives *Posting* and *Comment* with the app category *Social Network*, and the same named input motive *Search* with the app category *System* (a comparison with an app category is not clearly possible in this case. We have chosen the category System, as it contains some search apps).

Input Text Selection	LIWC MATCHING RATE M	LIWC MATCHING RATE SD	WORDS PER TEXT INPUT M	WORDS PER TEXT INPUT SD
App Category: Communication	48.32%	31.05%	16.43	38.30
input motive: Messaging	50.64%	30.33%	12.43	18.80
App Category: Social Media	31.59%	33.52%	12.87	39.60
input motive: Posting	38.82%	30.34%	12.84	19.00
input motive: Commenting	41.78%	30.35%	12.65	20.28
App Category: System	18.78%	31.67%	4.34	20.00
input motive: Search	13.02%	28.60%	2.30	6.80

Table 3.4 : Comparing characteristics of text inputs filtered by input motive (yellow background) and app category (green background). We compare mean and standard deviation for the two variables *matching rate* and *number of words per text input*.

In general, we found matching rates for messaging content being the highest (around 50% for both kinds of categorization), followed by social media content between 30% and the lower 40s. Search inputs, in general, match the dictionary rather badly (10% to 20%). This characteristic is not surprising, as messaging consists of rather natural language, whereas hashtags and more net-speech might enhance social media language. Search queries contain rather specific words (names, locations, etc.) that are not contained in the dictionary.

Filtering content by input motives instead of app categories could strengthen these characteristics: We found higher matching rates for messaging and social network content when selecting text input via input motives than when using the equivalent app categories (please find the effect sizes reported in Table 3.4). For the category *Search*, the matching rate behaves oppositely, i.e., the input motive *Search* shows a lower matching rate than the respective app category *System*. Standard deviations when filtering by motive category instead of app category were lower in all comparisons.

3.1.6 Discussion

3.1.6.1 RQ1a: Input Prompt Text Metadata Yields More Specific Context for Mobile Tying Data

With our study, we could successfully show that sensed UI metadata does help to put mobile text inputs into context. Regarding a text field's input prompt text allows

to derive the user's input motive. For interdisciplinary research this is of interest, as people behave differently depending on what kind of content they craft. For example, private messages have different characteristics than public social media posts [407].

The analysis of our study data showed that selecting text inputs via the input prompt text yields higher matching rates in the LIWC dictionary than when selecting via app categories for messaging and social media content. Although we did not compare modeling scores of exemplary dependent variables and psychological constructs (such as predictions of age and gender [228], affective states [229]), this implies that the yielded data is cleaner and of higher quality when selecting text inputs via the input prompt text. The opposite behavior for the input motive *Search* is not surprising and can be explained by the characteristics of search inputs: They are very short, usually just a few words (see Figure 3.3 for comparison) and consist of specific terms (e.g., companies, names, locations). They are, in general, more ambiguous and, thus, often not covered by closed-vocabulary approaches, such as the LIWC dictionary. Thus, lower matching rates for the input motive *Search* in comparison to the equivalent app category are not a bad sign but instead speak for a cleaner body of actual search inputs. To study search inputs using the presented approach, we recommend designing a dictionary specialized on the underlying research question first, such as Remus et al. [337] created for sentiment, or Cheng et al. [79] with who collected keywords indicating depressive states.

3.1.6.2 RQ1b: Sensed Contextual Information Enhances Behavioral Research Data

Understanding contextual factors of behavior is important for most research cases. Not-understood effects in data lead to a high variability and unexplained random effects. This makes the result of statistical tests less powerful and obstructs the overall data analysis. Furr [158] explain the relevance of context, especially when regarding language: Uttered statements technically are verbal behaviors. Their relevance regarding a psychological characteristic, such as psychology, highly depends on the context in which it was uttered. Understanding context is also helpful to design assessment methodologies, for example experience sampling design where users are questioned in-situ. Physical

context features have shown to explain a part of experience sampling compliance and missing-out behavior [336]. Our study, where we took input prompt texts as an example, showed that smartphone sensing can deliver such contextual information.

3.1.6.3 The Potential of Further UI Aspects

Our work is, to the best of our knowledge, the first that investigates the usage of the input prompt text as contextual variable about the user for research. Building on this starting point, we see more potential in input prompt texts that should be explored in future work. Furthermore, our approach can be generalized to regarding the surrounding UI of user input events, beyond solely regarding text input events.

Derive Context From Input Prompt Text In our work, we used the input prompt texts to categorize text inputs by the kind of input that developers expect. However, other kinds of categorizations are possible as well. For example, the input prompt text can be used to infer the language an app is configured to be in, which likely also is the input language.

Take Surrounding UI Into Account Besides the input prompt text, surrounding UI elements could also be of interest, such as the label of a text field or the input field's size.

UI Context for User Input Events Going beyond the here studied text inputs, our approach does apply to user input events in general. For all kinds of smartphone usage behavioral datathe surrounding UI and other contextual data could be taken into account.

3.1.6.4 Privacy Issues Restrict the Availability of UI Data

Our method is based on the Android accessibility services, which provide access to the full visible UI including texts, images, and user inputs. This data is highly security and privacy sensitive, why Android's policies only allow its usage for specific purposes, especially excluding deceptive purposes, or to work around privacy controls

and notifications[1]. Publishing an app that uses this approach in the app store would require to be declared through the permission declaration[2] and would only be allowed for “helping users with disabilities interact with your app.”[3]

These restrictions make it hard for researchers to use such data and deploy their research apps in a user-friendly way. Furthermore, access to privacy-sensitive data sources might be further restricted or withdrawn in the future. The reason for all this is that nowadays privacy protection concepts cannot deal with such data. A variety of information may be contained in user interface data. Current all-or-nothing consent mechanisms do not give users control over which aspects they are fine with being used and for which purpose. Thus, the currently best option for the operating system developer is to restrict the general access to such data heavily.

3.1.6.5 Service Privacy Fit

Mobile sensing use cases that are constructed mainly to fuel research with behavioral data unfortunately pose a rather bad service-privacy fit. While the benefit mainly lies on the researcher side, users by nature do not have a direct benefit, instead high privacy implications by giving away their data. The mostly prominent offline data processing makes them give away control, even worsening the privacy implications. Researchers try to counteract this by supplementing such studies with artificial benefits, such as monetary compensation of providing data feedback. However, in practice research app developers do hard creating a sufficient service privacy fit which leads to rather low adoption rates of such apps.

[1] https://support.google.com/googleplay/android-developer/answer/9888170?hl=en&ref_topic=9877467#accessibility, last accessed 2024-12-02

[2] https://support.google.com/googleplay/android-developer/answer/9214102, last accessed 2024-12-02

[3] https://developer.android.com/guide/topics/ui/accessibility/service, last accessed 2024-12-02

3.2 Sensing Data For Adaptive Applications

System developers and researchers aim to make our mobile devices more and more intelligent. Smartphones have incorporated personal assistants such as Google Assistant[1] or Siri[2], and smartphone apps reduce user interactions where possible with automated decisions and content, to make their usage as convenient as possible. Furthermore, recommender systems and proactive interventions have proliferated into many areas of daily life, such as (physical) activity recommendations to pursue better physical or mental health (e.g., [196]), or app suggestions based on one's current situation (e.g., [211]).

With machine-learning heavy research laying the technical foundations for more advanced prediction approaches in the past years (cf. [326]), HCI research can build on that to provide users a benefit thereof. However, all these approaches require that the model or agent understands the user's situation and is aware of their context. Contextual and interaction data [326], especially GUI hierarchy and user actions (e.g., [250, 425, 454]), contain a huge application potential.

In this section, we describe the prevailing discrepancy between a high usefulness of detailed smartphone data, and the limited availability due to privacy issues. We show recent developments from related work, and present three showcases that we have been working on: Context-adaptive keyboards, intelligent interventions for mindful smartphone use, and deep activity tracking. They all rely on sensing data, and thereby can hardly be implemented due to a lack of data availability. We argue that users would benefit if their data could be leveraged for intelligent applications in a privacy-friendly way.

Based on this discrepancy we argue that better privacy-enhancing systems need to be designed. We conclude that the prevailing information and consent mechanisms in smartphones are not sufficient, and throttle the proliferation of advanced intelligent and adaptive sensing applications. This section therefore follows the research question:

RQ1c *How can mobile sensing enable novel adaptive and intelligent smartphone app use cases that are to the users' benefit?*

[1] https://assistant.google.com/platforms/phones/, last accessed 2024-12-02
[2] https://www.apple.com/de/siri/, last accessed 2024-12-02

3.2.1 Novel Adaptive Intelligent Application Use Cases

Intelligence and adaptivity can be leveraged for the users' good in multiple ways. Based on reviewed literature, we compiled recent advancements in the field into the following categories.

Improved Background Behavior & Context Adaptivity OS developers improve their devices' self-management by adapting to the user's situation. Background processes, such as battery management, backup tasks, and other schedule-able background processes are usually not of immediate importance for the user and are thus flexible in their time of execution. The smartphones operating system tries to understand its user's usage patterns, in order to adapt performance and charging behavior for improved battery lifetime[1]. Similarly, gathered understanding of the user's context and behavior is used to schedule notifications, aiming for less interruption and distraction [239].

Personal Assistants & Novel Interaction Approaches With the recent rise of Large Language Models (LLMs), research on task automation concepts has increased. Here the user enters a prompt via voice or text input, and an agent interacts with the smartphone and its apps to complete the prompted task [411, 425]. Acting as a personal servant, such agents reduce the interaction with apps or even may replace it completely[2]. Personal assistants have proliferated built-in into the operating system, such as Siri suggestions. Chihani et al. [81] demonstrate how the composition of user-defined rules and context-awareness can facilitate application compositions, i.e. tasks where one smartphone app is not used individually, but a sequence of app usage is necessary.

Interaction Automation Research on model-based intelligent user interface adaptation follows multiple objectives, such as increasing efficiency (e.g., reducing task completion time or error rate), or improving subjective user satisfaction and providing a hedonic value [3]. Do and Gatica-Perez [117] develop an approach to predict the next used app, which can be used in concepts such as proposed by Moschelli [287] to

[1] https://9to5google.com/2022/09/15/android-adaptive-battery-explainer/, last accessed 2024-12-02

[2] https://www.telekom.com/en/media/media-information/archive/deutsche-telekom-frees-smartphones-from-apps-1060272, last accessed 2024-12-02

improve transitions between apps and simplify multitasking. In the concepts of Zhou and Li [454] and Lee et al. [250] the next click that the user might perform is predicted. They propose an overlay that auto-selects the target input element and thereby takes over this decision from the user and fastens the interaction.

3.2.2 Three Case Studies

In the following we present three case studies that we have proposed during my work. They heavily rely on mobile sensing data and yield a benefit for their users.

3.2.2.1 Adaptive Keyboard

Smartphone keyboards show a bar word suggestions, where the user are offered words that they might likely want to type next. While prediction algorithms that that predict which words will likely follow the typed text input advance constantly, little has changed regarding the interaction modalities. Keyboards present a set of words in the top row of the keyboard, which the user can choose from.

We propose two concepts how keyboard word suggestions could leverage contextual data to become more intelligent: First we propose to make the amount of displayed words dynamic, depending on the user's current activity. While being on the go, more word suggestions might be appreciated than in a steady situation. Second, we propose to let the user choose how a content is written instead of the exact words. For example, an LLM-based quick dial that offers different tones of a text could allow the user to adapt their text rapidly.

Context-aware number of word suggestions We propose to make the trade-off between occlusion by word suggestions and space for other screen contents dynamically. We developed a prototype based on the Android AOSP keyboard that can show between 0 and 3 lines of word suggestions at the top of the keyboard. While a button interface allows the user to increase or decrease the number of word suggestion lines manually, it also takes user's situation and context into account to learn the user's preferred choice.

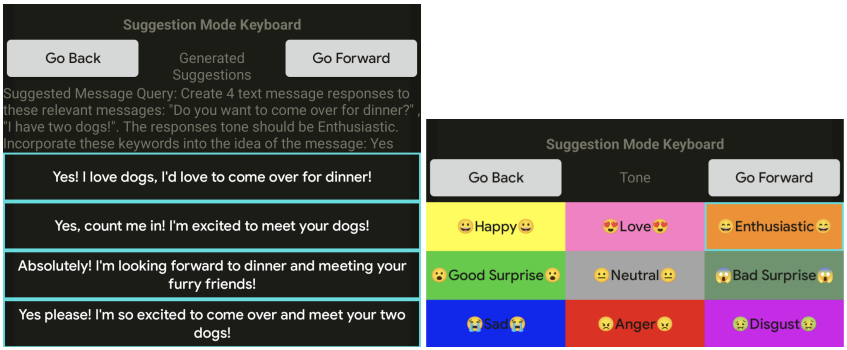


Figure 3.4 : Screenshots of two prototypes of an LLM-supported smartphone keyboard. *Left:* Response Generation with Prompt Adaptation, *Right:* Tone Adaptation

We motivate future work to deploy our prototype in the wild, to study two main questions: (1) How do users perceive the adaptable number of word suggestions and how does it influence their typing performance? (2) In which contexts is which number of word suggestions beneficial?

LLM-based content suggestions Large language models open a space for novel text generation approaches on smartphones. Instead of the user writing a text with AI supporting them by word suggestions, users can instead prompt the LLM to craft a text. Besides crafting a text from scratch, models can also be used to reformulate existing content in a user-defined style, for example to make a text more polite. We conducted two focus groups to come up with two concepts for how to leverage LLMs to improve smartphone text input interfaces:

Tone Adaptation This concept allows the user to reformulate an existing text in a different tone. They can choose from a set of message tones (e.g., happy, anger, bad surprise).

Response Generation with Prompt Adaptation When using the message history of an ongoing conversation, an LLM could be used to generate an appropriate response. Our focus group showed that users thereby prefer to get multiple response suggestions to choose from. Furthermore they indicated the desire to edit the prompt, for example to change some aspects of the generated responses and customize them.

3.2.2.2 Rabbit Hole Interventions

This section is based on the following publication:

Nada Terzimehic, Florian Bemmann, Miriam Halsner, and Sven Mayer. “A Mixed-Method Exploration into the Mobile Phone Rabbit Hole.” In: *Proc. ACM Hum.-Comput. Interact.* 7.MHCI (2023). DOI: 10.1145/3604241

The mobile phone rabbit hole (MPRH) describes rather over-the-top and, at times, prolonged digital content consumption compared to the user's initial intention [84, 263]. Given that the smartphone has become an ever-present companion, it is now feasible to, accidentally or not, drop into a digital rabbit hole at any given time and place. Research (e.g., [263, 264]) and society (e.g., [93, 389, 419]) have debated the negative rabbit hole-like effects, in particular in the digital well-being research area (e.g., [340, 404]).

Based on mobile sensing data from a two-week field study, we have shown that it is possible to predict users falling into a mobile phone rabbit hole. We reflect on the definition of the MPRH and discuss UI implications on how to communicate MPRHs, both in an intervening and preventive way. This helps users being more aware of their smartphone usage, and, if desired, break out of unwanted or excessive smartphone usage behaviors.

Predicting the Mobile Phone Rabbit Hole The prediction model relies on four types of mobile sensing data: 1) *smartphone sensor* data, which are internal smartphone sensors such as the accelerometer or proximity sensor; 2) *usage events*, such as the accessibility service events or Android app usage events; 3) *smartphone state information*, which refers to, e.g., the current ringer mode, the screen state or the internet connection availability and source; and 4) *smartphone events*, which include phone calls' or SMS' received.

We predict single rabbit hole sessions, i.e., develop an algorithm that can detect an ongoing rabbit hole right at a time or afterward. Therefore, we have split our dataset into a train, validation, and test dataset by participants, i.e., we assigned 15 users to the train set, 3 to the test dataset, and 3 to validation. Due to the huge class imbalance, we applied SMOTE oversampling [77]. After the initial model investigation, we selected a random forest as the model best-performing model. Then, we optimized the hyperparameters using with a grid search, tuning the parameters listed in Table 3.5.

Model Parameter	Optimization Range
n_estimators	5, 10, 100, 200, 500, 700
max_features	sqrt, log2, None
max_depth	4, 5, 6, 7, 8, 10, 12, 14, None
min_samples_leaf	1, 2, 4
min_samples_split	2, 5, 10
criterion	gini, entropy, log_loss

Table 3.5 : The model parameters and their optimization that was tried by a grid search. The values for session prediction are underlined identifying the best value of each parameter.

The model training with the identified best parameter configuration takes approximately 4 seconds on a commodity notebook. We implemented the prediction model using Python's sklearn library[1].

We investigated to predict whether a smartphone usage session is a rabbit hole or not, i.e., treating each session as an observation and the users' label on whether they did more than intended as the target variable. The chosen model optimization parameters, which we identified through a grid search, are underlined in Table 3.5. On the training dataset, with 15 participants, the model reached an accuracy score of 87.97%, and on the test dataset, with 3 participants, 64.97%. On the validation dataset, where we tested the model with 3 more yet unseen participants, the model's performance is 72.41%, which we consider as the model's actual performance. The precision for rabbit hole sessions thereby was higher (77%) than that of usual sessions (69%).

Analyzing our model's feature importance, we find that the features contributing most to the prediction of a rabbit hole session are related to app usage, precisely usage of apps of the categories *social media*, *system*, and *communication* (descending order by the impact on the model output magnitude). The number of clicks and click frequency show high importance, and also the number of sessions that happened beforehand in the previous 3 hours. Device settings rank rather low, only the WiFi status *connected* ranks high. Other settings, such as the ringer mode, show even less importance to the prediction. Calculating a conversion to time-relative features is beneficial for some features, esp. the number of usages of *system* apps and time spent in *social media* apps. To understand how these features contribute to the models' prediction result,

[1] https://scikit-learn.org/stable/index.html, last accessed 2024-12-03

ID	Expert's Research Topic	Gender	Age
E1	Health & Wellbeing at Work	M	31
E2	Mental Health & Wellbeing	F	27
E3	Mental Health & Wellbeing	F	32
E4	Technology-Mediated Communication	F	28
E5	Usable Privacy & Security	F	26
E6	Human-Robot-Interaction	M	28

Table 3.6 : Demographics overview of our focus group participants.

we created SHAP values and plots [265]. The SHAP values allow for a more in-depth analysis of how feature values influence the result. In Figure 3.5, we visualize the top 20 features for the prediction model as a SHAP plot. The distribution of the colored dots conveys how each feature value is distributed and which values contribute to which direction (i.e., push the prediction decision towards the result *rabbit hole* or *non-rabbit hole*. Spending time in *social media* apps is a strong predictor for rabbit hole sessions, same for *system* apps (red dots are shown on the right side only). A high number of scrolls in *system* apps and clicks in general also argue towards a rabbit hole. However, a generally high scroll frequency argues rather against a rabbit hole.

Design Suggestions to Communicate and intervene against the MPRH We conducted an expert focus group ($N = 6$), to collect ideas on how to effectively communicate the detection of an ongoing MPRH to the user, as well as brainstorm potential concepts for prevention and intervention. We are interested in ways of communicating that falling into an MPRH is likely to happen to affected smartphone users. We recruited six HCI experts (four female, two male, mean age = 28.6 years), i.e., people who do teaching and research in HCI. Table 3.6 provides an overview of the experts. All experts use a smartphone daily. We asked the experts to brainstorm ideas on how to communicate the likeliness of the user falling down an unwanted MPRH. We differentiated between the detection and prediction scenario, as well as the wrongful detection scenario. The goal of informing the user should be primarily to raise awareness that they are about to fall into a rabbit hole; E3 described it as *"a trigger for self-reflection."* The experts' group agreed that sudden notifications and pop-ups that consume the whole screen, such as iOS' screen-time notifications[1], should *not* be the way to go

[1] https://support.apple.com/en-gb/guide/iphone/iphbfa595995/ios, last accessed 2024-12-12

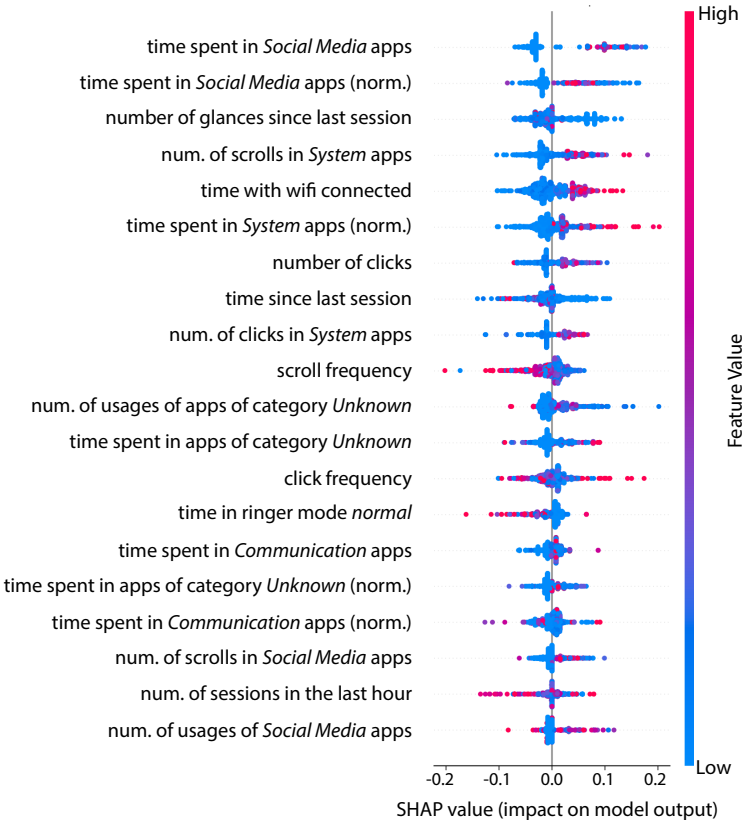


Figure 3.5 : Beeswarm SHAP plot, visualizing how the top 20 features contribute to our session prediction model. Each point indicates how an observation contributes to the model's output. A positive impact value pushes the prediction result towards deciding on a rabbit hole and a negative one against it. Features with the suffix *(norm.)* are normalized by the session length.

– these tend to become swiftly annoying and ignored by the user. Thus, four experts (E3-E6) proposed a reminder banner that blends in the user's current content context. For example, E3 proposed *"[a] reminder there so that you cannot miss it, but you can also easily just ignore it."* Blending in was an important discussion point – the rabbit hole is a continuum, and as such, the user should not be suddenly informed of the rabbit hole, but rather gradually, as the rabbit hole develops. This could give the user time to

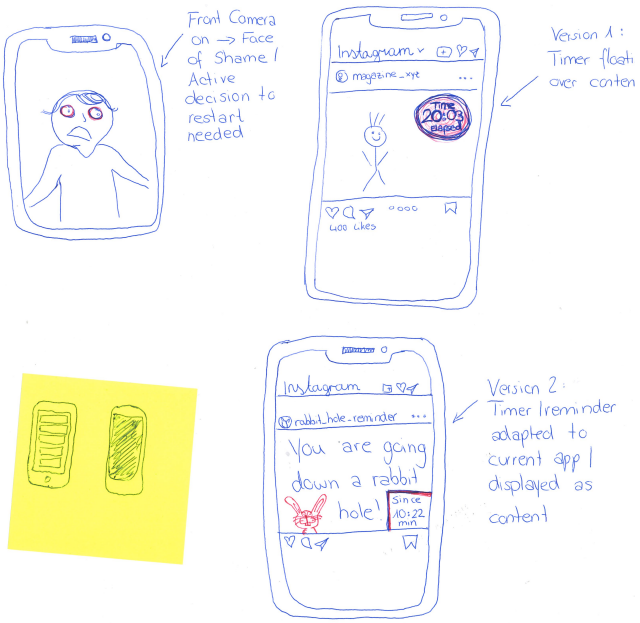


Figure 3.6 : Four resulting sketches from our experts' focus group on the question of communicating MPRH to the user. *From left to right: (1)* A user terrified of their image in the black mirror, after being in a MPRH. *(2)* Floating timer. *(3)* Quick shutting down of the screen once MPRH was detected. *(4)* Timer blended in content, as Instagram post.

prepare ending the interaction with their smartphone. Concrete examples suggested were presenting a timer (see Figure 3.6, left) or turning the screen gradually off (see Figure 3.6, right), similar to rendering a *tunnel* on the smartphone's screen. In the case of a textual prompt, the prompt could ask the user questions beyond their smartphone behavior. E3 and E4 proposed asking about user's emotional state, how they feel about this use session; reminding them of their initial intention and whether they have fulfilled it – if not, the smartphone could *"jump"* to the app of the initial intention; or proposing a contextually appropriate alternative activity instead, to exit the digital tunnel. E1 reflected at this point: *"How do we actually get out of rabbit holes now? [...] It's usually like, some time critical thing comes up that you have to do or you, like, finally convince yourself that, like, I could be doing something better every time."*

The smartphone could proactively and beforehand suggest an alternative activity either from user's daily ToDo List (E2) or randomly (E6) before the MPRH. Further discussion revolved around balancing effectiveness and annoyance – the more drastic ways are more effective (e.g., turning off the screen), but also more frustrating when they're wrong (e.g., the user might think the smartphone is broken). If an incorrect detection occurs or if the user wishes to explore further, certain participants suggested that the user should be given the option to continue their exploration. However, this would require the user to actively respond: *"[If] I wanted to be in the rabbit hole, I will just turn on my smartphone, which is okay, but I do have to do something active [as response]"* (E6). Another proposal involves a *"cute"* visualization of a rabbit – in order to reduce the annoyance factor – that needs to be taken out of a hole. E2 suggested the option for the user to indicate intentional rabbit hole desire beforehand, in order to prevent wrongful detection. In any case, experts enhanced the importance of the system learning individual patterns of use and adapting to those, as *"clearly, we all have, like, different [patterns of use]"* (E1).

The Importance of Contextual and Situational Data Our RHT app focused primarily on smartphone usage behavior, i.e., what users do on their phones. The focus group informed us that it takes an initial situation of the user (e.g., desire to kill time, intent to look up contextual information briefly) combined with a trigger (e.g., distracting notifications, recommendation algorithms) to lead to a rabbit hole session. To encompass these findings, future work could investigate sensing data with the aim of extracting rabbit hole-prone user contexts and situations. By combining data on location and time, the system could detect scenarios raised in the focus group, for example, killing time instead of going to sleep. We see potential in the prediction of rabbit hole-prone *situations*, rather than individual *sessions*. With our definition, we argue that the observed smartphone usage behavior is a property of the MPRH. In a rabbit hole-prone state (e.g., desire to kill time), users are prone to be caught by recommended content, which we comprehend as the high usage of entertainment and gaming apps in our data. The high occurrence of *system* apps in a MPRH, particularly launcher apps, expresses users often visiting the home screen. This might be due to the state of boredom and lack of inspiration on what to do next – in pursuit for the next dopamine shot.

With a greater dataset, our predictor model could enable more fine-grained situated monitoring of smartphone use, i.e., how smartphone use is situated in different everyday life contexts. These might build on top of existing self-monitoring visualization that, again, solely display smartphone use indicators ignoring the interplay of smartphones and users' everyday activities outside of it. Thus, future research could employ and compare the suggested approaches in its efficacy.

3.2.2.3 Intent Prediction

Albeit technically not being a sensor, one of the most used mobile sensing data types is app usage. This in most cases encompasses aggregated statistics about for how long and at which times an app has been used. More detailed information what the user did inside an app is used rarely for mobile sensing applications nowadays. While aggregated statistics are rather easy to evaluate and extract knowledge of, in-app behavioral data requires more sophisticated preprocessing. Current work that deals with in-app behavioral data either drops a lot of information during their preprocessing (e.g., [251], who only keep layout box positions and texts), or relies on tedious, manual labeling (e.g., [332]). Besides this practical issue, privacy issues makes such data hardly usable with current procedures.

Deep Activity Vectorization In this section, we propose an approach to vectorize deep activity data, i.e. sensing data about UI hierarchies and user input events. We developed this as a first step towards the realization of intent prediction, which enables app use cases such as quick dial interfaces or in-situ prefetching of results. A visual overview is given in Figure 3.7. The concepts overall idea thereby is to regard each *interface state* as one frame for a sequence prediction machine model. Interface states are thereby separated by user actions, i.e. decisions of the user on how to proceed, which are given to the system in the form of UI inputs (e.g., button clicks). The interface contents thereby are distinguished into two types:

1. **Informational Content** that contains information about the current state of the process that the user is pursuing, desired information as process output, and information illustrating the next options. These contents can consist of textual contents, graphics such as images, visualizations, and layout structure can be

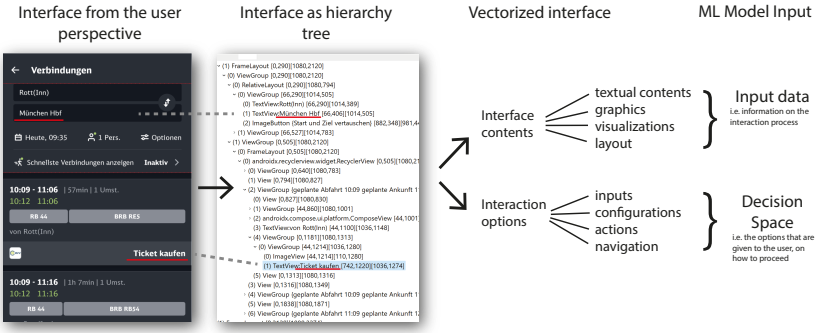


Figure 3.7 : Vectorization procedure for smartphone app interfaces. On the example of Android, an interface can be accessed as its UI tree hierarchy. We propose to extract relevant information thereof, distinguishing it into interface contents and interface options. Through this vectorization process, each interface state can be vectorized into one frame for sequence prediction models such as RNNs.

used to convey information. To vectorize these elements, past research has proposed approaches for the individual data datatypes. Li et al. [251] propose a layout vectorization approach. To include texts they use a BERT model [333], however for text vectorization a plethora of approaches have been yet developed, most interestingly for this context Word2Vec by Church [90].

2. **Interface Options** of which the user should choose in order to continue in the interaction process. These options can be single button clicks, or consist of multiple individual actions. For example, a user may need to fill out an input form such as visualized in the travel app example in Figure 3.7 where origin and destination need to be specified. Besides being also part of informational content, the offered options constitute the prediction target of an intent prediction model.

The Privacy Challenges of Deep Activity Data Albeit the proposed concept envisions novel, innovative application concepts, some work needs to been before this can come into practice. Besides the technical challenges outlined above, privacy issues need to be solved. Interface hierarchy data can contain any information the user may

enter into their phone and an app displays, including passwords, private messages and pictures. To make use of such data in practice, privacy-enhancing technologies need to be developed that are capable of working with data that:

- may contain a variety of content types, including but not being limited to text contents and images
- is able to deal with unforeseeable content structures
- manages to inform and provide control over data flows with a high frequency and resolution.

In contrast to other commonly used datatypes, it is unforeseeable which contents may be contained in UI tree data. Thus we need a solution that goes beyond listing users information to which they may grant or deny access.

3.2.3 Discussion

In the following, we discuss the benefits, challenges, and opportunities of sensing data.

3.2.3.1 User Benefits of Sensing Data

With the presented application use cases we have shown that users can benefit from the usage of mobile sensing data. All these approaches require that the model or agent understands the user's situation and is aware of their context.

Context-awareness is important for predictive use cases and adaptive interfaces, and access to detailed smartphone usage behavior enables proactive smartphone features that offload the user. Current implementations rely on data sources that are both originating from the end-user and other external sources [3]. In order to actually reduce the interaction of the user, mostly user-external sources have to be used: According to the review of Rachad and Idri [326], contextual- and interaction data are used by 3/4 of the studied projects, while user-provided feedback is collected only in 4% of the reviewed papers. Also for the adaptive keyboard use cases, a broad contextual understanding of the user's situation is essential. We hypothesize the main differences to originate from the user's physical activity (i.e. typing on the go vs. typing in a steady situation), and social contexts (i.e. whether the tying process interrupts a current social situation, or the user is alone and typing is their sole task) While physical

activity data yet is available to application developers (see e.g., the Google Awareness API on Android[1]), understanding social contexts is a yet difficult topic. Current research in social sensing leverages a variety of datatypes, including privacy-invasive datatypes such as microphone and call logs (cf. review of Zhang et al. [451]). Furthermore, access to the typed content is beneficial for adaptive keyboard cases. Especially when it comes to response generation, it is inevitable to process at least the current text input, even better would be to access also the past messages.

Among interaction data, user interface states and actions are important for mostly all projects (e.g., [250, 411, 425, 454]), especially access to apps' GUI hierarchy and user actions. Projects relying on classical contextual information sensed by the smartphone (e.g., [81, 117, 287]) are not capable of fulfilling specific tasks and going beyond app- or activity level predictions.

3.2.3.2 User-Oriented Use Cases Show a Good Service Privacy Fit

From the service-privacy fit perspective, mobile sensing use cases that produce a direct benefit for the data-originating user can be expected to have a rather good service privacy fit. Users immediately perceive the benefit of the data usage, and notice the outcome of disabling data access. This applies especially to proactive and predictive use cases such as the presented intervention and quick dial scenarios.

Nevertheless challenging may the reach of a good service privacy fit be with deep activity data. Deep activity tracking requires access to very detailed and in-depth data on what users do on their smartphones. It is hardly foreseeable for users what information such data consists of, as view hierarchy data can contain anything from text contents, over pictures, passwords, and data of third parties. Also a variety of aspects can be inferred, which is hardly estimate-able by users. This strong privacy impact opposes a likely rather strong service value, as users could directly see what benefit they can gain if that data is evaluated. However, as this data is not commonly used yet, we do not know much about users perceived service privacy fit with deep activity data.

[1] https://developers.google.com/awareness/, last accessed 2024-12-03

3.2.3.3 Challenges

As the outlined use cases have mostly not proliferated yet, uncertainties and challenges remain. Especially for complex data types, such as view hierarchies, sequences, and abstract contextual data no appropriate privacy enhancing technologies exist. Users is hardly enabled transparency and control of such data. Preprocessing of data on-device, which can drastically reduce the information gain to the necessary, becomes increasingly difficult with an increase in data complexity. While preprocessing of smartphone usage statistics are popular and even offered by the Android operating system, no widely applicable methods exist yet e.g., for view hierarchy data. A further challenge is introduced with the proliferation of LLMs. Nowadays most models run online, are a black-box, and it is uncertain whether and how data that is put in is incorporated into the model.

3.2.3.4 Opportunity: Interactive Explainability

In line with the high value on the service side in the service privacy fit, applications that yield a direct benefit for the user pose some specific opportunities. The directly visible benefit illustrates its user for what their data is used. Depending on the specific application it is possible to immediately convey which data enables which functionality, respectively allows to demonstrate a decrease in features with decreasing data access. This dependency could be used to enable transparency through interactive explainability, i.e. allowing the user to interactively experiences how an application output behaves with enabling or disabling specific data accesses. Furthermore, on-device (pre)processing of data often is easier with applications that create a benefit on the user side. While the nature of data collection for research usually is to build a database offline on a machine controlled by the researchers, user-oriented use cases may have their logic implemented in client-side code.

3.3 Sensing Against Societal Challenges

This section is based on the following publications:

Florian Bemmann, Carmen Mayer, and Sven Mayer. "Leveraging mobile sensing technology for societal change towards more sustainable behavior." In: *arXiv preprint arXiv: 2303.12426* (2023)
Florian Bemmann and Sven Mayer. "User-Centered Sustainable Technology Design: A Reflection on Human-Computer Interaction Research for a Sustainable Society." In: *International Conference on ICT for Sustainability* (2023)

Mobile sensing systems do not only have an effect on their direct users and app publisher. Being deployed at scale, the sum of individual effects has an influence on our society as a whole. When designing technology, it is thus important to regard also effects on the big picture, and to reflect on how technology shapes our environment. Ubiquitous technologies also accompany novel challenges to our societies. Scalable digital communication media such as social platforms proliferate rapidly, and change the way we consume information and communicate on a global scale.

Social media platforms give users a distorted impression on the public opinion and sentiment. Filter bubbles rather reinforce our current opinion, instead of contrasting it to other views and making us regard and consider different perspectives on a matter [120]. Digital footprint data thereby plays an essential role in regard of nowadays societal challenges: It contains a huge amount of (concealed) information about their users, what users are mostly not aware of. Using methods of psychometrics and psychological targeting [275], the data can be exploited for unethical purposes. Targeted advertisements that speak towards the user's fears, and subconsciously form their norms and values. Targeted advertisements based on digital footprint data can influence societies and poses huge challenges to our democracies [122]. For the challenge of climate change, ICT's scalability and fast proliferation poses novel opportunities. Technology makes our daily activities more efficient, and can support us in reaching our aims. However, especially information technology is also a large energy consumer, contributing substantially to green house gas emissions [1]. Furthermore, efficiency-increases are often eaten-up by rebound effects [175].

Besides these challenges, sensing data can also be used to tackle them. Regarding the challenge of climate change, a pro-environmental attitude in the general population

is essential to combat it. Society as a whole has the power to change economic processes through market demands and to exert pressure on policymakers - both are key social factors that currently undermine the goals of decarbonization [137]. Creating long-lasting, sustainable attitudes is challenging and behavior change technologies do hard to overcome their limitations. Environmental psychology proposes social factors to be relevant, a.o. creating a global identity feeling [330] and widening one's view beyond the own bubble. From our experience in the field of mobile sensing and psychometric data inferences, we see strong potential in mobile sensing technologies to implement the aforementioned goals. To shed light on the possibilities and risks that mobile sensing brings for our society, this section follows the research question:

RQ1c *How does mobile sensing at scale interplay with our society and its challenges?*

In the following, we outline how mobile sensing based technology can help to tackle nowadays societal challenges. We therefore first point out the current limitations that HCI is facing from literature. Then we join research from environmental psychology with the opportunities of HCI, so show mobile sensing based application cases that could support the fight against the challenges.

3.3.1 What SHCI Recently Did: Limitations of Behavior Change Technology

Having arised 15 years ago, sustainable human-computer interaction research (SHCI) has become a relevant field. It studies how computing technologies can be applied to limit environmental consequences and proliferate pro-environmental behavior [66]. Building on concepts of habit forming, self-optimization, and behavior change applications, HCI also investigated using such concepts to foster sustainable behavior. For example, in the domain of shopping and food consumption, there exist a variety of concepts, mostly based on recommender systems aiming to overcome the attitude-behavior gap. In-shop decision support systems provide information on specific products, e.g., food miles [210] or consumer-generated environmental impact information [286, 396]. Recommender-systems were developed to recommend sustainable products [395]. However, shopping decisions are subject to complex constraints like family dynamics and daily routines [91, 92] (e.g., the distance to the shop, preferences of all family members and required cooking effort.) Related systems have to fit into the broader

context of life in order to be used. Thus, recommender-systems that decide based on a limited set of criteria are no ideal solution [210]. Also mobility is a well-studied domain (e.g., [153, 424]), as mobility behavior is rather easy to track with ubiquitous devices and transportation is publicly known as major CO2 source.

Due to the limitations that recommender systems face, recent approaches try to push people more towards sustainable mobility [153] or foster sustainable consumption through self-reflection [41].

However, persuasive sustainable interventions have limited real-world impact because the main objectives against acting sustainably are external circumstances that cannot be overcome by persuasive technology [69]. Furthermore, achieved behavior changes of studied projects are often not long-lasting in the wild [185]. In their recent review, Bremer et al. [66] summarize the efforts and limitations of past SHCI research, and call for going beyond individual behavior change and rather aim for societal change.

While persuasive technologies in other domains, usually designed to directly improve an aspect of oneself (e.g., physical fitness, mental health), can directly track progress and report improvements, that is more difficult with SHCI related behaviors. One's impact on climate change cannot quantified as easily as improvements in one's fitness level. Regarding sustainable behavior, classical behavior change-supporting technologies face limitations in real-world applicability, above all a lack of "good reason to use" e.g., extrinsic motivation, (see the Technology Integration Model of Shaw et al. [366] for factors influencing continued use).

3.3.2 Environmental Psychology and the Power of Societal Change

Actual technology alone is not sufficient to combat climate change, societal change (that can be supported by technology) is at least as important [137]. They report consumption patterns and corporate responses to be the two social factors that still undermine the goals of decarbonization. Hereby the latter is indirectly controlled by the first (i.e. companies adapt to market demands). Behavioral- and environmental psychology try to explain why people do not behave sustainably even though they have an attitude towards it (i.e. attitude-behavior gap), or what counteracts people developing an environmentally friendly attitude.

Attitude-Behavior Gap Regarding consumer behavior, the main barrier towards actual sustainable behavior are hard circumstances like price, perceived availability, and convenience [17]. A lack of such extrinsic motivational factors come together with rather weak intrinsic motivations: Moral short-sightedness [18] and doubts whether one can make a difference as individuals throttle the intrinsic motivation of many people. Among the five obstacles towards far-sighted actions that Ascher [18] point out, especially selfishness and uncertainty play a role in our context. The effects of one's climate-negative actions are for western societies geographically far away (i.e. out of one's extended circle of selfishness) and the relationship is indirect, i.e. a concrete behavior does not directly lead to a concrete consequence.

Classical behavior change technologies (e.g., [149]) are thereby doing hard in making an actual change towards climate-friendly behaviors.

Pro-Environmental Attitudes In behavioral models, an attitude is a basis for behavior. Thus besides aiming for behavior change, the formation of a pro-environmental attitude among the population also is an important building block. Reese [330] argues that a common human identity, i.e. people regarding themselves as global citizens instead of part of some local group, could inform beliefs about environmental justice. Huber and Hilty [195] propose instead to leverage the *behavior-to-attitude link*. It is reported to be stronger than the vice versa link between attitude and behavior, although less studied yet. The behavior to attitude link can for example be observed when people are forced to life changes, e.g., when moving the location of home or workplace, in which associated higher flexibility towards pro-environmental change was observed [423].

Global Identity Perception People regarding themselves as global citizens instead of part of some local group, inform beliefs about environmental justice and lead to more sustainable norms and motives [331]. Although decision models say people rate their rational decisions from a moral perspective, immoral behavior often happens subconsciously. A social, physical, and timely distance from the effects of climate change lead to moral disengagement. Behavioral psychology explains its factors and proposes ways to overcome moral shortsightedness that could be applied in HCI and lead to more sustainability-oriented overthinking of rational decisions [34].

3.3.3 Application Concepts

Nowadays ubiquitous devices such as smartphones and -watches accompany their users throughout the whole day. We envision the following technologies as means to support societal change and implement approaches pointed out by environmental-psychology research in the previous section. Finally, we present application cases that we envision to help climate and society.

Mobile sensing technology supports these application concepts by providing **ubiquitous behavioral data**, i.e. the ability to access data on the user's behavior, context, and situation unobtrusively in the background [181, 182]. Common behavioral data encompasses but is not limited to device usage, and mobility behavior including the choice of means of transport (e.g., via Google's Awareness API), and mobile language use. Information on behaviors that cannot be directly sensed by the smartphone, such as consumption and nutrition behaviors, can either be gathered with journaling methods [401] (e.g., asking the user daily for their consumed amount of meat), via third-party devices or services (e.g., financial APIs that have access to purchases), or a semi-automatic approach combining both (e.g., taking a photo of each meal that is processed by image recognition) [41]. Most data is available immediately in the situation (in situ), allowing the user to follow their progress live.

Machine Learning based Inferences allow assessing non-directly measurable behaviors and attitudes, such as personality traits [384] and political orientation [224]. Explained decisions of models support users in reflecting on their data and identifying connections between and reasons for behaviors [39].

Data becomes especially powerful when it is collected at scale, put into context, i.e. comparing it with one's own historical data or with the data of others. Data of other groups of people can be collected either via **mobile crowd sensing** systems [161], derived from existing sensing datasets of past studies such as conducted by Schoedel and Oldemeier [356], or accessed via APIs. Such comparisons can help people to classify their behavior with the local/national/global average. People can thereby also be pulled out of their bubble, which is a strong measure towards a sustainable attitude as depicted hands-on in Section 3.3.3.1.

In the following, we interconnect the presented insights from environmental psychology with the specific capabilities of mobile sensing technology, to propose application concepts supporting societal change.

3.3.3.1 Extrapolation of Sensed Behavior: Becoming Aware of Own Behavior

Many behaviors that have an ecological positive or negative impact can be captured with smartphone sensing in situ, i.e. at the moment when it happens. Data on environmentally-relevant behaviors, such as mobility or consumption, can be used by applications to track their progress over time, or support behavior change [386]. A major factor limiting the proliferation of HCI towards sustainable behavior is the individual feeling of not having a higher-level impact. This limits intrinsic motivation and post-use evaluations, leading to non-adoption of technology. To overcome this issue, we envision an application that makes users conscious of their behavior in relation to others.

Show Environmental Impact if Everybody in Your *Country* Behaves as You at the Moment By taking the difference of the user's behavior to national average values, users could be made aware of which impact one has as part of a larger group. By distinguishing between people that (a) already take efforts to live environmentally friendly and (b) those who do not, it could be further pointed out which impact it would have if (a) engaged individuals would stop their engagement (corresponding to lacking motivation) and (b) further people could be convinced. This might foster a global identity feeling, which is a key factor to environmentally sustainable behavior [330].

Show Environmental Impact if Everybody in the *World* Behaves as You Do at the Moment A different effect might be achieved when comparing with global averages . From the viewpoint of members of western societies, even the behavior of environmentally engaged people is carbon intensive when compared with the global average. The awareness of this should hint people to that (a) further engagement is still necessary, and (b) helps perceived losses of quality of life (e.g., renunciation of air travel) from outside their own bubble. While in one's (social media) bubble it seems usual to fly several times per year, this is not the case when compared with the global standard. This view should help users regard themselves as global citizens and to judge their behavior regarding global standards.

General Design Considerations In general, such an application should be designed for passive use, i.e. the app giving the user information and food for thought occasionally when appropriate. Ambient narrative interfaces, such as visualization on the lock- and

home screen as proposed by Murnane et al. [288], are promising because users do not have to actively use them and research has shown that ambient information is easier to process [177]. Also augmenting the real world, for example with public displays [273] or AR augmentations should be considered.

3.3.3.2 Believable Agents

Conversational interfaces that express emotion and personality [319], could incorporate sustainability-oriented norms in their language. We are influenced by our impression of how our peers' and the society's opinion is. Especially, the perceived expectation of fellow people towards ourselves is known to have a strong influence. With technology, such as conversational agents and robots becoming more and more human-like, they might also gain the ability to impact ourselves through a kind of peer pressure.

3.3.3.3 Games

Video games cannot only fulfill entertainment purposes. They constitute an area of social life, can be used for learning and collecting experiences. Video games, in contrast to the real world, enable people to experience the effects of their actions in a try-and-error like manner, without facing the potentially serious consequences that the real world brings. In the context of SHCI games thereby are a valuable tool (cf. [226]). Users can learn through simulations (e.g., [318]) and experience how their acting influences the climate (e.g., [218]) and the society (e.g., [290]). Furthermore, video games can support triggering pro-social behavior, such as empathy Wulansari et al. [439]. Specific game designs or design aspects may also be evaluated to increase sustainable norms and values, for example by including green nudges [307]. Klammer [226] for example distinguish between effects on a *rational* level, i.e. making people understand relationships, an *emotional* level, i.e. making people care about it, and lastly the *action* level, where people actually take action against it. Games are a promising environment through their massive adoption and usage in society. They manage to immerse and engage their users more than many other media, and induce fun and enjoyment. People thereby are often intrinsically motivated to play a game, which solves the major issue of a lack of motivation that many other SHCI applications face. Besides classical desktop games that are played in a steady situation at home, mobile

ubiquitous games have emerged in the past years and shown to be successful. Mobile augmented reality games that take place in the real world, the most prominent example is Pokemon Go, integrate themselves into daily life and context. We expect especially this integration of game elements into the real world to be promising. Together with mobile sensing based context-awareness, all three levels of effects (rational learning, emotional caring, and acting) could be targeted in a way that integrates well and with low user efforts into daily life. We see strong potential in video games for the field of SHCI. Further research should investigate how the specific characteristics of games can be used in regard of nowadays societal challenges.

3.3.3.4 Context-Aware Decision Support

Although decision support systems have faced challenges in their real-world applicability recently, we would like to motivate further investigations in that direction. With the rapid progress in intelligent systems, especially the rise of large language models, decision support systems might become able to mitigate their current limitations.

3.3.3.5 Content Recommendation

Content recommendation systems, such as social media, news feeds, and search engines, contribute to the perception of our social bubble and norms. Today, they support current beliefs and perceptions instead of widening perspectives. This has mainly economic reasons. It is in the platform interest to rather show people content that they like i.e. sticks to their opinion, and controversial, rather extreme content by nature engages people which again makes users use the platform. SHCI needs to find solutions on how the negative effects of content recommendation systems with the thereby created filter bubble can be mitigated, in a way that has the chance to be adopted by platform operators. The workings and effects of social polarization are complex [143], and the design of such technologies thereby needs careful consideration of psychological and sociological effects. For example simply exposing people to other perspectives has rather shown negative effects: Exposing people to beliefs of opposing groups does not reduce, but even increase polarization [24].

3.3.4 Discussion

In this section, we discuss the promising future approaches and ethical aspects.

3.3.4.1 Real-World Interventions and Suggestions Need Context Awareness

Context-awareness is key to actually have an effect in SHCI. For games, recommender systems, and LLM based agents, awareness of the user's situation is essential to act appropriately. Without data on the user's situation, also novel approaches like LLMs cannot create an output that fits to the user's current situation and thus as real world applicability.

3.3.4.2 Ambient Information

While interactive systems were initially designed to have users explicitly interacting with technology, *ambient intelligence* or *implicit human computer interaction* integrates technology and interactions with it into our environment. Interactions happen in situ, more "on-the-fly," and thereby more natural and less interrupting [350]. Ambient intelligence can be found in various artifacts, such as conversational agents that are in our phones or are ready for use in smart home devices and on websites, robots, and also interfaces such as smartphone lockscreens or smart devices may contain ambient intelligence elements. Concepts such as the one Murnane et al. [288] show that information presented ambiently on the smartphone lockscreen has persuasive potential. Ambient information presentation especially has the advantage of bypassing the issue of motivational lack, as users automatically get in contact with it.

Context-Aware, Ambient Information could help users classify and compare their behavior with others and different social and global groups [43] by including them on smartphones' lock screens (cf. [288]) or smartwatches. Concepts that leverage ambient interfaces especially pose the advantage that they do not require active usage. Users stumble upon them passively, while following other usage intentions. Thereby such interface concepts can reach people who are not actively looking for or aiming at dealing with topics such as pro-environmental behavior.

3.3.4.3 Personality-Based Targeting: Unconscious Attitude Formation

Targeting content towards specific user groups has long existed especially in the context of the advertisement or election promotion campaigns, for example adapting ads by location, nowadays known as *macro targeting*. With the rising availability of more detailed user data, targeting procedures became more personalized and dynamic. From targeting ads to situations (e.g., work vs. leisure [26]) up to targeting content to an individual's personality, known as psychometric targeting [113, 275]. These individual targeting mechanisms are also known as micro-targeting [50]. Micro-targeted ads unconsciously influence their audience, by speaking to fears and other subconscious triggers. Cambridge Analytica demonstrated the power of such technology, by influencing a.o. the Donald Trump election and Brexit vote [122] with mass persuasion through targeted content based on social media data. The border between clearly unethical use cases of psychometric targeting methods, such as the raisal of people's fear supported by fake news in the Donald Trump campaign, and societally accepted uses, such as personalized advertisements on social media recommending products in one's area of interest, is a continuum. Research should discuss to which extent the application of psychometric targeting can also be used for the good in an ethical manner [35]. Barriers to sustainable behavior are diverse and depend on individual norms, education, and experiences. One's attitude can make an exemplary subdivision: Among people whose general attitude is in favor of sustainable behavior, the *attitude-behavior gap* describes reasons that hinder actual sustainable behavior. On the other hand, there are people whose attitudes are not in favor of acting sustainably at all. Both groups of people have to be targeted differently when designing systems supporting sustainable behavior. In the first case, it is promising to support people in their intended actions (e.g., lowering burdens of the behavior). However, in the latter case, persuasion of one's internal beliefs and attitude would have to go first. Targeting could encompass various kinds of content. Advertisements and pro-environmental campaigns in social media could be targeted, to approach the viewer's individual burden against sustainable behavior. The unconscious approach could thereby bypass limitations of conscious targetings, such as rebound effects and climate depression. Furthermore, it enables us to talk to audiences that are not inherently interested in the topic of climate change.

3.3.4.4 Ethics: Sustainable = Good?

What is a *good* purpose is a matter of perspective. While for many people it might be undebatable that fostering sustainability is a good aim and political popularism is not, however, outside of this bubble, for example from the viewpoint of a confident republican politician, it might be vice versa. As a basis for the previous discussion point, we need to discuss whether what is *good* can be defined at all. Is it ethically correct to try to convince people with our pro-environmental viewpoint? This discussion is relevant for most technology that we design. Although for attitude and behavior changing applications its influencing nature is obvious, nearly all of our technology has some kind of influence on its users.

3.3.4.5 More Interdisciplinary Research

To design real-world applications we always need to go beyond the field of sole HCI. Especially when we talk about human behavior and societal relations, collaboration with other research fields is necessary in order to build good solutions. In the context of ecological and societal sustainability, this are especially the fields of environmental psychology, sociology and behavioral psychology. We want to motivate research go more often beyond their own research bubble, and include perspectives, experiences and insights from other related fields.

3.4 Chapter Conclusion

With the presented mobile sensing use cases in this chapter, we have shown that mobile sensing workflows can generate valuable benefits for multiple stakeholders. Mobile sensing data collection methods enable researchers to collect high-level human behavioral data in the field unobtrusively, fueling interdisciplinary research. Smartphone usage data reveals bad usage patterns, such as the mobile phone rabbit hole, and can enable interventions that help users to prevent or interrupt these. Detailed behavioral and contextual data furthermore allows to develop novel adaptive applications, of which users gain a direct benefit. The more in-depth the available data describes the user's context, the better such applications can integrate into the user's daily life. Being deployed at scale, mobile sensing technologies can have an influence on whole social groups and societies. This poses the opportunity to use them for the common good, e.g., to fight nowadays societal challenges such as polarization and climate change, however, also brings risks. In order to design adaptive, user-supporting applications well, a fine-grain and as comprehensive as possible view of the user's state and situation has to be available to the system. Especially novel proactive AI systems need detailed contextual information. Without such contextual data, proposed actions and interventions may be irrelevant as they may occur in situations where they are not appropriate.

However, using such rich, contentful mobile sensing data is hardly possible yet. People hesitate to adopt data-heavy apps for privacy reasons, and operating system developers restrict the access to such data in order to protect their users. Research lacks general insights about users' perceptions of such data usages, especially on the perceived privacy implications of mobile sensing apps. Although it is generally known that mobile sensing is perceived as privacy invasive, little is known about users' concrete fears, concerns, and worries. Furthermore application designers lack evidence on which features raise more or less concerns, and how they could be mitigated best. An essential model to describe users' app adoption intention is the service privacy fit. The service value of a sensing app varies strongly across the here presented application use cases. While the use cases where users have a direct benefit show a high service value, mobile sensing studies, where by nature a researcher or organization is the

major stakeholder, show a low service value. However, concrete insights on which features of apps and other factors facilitate or hinder the adoption of mobile sensing apps do not exist yet.

Chapter Take Away

This chapter motivates mobile sensing-based applications for the good of multiple stakeholders. Detailed, rich and contentful sensing data is thereby an essential building block of advanced app use cases. However, privacy concerns arising with such data and access restrictions by the operating system hinder their proliferation.

4

The User Perspective on Mobile Sensing Privacy

This chapter reports on **the users' perspective of smartphone sensing privacy**: We study how users perceive their privacy in the context of mobile sensing-based smartphone apps, which **data practices** raise privacy concerns, and what **outcomes they specifically are afraid of**. We report **how users currently approach these issues**, and which **mitigation measures they envision** to improve their privacy situation. These insights constitute an important basis for the design of novel privacy-enhancing interfaces.

This chapter is based on the following publications:

Planned Publication: Florian Bemmann and Sven Mayer. "The Impact of Data Privacy on Users' Smartphone App Adoption Decisions." In: *Proc. ACM Hum.-Comput. Interact.* MobileHCI '24 MHCI (2024). DOI: 10.1145/3676525

Florian Bemmann, Maximiliane Windl, Tobias Knobloch, and Sven Mayer. *Users' In-Depth Privacy Concerns with Smartphone Data Collection*

In the previous chapter, we have seen that the better a device understands its user's behavior and context, the better an adaptive service can be. Therefore, very detailed, contentful data types, such as detailed smartphone usage behavior or text contents, are especially interesting. However, data-heavy applications hardly make it into published apps. Users hesitate to install apps that use a lot of data, and operating system developers restrict access to rich, contentful data to specific purposes only. Current research has already identified privacy concerns as the most important barrier to sensing app adoption. Yet, the underlying reasons are not understood, cf. [56, 80, 128]. We do not know much about the effects of individual data types, privacy-enhancing technologies, and other app characteristics concerning users' privacy decisions. Furthermore, there is a general lack of the user perspective of privacy. Existing research studied barriers to the adoption of smartphone apps that include passive sensing, e.g., [83, 348] and mobile sensing research apps [221, 338]. Moreover, Christin et al. [89] provide an overview of technical privacy issues and measures. A plethora of papers propose technical concepts to reduce security issues (e.g., [40, 256]). Today's literature concludes that privacy is the most significant social barrier to adoption [221, 266]. Still, the underlying mechanisms of such effects are unclear: To tackle user privacy concerns, a deeper understanding of what users actually are afraid of, which outcomes they fear, and which solutions they desire is necessary. Making an app technically safe and privacy-friendly does not go far enough to ensure user acceptance.

These aspects are important to understand to build privacy-enhancing technologies that match the users' desires. Designers of future data-using smartphone apps and mobile sensing systems need to know how to foster user trust and lower privacy concerns to reach satisfactory adoption rates. To gain insights towards this problem, we put up the overall research question for this chapter:

RQ2 *What concerns arise from mobile user quantification?*

To collect knowledge on the users' perspective of privacy, we conducted two online surveys ($N = 100$ each) and one interview study ($N = 20$). First, we conducted an online survey to get quantitative insights into how specific app characteristics affect

users' app adoption decisions. Depending on several app characteristics, we asked users about their willingness to install and use a new mobile app. We consider the data types used, privacy concerns, privacy-enhancing features around transparency and control, and the benefits that users expect from the app. Second, we conducted another online survey ($N = 100$) investigating (a) users' knowledge of general smartphone data practices and privacy-enhancing technologies, (b) what users are concerned about, and (c) how their concerns can be mitigated. Through additional interviews ($N = 20$), we enrich the mainly quantitative results with a more qualitative-driven user perspective. Using a mixed methods approach, we investigate contextual factors influencing privacy concerns and analyze concrete user concerns. In detail, we show the underlying reasons causing the concerns, the specific privacy issue and feared consequences, the involved actors, the actions causing users' privacy concerns, the especially dangerous data types, and possible mitigation measures.

Our first online survey confirmed that users' main decision criteria for or against app adoption are the trade-off of privacy and expected benefit. Moreover, we uncovered more details on the specific data types and benefits. Using information about wallet- and account information induces most privacy concerns, followed by contentful datatypes such as text messages and microphone data. On the other hand, the expected benefits around productivity best mitigate these issues. Especially apps that help pursue a productive daily life or offer a monetary incentive have a high app adoption likelihood. The second online survey together with the interviews reveals what users are concerned about precisely. Here, we found that users are most concerned about third parties getting access to their data and misusing it to steal money or harm them. Users fear financial loss, being manipulated, and criminal activities being targeted against them. Users are initially unaware of their concerns; however, concrete privacy concerns arise when they engage with the topic. They see the underlying privacy issue mostly in data shared with or stolen by third-party actors such as companies or hackers. Overall, we found a lack of transparency in all stages. To mitigate the users' privacy concerns, app developers should mind user-centered privacy, including transparency, control, and trust.

We go beyond existing literature by investigating smartphone users' privacy concerns in-depth by going further than merely finding that privacy concerns are an important issue. Our findings show which app design settings designers and devel-

opers must be especially careful about, such as users' fears. We steer future work of privacy-enhancing technologies and underline the importance of better smartphone consent mechanisms, especially for contentful datatypes.

Research Gap

People's privacy decision-making processes have been studied a lot in the past. There exists a plethora of mental models (e.g., [110, 148, 225]), each with its own focus explaining several aspects that affect users' decisions. What strategy a user chooses and how they weigh the multiple aspects, depends on personal and app-specific characteristics [118]. We know from present studies that opinions and experiences from other people play a significant role (i.e., reviews and ratings). However, other app characteristics, especially tracked datatypes, and contents, are said to be the main factors influencing data-sharing decisions but are not studied in detail yet [173]. But with more detailed data being used more and more and fancy models being on the rise, knowledge about the effects of usage of such data gains importance. Literature reports strong non-willingness to install a tracking-heavy app as soon as users know that behavior [47, 61, 338]. While a few works study the influence of extensive data tracking in general, research lacks evidence of the effects of data types specifically. Also, regarding app characteristics around transparency and control features, we do not know much about their effect on adoption. We put up the first two research questions:

RQ2a *Which app characteristics hinder or facilitate the adoption of a mobile sensing app?*

RQ2b *Which datatypes stop people most from adopting a mobile sensing app?*

Regarding related work on user's concerns we found many markers; however, when drilling down on specific issues, we found only very limited insights and explanations for the specific user concerns. For example, in studies where users are directly asked for privacy concerns, a diverse set of privacy and security issues is reported, but only a few concrete, underlying concerns were mentioned [152, 324]. Insights on underlying reasons and fears are often rather a byproduct of studies, e.g., in Frik et al. [152] who directly admit that concern mentions are imprecise and superficial. Colnago et al. [94]

define *privacy concern* as *"an expression of worry towards a specific privacy-related situation."* Reviewing the related work, we do not find sufficient insights, especially the criteria of *specific* is not yet satisfied.

With this work, we aim to understand the in-depth user concerns, the underlying reasons, and the users' perspective. Going beyond technical security and designing for the user is essential to overcome trust and adoption issues. End-users' concerns and expectations are not sufficiently considered in the design process of privacy characteristics and features in systems [21]. To accomplish privacy by design, designers, and developers should take a more user-centered approach and should put a stronger emphasis on involving users' views and feedback. Therefore, it is important to understand the users' current concerns, so that researchers and developers can design appropriate user-centered solutions. Thus, To reach a better foundation for the development of future privacy-enhancing measures in smartphones, we compiled the following research questions addressing three research areas: *Understanding People*, *In-Depth Privacy Concerns*, and *Solutions to Mitigate Concerns*.

People's Understanding and Knowledge About Data Logging and Usage (RQ2c) In the context of apps, it is known that users weigh perceived risks against benefits when deciding for or against the installation [437]. Thus, to make an informed decision, users have to be knowledgeable. Moreover, from research on privacy policies in the context of online services, it is known that users barely read them [277, 294]. However, without understanding how an app and its privacy protection measures work, it is difficult for users to reach a low level of concern. Thus, we pose:

RQ2c *What is people's level of knowledge regarding privacy and security practices of the data collected through their smartphones?*

Users In-Depth Privacy Concerns (RQ2d) While RQ2d will give an understanding of users' preconceptions, we lack crucial insights about the actual concerns. So far, researchers have only investigated other domains, such as IoT [324], online advertisements [375], or smart homes [434]. Therefore, we investigate the users' privacy concerns of mobile sensing apps in-depth with:

RQ2d *What are people's detailed privacy concerns and feared real-world consequences of smartphone privacy issues?*

	Research Question	Paper	Chapter
RQ2	What concerns arise from Mobile User Quantification?		Chapter 4
RQ2a	Which app characteristics hinder or facilitate the adoption of a mobile sensing app?	[44]	Section 4.1
RQ2b	Which datatypes stop people most from adopting a mobile sensing app?	[44]	Section 4.1
RQ2c	What is people's level of knowledge regarding privacy and security practices of the data collected through their smartphones?	[48]	Section 4.2
RQ2d	What are people's detailed privacy concerns and feared real-world consequences of smartphone privacy issues?	[48]	Section 4.2 and Section 4.3
RQ2e	What are solutions to mitigate privacy concerns from the users' perspective?	[48]	Section 4.2 and Section 4.3

Table 4.1 : Overview of the studied sub research questions of RQ2.

Solutions to Mitigate Concerns (RQ2e) To complement the insights on concerns and their influencing factors, we investigate the user's perspective of how concerns could be mitigated with:

RQ2e *What are solutions to mitigate privacy concerns from the users' perspective?*

4.1 Study I: Privacy Issues Hinder the Proliferation of Sensing Apps

4.1.1 Methodology

To quantify the effects that app characteristics, especially logged datatypes and contained transparency and control features, have on smartphone users' app adoption intention, we conducted a large-scale online survey. The questionnaire consisted of three phases: 1) demographics, 2) effects of app characteristics on app adoption, 3) differences between individual datatypes. We provide the questionnaire in the Supplementary Material. We describe our results descriptively and run comparative tests, and finally, we discuss the implications of our results on future human-centered smartphone app privacy design.

4.1.1.1 Survey Design

In this section, we show the structure of our survey and explain how we came up with our questions and assessed items.

Part 1 - Demographics In the beginning, we asked for participants' smartphone usage to confirm their study eligibility and assessed demographics (country, gender, age, education, occupation). To classify our sample regarding their technology and privacy predisposition, we also assessed affinity for technology interaction (ATI) [151] and the IUIPC questionnaire [268].

Part 2 - Factors on App Adoption We assessed participants' app adoption intention, depending on several app characteristics and perceptions. We assessed each aspect with multiple items that we derived from literature-based constructs and calculated a score value for each aspect. We introduced the participants by telling them that we were interested in their decisions for or against installing and using a new smartphone app. All questions began with *I usually install an app on my personal smartphone...* following the adapted item. For example, *...if I feel I have control over my personal information that has been released* (one of the three items on transparency), or *... that requests sensitive personal information* as one of three items constituting permission sensitivity. Where appropriate, we opted for the more extreme wording of an item (i.e., using reinforcements such as "very"), as people in general agree that data is sensitive and risk-related. Thus, an extreme formulation yields better distribution in the responses [109]. We presented all statement questions using a slider ranging from *Strongly disagree* to *Strongly agree* on a 100-point scale without ticks and default selection (cf. [272, 335] who have shown that sliders lead to more precise responses). To ensure high data quality, we included attention checks as a slider item, which had to be moved to the very left or right at the end of each phase. We assessed the following potential factors of app adoption:

- **Perceived Permission Sensitivity** Perceived permission sensitivity describes *the level of discomfort users perceive when an app requests certain permission to control their mobile devices and use of their personal information* [171]. We use this construct to estimate how much impact mobile sensing data usage

overall has on the app adoption intention. We adopted the items of Gu et al. [171] (SENS1 - SENS3), which they developed and validated with a confirmatory factor analysis (CFA) [150].

- **Benefit Expectancy** We measure the extent to which a user believes they will benefit from installing an app with the construct of *benefit expectancy*. We use three items that Hsieh and Li [192] adapted from Venkatesh et al. [406] and Lai and Shi [240], and adapted them to our context's wording.
- **Transparency Features** Privacy-enhancing technologies that offer transparency to the users are known to affect users' app adoption decisions [47]. We assess the three transparency aspects *data collection*, *process transparency*, and *data use transparency* through three subscales from Agozie and Kaya [7].
- **Control Features** Equivalent to transparency, we also assess the impact of control features. We base five items on those of Xu et al. [442].
- **Service - Privacy Fit** It describes whether the service of an app matches its requests [192, 197], from a user's perception. To assess the effect of the service-privacy fit on app adoption, we adapted the items used by Hsieh and Li [192], who adapted items on task-technology fit [168] from Yang et al. [444] and Laugesen and Hassanein [246].
- **Privacy** We assess the effect that potential privacy concerns have on app adoption with four items adapted from Gu et al. [171].
- **User Ratings** We create an item that asks for the relevance of app store ratings.
- **Trust in Publisher** We assess trust in the app and its publisher with adapted items from Duan and Deng [125], who adapted items on trust in system [292] and trust in organization [13].

Part 3 - Effects of Specific Characteristics In the third part of our survey, we regard (1) datatypes, (2) publisher, and (3) personal benefits in more detail. We gathered a list of datatypes (respectively publisher types and personal benefits) from related work, and let participants rate them individually. Thereby, we create a more nuanced understanding of how these three factors affect an app adoption decision in detail.

Datatypes To compile a list of datatypes, we reviewed all Android permissions[1], accessibility service event types[2], and iOS permissions[3], and grouped them into human-understandable data types. We distinguish between read access and permissions that request write access or the ability to perform actions. For each permission, we ask users about (1) its perceived sensitivity and (2) potential risk. All permissions are listed in the Supplementary materials.

Publisher We let our participants rank four types of app publishers by how much they trust them to protect their privacy. We reviewed publishers in the Android and iOS app stores and found it comprehensive to cover *university*, *governmental organizations*, *companies*, and *non-profit organizations*.

Benefit To break the expected benefits down in more detail, we collected specific benefits from related work. We used items based on Jung [208], formulated based on their code name and examples. We added *monetary incentive*, which is mentioned by Malik et al. [269] but not included in the codes of Jung [208]. We report on this rating separately and do not include it in the score on benefit expectancy.

Open Question As the last element, we added an open text field question, where participants could enter any feedback about the survey.

4.1.1.2 Pilot Testing

We piloted the study with 20 participants. We ensured that we received an appropriate data distribution, and checked for critical comments in the final open feedback question that could have hinted towards issues with understandability. Our pilot test did not raise any issues.

[1] https://developer.android.com/reference/android/Manifest.permission, last accessed 2024-12-03

[2] https://developer.android.com/reference/android/accessibilityservice/AccessibilityService, last accessed 2024-12-03

[3] https://developer.apple.com/documentation/bundleresources/protected-resources, last accessed 2024-12-03

4.1.1.3 Procedure

We implemented the questionnaire in the survey tool Qualtrics and recruited participants through Prolific. We balanced the participant pool by gender, age, and country of residence and required participants to speak English fluently. We rewarded participation with 3£ as the study took approximately 20 minutes.

4.1.1.4 Participants (Survey Part 1)

We recruited 100 participants (47 female, 50 male, and 3 non-binary), aged 18 to 66 ($M = 31.9$, $SD = 10.1$). Most participants were either full-time (50) or part-time employed (24). A third (34) were students, and half held a university degree (35 had bachelor's degrees, and 20 had master's degrees). Most participants lived in Poland (13), South Africa (11) and Italy (10). To assess our sample's affinity with technology, we used the affinity for technology interaction scale (ATI) [151]. Its scale ranges from 1 (least affinity for technology) to 6 (highest possible affinity). Our sample had an average score of around 4 ($M = 4.13$, $SD = 0.82$). This indicates a tendency towards a higher technology-affine sample than the average population. According to the classification of Franke et al. [151], the ATI of an average population is to be expected at around 3.5, with high ATI samples around 4. Regarding the questions on perceived information privacy, our participants rated *Awareness* on average with 6.19 ($SD = .76$), *Control* with 5.83 ($SD = .87$), and *Collection* with 5.70 ($SD = 1.13$) (higher scores correspond to higher privacy).

4.1.2 Results

Participation took approximately 17 minutes to complete the study. We run the statistical evaluation in Python and R. Moreover, we applied non-parametric tests when the normality was violated.

4.1.2.1 Factors to App Adoption (Survey Part 2)

In the second part of our survey, we assessed the relevance of a selection of factors to app adoption intention. For each factor's items, we calculated the mean value over its items, constituting a score value. We inverted item values where necessary so that a higher score indicates higher app adoption intention for each factor. All scores

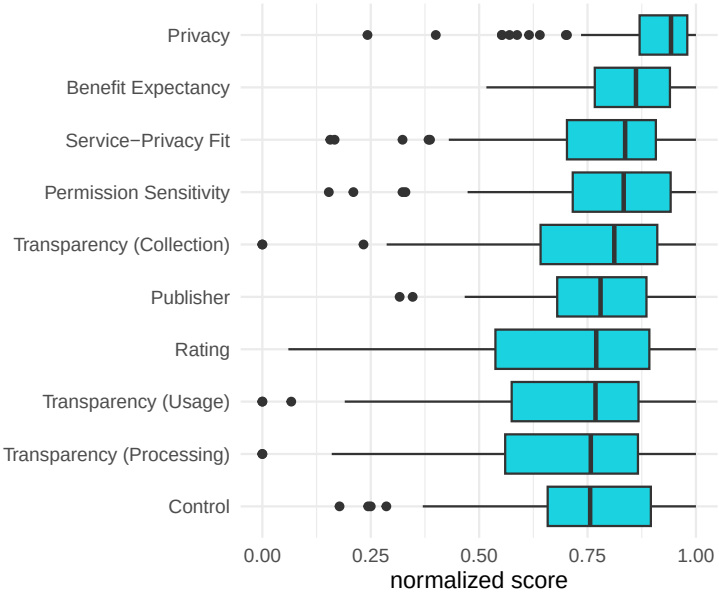


Figure 4.1 : Ratings of app adoption intention, for ten factors.

were normalized to a value range from 0 to 1. The obtained scores are generally high, constituting a distribution that is shifted towards the upper end of the scale. As our data thus is not normally distributed, we regard the median instead of the mean in the following.

The strongest positive effect on app adoption intention was indicated for the factor *Privacy*, i.e., if users perceive an app as being privacy friendly ($Mdn = .94, SD = .14, min = .24, max = 1$). Thereafter, follow *Benefit Expectancy*, i.e. users tend rather to install an app if they expect to have a benefit thereof ($Mdn = .86, SD = .12, min = .52, max = 1$). Thereafter follow the perceived *Service-Privacy Fit* (i.e. whether users perceive that the app's privacy invasions are appropriate regarding its provided service), *perceived permission sensitivity* and *Transparency about data collection* all with median scores above 0.8. With a median between 0.7 and 0.8 follow the type of *Publisher*, the app's *Rating*, and *Transparency about Data Usage*, *Transparency about Processing* and lastly, the presence of *Control features*.

					Post-hoc test								
factor	**Mdn**	**SD**	**Min**	**Max**	Privacy	Benefit Expectancy	Service Privacy Fit	Permission Sensitivity	Transparency (Collection)	Publisher	Rating	Transparency (Usage)	Transparency (Processing)
Privacy	0.942	0.137	0.242	1	-	-	-	-	-	-	-	-	-
Benefit Expectancy	0.862	0.119	0.517	1	.057	-	-	-	-	-	-	-	-
Service Privacy Fit	0.837	0.182	0.157	1	0	.764	-	-	-	-	-	-	-
Permission Sensitivity	0.833	0.183	0.153	1	.001	1	1	-	-	-	-	-	-
Transparency (Collection)	0.812	0.218	0	1	0	.064	1	1	-	-	-	-	-
Publisher	0.780	0.157	0.317	1	0	.035	1	1	1	-	-	-	-
Rating	0.770	0.241	0.060	1	0	.002	1	.566	1	1	-	-	-
Transparency (Usage)	0.768	0.245	0	1	0	.001	.065	.136	1	1	1	-	-
Transparency (Processing)	0.758	0.239	0	1	0	0	.467	.049	1	1	1	1	-
Control	0.756	0.199	0.178	1	0	.005	1	1	1	1	1	1	1

Table 4.2 : Descriptive statistics of the rated app adoption intention given that the respective feature is present in an app (left side). On the right, we show the p values of a pairwise Wilcoxon rank sum test (Bonferroni adjusted), for which we conducted post-hoc tests of a Friedman test.

The score on *Perceived Privacy* turned out to be significantly higher than all other factors except *Benefit Expectancy* (Friedman rank sum test with post-hoc pairwise Wilcoxon rank sum tests; $X^2(9) = 128.21, p < .0001$). *Benefit Expectancy* also shows a significant difference with most factors. Besides, the only significant difference was detected between *Permission Sensitivity* and *Transparency about Processing*.

4.1.2.2 Factors in Detail (Survey Part 3)

Datatypes In four blocks, we asked our participants to rate each datatype's (1) permission sensitivity and (2) perceived risk; separately for permissions that depict read access and permission that grant write access / perform some action. The rating was done on a continuous slider, yielding values between 1 and 100. The two measures *sensitivity* and *potential risk* show a significant moderate to strong correlation to each other ($r(98) \in [.49; .79]; p < .0001$), except for *reading wallet information*, whose correlation is only weak ($r(98) = .35, p < .0001$), cf., classification of Schober et al. [354]. Therefore, we will report on both measures together. Detailed independent values can be found in Table 4.3.

	Sensitivity		Potential Risk		Pearson Correlation		
Item	**Mdn**	**SD**	**Mdn**	**SD**	**r(98)**		**CI**
Read Data							
wallet information	99.5	17.5	100	20.1	.35	***	[.17;.51]
account information	94	21.8	94	26.3	.53	***	[.37;.66]
text messages and calls	91	27.2	90.5	28.4	.61	***	[.47;.72]
microphone data	89	30.4	87	30.2	.65	***	[.52;.75]
camera data	87	26.7	82	28.9	.66	***	[.54;.76]
location	84.5	26.3	86.5	28.6	.59	***	[.45;.71]
contacts	82.5	28.7	85	28.2	.64	***	[.50;.74]
screen contents	70	27.7	75	28.7	.62	***	[.48;.73]
data from smart home devices	65.5	30.3	67.5	30.2	.58	***	[.43;.69]
keyboard typing data	63	31.5	67.5	33.0	.64	***	[.50;.74]
notifications	59	30.5	51	29.9	.49	***	[.32;.62]
calendars and reminders	58	30.3	53.5	30.4	.66	***	[.53;.76]
body sensors and health data	52.5	31.8	50.5	31.2	.70	***	[.58;.79]
interactions and touch behavior	51	28.3	54	27.5	.63	***	[.50;.74]
phone state	51	30.5	55	29.0	.63	***	[.49;.73]
physical activity data	50.5	27.8	49	27.0	.62	***	[.48;.73]
motion data	48	26.6	44.5	26.7	.59	***	[.45;.71]
usage statistics	47.5	26.6	46	30.0	.73	***	[.62;.81]
music library	22	25.3	22.5	24.4	.59	***	[.44;.70]
Write Data							
send text messages	92	20.3	89	21.9	.73	***	[.63;.81]
edit contacts	89	22.8	81.5	24.0	.69	***	[.57;.78]
install apps and packages	89	24.3	88	24.6	.66	***	[.54;.76]
add, edit, and delete files and media	87	24.4	88	25.8	.71	***	[.59;.79]
change my phone's state	67.5	31.8	63	31.0	.72	***	[.61;.80]
access the internet	67	31.6	68	30.9	.79	***	[.71;.86]
edit calendar entries and reminders	65	27.9	65	30.0	.75	***	[.66;.83]
send me notifications	49.5	30.6	44	30.3	.72	***	[.61;.81]
edit my music library	40.5	31.3	33	31.1	.74	***	[.63;.81]

Table 4.3 : Statistics of participants' sensitivity ratings and potential risk for various read and write datatypes. Pearson's correlation statistics show how the two assessed factors *sensitivity* and *potential risk* correlate (*** p<.001).

Read Access The highest rank datatypes about *wallet* and *account information*, followed by the contentful datatypes *text messages*, *microphone data*, *files and media*, *camera*, *location*, and *contacts*. They all show median rated permission sensitivity of at least 82.5 and potential risk above 82.0. After a gap of 12.5 points on the sensitivity scale and 7 points on the potential risk scale, users rank *screen content*, *keyboard typing data* and *data from smart home devices* with at least 63 points regarding sensitivity and 67.5 potential risks. For potential risk, we see another gap between the remaining datatypes. All do not reach a higher median than 55, while that is not

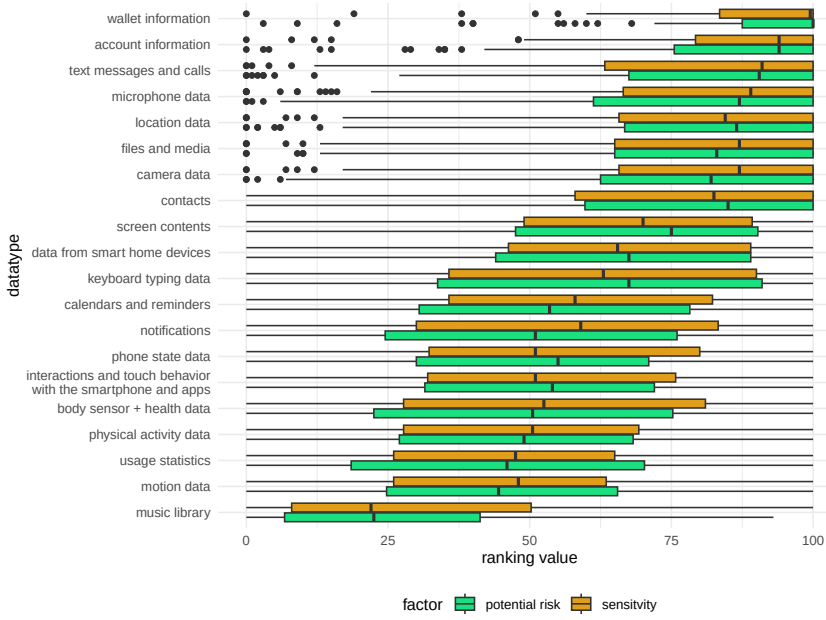


Figure 4.2 : Participants perceived sensitivity and potential risk of different read-only permissions.

existent for sensitivity (next-highest datatype at 59). The last and by far lowest ranked datatype is *music library* with a median sensitivity of 22 and a median potential risk of 22.5. The second-last datatype ranks at least twice as high for both scales.

Write Access Regarding write access permissions, for both scales, the four permissions *send text messages*, *install apps*, *add, edit and delete files*, and *edit contacts* rank the highest, showing a gap between 87 and 67.5 points in sensitivity respectively 81.5 and 68 points in potential risk. The following three datatypes *change my phone's state*, *access internet*, and *edit calendar entries and reminders* were ranked between 67.5 and 65 in sensitivity, respectively 68 and 63 points in potential risk. With some distance, *send me notifications* and *edit my music library* are at the end of the spectrum with a sensitivity of 49.5 and 40.5 points and a potential risk of 44 and 33.

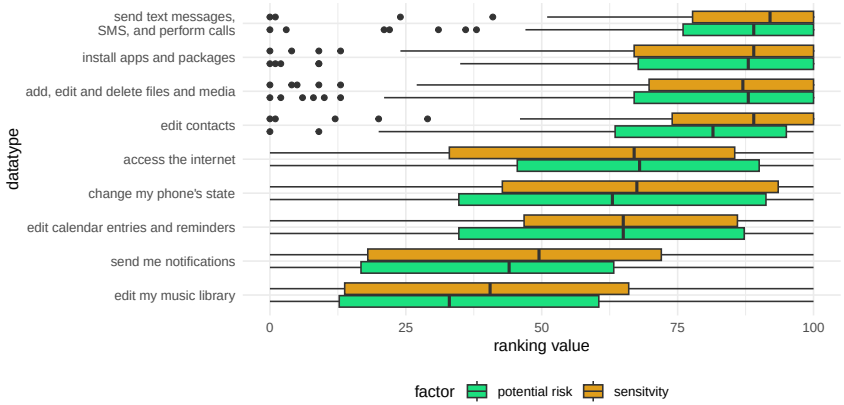


Figure 4.3 : Participants' perceived sensitivity and potential risk of different write-access permissions.

Differences Between Read and Write Permissions For all datatypes that can be paired to a respective read and write variant, we compared the participant ratings between the according read and write permissions. Write permissions were rated more sensitive and imposing higher potential, except for *notifications*; there users did rate oppositely. Differences are significant for both ratings of phone state ($p < .01$), notifications ($p < .05$), music ($p < .01$), the sensitivity of text messages and calls ($p < .05$) and contacts ($p < .05$), and the potential risk of calendar and reminders ($p < .05$) (paired Wilcoxon Signed rank tests).

Publisher Regarding the app publisher, users rate apps published by universities as most likely to be adopted ($Mdn = 76, SD = 26.9$), and non-profit organizations as second ($Mdn = 61.5, SD = 28.0$). Thereafter, governmental organizations ($Mdn = 52, SD = 30.1$), and the lowest app adoption intention were rated to companies ($Mdn = 35.5, SD = 28.1$). The best-rated option *university* and the last option *company* each differ from all other options significantly (Friedman rank sum test with post-hoc pairwise Wilcoxon rank sum tests; $X^2(3) = 79.02, p < .0001$).

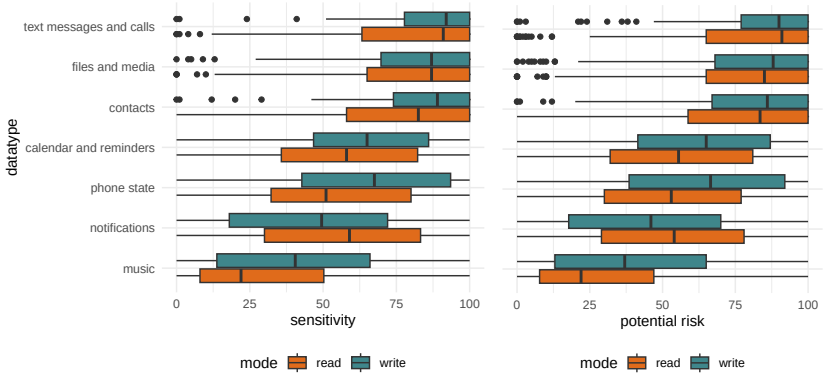


Figure 4.4 : A pairwise comparison of read and write permissions show that users rate write permissions more sensitive and impose higher potential risk than their read equivalent, except for notification access.

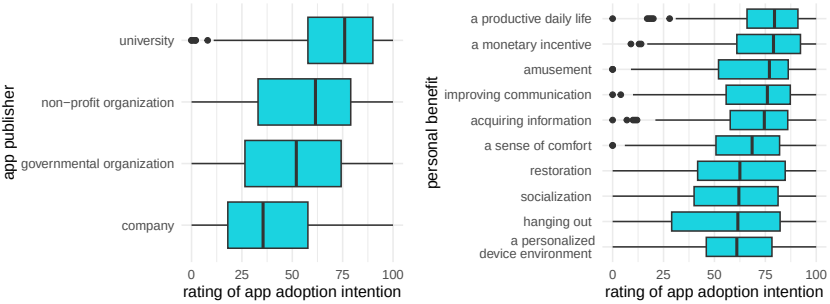


Figure 4.5 : While participants indicated clear differences in their willingness to adopt an app between different publishers, the differences between various personal benefits are rather low.

Personal Benefits The ratings of personal benefits regarding app adoption intention indicate that participants' app adoption intention does not differ that much by which personal benefit an app provides. The medians of all ten rated benefits range between 79.5 and 61 and thus show less spread than the other factors that we assessed. The top five benefits are goal-oriented, rather productive benefits (*productive daily life* ($Mdn = 79.5, SD = 22.1$), *monetary incentive* ($Mdn = 79, SD = 23.3$), *amusement* ($Mdn = 77, SD = 25.4$), *improving communication* ($Mdn = 76.5, SD = 22.7$) and *acquiring*

information ($Mdn = 74.5, SD = 24.4$)). Less concrete and more leisure-focused benefits show up at the end of the rating: *Sense of comfort* ($Mdn = 68.5, SD = 24.1$), *restoration* ($Mdn = 62.5, SD = 27.7$), *socialization* ($Mdn = 62, SD = 27.0$), *hanging out* ($Mdn = 61.5, SD = 30.7$), and *personalized device environment* ($Mdn = 61.0, SD = 25.3$).

4.2 Study II: Online Survey: The User Perspective of Privacy Concerns, Fears, and Mitigation

We first conducted a large-scale online survey to gain quantitative insights into our research questions. The questionnaire consisted of three phases: 1) demographics and knowledge, 2) understanding users' concerns in general, and 3) specific concerns and envisioned mitigation measures. We provide the questionnaire in the Supplementary Material.

4.2.1 Survey Design

We presented all statement questions using a slider ranging from *Strongly disagree* to *Strongly agree* on a 100-point scale without ticks and default selection, cf. [272]. To ensure high data quality, we included attention checks as a slider item, which had to be moved to the very left or right at the end of each phase.

Phase 1: Demographics and Knowledge This phase consists of six blocks: 1) demographics, 2) participants' general privacy perception (IUIPC questionnaire [268]), 3) technology affinity (ATI scale [151]), and 4) a set of self-constructed free text items on which smartphones the participants own and which mobile sensing apps they are familiar with. Here, we also introduced a definition of mobile sensing apps. Afterward, we had 5) a self-constructed set of items on the knowledge and understanding of 4 privacy-enhancing measures occurring in mobile sensing systems (encryption, anonymous data collection, hashing, remote server). For each concept, the participants had to indicate for three statements, whether they were true or false. We randomized the

order of the questions to prevent order effects. The last item in this first phase is 6) one self-constructed item about how much users familiarize themselves with the privacy implications before installing an app.

Phase 2: Understanding Users' Concerns In the second phase, we openly asked about the users' concerns. To not bias the participants, we deliberately asked open questions before letting them rate items that tackled specific aspects. The open questions asked the participant to name one specific concern, define what exactly they are afraid of happening, which situations, datatypes, and involved actors they considered particularly concerning, and how they envisioned their concerns to be mitigated. This phase was implemented as an optional loop so that participants could potentially express multiple concerns.

Phase 3: Specific Concerns Afterward, four mobile sensing app use case scenarios were presented in a randomized order, namely *Ambient Noise App*, *Navigation App*, *Sports and Fitness App*, and *Travel Advice App* (see Supplementary Material for a scenario description). For each scenario, we asked questions on the general concern, familiarity, perceived usefulness, and envisioned concern mitigation options. After doing this for all four scenarios, we again went through the scenarios and asked for concerns regarding specific aspects compiled from Windl and Mayer [434] and Barbosa et al. [27], such as third-party data access or profile building. We further randomized the order of the four scenarios.

4.2.2 Pilot Testing

We piloted the study with 20 participants, including a full qualitative data analysis. We ensured that all questions were understandable, the questionnaire was working well technical-wise, and that we received the desired kind of responses. After analyzing those 20 pilot responses, we made significant changes, especially to the open questions of the second phase. Moreover, we tested and discussed the design of the self-constructed set of items on knowledge and understanding of privacy-enhancing technologies with three researchers from our lab who were not involved in the project. We thereby ensured that the items were not formulated misleadingly and were solved correctly for people who were familiar with the presented concepts.

4.2.3 Procedure

We implemented the questionnaire in the survey tool Qualtrics and recruited participants through Prolific. We balanced the participant pool by gender, age, country of residence, and occupation and required participants to live in Europe and speak English fluently. We rewarded participation with 3.34£ as the study took approximately 25 minutes.

4.2.4 Participants

We recruited 100 participants (48 female, 51 male, and 1 non-binary), aged 19 to 72 ($M = 31.5$, $SD = 9.2$). Most participants were either full-time (49) or part-time employed (14). A third (31) were students, and half held a university degree (34 had master's degrees and 27 had bachelor's degrees). Most participants lived in Poland (17), France (11), Portugal (8), and Hungary (8). To assess our sample's affinity with technology, we used the affinity for technology interaction scale (ATI) [151]. Its scale ranges from 1 (least affinity for technology) to 6 (highest possible affinity). Our sample had an average score of around 4 ($M = 3.87$, $SD = 0.97$). This indicates a tendency towards a slightly higher technology-affine sample than the average population. According to the classification of Franke et al. [151], the ATI of an average population is to be expected at around 3.5, with high ATI samples around 4.

4.2.5 Data Analysis

We preprocessed the questionnaire data with Python and imported the free text answers into ATLAS.ti for coding. Two researchers independently coded the first 20 participants. We then discussed the coding and revised these participants' codings before one researcher coded the remaining participants. With all participants coded, three authors met in person to discuss the codes and formed initial themes. We iteratively reworked the codes and themes in multiple sessions by comparing the coded snippets across all themes. The final coding consists of 374 distinct codes organized into 53 code groups. Each code expresses a specific aspect (e.g., *take out a loan*), while code groups categorize them to a broader level (e.g., *financial loss*).

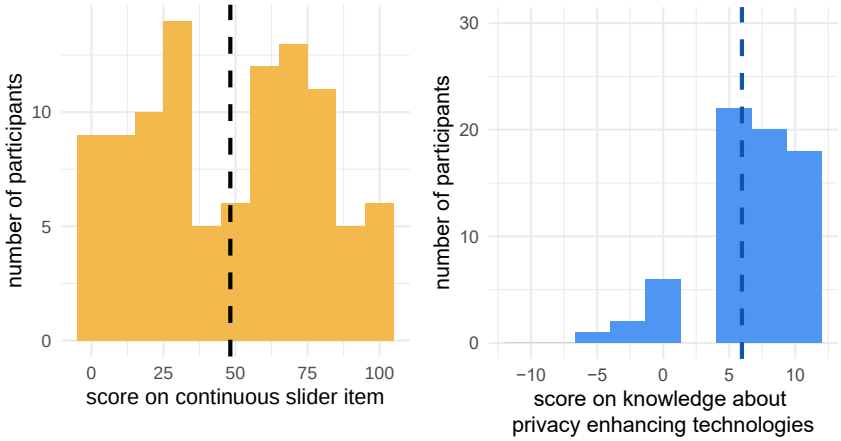


Figure 4.6 : Distribution for our two measures of engagement with privacy information (left) and knowledge about privacy enhancing technologies (right). The dashed lines represent the means of each measure.

4.2.6 Results

In this section, we present our results along with our research questions. We start with general concerns and influencing factors before we describe the detailed types of concerns expressed by participants. We show how concerning our participants rated several aspects of sensing applications and what they imagined to mitigate their concerns.

4.2.6.1 User Knowledge (RQ2c)

In our quiz items that assess how knowledgeable and informed users are about technology, privacy, and security in the smartphone context users are, participants mostly reached 6 points ($M = 5.62$, $SD = 3.34$; scale [-12;12], the expected random response is 0). The score distribution is skewed towards the right and not normally distributed, see Figure 4.6.

We also asked the participants how much they made themselves familiar with the data practices before installing new apps. On a continuous scale between 0 and 100, the average answer is in the middle ($M = 48.22, SD = 30.73$). Taking a look at the

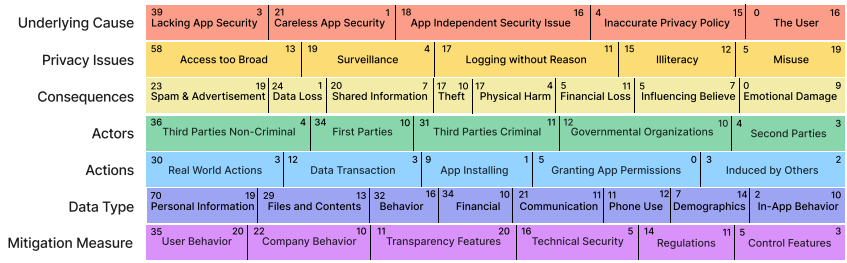


Figure 4.7 : Our code groups that underlay the seven themes of our privacy concern model. Numbers in the upper left corner of each code indicate the number of *online survey* participants expressing the code, the number in the top right corner the number of mentions in the *interviews*.

distribution (cf. Figure 4.6) , we found a gap in the middle, i.e., very few participants replied with values around 40. Additionally, we found two peaks, one at around 30 and one at around 70. Thus, we identified two types of users: Those who care very little and those who care rather much about the data handling practices of an app.

4.2.6.2 Privacy Concerns in the Sensing Data Pipeline (RQ2d)

A comparison of concern ratings regarding given aspects assessed both in the beginning and end of the survey revealed significantly higher concerns in the second assessment ($M_{before} = 39.09$, $M_{after} = 58.16$; scale [0;100]; Wilcoxon signed-rank test: $p < .001$). This shows that people are not fully conscious of their concerns initially, instead, concerns arise while dealing with the topic.

In the questionnaire, we presented each user with four scenarios of using a mobile sensing app. Designed in a 2×2 factorial design, they incorporate the two variables *approaches* of mobile sensing systems (people-centric vs. environmental-centric) and *involvement of people* (participatory sensing vs. opportunistic sensing) proposed by Laport-López et al. [244].

We found significantly higher concerns for opportunistic sensing (i.e. passively collected data) ($M = 43.1$, $SD = 33.8$) than for participatory sensing (i.e. actively entered data) ($M = 35.1$, $SD = 29.9$). Furthermore, environment-centric data ($M = 44.1$, $SD = 33.4$) yielded higher concerns than people-centric data ($M = 34.1$, $SD = 30.1$). We conducted a two-way non-parametric ART ANOVA [436] for *Approaches*

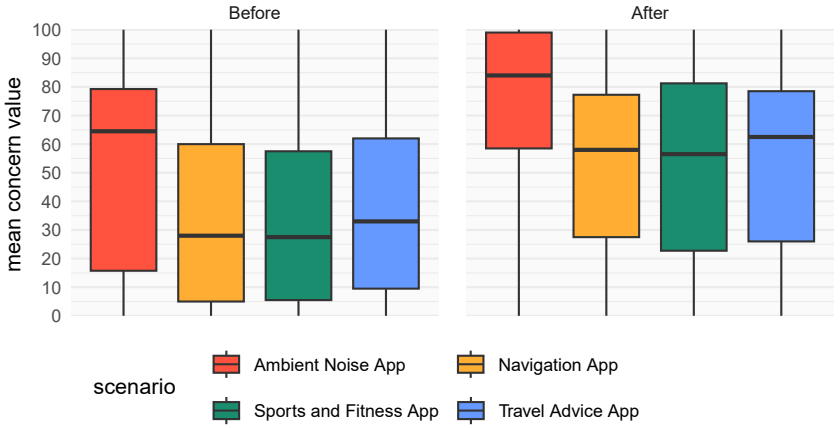


Figure 4.8 : Reported privacy concerns regarding four presented mobile sensing scenarios.

vs. *Involvement*. We found both main effects, *Approaches* and *Involvement*, are statistically significant different, $F(1,396) = 9.461$, $p < .01$ and $F(1,396) = 6.214$, $p < .05$, respectively. Moreover, we found a statistically significant interaction effect $F(1,396) = 4.954$, $p < .05$. We found significant interactions between those two variables and the users' privacy concerns.

Regarding the results by individual scenarios, the reported privacy concerns were clearly the highest for the scenario *ambient noise app*. The *navigation app* and *sports and fitness app* were judged as least privacy concerning. The scenario *travel advice app* was rated a bit more privacy concerning than the latter two. The concerns per scenario are visualized in Figure 4.8.

We found that *data misuse* ($M = 72.68$), *3rd party data access* ($M = 71.79$) and *data getting stolen* ($M = 69.45$) were the most concerning aspects. *3rd party data access* was rated significantly more concerning than *2nd party data access* and *1st party data access* using Kruskal-Wallis rank sum test and Dunn's Test [296]; see Figure 4.9 and Table 4.4. Furthermore, we found significantly lower concerns for *local data processing* and *local data storing* in comparison to their global alternatives.

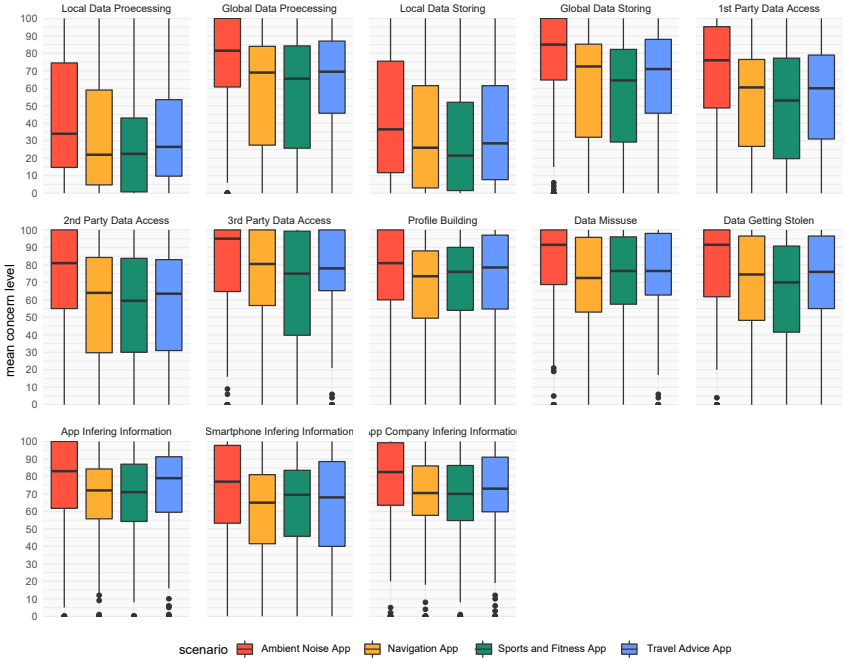


Figure 4.9 : The rated concern level of specific privacy-threatening aspects of mobile sensing apps, regarding four mobile sensing app usage scenarios.

4.2.6.3 Qualitative Analysis: Users' Privacy Concerns Regarding Smartphone Data (RQ2d)

Our thematic analysis revealed seven overarching themes that describe the scenarios evoking privacy concerns. Figure 4.10 gives an overview of how the themes connect: An UNDERLYING CAUSE (e.g., a weakly secured server) triggers a PRIVACY ISSUE (for example, data being stolen from that server by hackers). Privacy issues provoke a (REAL WORLD) CONSEQUENCE (for example, the stolen data being leveraged to withdraw money from an online banking account). A PRIVACY ISSUE is caused by ACTIONS, involves ACTORS (e.g., hackers, companies), and affects specific DATA TYPES.

MITIGATION MEASURES can be employed to tackle a PRIVACY ISSUE. Figure 4.7 shows an overview of all themes and their affiliated code groups. In the following, we describe the themes in more detail.

Underlying Cause Most (82) users see the fault for privacy issues in the app companies. Of these, most (49) users believed security issues occurred without companies' malicious intent but due to *Lacking App Security*. Users described scenarios, such as data leaks, security breaches, or hacker attacks. However, several (28) users also believed that privacy issues are caused by companies not employing sufficient measures (*Careless App Security*), as P63 stated: *"I am concerned about lack of proper care from a company."* Moreover, several participants (19) also mentioned app-independent security issues, such as *"viruses and spyware"* (P89). Finally, five participants also mentioned *Inaccurate Privacy Policies* as a trigger for privacy concerns, as P35 explained: *"[I am] concerned that sometimes terms of privacy do not tell everything about the use of my data. That maybe they are lying about [...] what happens to my data."*

Privacy Issues Most (80) participants were concerned about too many people having access to their data. Precisely, data theft (28) and their data being sold (24) were mentioned. For example, P1 is concerned that *"the company that had my data could have sold it, or another way would be that they were hacked and the database was compromised."* Data theft and trade are followed by the feeling of being surveilled (23), for example, by tracking their location. Logging data without reason was also frequently mentioned (20), which included the collection of personal histories, constantly logging in the background, or accessing data sources that the user did not authorize. Participants

	KRUSKAL WALLIS			DUNN'S TEST	
	chi-squared	*df*	*p*	*Z*	*p*
1st party data access vs. 2nd party data access				−1.471	1.
1st party data access vs. 3rd party data access				−4.794	<.001
2nd party data access vs. 3rd party data access	228.36	12	<.001	−3.323	.070
Global Data Storing - Local Data Storing				6.667	<.001
Global Data Processing - Local Data Processing				6.736	<.001

Table 4.4 : The two-way F-statistics of users' privacy concerns, regarding different types of threats. P-values of Dunn's test are Bonferroni adjusted.

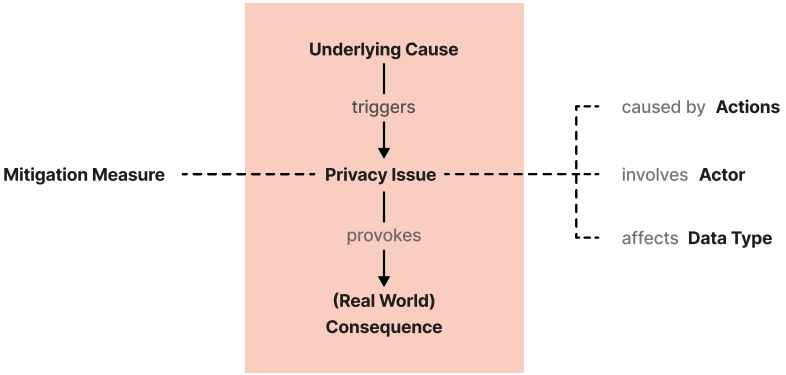


Figure 4.10 : Our privacy concern model. Privacy issues are at the center, triggered by causes and leading to consequences.

also mentioned concerns caused by not knowing what an app is doing (18), such as using or sharing data without their knowledge. Data misuse by the company, e.g., for profit, was mentioned least often.

Consequences Most (28) participants feared not necessarily harmful but annoying consequences such as personalized advertisements or being manipulated regarding their shopping behavior. Participants also frequently (25) mentioned consequences that represent a loss of control, for example, that big datasets about them could be gathered or that their identity is stolen and misused, leading to further consequences such as *"[..] the possibility of identity theft."* (P7). Moreover, several (21) participants feared that their data might get publicly available, or even directly affect participants' lives beyond the online world through theft (19), or even physical harm (18), such as stalking. A different aspect of consequences with real-life impact is financial loss (6), as P66 describes *"[...] my financial data. Let's say I did not pay one month of my mortgage. This info can then be spread all over the globe, and I will not be able to get a loan from a bank [...]."* Lastly, manipulation was also mentioned by a few (5) participants.

Actors Participants most frequently (46) mentioned third parties without criminal intent, such as advertising companies, when asked who triggered their privacy concerns. The second most frequently (38) mentioned actor was the first party, i.e., the app

company. Interestingly, 35 participants fear that third parties with criminal intent, such as hackers, pose a danger to their privacy. Our participants also named government organizations (15), and secondary parties (4), such as phone manufacturers, as actors evoking privacy concerns.

Actions Participants mentioned most frequently (42) that actions "in the real world" trigger their concerns, especially the use of public WiFi networks. Users also expressed concerns during data transactions (14), i.e., while documenting, viewing, or working with data. When interacting with an app, most users (10) were more concerned when installing an app (10) than when being asked to grant permissions (6). Lastly, participants were also concerned about other people causing a privacy issue for them (4).

Data Types The most frequently mentioned data type was personal information (131), such as login details, phone numbers, or home addresses, followed by files and content (37), such as photos. Moreover, participants often named behavioral data (36), such as location and habits, and financial data (34). Finally, participants mentioned communication protocols (27) (e.g., WhatsApp messages), phone use (13) (e.g., screen time), demographics (8) (e.g., age and gender), and in-app behavior (3) (e.g., the time they spent in an app) least frequently.

4.2.6.4 Solutions to Mitigate Privacy Concerns (RQ2e)

Interestingly, the most mentioned measures to reduce their concerns were changes in their own behavior (mentioned by 45 participants). Most ideas are related to using an app less or not at all, blocking permissions, or only choosing trusted companies. Behavior changes by the app company were mentioned second most frequently (25). Most suggestions were related to data minimization: Apps and companies should ask for and use less data (e.g., *"companies not requiring as much data and not tracking online activity,"* P33), and store data only for the necessary amount of time. Next, we recognized a desire for more transparency features (18) (better understandable privacy policy, more clearly stating what, when, and how data is used) and the wish to have more control over what happens with their data. For example, P35 expressed the desire to learn *"how the app works and collects [their] data."* Suggestions for technical

security (18) included safer storage and transmission of data, e.g., by encryption or processing and storing data locally. Regulatory measures are often mentioned (16), here, participants desired more laws for data safety, global standards, and institutions that enforce the rules. Least frequently (6) were statements that we categorized as control features, e.g., the ability to turn off data access.

4.2.7 Summary

Our survey shows that users generally know about privacy-enhancing technologies (RQ2c). However, we found that they are ambivalent about informing themselves about the privacy aspects of new apps and denounce that they are not well informed. The online survey's qualitative part revealed themes around the topic of smartphone privacy and gave us an impression of which aspects are relevant to users (RQ2d). Quantitative questions confirm that third-party data theft and misuse concern users the most. Furthermore, the quantitative questions revealed high concerns, especially for passively sensed data on contextual variables. However, due to the nature of an online survey, it is hardly possible to get a detailed understanding of users' concerns and to understand how they could be mitigated. Therefore, we decided to supplement these insights through interviews to gain a more in-depth understanding. This will especially help answer RQ2d (in-depth concerns of the users) and RQ2e (mitigation measures).

4.3 Study III: Interviews: Users' Privacy Concerns, Fears, and Envisioned Mitigation Approaches

We conducted an interview study to get a more in-depth understanding of the patterns that participants came up with in our online survey. In semi-structured interviews, we dig deeper into the users' concerns and perceived mitigation opportunities. Furthermore, we want to confirm our privacy concern model (cf. Figure 4.10) with a second participant sample.

4.3.1 Procedure

We decided to conduct semi-structured interviews with a guideline whereby the order of the questions does not have to be strictly followed, and the interviewer can ask follow-

up questions whenever appropriate [191]. The guideline contained three key topics: Users' knowledge level, users' concerns and fears, and mitigating factors, including knowledge of protective measures. Each section had three to five questions, and we also added possible follow-up questions. We designed the questions open-ended [427], and we explicitly noted that the questions did not suggest particular answers so that each participant could reflect on their knowledge or opinions without biases. When developing the guideline, we first formed the topic areas before we formulated the concrete questions. In the next step, we critically reviewed these questions and reformulated them whenever necessary. Thus, the questionnaire creation was roughly based on Helfferich [186].

At the beginning of the interview, participants had to fill out an online questionnaire to assess their demographics, affinity for technology interaction (ATI) [151], and their individual information privacy concern level using the IUIPC questionnaire [268] and three items adapted by Prange et al. [323] based on Malhotra's causal model [268]. In addition, we formulated three statements on concerns about smartphone data collection and disclosure (i.e., general concern about smartphone data collection and second and third-party sharing), where participants had to indicate their level of agreement on a continuous 100-point slider, ranging from "strongly disagree" to "strongly agree." We recorded the interviews and compensated the participants with 5€.

4.3.2 Data Analysis

Each interview took, on average, 22 minutes and 33 seconds ($SD = 9m20s$), resulting in 7.5 hours of audio material. The interviews were transcribed using the transcription software Trint[1]. Afterward, we proofread all transcriptions and corrected any errors. We analyzed the interviews using ATLAS.ti and thematic analysis [63], meaning that three researchers first independently open-coded two interviews. We then met to discuss our codes, resolve ambiguities, and form a joint codebook. One researcher then coded the rest of the interviews, after which a fourth researcher joined to form code groups and overarching themes through multiple rounds of hour-long discussions.

[1] https://trint.com/, last accessed 2024-12-06

4.3.3 Participants

We recruited 20 participants, half through the university's mailing list and half via convenience sampling. Through this, we hoped to recruit a more diverse sample, including different professions and age groups. The participants from the interview study are, on average, older than our first sample from the online survey, i.e., closer to a societal average ($M = 39$, $SD = 20$, $min = 18$, $max = 82$). Nine participants were full-time employed, eight were students, two were retired, and one currently underwent training. Their affinity for technology interaction is slightly below average ($M = 3.55$, $SD = 1.04$, cf. the classification of Franke et al. [151]). Regarding the questions on perceived information privacy, our participants rated their Awareness on average with 6.17 ($SD = .95$), Control with 5.32 ($SD = 1.15$), and Collection with 5.38 ($SD = 1.15$) (higher scores correspond to higher privacy). Participants indicated medium general concern about their smartphone collecting their information ($M = 63.60$, $SD = 27.45$), a little higher concerns about data being shared with second parties (i.e., the device manufacturer or operating system developers) ($M = 70.0$, $SD = 20.81$), and rather high concerns on data being shared with third parties ($M = 80.95$, $SD = 20.15$).

4.3.4 Results

We mapped the interview citations to our privacy concern model codes developed based on the survey results (cf. Figure 4.10). The interview data fits well with our model and we could rediscover all codes. Additionally, we added the new code *The User* to the code group *Underlying Cause*, and *Circumstances* to *Mitigation Measures*. *The User* entails statements of users seeing themselves as the cause of privacy issues, a pattern we did not find in the online survey. *Circumstances* include mitigating factors that are passive factors instead of actively performed measures.

4.3.4.1 Underlying Causes (RQ2c)

Regarding underlying causes of their privacy concerns and present privacy issues, interview participants majorly mentioned topics that affect the current privacy information mechanisms (code group *Inaccurate Privacy Policy*) and issues that they see among themselves (code group *The User*).

Inaccurate Privacy policies. Users mention that they would like to know more about what happens with their data, but the given information mechanisms make it hard for them. Our participants especially criticized that privacy policies are too long and hardly understandable. Moreover, P13 explained that they *"tried to read privacy statements, but it's a lot of text."* Even when users overcome the issue of time, they are not satisfied, as P18 describes in their experience: *"Sometimes, I also read the explanation, but all the conditions are unclear."* P10 even accuses companies of deliberately hiding details, saying *"that is of course somewhere already intentional that you make it so complicated [...]"*. P9 formulates precisely what they would prefer, namely *"not such a mega long text, but one that is so probably presented in bullet points or so"* aiming for knowledge about *"how this data is processed and whether it is passed on to third parties."* Many participants admit that they usually do not really read but blindly accept privacy policies.

The User Our participants saw the most common causes of privacy issues among themselves. Due to the current weak informed consent mechanisms, participants expressed unknowingness 22 times. In 19 quotes, they describe that for them, *comfort outweigh concerns*, resulting in the user being the cause for privacy issues. Eleven quotes even indicate that users reached *resignation* on the topic of privacy. They lack a general feeling and understanding of what happens with their data behind the scenes: *"These Internet giants, which I can't assess at all, and which are like a black hole for me, and where I do not know at all what they are doing with it and what they are capable of when it really matters"* (P5). Besides what happens with their data, participant *"do not know how many years that the data will be stored"* (P9), and wonder *"to what extent that then saves in the long term"* (P3). What kind of data gets logged and processed by default, and to what users agree simply by purchasing a device, are also unclear. Furthermore, the reasons for data usage are often unclear, meaning that data logging often lacks a justification, as P9 explains: *"I can not understand that [why location is requested] and I have no idea why the location is then requested and therefore I click on deny because I have no idea"* (P9). Our interviews indicate that this lack of justification leads to both concerns and an increased tendency to deny data access. In general, we found a perceived lack of control. P19, for example, states that *"I do not think we have any influence at all, because we do not know what's in the technology"*.

However, people do not feel alone with these issues and do not see themselves as the problem; rather, they think that most people face similar problems, as P16 describes: *"I do not think I'm the only one who knows so little about it"*.

Security Issues We aggregated the mentioned security issues into the code groups *Lacking App Security*, *Careless App Security*, and *App Independent Security Issues*. Participants mentioned only very few technical security issues with apps, and if so they were vague. They believe that the risk of hackers accessing resources can hardly be ruled out. Besides these few technical concerns, participants mentioned many soft causes involving app-providing companies. They often criticize that they are forced to disclose their data to use a service, making them feel powerless: *"So I do not have the feeling that I have any influence on it"* (P1). Moreover, a lack of trust in the companies is omnipresent: *"I do not think that this is one hundred percent certainty, that only data that you agree to be used is actually collected"* (P14).

4.3.4.2 Privacy Issues (RQ2c-d)

Illiteracy As a result of the previously observed dissatisfaction with privacy information, illiteracy is a central theme that we found among the participant's privacy issues, e.g., P7 stating *"sometimes it's not quite clear to me exactly which data is being provided."* The lack of understanding evokes skepticism. Our interviews show that people would be less concerned if they were better informed about what happens with their data.

Third Party Data Sharing Concerns arise, especially around whom data is shared with and for what purposes. Disclosure to third parties is the most mentioned privacy issue. *"There is the discomfort of not knowing how it will be used and, above all, to whom it will be passed on"* (P19). People believe that companies might sell their data to others and are even afraid that their data will become publicly available. Besides deliberate disclosure by companies, our participants also believe in data getting stolen frequently. Such concerns were mentioned half as many times as the aforementioned concerns: *"There are so many, these data leaks, these bank data leaks or PayPal data leaks"* (P4).

Misuse Participants mentioned concerns about misuse, especially the creation of profiles: *"I actually do not want user profiles to be created about me and all the data, so all this data ends up coming together in a user profile"* (P3). With such user profiles *"it may also be possible to read off interests, attitudes, and the like. I would subsume that under the term personality profile."* Thereby participants are especially afraid of technology's ability to reveal information that is uncomfortable or not even known by themselves. For example, P8 envisioned a case of the smartphone being aware of one's pregnancy earlier than the woman herself: *"There is a teenage girl who was pregnant, but she did not know that she was pregnant and googled her symptoms. And then some company sent her a sample pack of Pampers."*

Acceptance of Data Logging The acceptance for data usage is generally quite low, as users believe data is logged without valid reasons. Overall, they dislike data getting logged: *"On the whole, I find it generally bad that data are collected"* (P12). Some participants think beyond themselves and complain that information on other people is also logged without their consent. *"I do not think that's okay because what can my friends do if I agree and then their data is passed on?"* (P16).

4.3.4.3 Consequences (RQ2d)

After looking at privacy issues and their underlying causes, we were interested in what *consequences* users actually are afraid of. Consequences can be real-life events that might happen as an effect of privacy issues or outcomes in the digital world that affect the user. As major topics concerning the user, we mostly found real-world consequences like *financial loss* and *influencing beliefs*. The participants further mentioned criminal activities like *Theft*, *Physical Harm*, and *Shared Information*. Furthermore, *Emotional Damage* and *Data Loss* were mentioned.

Financial Loss Participants are afraid of *"that your bank account is emptied"* (P12) and *"that you just get bills that you have to pay because you ordered goods but did not get them"* (P12). Furthermore, participants envision that activity and health data could have an influence on their credit-worthiness and insurance rate. While users find direct monetary loss to be likely through hacking, they expect that insurance companies would rather buy data from app companies to fuel their risk assessment.

Manipulation Our participants are aware of procedures that make use of user data to *influence beliefs* and manipulate the behaviors of their users. They see risks in derived profiles being used to subconsciously influence one's beliefs, especially regarding political opinions, shopping behavior, and increased device usage: *"Manipulations happen in people unconsciously. And I think this is a very difficult and dangerous topic for our society"* (P7). They mention political manipulation more often than other factors like shopping, and stress that the impact is more severe: *"I find it difficult when I am manipulated in a way to vote for a party that wants to come to power. I find that has a very different effect than if I buy the top from the brand because it was advertised to me"* (P7). P8 generalizes this to the issue of filter bubbles: *"What's problematic is when you then use that to reinforce some opinions, or spread certain content more. So I do not know, for example, if you think about the US elections or all these fake news scandals, that you get the same information over and over again instead of somehow getting a broader picture of it."* On the bigger picture, P3 mentioned the concern of *"a change in the political landscape."*

Criminal Activities Participants mentioned many criminal activities they believed to be enabled by privacy issues. For instance, having one's location data *"they see exactly when you are not at home, [and] could take advantage of that to break into your house"* (P17). Identity theft is an issue that some participants came up with, envisioned for various use cases. Beyond buying things on one's behalf, which is closely related to the aforementioned issue of financial loss, impersonation was brought up: *"They naturally use your data to take your entire identity and then use it to either go shopping or impersonate you"* (P12). P10 is afraid of being dragged into criminal activities by their accounts being misused: *"you are pulled into some criminal stories and have no idea at all about it, because your email was tapped."* P7 even envisions that *" my whole identity could be erased and someone else could take my identity. Now, to put it bluntly, if there are photos, if my whole character can be recreated, if you want to have me away from society, you can achieve that through that."* Further criminal points of attack are fraud via telephone or postal mail. Forged documents could be created by using signatures and photos. As a less sophisticated, but nevertheless annoying

consequence, participants mentioned *spam and advertisement*: *"Consequences are simply that you are spammed too much by companies"* (P9). Spam can result in fraud by requesting payments or asking to enter login credentials.

Shared Information Participants mentioned the vague concern of being spied out: *"And so there can even be spy features in there that we do not know about, that we do not even know what the purpose is of collecting and evaluating this message somewhere"* (P19). This is especially concerning as it may happen in the background without the user's awareness. *"That's just this paranoia that you always have a little bit. I have [camera and microphone] disabled, but maybe it still works somehow, maybe that's running in the background"* (P20).

4.3.4.4 Actors, Actions, and Datatypes (RQ2d)

Actors In contrast to the online survey, participants mentioned criminal third-party actors more often than non-criminal third-party actors. Participants are especially afraid of hackers gaining illegal access. Second-party actors, such as the device or OS developer company or big tech companies were only mentioned a couple of times. Governmental organizations, such as governments, parties, or the police, were mentioned often. Not surprisingly, the first-party company, i.e., the developing company of an app or platform, sometimes raises concerns in our participants.

Actions Participants expressed most concerns during real-world actions and active data entry. Having their phone with them yielded concerns during arbitrary activities. Situations where people have to enter data into an application manually also raise concerns about the usage of services. Here, participants brought up examples about photos, contact management, and search. The lifecycle steps of an app, i.e., installing it and granting permissions, were mentioned only sporadically.

Datatypes In the interviews, participants often mentioned *Personal Information* as a matter of concern. Many stayed rather general and did not specify this further, while concrete aspects like interests, contact information, and name and phone number were

mentioned. Among behavioral data, everything that is related to one's location was a big matter of concern. More device-related data like files, contents, and communication details were also mentioned but did not stand out.

4.3.4.5 Solutions to Mitigate Concerns (RQ2e)

User Behavior Most mentioned mitigation measures evolve around *user behavior*, i.e., actions that users themselves can and should do. This mostly includes active non-disclosure of data, for example by denying apps' permission requests, disabling data sources in the device settings (e.g., disabling location or even being in flight mode), leaving input fields empty where possible, or otherwise entering falsified data, as P2 describes: *"by disclosing as little data as possible, so that if I have to give personal data somewhere, I just give a false date of birth, a false name and so on."* As it is not always possible to use a service without giving data, participants mention the non-use of services or devices as their last option to protect their data. Some participants also described that they do some tasks only on their laptops instead of on their smartphones: *"When I log in to Paypal, it is not on my cell phone most of the time, but rather on my laptop"* (P7). Others describe that they leave their smartphone at home at some times which makes them feel less surveilled, or do not use some kinds of devices at all. Especially smart home devices: *"So that's why I do not have Alexa in the house either"* (P17). Some participants admit that better password management or using separate email addresses would be beneficial. They suggest making use of data deletion options more often, for example, clearing cookies, and histories, or actively requesting the deletion of their collected data with app companies.

Transparency The two most frequently mentioned concern mitigations were (1) transparency on the data usage and (2) using data for a purpose that users find beneficial. The latter is for example *"(market-)research"* (P12, P8), *"location for location-based services"* (P3, P9), and in general use cases where users see a direct benefit for them by granting access to data. They admit that *"a lot of data is also necessary for the services that are offered there"* (P13). Continuing closely on the results of bad transparency through privacy policies (cf. Section 4.3.4.1), users would like to know what data is logged, would like to get insights into created profiles, and wonder for how long data is stored. What is happening to their data is also interesting, i.e., how it is

processed and whether it is passed on to third parties. They lack knowledge of reasons for data processing and possible effects. Participants mention many suggestions on how these information needs could be satisfied: Information on data usage should be *"as concise and informative as possible and as understandable as possible"* (P5). To improve current consent mechanisms, P13 suggests visual means to convey what happens: *"Sometimes, it would be nice if there was a bit of a simple overview. Somehow no idea, little pictures, little pictures that tell me what happens with my data. I would find that practical."*

Regulations An also frequently mentioned suggestion to improve the users' privacy situation are *regulations*. While P13 formulates it rather soft *"there should be standards that encourage companies to handle data responsibly and not sell it to third parties"*, P15 sees a chance in strict laws: *"if there were clearer laws about how those could be used, if I knew, okay, they can use them, but if they're used past a certain point, then it's sort of illegal."* To avoid loopholes, regulations would ideally be on a supra-national level. P16 believes that *"it would, of course, be great if something like this were regulated throughout the EU [...] because that would be more effective, I imagine."* Participants thereby emphasized that *"[governmental agencies] then have to monitor it accordingly"* (P5). As users do not trust governmental organizations to stick to such rules and control themselves, *"then there might have to be external control systems to control that"* (P5).

Company Behavior and Technical Security Fewer participants envision that behavior changes by the companies and the application of technical security measures could mitigate their privacy concerns. Wishes include simply collecting less data (P5, P7, P2), deleting data as soon as possible (P2, P4, P7), and not selling or disclosing data to others (P4, P6, P10, P13). An idea of P19 includes managing data in a similar way as it is done with artistic content. They propose that *"when data is resold, I think it's right that, as in the case of copyright, you should also be involved, so to speak. And not to get rich, but simply to limit the whole thing a bit"*. Purely technical security measures such as encryption (P13, P17), two-factor-authentication (P18, P20) or very general *"more security"* (P12) were mentioned only sporadically.

Circumstances Also, sporadically, some participants said that their current living circumstances contribute to mitigating concerns. P1 mentions being less concerned *"because I have not yet had any very bad experiences,"* and P2, P3, and P20 mention that they are less concerned due to our solid political system. Also, a couple of participants believe that *"on an individual level, no one bothers to get our personal data, my personal data, because there is simply too little to get"* (P11) what contributes to mitigated concerns.

4.4 Discussion: The State of Mobile Sensing Privacy

With these findings we go beyond existing literature, by investigating smartphone users' privacy concerns in-depth rather than merely finding that privacy concerns are an important issue. We summarize the implications of our studies, and map them to this chapter's research question, in the following.

4.4.1 RQ2a: The Ratio of Privacy and Benefits Mainly Decides App Adoption Behavior

Our ranking of the importance of factors for app adoption (see Section 4.1.2.1) shows that the ratio of privacy and benefit matters most for potential users. The top 4 rated factors to app adoption all relate to privacy and benefits. In general, that aligns with past work where related constructs are studied, such as the privacy calculus (e.g., [148, 216]). To gain a deeper understanding of the two aspects *privacy* and *benefits*, we let users rank them more specifically in Section 4.1.2.2, to go beyond the insights of existing literature.

4.4.2 RQ2b: Users Are Most Concerned About Leakage of Contentful Data and Actions on Their Behalves

Identity- and Financial Theft *Wallet information* and *account information* rank highest on our scale of perceived sensitivity and potential risk. These are not of an informational nature that reveals something about their user, but rather can give third parties access to one's resources and may enable identity theft. This is in line with *send text messages*

being rated as the most concerning writing-datatype, which also depicts some sort of identity theft. We conclude that users are generally most afraid of outcomes that affect their financial status and online identity.

Misuse of Contentful Datatypes Besides the identity-related datatypes, we see data types that are contentful, i.e. user-centric information that may contain private topics, rank high. Namely, these are *text messages and calls*, *microphone data*, *files and media*, *camera data* and *screen contents*. Within these also occur *location data* and *contacts*, which, however, are rather an observed property of the user respectively actively entered collection of information. The actual information value contained therein is diverse and can range from rather worthless ambient noise to personal private conversations.

4.4.2.1 Productive Benefits Over Leisure Stuff

In the ranking of the effect that several benefits that an app provides have on the app adoption intention, we see that productive and use-oriented benefits make a stronger impact on app adoption than fun and leisure-oriented apps. *A productive daily life* as app purpose even surpases *a monetary incentive*. The rather abstract purpose *amusement* is the only non-productivity benefit that was rated to the upper half of the spectrum by our participants.

4.4.3 RQ2c: Users Are Knowledgeable in General but Lack Information About Concrete Apps

Summarizing the online survey and interviews, we found that the major issue regarding knowledge about smartphone privacy is a lack of information about the data practices of apps. Users' general knowledge and understanding of how smartphones and privacy-enhancing technologies work are rather good in our sample. 90% of our participants could answer most of our quiz items correctly. However, it has to be noted that the sample of our survey was rather young and reported an above-average ATI score.

Regarding concrete apps, participants mentioned that they are often not clearly aware of what data apps are logging and what happens with their data afterward. They are uncertain about what data apps can obtain from the device. Privacy policies are

criticized for not being understandable. In our quantitative assessment of engagement with privacy information mechanisms, we found that users behave ambivalent; while some people claim to engage a lot and try to inform themselves, others do not.

People's concerns were partially vague; they mentioned many uncertainties and things they were unsure about but "have heard about it." We see that this presence of uncertainties and the laborious nature of informing oneself about privacy leads to reduced motivation to do so. This negative feedback loop finally reduces trust and transparency and increases concerns. Thus, we conclude on RQ 2c that **the users' general knowledge of smartphones and privacy-enhancing technologies is good, but they lack information and understanding of apps' data practices**.

4.4.4 RQ2d: Users Fear Uncertain Data Incidents That Affect Their Real-World Lives

The quantitative and qualitative results of both studies agree that users are most concerned about third parties stealing and misusing their data. Users are well able to name concrete consequences and are rather afraid of events happening in the real world, such as financial loss and burglary. They also came up with non-obvious, far-away consequences like manipulating (political) beliefs based on created personality profiles. Moreover, passively sensed data (e.g., location, audio) thereby concerns them more than actively entered data (such as personal information entered in a form). Data processing on global, remote servers concerns them more than local.

The users' concerns evolve mostly around uncertainties and demand better transparency and control. We also found that users who know and understand more about privacy-enhancing technologies have greater concerns. This finding aligns with the literature [157] that found that users initially ignore privacy but show growing concerns when they become aware of the possible consequences. Furthermore, our results also reflect the findings of Coopamootoo and Groß [97], who argue that responses are heavily biased by the cognitive processes that are triggered by the assessment methodology.

We conclude on RQ2d: **Smartphone users are concerned about their data being misused by unknown third parties, and being used against them with negative implications on their lives.**

4.4.5 RQ2e: Mitigation of Privacy Concerns: User-Centered Privacy Measures

In **RQ2e**, we probed users on how they think their concerns can be mitigated. Our participants suggested various measures that they can take themselves to overcome privacy risks, such as changing their own behavior or measures they can easily understand. Technical measures, such as improving app security, encryption, and stronger authentication, were sporadically mentioned. We see that user-centered measures are more relevant than technical ones – thus educating users. As users cannot fully understand technical security measures and do not have proof of effect, the yielded trust benefit in an app is limited. Furthermore, users' views, concerns, and expectations have to be incorporated more [21]. **Thus, this is a clear call to incorporate user-centered privacy in apps, to increase users' trust in the apps they use.**

4.4.6 Lack of Trust: Call for Regulations

Transparent information mechanisms and control features require trust in the information being true. Respectively, control features affect what they are claiming. However, in the interviews, we noticed that participants were skeptical about whether companies are actually honest. Participants accused companies of incompleteness of provided information, not implementing what they promised, and a general reluctance to act for the users' good. Hence, a major group of suggested mitigation measures are regulations, as a way to force the companies and guarantee user-friendly behavior. In our first survey on app adoption intention we saw huge differences between different types of app publishers, with companies scoring worst. Thus, (tech) companies should **work on their reputation and find ways to guarantee privacy**. Role models from the technical perspective could be approaches like differential privacy (cf. Zhao and Chen [453]) or digital signature approaches.

4.4.7 Users Do Hard Implying Privacy Risks of Abstract Data Types

From the literature, we know that users do hard estimating which high-level information about them can be inferred from low-level data, such as raw sensor values [179, 235]. Our study shows that this also applies to less abstract data types: Interestingly, users rated potential risk and sensitivity of *screen contents* only on rank 9, and thus lower as,

e.g., *text messages* or *files and media*. This shows that, although it might be logical for most users that screen content can also contain textual content, it is perceived as less sensitive at first sight, as the informational content is less abstract.

4.4.8 A Lack of Appropriate Privacy Enhancing Technologies throttles Mobile Sensing Applications

Our studies reported that privacy issues are a major barrier to the adoption of mobile sensing smartphone apps. Especially, deep, contentful data was rated critically, as users can hardly estimate its information gain and what can potentially inferred from it. Transparency and control are important factors that can mitigate concerns, however current privacy enhancing mechanisms can hardly deal with complex, contentful datatypes.

A lack of transparency is the major overarching topic that goes through all three studies. It came up in all code groups, respectively, stages of our privacy concern model. Also, Transparency was indicated an increasing app adoption intention in the first survey. Uncertainty and a lack of information, knowledge, and understanding constitute the cause of many privacy issues and feared consequences.

Strongly tied to the aforementioned lack of transparency is also the desire for more control. It is not as present in the users' minds yet as transparency is, but when digging deeper into interview responses, we saw that they feel like they are not the owners of their data anymore. The high relevance self-centered privacy issue mitigation measures that users have reported further underlines the lack of control. Users see themselves forced to restrict their own behavior, as last measure of impact.

Users are especially interested in what their data is used for, and would like to control it also after data left their devices. Both desires of Transparency and Control are hard to fulfill in the context of contentful data, the application potential of which we have envisioned in Chapter 3. This contradiction currently reflects by smartphone OS developers heavily restricting access to such data, which in turn throttles novel innovative mobile sensing data based applications.

4.4.9 Motivation of Novel User-Centered Privacy-Enhancing Technologies

During our research we found that current privacy-enhancing technologies, i.e. information and consent mechanisms, are not satisfying. Based on the demands from Chapter 3 and the privacy issues found in Chapter 4 we motivate the following requirements for the design of novel privacy-enhancing technologies.

Solutions for Contentful Data Current privacy consent mechanisms are rather limited in their options. Usually, one can only completely grant or deny access to a datatype. Thereby all datatypes are treated similarly, without considering the inherent differences that they pose. For example, for location, different granularities could be offered, or for microphone data, the user's context may be considered. The span of inherent information is wide, and with nowadays all-or-nothing permission concepts such data cannot be used in a privacy-respectful way. Our study thereby motivates the design and evaluation of novel privacy-enhancing technologies, especially for contentful datatypes. Not only app developers but also users would benefit: Besides gaining increased privacy and mitigating risks, they would benefit from more adaptive and intelligent apps fueled by rich, detailed data.

User-Centered Privacy Privacy-enhancing technologies should keep the user in the loop. Sole technical security measures do not make a privacy difference from the user perspective. Users has to be conveyed the feeling of understanding what happens and, in case they feel the need, the ability to control processes related to their data anytime. Most users are more satisfied when they understand what happens (although many users initially cannot express the demand), and regarding we have a clear lack. Users have to be kept in the loop of their data processing, from logging until (remote) processing.

Explain Data Inference Potential Related work and also our study results have shown that users do especially hard estimating the privacy implications that data inferences may have. Approaches have to be developed that convey what kind of hidden information lies in their data. This is challenging as one by nature cannot know what is or will in the future be possible. Furthermore, knowledge discovery processes

mostly happen offline and with an individual's data being merged into a larger dataset encompassing many users. Thereby the processes become more complex to explain and are out of the scope of control of the user.

Design for Different Audiences We saw that the spectrum of how much people care about privacy is somewhat polarized. While some people really care and can be hardly satisfied, others do not care at all and are not willing to spend time and effort into privacy-related tasks. Future research should consider designing differently for these two audiences.

Transparency Thus, the most desired mitigation measures evolve around increasing transparency: Better consent mechanisms and precisely short and concise information that highlights the essential aspects were mentioned by nearly all participants. It has to be easy to understand and could be augmented with visual elements that help to understand the flow of one's data. Advantageous would also be to make the aim and purpose of data processing clear. Although we saw a generally low acceptance for data logging and processing, our interviews have shown that people mostly agree with data usage if the purpose is clear and, in their eyes, meaningful. Thus, we conclude that **system designers and developers should put an emphasis on concise, transparent, and understandable information and consent mechanisms**.

Control However, having direct control over one's data from logging to processing poses an effective concern mitigation measure. We saw that most mentioned mitigation measures evolve around things that the users themselves perform, and rather few mitigation measures the data processing opponent or other third parties. We argue that **users should be offered full control over their data and what happens with it at all times**. reduces the perceived concerns.

Not-In-Situ Solutions Privacy-decisions are usually made in-situ, i.e. permission requests are prompted in that moment when an app actually needs access to some data. On the one hand contextualizing privacy decisions has shown to be good (e.g.,

[433] in the context of web consent forms), on the other hand users in-situ have a aim or task they want to fulfill, which decrease their motivation to cope with privacy decisions right in that moment.

4.5 Chapter Conclusion: The State of Mobile Sensing Privacy

To shed light on the concrete outcomes that users are afraid of happening due to privacy issues, and to see which factors play a role in mobile sensing app adoption decisions, we conducted two online surveys and one interview study. We overall found that users are concerned of outcomes that affect their real-world lives. That are especially stolen wallet- and account information which may lead to financial loss or identity theft. Rich datatypes such as text messages and microphone data are deemed dangerous due to the large and varying amount of content. Concerns are mitigated if users expect a benefit alongside the data, especially benefits towards their productivity. In general, users are initially unaware of many issues, until they deal with the topic of data privacy. For rather abstract datatypes they do hard estimating the contained information gain, and they can hardly judge which further information prediction models might infer from their data. They see the underlying issue mostly in data shared with or stolen by third-party actors such as companies or hackers.

With these findings we go beyond existing literature, by investigating smartphone users' privacy concerns in-depth rather than merely finding that privacy concerns are an important issue. Privacy-enhancing technologies are needed, that tackle these issues. Especially more user-centered privacy design, control throughout the full data processing pipeline, and solutions to convey trust in technical privacy and security measures are necessary.

Chapter Take Away

Current privacy-enhancing technologies are not sufficiently user-centered. Measures that improve privacy and security from a technical perspective do not necessarily also affect the user's privacy perception. Rich and contentful data, processed through machine learning, poses special challenges. Users do hard estimating the information inherited in such data and understanding prediction procedures. Novel approaches to provide transparency and control to users need to be researched.

5

How Can We Improve User Privacy, Without Obstructing the Data's Output?

In this chapter, we propose **solutions to improve the smartphone users' privacy** regarding mobile sensing applications, while **keeping the data usable** for the application's purpose. Building the basis for most privacy-enhancing interfaces, we first study the **effects of transparency and control** in smartphone apps on users' privacy perception and application adoption behavior. We present **four approaches towards improved user privacy**, three of them were evaluated in-the-wild yielding insights on their real-world effects.

In Chapter 3, we have seen that there is a need for more in-depth sensing data, in order to enable innovative use cases. However, Chapter 4 has shown that privacy issues throttle the adoption of smartphone apps with mobile sensing technology. They raise privacy concerns among the data generating user, and raise concerns among OS developers about their users security and privacy. The underlying problem is that

current privacy-enhancing technologies, i.e. information and consent procedures and features, are not sufficient. The high information gain of mobile sensing data is opposed by a lack of transparency and control. The concerns on both user and developer side lead to lower app adoption and availability of mobile sensing data. Thereby, fewer innovative applications proliferate in the wild. In this chapter, we first study the influence that transparency and control have on users' app adoption decision. Based on the main insight that transparency always should be accompanied by control, we study three concepts that provide transparency and control over mobile sensing data. In Section 5.2 we investigate the potential of on-device pre-processing and log data review in the context of mobile language data. Section 5.3 presents a concept providing fine-grained permission control on smartphones through sliders. Through an interactive machine learning interaction concept that we present in Section 5.4 we finally propose solutions regarding prediction models and data inference. With the presented concepts, we show that well-designed transparency and control features can facilitate the proliferation of mobile sensing data. The user-desired transparency helps them being informed and understand what happens, however to actually mitigate privacy concerns and increase app adoption the availability of control features is key. How both affect the users is thereby highly individual, and varies across users. Overall, this chapter's studies adhere to our third research question:

RQ3 *How to improve privacy without obstructing usability?*

	Research Question	Paper	Section
RQ3	How to improve privacy, without obstructing usability?		Chapter 5
RQ3a	How do transparency and control in a privacy dashboard affect the number of users adopting and dropping out of a passive mobile sensing app?	[47]	Section 5.1
RQ3b	How do transparency and control in a privacy dashboard affect the awareness of and knowledge about the data logging?	[47]	Section 5.1
RQ3c	How do transparency and control in a privacy dashboard induce behavior change and self-reflection and thus the logged data of a passive mobile sensing app?	[47]	Section 5.1
RQ3d	How do transparency and control in a privacy dashboard affect a passive mobile sensing system user's privacy concerns and trust?	[47]	Section 5.1
RQ3e	How does on-device preprocessing of mobile language data affect users privacy perception?	[40]	Section 5.2
RQ3f	What are helpful sub-steps for fine-grained data control?	[46]	Section 5.3
RQ3g	How do we deliver the additional control that is usable?	[46]	Section 5.3
RQ3h	How does it perform compared to the existing Android permission UI?	[46]	Section 5.3
RQ3i	How can interactive machine learning yield transparency regarding prediction and inference potential of mobile sensing data?	[38]	Section 5.4

Table 5.1 : Overview of the studied sub research questions of RQ3.

5.1 The Influence of Transparency and Control Features

This section is based on the following publication:

Florian Bemmann, Maximiliane Windl, Jonas Erbe, Sven Mayer, and Heinrich Hussmann. "The Influence of Transparency and Control on the Willingness of Data Sharing in Adaptive Mobile Apps." In: *Proc. ACM Hum.-Comput. Interact.* 6.MHCI (2022). DOI: 10.1145/3546724

Previous research proposed the concept of "consent as a process" to tackle privacy issues, including making data logging processes transparent and giving the users control over their data [179]. In our studies presented in Chapter 4, we also found a demand for transparency and control; although transparency was more often requested than control. However, the effects of transparency and control are rarely studied yet.

In the context of privacy dashboards [187], studies have been conducted with diverging results. While some indicate positive results [397], other studies revealed none [213] or even contradictory effects [203]. A promising direction is supplementing transparency with control which has already been shown to mitigate the adverse side effects of transparency [141, 187]. For example, increased transparency decreased

trust and willingness to share data [206, 343]. However, those studies were either conducted in the domain of actively donated data [187], conducted as vignette studies [213, 221] or did not evaluate transparency and control independently [213, 397]. Thus, developers of mobile sensing applications in industry and research cannot build on insights on the effects of transparency and control features. Regarding privacy dashboards as a means to implement transparency and control, the literature showed that a well-informed design is needed to avoid adverse effects [141].

In this study, we investigated the effects of transparency and control features on smartphone users app adoption decision and perceived privacy. On the example of a privacy dashboard, we quantified conversion rates depending on transparency and control as two individual factors.

The findings of this study help us to steer the design of future privacy-enhancing technologies. We find which features actually make a difference regarding app adoption, and which aspects have to be considered with care. Thereby these results constitute the foundation for the studies presented in the following.

5.1.1 Background: Transparency and Control Through Privacy Dashboards

In this section, we introduce privacy dashboards as a means to tackle the privacy issues and implement the privacy tool suggestions derived from the previous section. We first define privacy dashboards, describe existing privacy dashboard projects, and finally give an overview of work on transparency and control in privacy dashboards.

The privacy dashboard is a common privacy design pattern [121]. Other privacy design patterns are, for example, the *Personal Data Table* and *Privacy Policy Icons* [372]. Privacy dashboards make users aware of the data services have collected about them. They should provide successive summaries of the collected data and give an easily understandable overview [121, 457]. For this, they can use demonstrative examples, predictive models, visualizations, or statistics. Additionally, a privacy dashboard can provide control options and privacy settings to empower users to control the processing and collection of future data, cf. [121, 457]. Especially actions like deletion and correction of data are highlighted. Finally, privacy dashboards should give an overview rather than presenting every detail of possibly thousands of data items [372].

Privacy dashboards are spreading in practice, for example, the Google Privacy Dashboard, and have become subject to research. They were examined as a European General Data Protection Regulation (GDPR), compliant alternative to consent forms [54]. They were studied as tools to give users a sense of what data is collected and inform the user instead of listing every detail [20, 213]. Raschke et al. [329] implemented a comprehensive privacy dashboard that adapts the newsfeed concept from social media. The dashboard incorporates transparency and control features so that users can view the collected data and learn about the purpose by obtaining information about involved processors, requesting rectification or erasure of each data item in the timeline, or reviewing and withdrawing the consent for each individual data type.

Privacy dashboards are a tool to implement the principles of *notice and choice* [346, 438], often through features that provide transparency and control [353]. Transparency and control have long been studied in the context of mobile systems. Permission popups force mobile apps to provide transparency and control about what data an app can access [144]. However, the context is limited [410]. Permission popups lack appropriate information and contain hardly understandable terms, making it hard for users to grasp the implications of granting permission [217]. Also, the amount of information conveyed to the user leaves space for improvement [53, 144]. In contrast, interfaces that provide more detailed information on what happens with the data increase user confidence [402]. In the web context, similar issues are proliferated. Privacy policies are long and hard to understand and, thus, often ignored [294]. In addition, they fail to provide sufficient transparency to the user [53]. Here, consent popups may even be designed to nudge users towards illegal configurations [293].

Permission popups offer transparency and control before the data logging happens. In contrast, privacy dashboards take effect afterward. The retrospective approach has the advantage that the user can be informed about what has actually been logged. Transparency and control features, incorporated through privacy dashboards, have shown positive effects: The Google privacy dashboard [141] and a dashboard for online shopping [187] increased user trust. However, this is only valid for raw data: In the study of Herder and van Maaren [187], showing derived data increased perceived privacy risk and reduced user trust. Perceived risks and trust may lead to fewer people using a service, not sharing required data, or dropping out early. For example, in vignette studies, participants indicated to prefer using a service that provides transparency

Paper	T	C	Finding	Context	Method
Keusch 2019 [221]	✗	✓	An option to switch off data collection would significantly increase the willingness to participate	mobile sensing studies	vignette study
Tsai 2011 [397]	✓	✗	Transparent privacy information policies increase the usage of the service and even justify higher prices	online purchasing	between-subject lab experiment
Elevelt 2019 [134]	✗	✓	74% of the participants continued sharing their GPS data although they could have opted out	smartphone sensing	Longitudinal diary study with sensing (enabled by default, opt out possible)
Awad 2006 [20]	✓	✗	Acceptance of tracking and effects of transparency vary with personality	personalization in online shops	survey
Karwatzki 2017 [213]	✓	✗	Transparency has no positive effect on willingness to disclose information	personalization of event finding online service	evaluation of screenshots in 2x2 design
Farke 2021 [141]	✓	✓	Transparency reduces concerns, control features are used rarely	Google Privacy Dashboard	interview after guided usage
Herder 2020 [187]	✓	✓	Control and transparency on raw data increase trust, while transparency on derived data decreased trust but increases concerns	online purchasing	vignette study with assessment of behavior intention
Schnorf 2014 [353]	✓	✓	Depending on the user predisposition, transparency can also raise anxiety	inferred user interests	survey
van Kleek 2017 [402]	✓	✗	Transparency on what happens with their data increases user confidence when deciding for an app	mobile app choice	prototype study

Table 5.2 : Studies on the effects of transparency (T) and control (C) in privacy dashboards, their context, methodology, and findings. Most studies were conducted with vignette or survey methodologies, while evaluations of real behavior are rare.

over the logged data [20, 221, 397] or an option to switch off the data collection [221]. However, while control features show a positive effect, they are only seldomly used. In the studies by Farke and Elevelt only a quarter of the participants involved had already used or indicated a willingness to use such features in the future [134, 141].

While the previously reported studies in the contexts of webshops, surveys, and personalization of online services agree that the provision of decision-relevant information is positive [456], the literature is contradictory in the context of sensing data [213]. Here, transparency increases privacy concerns resulting in less data disclosure [206].

This inverse impact of transparency features can also be found in studies on personalized advertisements [422] and inferred user interests [353]. Also, the reaction to transparency features depends on the user's privacy predisposition [353]. To give an overview, we show studies evaluating the effects of transparency and control with their context, methodology, and findings in Table 5.2.

5.1.2 Research Questions

Incorporating privacy-enhancing features that provide transparency (e.g., [20, 221, 397]) and control (e.g., [221]) have positively affected users' trust and privacy concerns. Furthermore, vignette studies indicated positive effects on the usage of such services [213, 221]. However, it is unclear whether those effects hold in a real application, especially in the light of the Privacy Paradox that indicates discrepancies between behavioral intentions and real-world behavior, cf. [5, 292]. Further, most research has been conducted in the context of online shops or personalized adaptive services and was fueled by data actively provided to the system by its user. In contrast, only a few works exist in mobile passive sensing applications, where data is collected without the user's active involvement. And if so, studies used vignette methodologies and only assessed user intention via self-reports instead of actual behavior. Yet, to the best of our knowledge, no study measured the effects on participation rates and app usage, the resulting data (i.e., gaps), the privacy concerns, and trust in mobile sensing apps equipped with a privacy dashboard while treating transparency and control features as two independent factors. To address this gap, we define the following research questions:

RQ3a *How do transparency and control in a privacy dashboard affect the number of users adopting and dropping out of a passive mobile sensing app?*

RQ3b *How do transparency and control in a privacy dashboard affect the awareness of and knowledge about the data logging?*

RQ3c *How do transparency and control in a privacy dashboard induce behavior change and self-reflection and thus the logged data of a passive mobile sensing app?*

RQ3d *How do transparency and control in a privacy dashboard affect a passive mobile sensing system user's privacy concerns and trust?*

We compared four privacy dashboard variants to investigate these research questions: Transparency and control, either transparency or control, and a baseline variant without both. This 2×2 factorial design allowed us to evaluate the effects of transparency and control independently. In addition, we deployed the privacy dashboard as part of a passive mobile sensing app in the wild ($N = 227$) to overcome the limitations of related work that often relied on vignette studies. In the beginning, we did not tell participants that the privacy dashboard was the study's primary objective to be able to measure natural, unbiased behavior. We started with a preliminary survey to operationalize transparency and control and decide which features to implement in the dashboard.

5.1.3 Preliminary Survey

We first assessed which transparency and control features are important to users to inform the development of our privacy dashboard. Then, we aimed to incorporate only essential features to not overwhelm the users. We conducted a survey ($N = 118$) to determine which transparency and control features are most important for users. Therefore, we presented our participants with a vignette of a hypothetical mobile sensing scenario and asked how likely they were to participate in that study. The design of the vignette study, including the questions, is adopted from Keusch et al. [221]. The participants rated a set of privacy dashboard transparency and control features on a 5-point Likert scale for their likeliness to increase their willingness to participate. We collected the presented features from related work by constructing a list of features these papers used to implement transparency or control in their privacy dashboards [221, 329]. We further asked them to select a minimal set of features that they would like to have in a hypothetical mobile sensing application. Finally, we incorporated all features desired by most of the users into our privacy dashboard design.

5.1.3.1 Participants

The sample consisted mainly of students (62%) and employees of IT companies (20%). Their mean age was 27 years ($SD = 8.4$). The following insights are based on 115

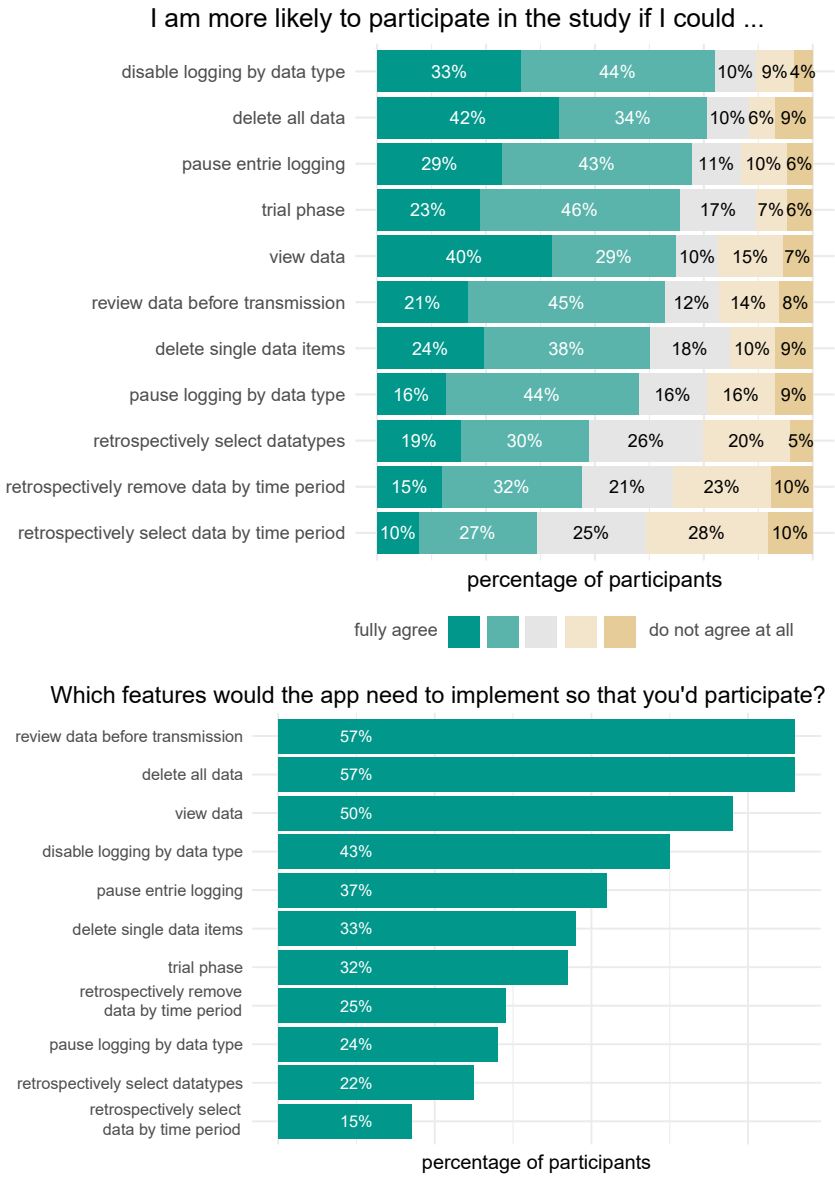


Figure 5.1 : Results of our preliminary survey to inform the privacy dashboard design. We implemented all features where the majority of the participants fully agreed or rather agreed that it would increase their likelihood to participate (a).

participants (3 participants stated in the first question that they would participate in the vignette study in any case, and thus we could not ask them about features to improve their participation likelihood).

5.1.3.2 Results

Detailed control features were the most desired. Disabling data logging ranked high in both questions, whereas fine-grained control (disabling single data types, 77% agreed) was more desired than disabling the full logging (72% agreed). They also desired the option to delete all data (76% agreed) and single data types (60% agreed). Further, the transparency features to view the recorded data (69% agreed), and a trial phase (70% agreed) were rated to increase participation likelihood. Participants rated retrospective features as less critical. Less than half of the participants stated that retrospectively deleting data of specific periods (47% agreed), exclusively allowing data of specific periods (37% agreed), or removing data of single data types (49% agreed) would increase their likelihood of participating.

5.1.4 Study

We conducted the study with the PhoneStudy app[1] and added a privacy dashboard to understand the impact of transparency and control. We evaluated the mobile sensing privacy dashboard in a 2×2 factorial study design, with the presence of (a) TRANSPARENCY and (b) CONTROL features as the independent variables. Thus, each participant was assigned to one of the following four conditions:

- ***Baseline***: The mobile sensing app without any possibility to view logged data or control the logging
- ***Transparency Features***: The app with a privacy dashboard that allows users to view the data but not delete or pause data entries (Figure 5.2 1-4)
- ***Control Features***: The app with control features where users can pause the data logging (Figure 5.2 5-7)
- ***Both***: The app with both a dashboard to view data and control features to pause logging and delete data entries (Figure 5.2 1-8).

[1] https://phonestudy.org/, last accessed 2024-12-06

This modular design allowed us to turn single features on and off depending on a user's study condition. Following our research questions, we can hence study the behavior of users who were exposed to one of the two factors independently and compare how our dependent variables behave. Hence, in contrast to studies from related work, we can evaluate the effects of transparency and control features individually.

5.1.4.1 Apparatus

The PhoneStudy app collects the sensing data, incorporates the privacy dashboard, and implements study management features such as prompting questionnaires when desired by the researcher. It was available for Android (version 6 or above) and was distributed as an APK file. We implemented the privacy dashboard as a web app built with the React JavaScript Framework to make it as reusable as possible for other applications. For this study, it was seamlessly integrated into the sensing app using an Android WebView. WebViews allow direct communication between its JavaScript environment and the native Android code[1], such as displaying and deleting data from within the dashboard. The control features to pause the data collection were implemented via a native Android screen. The sensing app communicated with a central server to control the study flow and gather the data. This also allowed the distribution of the questionnaires at specific study stages via push notifications. The questionnaires were implemented in the online survey platform SoSci Survey[2]. Thus, we could run our study completely remotely.

We made the following design decisions for the privacy dashboard: 1) We only show raw data but no aggregations and interpretations. Raw data is the first step in every data processing workflow and is thus present in every mobile sensing application. Furthermore, the effects of aggregations and interpretations are highly dependent on their precise design and implementation. Since we wanted to keep our results generalizable and independent of one specific application case, we forward such analyses to future work. 2) Informed by the results of the preliminary survey, we implemented all features, where the majority of the participants fully or rather agreed that they would increase their likelihood to participate. We list the resulting set of

[1] https://developer.android.com/reference/android/webkit/WebView, last accessed 2024-12-06
[2] https://www.soscisurvey.de, last accessed 2024-12-06

Feature	Implementation	*T*	*C*	*B*
Disable logging by data type	In the *Settings* screen **(5)** one can disable the logging via data type specific toggles.	✗	✓	✓
Delete all data	Is already implemented in a dedicated tab to fulfill the privacy regulations.	✓	✓	✓
Pause entire logging	In the *Settings* screen all toggles can be turned off, and a pause duration in hours can be specified **(6)**.	✗	✓	✓
Trial phase	The first data items are uploaded after 24 hours, thus one can uninstall the app within the first day without data being transmitted.	✓	✓	✓
View data	In the *View Data* tab a timeline visualizes all logged data items **(1)**. Additionally a detail view **(2)** which is accessible via a context menu **(8)**, raw data view **(3)**, and explanations **(4)** about what can be inferred from the data are provided. The timeline can be filtered by datatype and timerange **(9)**.	✓	✗	✓
Review data before transmission	Data is uploaded only after 24 hours, thus one has time to view and withdraw before transmission by deleting all data of the current filter selection **(7)** or single data items via an item's context menu **(8)**.	✗	✗	✓
Pause logging by data type	The toggles in the *Settings* screen allow to set a pause duration **(6)**.	✗	✓	✓

Table 5.3 : In this table we show the features derived from our preliminary survey and requirement analysis (left column) and matches it to how each feature is implemented in the dashboard (column *Implementation*). The three rightmost columns denote in which of the experimental conditions each feature is present: Transparent Features (T), Control Features (C), and Both (B).

features in Table 5.3. 3) The privacy dashboard has two main screens. A *view data* screen that mainly implements the *Transparency Features* and a *settings* screen that contains most control features. Figure 5.2 shows a visualization of the structure, and Table 5.3 explains how the agreed-on features are incorporated. 4) The timeline concept of the *view data* screen is informed by the privacy dashboard of Raschke et al. [329]. Each data item is listed in chronological order, beginning with the newest. In contrast to theirs, our dashboard is optimized for smartphone screens, i.e., controls to configure filters like time range, data type, etc., are hidden in menus. While deleting data is supported in the according study condition as well, we did not include features to rectify erroneous data. Multiple items of the same type in a row are collapsed (e.g., "9 more app usages") but can be expanded on demand. In general, all views follow Google's design standard *Material Design*[1].

[1] https://m3.material.io/, last accessed 2024-12-06

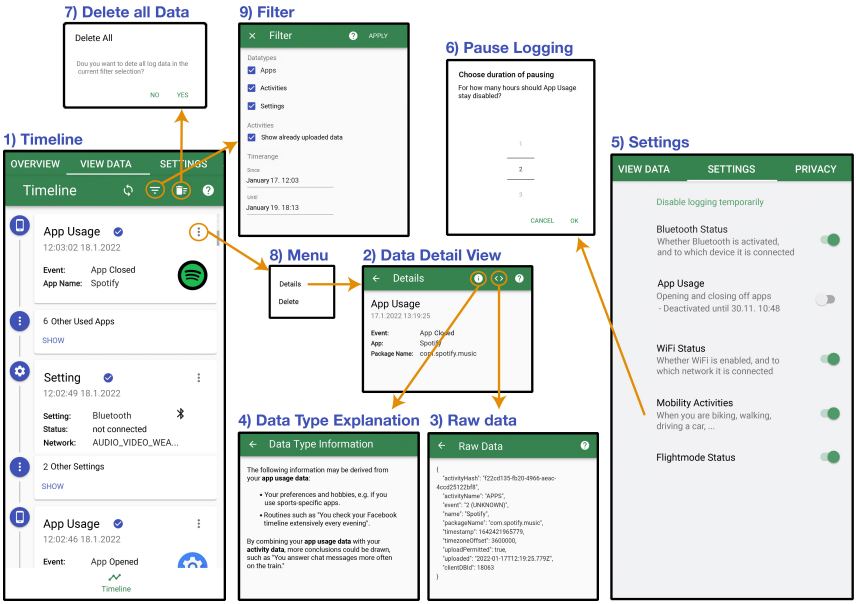


Figure 5.2 : The UI of our privacy dashboard is structured into two main components: The timeline view (1-4) that offers transparency, and the settings features (5-7) that implement control over the data logging.

5.1.4.2 Procedure

We used convenience sampling to recruit our participants (via email lists, social media, and Slack). The advertisement contained only a little information. We merely advertised it as a study on smartphone usage in daily life. For further information, the ad referred to an onboarding questionnaire to reduce the risk of a hidden selection bias by privacy disposition (DTVP) [213]. If people drop out in the onboarding questionnaire instead of the study ad, we could count them. When opening the onboarding questionnaire, users were randomly assigned to one of the four study conditions. Then the study details (e.g., mobile sensing app has to be installed, data is logged, study duration) were introduced to the potential participants. In addition, for the non-baseline condition, we advertised the respective privacy features prominently. Via this onboarding procedure, we could retrace how the transparency and control features already influenced the decision to install the mobile sensing app. Thus, monitoring the "interest in the study."

The participants could download the app via a QR Code or a link. In the non-baseline condition, the setup process started with an intro slider where the respective privacy features were again advertised. Afterward, participants had to walk through a four-step setup process to accept the app's privacy policy and grant the necessary system permissions. The app then summarized the study procedure. Finally, the app prompted a link to the pre-study questionnaire (see Section 5.1.4.3 for the instruments).

Participants should then keep the app on their phone for seven days while data was passively logged in the background. Our app logged smartphone behavior (i.e., opening and closing apps), connectivity status (i.e., wifi and Bluetooth status), and high-level activity data (like walking, biking, or running[1]). After two days, the app reminded the participants about the transparency and control features via a notification. After seven days, the app prompted the post-study questionnaire via a notification (see Paragraph 5.1.4.3 for the instruments). At the end of that questionnaire, participants chose their compensation, and the study was finished. We compensated participants for their participation with either 15€ via PayPal or a respective amount of study points that can be credited at our university[2]. We visualize the study procedure in Figure 5.3. According to the ethics approval procedures at our faculty, we assessed our study with the ethics committee's questionnaire. As a result, we concluded that it was not ethically questionable and forwarded the completed questionnaire to the ethics committee.

5.1.4.3 Measurements and Logs

The collected data consists of three parts: 1) a pre-study questionnaire in the app, 2) the app usage data, and 3) a post-study questionnaire. We chose this study design to observe changes in behavior and knowledge. Moreover, some measurements need to be collected before to get unbiased insights (e.g., the prior privacy experience), and others can only be collected after using the dashboard (e.g., behavior change and self-reflection).

[1] Retrieved via the Google Awareness API activity recognition, https://developers.google.com/awareness/android-api/snapshot-get-data#get_the_current_activity, last accessed 2024-12-06

[2] The participation in the study is still anonymous, and data required for compensation and study credits is kept independent of the study sensing data and questionnaire answers.

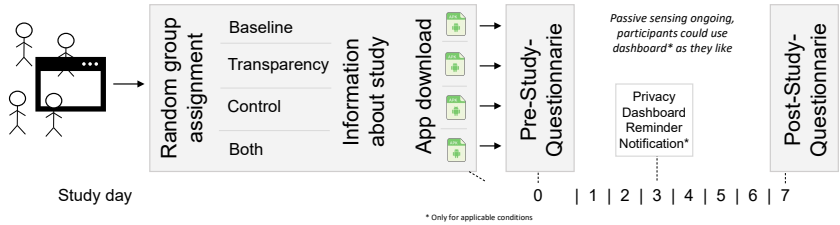


Figure 5.3 : A flowchart visualizing the procedure of our study. Potential participants were recruited with a sparse study description (i.e. not mentioning that a mobile sensing app is involved). When they clicked on the onboarding link which was realized with an online questionnaire tool, they were immediately randomly assigned to one of the four study conditions. Afterward, the full information about the study was presented, mentioning the privacy dashboard in the applicable conditions, and the condition-specific Android app could be downloaded. After the installation participants had to fill out the pre-study questionnaire, on day seven the post-study questionnaire.

Pre-Study Questionnaire After installing the app, the pre-study questionnaire was prompted via a notification which took approximately 5 minutes to complete. There we assessed participants' prior privacy experience and how good they were informed about what data was logged during the study and what happened with their data:

- Prior privacy experience: Contributing to RQ3d, we used the construct collection of Degirmenci et al. [110], which consists of adaptions of the items about prior privacy experience from Xu, Gupta et al. [441], computer anxiety by Stewart and Segars [387], perceived control by Xu, Teo et al. [442], and app permission concerns by Smith et al. [376].
- Knowledge about data logging: A set of self-constructed items constituting a data logging quiz. It consists of 12 statements about logging, e.g., " The PhoneStudy app is logging my precise location (GPS coordinates)." The participants had to rate each item whether it was true, false, or they did not know. We use this information to answer RQ3b.
- Knowledge of how data is processed: Also corresponding to RQ3b, we constructed another set of items constituting a data processing quiz. It consists of 10 statements on where the data is processed (e.g., "The collected data is never leaving my smartphone"), who has access (e.g., "My collected data is accessible

for everybody on the internet") or anonymization (e.g., "The collected data is anonymous, i.e., cannot be connected to my real-world identity"). Equivalent to the items on knowledge about data logging, participants had to rate whether each statement was true, false, or they did not know.

App Usage Logging The PhoneStudy app tracks its usage to determine how participants used the privacy dashboard and control features. We track lifecycle events of the PhoneStudy app screens and detailed usage of the privacy dashboard (which log items were visible, if the detail view was clicked, if a filter was applied). Furthermore, usages of the control features were logged (e.g., which data items were deleted, when the logging was paused). Among others, we need this information to investigate RQ3c, as the usage of control features directly affects the resulting log data.

Post-Study Questionnaire At the end of the study period, another questionnaire was prompted via a notification. Here we assessed whether the privacy perception and knowledge about data logging and processing changed during the study, how the control features were used (if present), which effects the app usage had on the participants, and whether the app was usable in general. The post-study questionnaire took approximately 8 minutes and ended with the choice of compensation. In detail, the questionnaire inquired about the following factors:

- Prior privacy experience (repetition, cf. Section 5.1.4.3)
- Knowledge about data logging (repetition, cf. Section 5.1.4.3)
- Knowledge of how the data is processed (repetition, cf. Section 5.1.4.3)
- Data deletion behavior (condition *Control* only): Freetext and slider items on how many data items participants deleted in total, how many of which data types, why they did it, and in which situations (RQ3c) (see Table 5.4).
- Logging pausing behavior: Freetext and slider items on how often the logging was paused in total, how often for which data types, why they did it, and in which situations (RQ3c) (see Table 5.4).

- Behavior change and self-reflection: Four self-constructed items, inspired by the Technology-Supported Reflection Inventory of Bentvelzen et al. [51], on how the app usage affected smartphone usage, real-world behavior, and self insights. Additionally, a free text item on what changed and why (RQ3c).
- Logging awareness: Four self-constructed items on how aware participants were of the logging, whether this awareness influenced them, and if yes, what and why (RQ3b).
- Usability and other comments: Finally, we assessed the UEQ item groups on attractivity, perspicuity, and stimulation [361]. Furthermore, participants could enter any comments or remarks on the app and study in a free text field.

5.1.4.4 Data Analysis

The Android app sends its log data to our central server, where the data of all users is collected. The questionnaire data, which we initially stored at a university server, is also imported here. The raw data is not exported to the researchers' local computers for privacy and security reasons but instead analyzed on the server. Therefore we use an RStudio Server[1] instance running the statistics language R at version 4.1.3. We provide the preprocessing script files and the aggregated data in a Jupyter Notebook[2].

5.1.5 Results

In the following, we present the results of our evaluation of the privacy dashboard's transparency and control features in the wild. Each of the following subsections corresponds to one of our research questions: We show how transparency and control features affect the app installation rate, usage, and dropout (RQ3a); evaluate which effects on privacy concerns and trust are raised by both aspects (RQ3d); how aware users are about the logging and whether the knowledge about the logging differs (RQ3c); and whether induced behavior change and self-reflection could be noticed, thus the resulting data is influenced by the two factors (RQ3b).

[1] https://www.rstudio.com/products/rstudio/#rstudio-server, last accessed 2024-12-06
[2] https://github.com/mimuc/mobilehci22-transparency-and-control, last accessed 2024-12-06

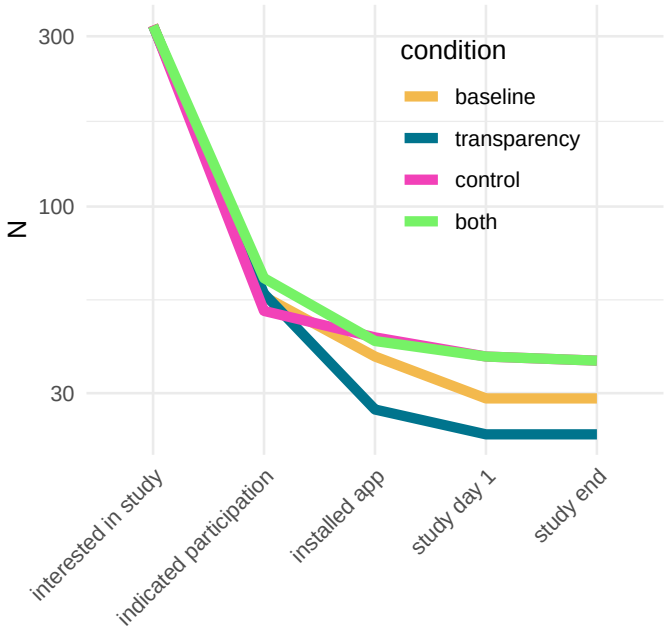


Figure 5.4 : Participation rates throughout our mobile sensing study. Users with the control features were significantly more likely to install the app, whereas users with the transparency features were significantly less likely to do so.

5.1.5.1 Mobile Sensing App Usage

In this section, we show which influence the experimental conditions had on how many participants installed our mobile sensing app, how long they kept it on their phones, and how much they actively used the app with its respective privacy features. These objectives correspond to RQ3a.

In total, 1286 potential users opened the onboarding questionnaire through our study advertisement. Already here, they were equally assigned to the four experimental conditions. Of those, 17.7% (227) finished the onboarding questionnaire and indicated a willingness to install the app, roughly equally across the conditions ($\chi^2(3) = 1.607$, $p = .658$). Of those who indicated a willingness, 66.1% actually installed the app and granted the required system permissions. There was a significant difference between

users who indicated participation and those who installed the app ($\chi^2(3) = 16.557$, $p = .0009$). Post hoc comparisons revealed that users with the *Control Features* installed the app, significantly more often (84.3%), while significantly fewer users with *Transparency Features* did so (47.4%). After having installed the app the dropout was comparatively low. 85.3% of those who installed the app kept it for at least one day ($\chi^2(3) = 3.674$, $p = .299$), 84.0% for 7 days (until end of the study) ($\chi^2(3) = 1.390$, $p = .708$).

Dashboard Usage: Factor Transparency Throughout the 7-days study, the users who had the possibility to view their data in the privacy dashboard (*Transparency Features* and *Both*) did so on average 14.60 times. A Mann-Whitney U Test ($W = 631.5$, $p = .430$) showed no differences between the conditions ($M_{Transparency} = 14.74$, $M_{Both} = 14.45$). The distribution of usage frequencies (visualized in Figure 5.5) is common, with a peak at around ten times, a set of more frequent users that opened the dashboard between 20 and 50 times, and a few outliers with up to 80 usages. Due to the data not being normally distributed according to a Levene test, we used the nonparametric Mann-Whitney U test instead of a standard t-test.

Dashboard Usage: Factor Control Those who had the option to control the logging, i.e., turning the logging of specific data features off or deleting data entries, made only rarely use of that. The average user paused the logging of one datatype 1.8 times

	Baseline		*Transparency*		*Control*		*Both*		**Total**	
Study Phase	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Interest in Study	322	100.0%	321	100.0%	322	100.0%	321	100.0%	1286	100.0%
Indicated Participation	56	17.4%	57	17.8%	51	15.8%	63	19.6%	227	17.7%
App Installed	38	67.9%	27	47.4%	43	84.3%	42	66.7%	150	66.1%
Study Day One	29	76.3%	23	85.2%	38	88.4%	38	90.5%	128	85.3%
Study End	29	100.0%	23	100.0%	37	97.4%	37	97.4%	126	98.4%

Table 5.4 : The number of participants throughout each study stage. The relative values relate to the stage before, i.e. report how many users continued since the previous stage.

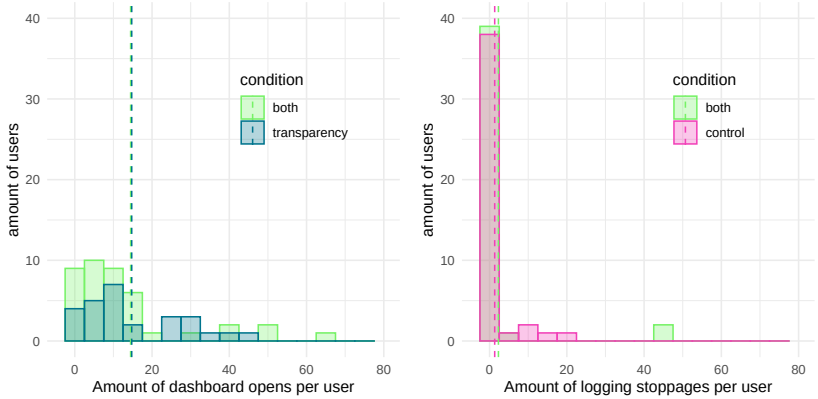


Figure 5.5 : Histograms visualizing the average usage frequency of the provided privacy features per user. The dashed lines shows the group mean. Users with the factor transparency (a) used the privacy dashboard on average 14.74 (*Transparency Features*) resp. 14.45 (*Both*) times. The features of factor control (b) in contrast were used very rarely, on average 1.33 times (*Control Features*) resp. 2.29 times (*Both*).

($M_{Control} = 1.33$, $M_{both} = 2.29$; Mann-Whitney U Test: $W = 960$, $p = .395$). The frequency distribution shows that the majority of the users did not use that feature at all (74 out of 85 users in *Control Features* and *Both*, respectively 87.06%).

Participants in the condition *Both* additionally could delete data entries from within the dashboard. This feature was used even more rarely by only 3 out of 42 users. In total, 15 data entries were deleted.

5.1.5.2 Logging Awareness and Knowledge

We assess how much the participants know about (1) what happens with their data and (2) what data is logged during the study to answer RQ3b. Therefore, we used items designed as a data logging quiz and data processing quiz that had to be answered in the pre-study and post-study questionnaires.

Data Understanding The data understanding quiz assessed knowledge on what happens with the data that the participants' app collects during the study, i.e., who has access to it, where it is processed, and where it is stored. The participants had

	Baseline	*Transparency*	*Control*	*Both*
Data Understanding - What happens with my data?				
Pre-Study Questionnaire [0;10]	6.55	7.91	6.54	8.08
Post-Study Questionnaire [0;10]	6.63	7.13	7.21	8.44
Difference Before/After [-10;10]	+0.296	−0.818	+0.559	+0.333
Logging Knowledge - What is logged?				
Pre-Study Questionnaire [0;12]	5.72	6.27	6.05	5.83
Post-Study Questionnaire [0;12]	6.26	6.61	6.5	6.56
Difference Before/After [-12;12]	+0.704	+0.182	+0.529	+0.788

Table 5.5 : Participants answered two groups of quiz-like items to assess their knowledge of (1) what happens with their data, and (2) what data is logged. While the latter did not show differences between the conditions, we noticed significantly higher knowledge of what happens with the data among participants that were using either the transparency or control features.

to check for each of the ten statements whether they were right or wrong. For each correctly rated statement, one gained 1 point; otherwise, 0. Thus, each participant reached a score between 0 and 10. We performed a Shapiro–Wilk test which showed that the data is not a normal distribution; thus, we used the nonparametric Aligned Rank Transformation (ART) ANOVA [436].

We first assessed the pre-study questionnaire. Here, users of *Both* and *Transparency Features* scored on average higher than those of *Baseline* and *Control Features* (see Table 5.5). The ART ANOVA shows a statistical significance for the main factor TRANSPARENCY. However, there was no statistically significant difference for the main factor CONTROL, nor was there an interaction effect, see Table 5.6.

After using the app for seven days (post-study questionnaire), users of the *Transparency Features* had a lower mean score than before (-0.818), while the other conditions on average increased (*Baseline*: +0.296, *Control Features*: +0.559, *Both*: +0.333), see Table 5.5. The ART ANOVA showed again a statistically significant difference for the main factor TRANSPARENCY; however, not for CONTROL or an interaction effect, see Table 5.6.

Important for app developers is if the app – the transparency and control features – impacts the user understanding. Therefore, we run a third ART ANOVA on the change in score (differences *start* − *end*). Here, the ANOVA could not reveal any difference; see Table 5.6.

	Analysis of variance (ANOVA) using ART [436]							
			TRANSPARENCY		CONTROL		T × C	
	dfn	*dfd*	*F*	*p*	*F*	*p*	*F*	*p*
Data Understanding - Pre	1	120	18.606	<.001	0.031	.861	0.699	.792
Data Understanding - Post	1	114	12.818	<.001	2.693	.104	0.153	.697
Data Understanding - Differences	1	112	3.135	.079	1.707	.194	0.506	.478
Logging Knowledge - Pre	1	120	0.211	.647	0.213	.645	0.822	.366
Logging Knowledge - Post	1	114	0.096	.758	0.131	.718	0.412	.522
Logging Knowledge - Differences	1	112	0.101	.751	0.341	.560	1.733	.191
I learned new things about myself	1	111	0.137	.712	0.030	.863	1.543	.217
I learned new things about my behavior	1	111	0.001	.978	0.227	.634	3.590	.061
I have changed my smartphone usage	1	111	0.475	.492	0.957	.330	0.169	.682
I have changed my behavior	1	111	0.677	.412	0.879	.351	0.025	.874
Perceived control	1	117	2.598	.11	9.358	.003	2.679	.104
App permission concern	1	117	0.004	.949	3.709	.057	1.14	.288
Perceived surveillance	1	117	0.374	.542	1.928	.168	2.85	.094
Perceived intrusion	1	117	0.824	.366	3.41	.067	9.441	.003
Permission Acceptance	1	117	1.114	.293	3.18	.077	1.89	.172
UEQ Score	1	111	0.399	.529	0.521	.472	0.748	.389

Table 5.6 : The two-way F-statistics of our two factors *Transparency* and *Control*, and their interaction effects.

Logging Knowledge Similarly, as with the data processing quiz, we assessed how much the participants knew about what the app was logging in the data logging quiz. Here we let them rate 12 items about the logging of data types (e.g., is GPS location logged raw? Are phone calls recorded?). The scores were, in general, lower than for the data understanding items, especially since the scale had a higher range (0 to 12). We could neither find any significant differences between the conditions in both pre-study- and post-study questionnaires nor significant effects between the factors, see Table 5.6.

Comparing the scores before and after the seven-day study we found a slight improvement over all conditions, again with *Transparency Features* users showing the lowest (+0.182). *Both* shows the highest increase (+0.788), closely followed by the *Baseline* (+0.704). *Control Features* users increased their logging knowledge by, on average, 0.529 points. We could not find any statistically significant effects using an ART ANOVA [436]; see Table 5.6.

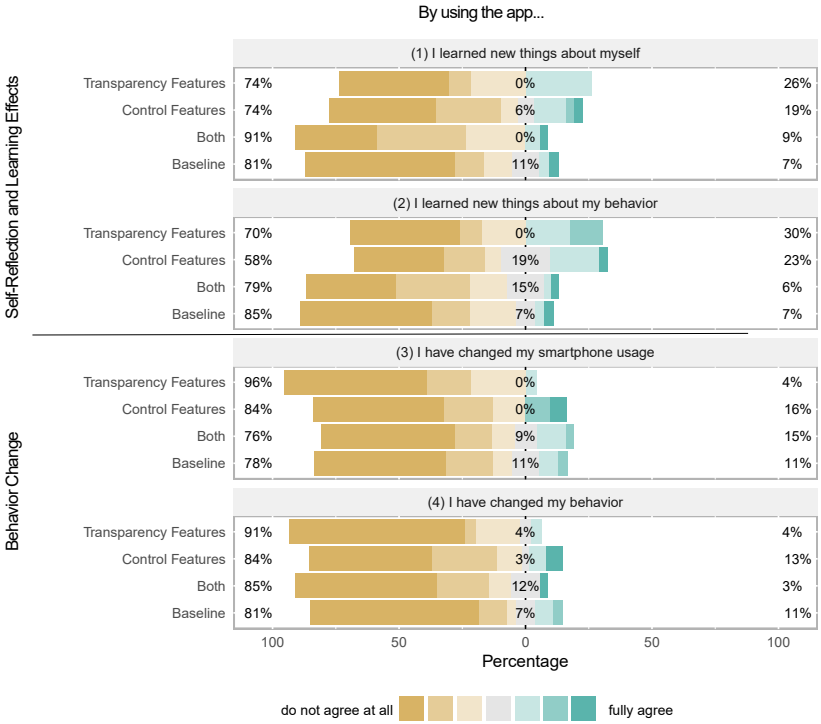


Figure 5.6 : Users of the conditions *Transparency Features* and *Control Features* reported slightly higher learnings about themselves and their behavior. However, none of the effects was significant, and no effect could be observed regarding self-reported behavior changes induced by our privacy dashboard.

5.1.5.3 Behaviour Change and Self-Reflection

The items on behavior change and self-reflection correspond to RQ3c. We included four items inspired by the Technology Supported Reflection Inventory (TSRI) in the post-study questionnaire [51] to assess self-reflection and learning effects (*learnings about myself* and *learnings about my behavior*) and behavior change (*change in smartphone usage* and *change of behavior*) induced by our privacy dashboard. Slightly more users of the conditions *Transparency Features* and *Control Features* agreed

that they had learned something about themselves and their behavior; however, there are no statistically significant differences, see Table 5.6. Moreover, no differences are visible between changes in smartphone behavior and real-world behavior change, see Figure 5.6 and Table 5.6. Furthermore, the free-text responses did not reveal differences regarding transparency and control features. Although users mentioned gaining insights during the usage (e.g., that they are using their smartphone too much, high social media usage, or unlock it unnecessarily often) and reported behavior changes (more conscious phone usage, reduction of screen time, and unlocks), we found no relation to the presence of transparency and control features. The sole presence of the logging app had more effect than the privacy dashboard.

Bayes Factor estimates returned values for BF_{0+} below $1/3$, which according to the classification scheme by Jeffreys [339], provides moderate evidence for H_0. This means that the data is more than three times more likely to occur for a system where the factors TRANSPARENCY and CONTROL have no effect than for one where the privacy dashboard triggers learning effects and behavior change. For all four measurements, except for *learnings about my behavior* BF_{0+} is below 0.1. This donates strong evidence that H_0 (no influence) is 10 times more likely (*learnings about myself*: $BF_{0+} = .089$, *learnings about my behavior*: $BF_{0+} = .196$, *change of smartphone usage*: $BF_{0+} = .085$, *change of behavior*: $BF_{0+} = .085$).

5.1.5.4 Privacy Concerns and User Experience

According to RQ3d, we evaluated privacy concerns about the usage of our app at the end of the seven days of app usage. Therefore we used the item collection by Degirmenci [110], which is tailored to the use of mobile devices. It is structured into five subscales. We tested the effects of the factors transparency and control via two-way ANOVA tests for normally distributed data, the nonparametric ART ANOVA otherwise. The results are plotted in Figure 5.7.

The factor control showed a significant positive effect on the ratings about perceived control, see Table 5.6. Users of the condition *Control Features* reported the highest scores for perceived control ($M_{Control} = 21.0$, scale range: [5;35]), followed by the condition *Both* ($M_{Both} = 20.5$) and *Baseline* ($M_{Baseline} = 19.1$). Users of the condition *Transparency Features* scored lowest ($M_{Transparency} = 14.6$). In the ratings for **app permission concern**, slightly lower scores were reported for the conditions *Control*

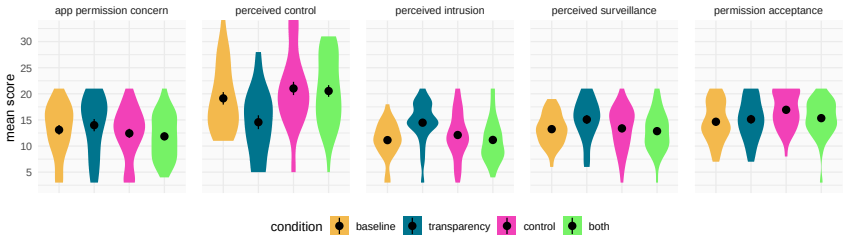


Figure 5.7 : Scores of the items on prior privacy experience [110]. Users that were offered transparency features were in general most concerned, even more than those who did not have the privacy dashboard at all. Control features could to some extent mitigate those concerns, and the both conditions scored equally and for some items better than the baseline condition without any privacy dashboard features.

Features and *Both*; however, the factor control does not reach the significance level of $p < .05$. For the **perceived surveillance** scale, the condition *Transparency Features* reports slightly higher values; however, again not significant. **Perceived intrusion** also shows the highest scores for users of the condition *Transparency Features*, followed by *Control Features*, which scored above average. Although the individual factors do not reach significance, we found significant interaction effects. The **permission acceptance** is highest in the condition *Control Features*, with the other conditions ranging equally. None of the factors was statistically significant.

UEQ scores for Attractiveness, Perspicuity and Stimulation were rather bad, according to the benchmark intervals of Schrepp et al. [361]. The condition *Control Features* scored highest ($M_{Control} = 0.555$), followed by *Both* ($M_{Both} = 0.334$) and *Transparency Features* ($M_{Transparency} = 0.311$). The baseline condition without any transparency or control features received the lowest scores ($M_{Baseline} = 0.288$). However, we could not reveal that the factors are statistically significantly different, see Table 5.6.

5.1.6 Discussion: Transparency and Control

RQ3a: How do transparency and control in a privacy dashboard affect the number of users adopting and dropping out of a passive mobile sensing app? Our results

show that transparency and control have different effects. While transparency led to fewer users actually installing the mobile sensing app, control, in contrast, increased the number of users. The dropout, later on, seems not to be affected.

The observed app usage rates across the study phases are common for mobile sensing studies. The literature agrees that once a potential participant has agreed to participate, it is unlikely that they quit their participation early [201, 282]. The low conversion rate of 10% is also common. For example, Kreuter et al. [234] report a revocation of 88% in a comparable mobile sensing study. The *Control Features* having the highest participation rate and best concern and trust rates in opposite to *Transparency Features* aligns with the literature. We can confirm the finding of Schnorf et al. [353] that control does not lead to less trust. Also, we support the recent work by Farke et al. [141]. They studied the "Google Privacy Dashboard" and emphasized the importance of control. Transparency features alone rather deter the users. They become aware of what is logged, which - without control features - might make them feel like their data is not in their hands anymore. From a trust and concern perspective, this is even worse than not providing any transparency. We argue that in the *Baseline* condition, the users instead experience a sense of security due to unknowingness, which seems to be better than knowing what happens in detail but being unable to control it. However, it is interesting that the biggest difference between the groups, with significant effects for both factors control and transparency, did not occur while using the app with the users' data but during the onboarding phase. This means that the main difference was not made by whether the users could view, delete and pause their data, but by the advertisement of the privacy-enhancing features in the onboarding process. This aligns with the usage statistics of the control features: The sole presence of control features made users feel better protected, although they only rarely made use of them. Therefore, we conclude that the screenshots during onboarding presenting the transparency features, including demo data, made the difference. We argue that they have better conveyed what is logged than the sole privacy declaration.

RQ3b: How do transparency and control in a privacy dashboard affect the awareness of and knowledge about the data logging? We found that transparency is the influencing factor: Users with this factor could better recall the information.

In contrast, users without it caught up during the one-week app usage. While the knowledge increased for both factors, it increased the least for transparency. The awareness of and knowledge about what data is logged, however, was not affected by transparency nor control. During the first assessment in the pre-study questionnaire, both the knowledge about what data is logged and what happens with the data had the lowest scores in the condition *Baseline*, i.e., when no privacy-enhancing features were available. The reason might again be the screenshots with demo data, i.e., the control features in the onboarding process, which as a side effect, seem to convey what the app is doing. However, during the one-week usage period of our mobile sensing app, the scores in *Transparency Features* behaved significantly differently. They decreased for the data understanding items, while all other conditions increased, and reported the lowest increase of all conditions for the logging knowledge items. The reason for the low improvement of the logging knowledge scores might be caused by the presentation: The control features screen in the conditions *Control Features* and *Both* provided an overview over all logged datatypes, whereas the main view in condition *Transparency Features* consists of a timeline view. As a result, rarely logged data types (e.g., Bluetooth settings changes) were likely not seen by users who used the dashboard rarely and did not scroll down much. In future systems, we recommend not only providing a strictly chronological order (timeline view) but also a grouped view where at least one entry of each data type is presented prominently. We suspect self-reflection effects regarding the increase of the data understanding scores, which were present in all conditions except *Transparency Features*. For example, since changing logging settings and deleting data requires an active decision, we suspect that this might have made people think more about what the app does.

RQ3c: How do transparency and control in a privacy dashboard induce behavior change and self-reflection and thus the logged data of a passive mobile sensing app? After having used the app, we asked the participants about the learning effects and behavioral changes induced by the app. If they at least rather agreed on having learned something or changed their behavior, we further asked them to describe the effect. Interestingly equal self-insights and behavioral changes were mentioned across the study conditions. We conclude that not the presentation of the data led to those effects but the sheer presence of the mobile sensing app. Thus, it does

not make a difference whether a privacy dashboard is included or not. Our Bayes factor analysis further supports that an effect by the privacy dashboard itself is very unlikely. The control features implemented in the dashboard, i.e., pausing the logging and deleting logged data, were used only rarely. Besides a few users who used those features regularly, the vast majority did not make use of pause or delete features at all. Concluding on **RQ3**, we did not find any indicators that the log data is influenced by a privacy dashboard incorporating transparency and control significantly. However, we are aware that self-reported measurements, as we used them to assess self-reflection and behavior change, do not provide full evidence. We encourage future research to conduct studies that measure actual behavior in the wild.

RQ3d: How do transparency and control in a privacy dashboard affect a passive mobile sensing system user's privacy concerns and trust? We found that the factor control showed significantly higher scores in perceived control which is not surprising and confirms the effectiveness of the control features, albeit not used frequently. The app permission concern shows a similar (inverse) trend; users who had the option to control the logging might, thus, be more willing to grant the app permissions. The results of the permission acceptance items behave accordingly, supporting this conclusion. Furthermore, we found high scores in perceived surveillance and intrusion in the condition *Transparency Features*, but not in *Both*. This aligns with our findings from the app usage and dropout rates: *Transparency Features* should always be accompanied by the ability to control the logging.

5.1.6.1 Future Work on Privacy Dashboards

In the future, we recommend that more focused investigations would need to be conducted to obtain detailed insights on behavior change induced by privacy dashboards. We did the first step by investigating RQ3. However, measuring behavior change in in-the-wild studies is difficult and self-report scales as we used it can be biased [119].

Our privacy dashboard presented only raw data to the users. We deliberately omitted any aggregated or inferred data to keep our study setting generalizable and neutral. However, nearly every real-world application does some processing to use the data. Following the "consent as a process" approach [179] and guidelines for privacy in big data systems [274], the data processing steps following the raw data collection should

be incorporated into privacy dashboards. Therefore, dedicated research becomes necessary that studies the effects of aggregated and derived data in privacy dashboards. Current research is contradictory. Here, Herder et al. [187] report increased trust and decreased perceived risks by derived data, while Rudnicka et al. [343] hypothesize that transparency about the derived data might make people less fearful.

Privacy perception is a very individual construct. The studies of Schnorf et al. [353] distinguished different groups of users by their privacy-related predisposition. Also, Awad et al. [20] reported that the relation between transparency and the resulting effect depends on the user: People who desire transparency are less willing to be profiled. Thus, for them, trust decreases, and concerns increase with the provision of transparency than for people who have a weaker desire for privacy. Future privacy dashboards could make use of this and adapt to their user. Therefore, researchers could extend our RQ4 and investigate different kinds of privacy dashboards depending on the individual user's predisposition towards privacy perceptions.

5.1.6.2 Take-Aways Regarding Transparency and Control in Mobile Apps

Study Take Aways

For the design of information and consent UIs we take away the following key points: **Transparency should always be augmented with control.** Sole transparency rather raises concerns but leaves the users with their concerns. Users should be **offered** control, rather than forcing them to exert it. This means, interfaces should rather have a character of **being available on-demand, instead of forcing users** to make privacy decisions in a specific situation. The **privacy-enhancing features should be advertised** prominently. Users rarely actively look for privacy. Giving them the ability to inform themselves about the happenings and take over control passively is essential.

5.2 On-Device Preprocessing of Mobile Language Data

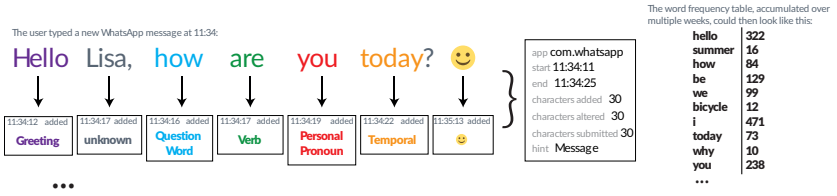


Figure 5.8 : Overview of our data logging method to facilitate privacy-respectful studies of language use in everyday mobile text communication: Text entered by a participant (e.g., in a chat app) is abstracted to avoid revealing private content to the researchers while still catering to a wide range of common research interests. *Left:* Our *Word Categorisation* concept maps a predefined set of words to categories (e.g., "Hello"→"Greeting"). Moreover, *Custom Regex Filtering* allows for flexibly logging predefined strings, such as emojis. *Centre:* Metadata about the keyboard session is logged as well. *Right:* Our *Whitelist Counting* concept logs total usage counts for words in a predefined whitelist.

> This section is based on the following publication:
>
> Florian Bemmann and Daniel Buschek. "LanguageLogger: A Mobile Keyboard Application for Studying Language Use in Everyday Text Communication in the Wild." In: *Proc. ACM Hum.-Comput. Interact.* 4.EICS (June 2020). DOI: 10.1145/3397872

Collecting data on language use is a key challenge central to all such work. Researchers either examine existing corpora or they have to collect data from computer-mediated communication (CMC) applications on their own [407]. This comes with an inherent methodological challenge: There is a difficult tradeoff between a) limiting data collection to respect participants' privacy, and b) collecting comprehensive, unbiased and natural data to answer open research questions. Recent studies highlight this: For instance, Rosenfeld et al. [342] report that reluctance to participate was a key challenge for their study of communication via WhatsApp, despite emphasising encryption and anonymisation.

Some studies aim for control in the lab and/or with writing tasks (e.g., "[w]rite an e-mail [...] explaining that you will not be able to take the next exam." [124]). This avoids observing private messages yet clearly limits the data to the given topics. To

collect actual everyday messages, researchers have asked people for retrospective submissions of selected chat-logs. However, this data may be biased, for example due to capturing only one channel (e.g., one app, such as WhatsApp) and participants' criteria for selecting the submitted chats [154, 380, 400, 407].

It can also be difficult for researchers to avoid looking at private content during analysis of natural text logs (e.g., when thematically coding statements). This also extends to third-parties (e.g., chat partners in an app), who might not have consented to the study at all. As examples of private content, full chat logs may contain sensitive information like names, phone numbers, intimate conversations, passwords, and financial data.

We present a novel method and tool to address this challenge and support researchers in running studies of language use in the wild: LanguageLogger, a mobile keyboard app, allows researchers to log useful abstracted language data from text entered on participants' smartphones (see Figure 5.8). We also provide this functionality as a logging module for integration into other research apps. Concretely, we employ three text analysis methods in a novel way, namely for *text abstraction that runs directly on participants' phones*. Hence, our app never reveals raw text to the researchers yet caters to many research interests. In brief, our text abstractions are:

- *Whitelist Counting:* The number of occurrences is logged for each word in a predefined whitelist, for example *"conference: 14", "paper: 25"*.
- *Word Categorisation:* Entered words are mapped to categories configured by the researchers. For instance, *"Jane is happy"* might be logged as *[name, verb, positive adjective]*.
- *Custom Regex Filtering:* Strings that match predefined regular expressions are logged as-is, or as an event indicating the occurrence of a match, but not the string itself. For instance, this can be used to log emoji use.

Note how these abstractions not only avoid recording potentially private text content but also pre-process the data in a way that many researchers require anyway. We provided the user transparency and control over the data collection: In a web-based dashboard they could view samples of their logged data to get an impression of the workings of the preprocessing. Control over the data logging is enabled through a button that is always present in the keyboard, which allows disabling the logging

temporarily. The contribution of this study is two-fold: We propose a text abstraction process for mobile typing data, that, for the best of our knowledge, has the best trade-off of privacy-friendliness and richness in gained data yet. Second, we present insights from a user study on the effects of incorporated transparency and control mechanisms. The contribution to this thesis is especially the report of how these measures affect users' perception of privacy. This project integrates into this thesis through the following research question:

RQ3e *How does on-device preprocessing of mobile language data affect users privacy perception?*

5.2.1 Concept Development Process

Overall, we employ a keyboard app with three text abstraction methods. Here we describe our concept development process.

5.2.1.1 Defining Foundations for Supporting Privacy & Trust

At the outset, our need for a new data logging concept was motivated by our practical experiences with several interdisciplinary studies on data collection in the wild. We found that trust in such field study setups is more complex than a yes/no decision with typical informed consent and approvals. Thus, taking additional measures to make a study setup more privacy-respectful in our view is a highly worthy investment. This also follows the foundations presented in related work on privacy-aware keyboard logging [71]: The goal is to support a trustful relationship between participants and researchers with a concept and tool that records relevant data without revealing private content.

Note that, as in the work by Buschek et al. [71], this assumes a generally trustworthy foundation: We do not claim full protection against malicious intent (e.g., researchers actively trying to spy on their participants) since ethical research standards still need to be upheld by the researchers and in more fundamental, institutional ways – and not exclusively by a keyboard app.

5.2.1.2 Assessing Logging Requirements for Research

To support a wide range of research interests we analysed the literature in detail. Our goal was not to conduct an exhaustive survey but to identify the main study approaches, as presented in the related work section. We analysed the work with regard to requirements for both logging procedures and the resulting logged data to answer these questions: 1) Which data is observed (e.g., chat messages)? 2) Which measures are computed on said data (e.g., word counts)? 3) Which privacy measures/challenges are reported (e.g., participants reviewing messages)?

In addition, we ran an interdisciplinary workshop on logging language use and a series of subsequent discussions with four researchers from HCI, Psychology and Statistics. This was organised in the context of a joint long-term research project and study planning, thus eliciting real methodological needs. We also discussed our literature analysis with this group. We aggregated the following requirements for our mobile tool:

- **Integration in everyday life** To observe natural everyday language without bias the logging tool must be incorporated in people's usual environment/systems.
- **Differentiation between shared-public and private content** Language differs between personal messaging and publicly shared content. Thus, our tool should allow for assessing or filtering for the context of writing (e.g., in which app).
- **No extra efforts for participants** To enable long-term deployment in the wild with reasonable effort, the system should not demand extra work from participants, such as reviewing logs.
- **Device-wide logging – but none for chat partners** To yield a comprehensive corpus and respect privacy, the tool should be able to process all mobile writing of the participant, but none of other people (e.g., chat partners).
- **No access to raw text** To respect privacy, no one except the participant should be able to access the raw typed text.

5.2.1.3 Developing a Logging Strategy: Text Abstraction

From our requirements analysis it became evident that in the vast majority of studies the actual raw text content is ultimately rarely needed for the conducted text analysis.

Text Analysis	Study Goal	Data	Covered by
Word-based: no. of words/textisms (e.g., "lol")	Use of textisms in formal vs informal comm. [124]	Email written in study task	Whitelist Counting
Category-based: no. of words per pre-defined category (LIWC cat. [311])	Associations of word use and personality [446]	Public blog articles	Word Categorisation
Category-based (LIWC [311]); word-based: psycho-linguistic stats on counts of words in a 150k dictionary (MRC [432])	Associations of tweets and personality [166]	Public twitter data	Word Categorisation, Whitelist Counting
Message-based: no. of words/characters; keyboard-based: use of auto-correction & suggestion	Use of WhatsApp / chat behaviour [400]	WhatsApp chat histories	Keyboard Sessions, Further Logging
Word-based: no. of words/textisms; interaction-based: texting speed/frequency	Relating texting speed/frequency with language & literacy measures [303]	Mobile texting (in given study task)	Whitelist Counting, Keyboard Sessions, Further Logging
Non-textual cues: counting emojis	Associations of emoji use and personality [252]	Public twitter data	Custom Regex Filtering

Table 5.7 : Examples of common text analysis methods in research on (mobile) language use, along with example studies and data sources. The last column indicates which of our text abstractions and logging features cater to each analysis. Overall, the table illustrates that we address a wide range of research interests: We enable these analyses for data from everyday (personal) mobile text communication, while avoiding that people have to share actual raw text with the researchers. Note that some measures were self reported (e.g., use of auto-correction [400], texting frequency [303]), whereas we enable quantitative logging of such data.

For example, text is often analysed in abstracted and aggregated ways, such as the number of words per category. Table 5.7 lists key examples to illustrate this. Therefore, a viable approach to facilitating both participants' privacy and the researchers' workflow is to apply these text data processing steps directly on the device, to then only share the resulting abstracted data. A closer look motivates three particular types of text abstractions:

Words are counted: For instance, the number of occurrence of words or word categories is analysed with regard to associations with the author's personality [166, 446] or to compare language use between contexts (e.g., formal vs informal [124]).

Words are categorised, that is, mapped to pre-defined categories or values: For example, Herring et al. [188] was interested in 'female' and 'male' stylistic words. There also exist generalised categorisations that are widely used across studies, for example the LIWC dictionary [311]. "Categories" can also be numeric, such as SentiWS by Remus et al. [337], which assigns sentiment scores to words.

Specific terms are measured: Some studies analysed aspects of text that require more flexibility. Typically, this involves text elements not captured by categorisations such as LIWC [311]. However, these are still analysed in an abstracted way, namely as occurrences or counts. A prominent example are recent analyses of emoji use [252, 409].

5.2.2 Final Logging Concept

We next describe our final concept, that we then applied in the field study, in detail.

5.2.2.1 Text Abstractions

First, we describe how we realized the three kinds of text abstraction motivated above. For a visual overview, see Figure 5.8.

Whitelist Counting For this text abstraction method, the researchers define a list of words before the study. Whenever one of these whitelisted words is entered this is counted by the system. For example, consider a study that looks at the use of positive words such as *good, happy, great* and thus includes these in the whitelist. The sentence "Thomas is happy" would then be logged as "happy: 1." Other words, such as the name "Thomas", are not recorded. These counts are summed up, such that at the end of a study, the researchers in this example might get a table such as "good: 123; happy: 456; great: 789." It is easy to see that this does not reveal private messages as long as enough text is logged overall. We quantify this with our experiments.

Word Categorisation Here, researchers define a mapping from words to categories (e.g., "happy: positive emotion"). Words are then mapped to categories and the system

only logs that those categories occurred (we call these *word events*). If a word is not included in the mapping dictionary, the category *unknown* is recorded. As metadata, category log entries also have a timestamp, the app, a flag indicating if the word was added/edited/removed, and a reference to a keyboard session. In case of an edit, the category before and after the change is logged. Overall, entered text is thus logged as a sequence of such word events. For example, when a user types "I am happy" in a WhatsApp message this might be logged as three word events: *personal pronoun, verb, positive emotion*. This does not reveal the raw text as long as categories are not too specific (i.e. in an extreme case where each word is its own category this would log raw text). Again, we quantify this in our experiments.

Custom Regex Filtering We integrate a regular-expression matcher to allow for logging custom patterns that would be hard to specify in a list of terms, as for the other two abstraction methods. Entered text is checked against the configured matcher(s). In case of a match, the system can either 1) log the matched string as-is or 2) just the occurrence of a match. For example, this could be used to count how often people enter phonenumbers (without logging them), or to log emojis via the emoji unicode range: "Call me at 004915778948140 😁" would then be logged as two *regex events*, that is, "phonenumber" and "😁", plus metadata (e.g., timestamp, app). By the way, if you read this and send a text message to the aforementioned number, I will invite you to a free lunch.

5.2.2.2 Lemmatisation

Before applying the text abstraction concepts described above, each word can optionally be lemmatised. For instance, *went, going* would be replaced with the root word *go*. This is useful since common dictionaries such as LIWC [311] do not contain inflected forms.

5.2.2.3 Keyboard Sessions

We define a keyboard session to start when the keyboard opens and end when it is closed. We store metadata for each session: Number of characters added/altered, name of the app in which the text has been entered, hint-text/label of the text field, timestamps of the session's start and end.

5.2.2.4 Further Logging

Further logged data includes the use of word suggestions and auto-corrections (occurrences, not words). The app can log available Android sensor data which was also logged by related work [71], such as touch data (not text revealing), device orientation, accelerometer, gyroscope, and so on.

5.2.2.5 Logging Exceptions

Following related work [71], we realised two logging exceptions: First, we never log fields for passwords, logins, addresses etc. This uses Android field types. While we cannot guarantee that every field is marked in this way by developers, many apps do so since it is essential for accessibility. Moreover, note that strings like passwords and names of accounts/people are not part of our whitelist dictionary and thus never logged as text anyway. Second, as the related work, our keyboard UI shows a small lockpad icon that allows people to pause logging.

5.2.3 Implementation as an Android Module and Keyboard App

We created a reusable and remotely configurable Android keyboard application, LanguageLogger, as described next.

5.2.3.1 Overview of System Modules

For reusability, the LanguageLogger app project consists of four modules: One is the keyboard, the other three realise the data processing. Each module represents one layer of functionality. This loose coupling allows easy integration of the implemented functionalities (or a subset thereof) into other applications, such that our logging methods can be used with other studies and apps, not only as part of our keyboard.

The embedding into Android applications ensures that raw text content does not leave the client device. The dependencies between the four modules are visualized in Figure 5.9. Regarding the "flow" of the data, the application module provides the raw source data and passes it on to one or both of the processing modules. The resulting data then is returned to the application module. Thus no data is exchanged between the other LanguageLogger modules implicitly. In the following four subsections we present the four modules in more detail, each representing one layer of functionality.

5.2.3.2 Application Module (here: Keyboard)

The Application Module constitutes the "host app" that uses LanguageLogger logic to process its logdata. Here we build on the keyboard app of Buschek et al. [71], which in turn uses Google's Android Open Source Project (AOSP) Keyboard[1].

The host app module includes the other required modules as local library module dependencies[2]. A module's functionality can then be used by instantiating the respective class and calling their methods. The result is either provided as return value (synchronous operation) or can be obtained by implementing a callback and passing it with the method call for asynchronous operations. For flexibility, the storage or transmission of the resulting log data also has to be implemented in this host app module; in the case of our keyboard app implementation this includes the transmission to the LanguageLogger server. We visualize the dataflow between the modules in Figure 5.10.

5.2.3.3 Base Module

Functionality that is required by multiple other modules is located in the base module, to avoid redundancy. This includes a REST client that matches our LanguageLogger server implementation, and utility classes for handling console logs.

[1] https://android.googlesource.com/platform/packages/inputmethods/LatinIME/, last accessed 2024-12-06

[2] https://developer.android.com/build/dependencies#dependency-types, last accessed 2024-12-06

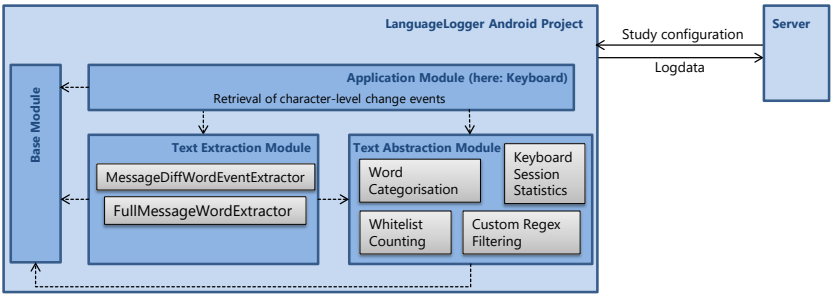


Figure 5.9 : Overview of the main architecture of our app: Dashed arrows indicated Android module dependence. The preprocessing (*Text Extraction Module*) and abstraction logic (*Text Abstraction Module*) are separated from the host application (*Application Module*). All modules are loosely coupled, thus it is possible to use the LanguageLogger logic in other Android applications as well.

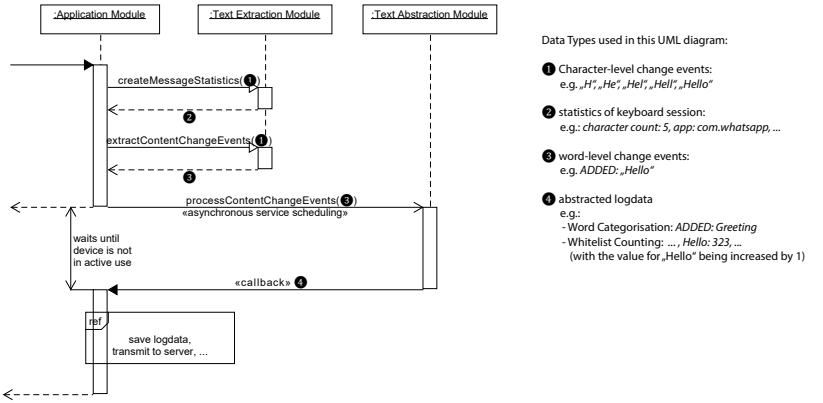


Figure 5.10 : An UML sequence diagram visualizing the dataflow between the LanguageLogger modules. Which types of data are passed is indicated by the numbers in the filled circles, which are explained on the right. Typing the word “Hello” is used as exemplary user action.

5.2.3.4 Text Extraction Module

The structure of the input data collected by applications depends on the data source. Thus, this module servers as a layer to separate text abstraction logic from source-specific preprocessing steps.

ContentChangeEvents Language log data typically arises in the form of character-level logs (e.g., when observing user interactions). In the case of our keyboard, for example, events consist of the text field content after each keystroke: For instance, a user typing the word "Tom" in a sentence "Hello Tom" yields the events *"Hello T"*, *"Hello To"*, *"Hello Tom."*

However, to apply our text abstractions, we need events on word-level, such as in this case *ADDED: "Tom"*. In the following, we refer to such an event as a *ContentChangeEvent*.

In particular, we distinguish two types of such events, fitting the two types of Word Categorisation described in Section 5.2.2.1:

- *ADDED, CHANGED, REMOVED, SPLIT, JOINED*, represent user actions (on words) that happened in temporal order during a typing session.
- *CONTAINS* corresponds to the presence of a word in the final submitted text (e.g., a text message entered in a chat app).

These *ContentChangeEvents* are created from incoming character-level events using the class *MessageDiffWordEventExtractor* (for events of type *ADDED, CHANGED, REMOVED, SPLIT, JOINED*) and *FullMessageWordExtractor* (*CONTAINS*). This works synchronously, as the processing is not computationally expensive.

FullMessageWordExtractor This can be implemented in a straightforward fashion: It takes the text and splits it into single words with a regular expression. In particular, this expression matches sequences of word characters (e.g., in our implementation for German: a-zA-Z0-9äöüß'), enclosed by non-word characters (everything except the word characters).

MessageDiffWordEventExtractor The extraction of *ADDED, CHANGED, REMOVED, SPLIT* and *JOINED* events is more complex. The general algorithm, in short, works as follows. For each character change event:

- Split the text into words, with the regular expression mentioned above.
- Identify the word that has changed or was added/removed: Compare each word in the content after the character change with the word at the same index before the character change.

All consecutive character change events affecting the same word are collected. Comparing the content before the first and after the last character change event of such a sequence yields the overall change that the user performed.

However, there are many edge cases of user actions that cannot be tackled without significantly extending this general procedure. To name just a few:

- Adding a new word in between existing ones: We observed users technically doing this by appending or prepending a word to an existing word and typing the separating space character only at the very end. For example, the sequence *"Hello how are you", "Hello Thow are you", "Hello Tohow are you", "Hello Tomhow are you", "Hello Tom how are you"* would, without special consideration, result in *CHANGED: "how"* $\rightarrow$ *"Tomhow", SPLIT: "Tomhow"* $\rightarrow$ *"Tom, how"*. However, the intended correct log should just be: *ADDED: "Tom"*
- Pasting and auto-correction: If users paste text sequences or auto-correction changes complete words multiple characters (possibly at different positions) may change with one change event published by the system, whereas the general algorithm described above expects just one character to differ.

We conducted early user tests, including raw content alongside the extracted words, to identify as many of these edge cases as possible. Thereafter, we used test-driven development to improve our algorithm, adding unit tests for each newly identified edge case. In this way, we extended our first general algorithm to a carefully grown decision tree.

5.2.3.5 Text Abstraction Module

This module implements the language processing of Section 5.2.2 in a resource-efficient manner that does not disturb the user. In contrast to the work performed in the Text Extraction Module, the Text Abstraction Modules is more computationally intensive: The word lists and category mappings defined by the researchers are loaded into working memory and are repeatedly queried. Thus, the required memory and computation time depends on the study configuration and can become quite large (e.g., the German dictionary we deployed in our study contains 300,000 words). We implemented the following measures to still keep the processing unobtrusive for the user:

- Load one list after another: The jobs are not parallelized and word lists are not combined to one single list. This lowers the peak memory requirement.
- Implementation as Android AsyncTask[1]: The work is performed on a background thread, avoiding blocking the calling thread.
- Resource-aware scheduling with Android JobScheduler[2]: The AsyncTask is wrapped by a JobService, that is scheduled to launch when the device is not in active use.

If lemmatisation is activated it is applied in this module. In our implementation, we used the TreeTagger software of Schmidt et al. [349], but this can be flexibly changed. All services for the text abstraction methods can be scheduled with one method call, through the class *RIMEInputContentProcessingController*. Due to the underlying asynchronous implementation, callbacks have to be implemented to obtain the resulting data.

Furthermore, the Text Abstraction Module contains a service class to calculate aggregated statistics of keyboard sessions (*KeyboardMessageStatisticsGenerator*). For example, this includes calculating the number of characters entered in the keyboard session and the keyboard session start/end time. Other applications could use this logic to get similar statistics for their respective unit of reference (e.g., for notification logging this would compute statistics per notification).

[1] https://developer.android.com/reference/android/os/AsyncTask, accessed 2024-12-12
[2] https://developer.android.com/reference/android/app/job/JobScheduler, accessed 2024-12-12

5.2.3.6 Backend

We implemented a backend application communicating with the mobile app through a REST interface. The backend fulfills two main objectives: Configuration of the text abstractions, and storage of the logdata.

Text Abstraction Configuration The text abstractions presented in Section 5.2.2 can be configured in a system of phyical and logical configurations: Researchers upload a word-to-category mapping (for Word Categorisation) and a wordlist (for Whitelist Counting), as a physical mapping/list. One or multiple physical lists can then be combined into a single logical configuration. These logical configurations are then used on the mobile device. This separation allows for easy changeability, for example, in the case that a single dictionary needs to be replaced or updated within the context of a larger study configuration. Moreover, Custom Regex Filtering is configured by entering regular expressions in the backend. The LanguageLogger app downloads these configurations from the backend during the setup process.

Storage of Log Data The log data that the backend retrieves from the mobile devices is stored in a relational database, implementing the relations between keyboard sessions and extracted categories. Each category belongs to a keyboard session. Custom Regex Filtering logs are also treated as categories, as their data structure is similar. The whitelist word counts are collected in an absolute frequency table, encompassing the whole study duration.

5.2.4 User Study

We ran a field study as an example deployment of our text abstraction methods and tool. Our aim was to test the methods and app and to assess the users' views of the text abstractions.

5.2.4.1 Apparatus: Keyboard App, Questionnaires, Web UI

For the study we set our keyboard to this example configuration: For Word Categorisation, we used the common LIWC dictionary [311]. For Whitelist Counting, we used the DeReWo Lemma List published by the Leibniz Institute of the German Language

(IDS) [230]. For Custom Regex Filtering, we specified matchers for emojis. These are example choices for this first deployment, motivated by the literature. Our tool is flexible and easy to use with other dictionaries, whitelists, and so on. Apart from our keyboard, we used the following components:

A *web interface* showed participants examples of their own logged data (e.g., word categories, counts). It was used as part of the post-study questionnaire and the interviews.

A *pre-study questionnaire* explained the three text abstraction concepts in detail, along with illustrated examples (similar to Figure 5.8). This not only served as an introduction but also as part of the informed consent procedure. The questionnaire also assessed people's attitudes towards privacy in general, using the item sets of Buchanan et al. [70]. Finally it asked for demographics and provided instructions to install our Android app.

A *post-study questionnaire* asked about 1) potential influences of the keyboard and logging on usability and experience, 2) perceived privacy protection, and 3) feedback on app and study. Moreover, the questionnaire linked to the web UI and thereafter again asked about perceived privacy protection.

5.2.4.2 Participants

We recruited 20 participants (12 female, 7 male, 1 prefer not to disclose) via newsletters and social media. Their mean age was 24.5 years (range 19 - 34). They received 5€, plus 5€ for an optional interview (which 4 people did). Using the items of Buchanan et al. [70], people's self-reported attitude regarding privacy concern, general caution, and the importance of technical protection was slightly above neutral.

5.2.4.3 Procedure

Participants first filled in the pre-study questionnaire. Then they installed our app on their phones and set it as their default keyboard. After two weeks of use, they filled in the post-study questionnaire, including the web interface that showed examples of their logged data. Finally, we invited participants to semi-structured interviews to get a more detailed picture on the points from the questionnaires. For example, we asked

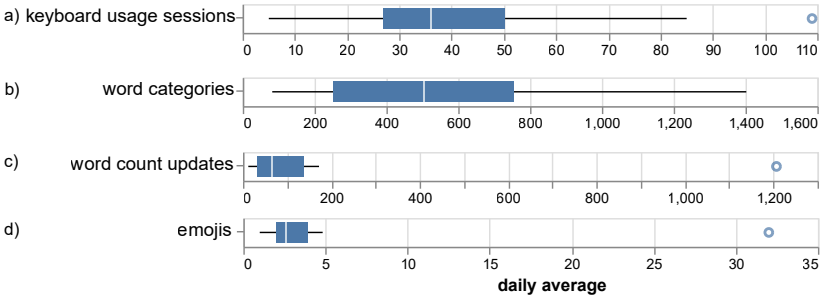


Figure 5.11 : Overview of our study data, averaged per day and person: (a) Number of logged keyboard sessions, (b) word events, (c) increases of word counts (Whitelist Counting), and (d) regex events (here: emojis). Each boxplot shows the median, upper- and lower quartile, min/max whiskers and outliers.

about their impressions of the study and app, perceived privacy protection with the text abstractions, and further feedback. Interviews were audio recorded to facilitate later analysis.

5.2.4.4 Results

For this report, we define a *word event* as entering a word, editing a word, or deleting a word. We define a *keyboard session* as the time from opening the keyboard to closing it.

Overview of Logged Data We logged a total of 12,318 keyboard sessions. On average, a session had 10.16 (SD 6.39) word events. Averaged per day and person, we logged 42.86 (SD 26.69) keyboard sessions, 539.56 (SD 345.61) categorised words by Word Categorisation, 172.65 (SD 266.01) count updates by Whitelist Counting, and 4.42 (SD 6.98) emojis by Custom Regex Filtering (also see Figure 5.11). Regarding typing context, the top five apps were: WhatsApp (6,455 keyboard sessions), Chrome browser (940), Instagram (435), Google Quicksearch on the homescreen (404), and Telegram messenger (376).

Perceived Privacy through Text Abstractions (Post-Study Q.) Our post-study questionnaire included the five-point Likert item *"I feel protected by this kind of data*

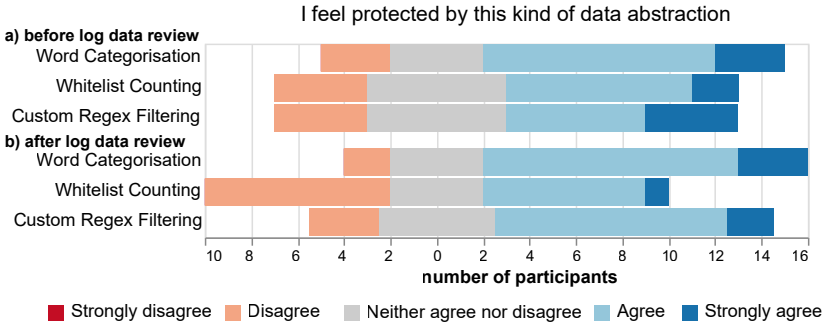


Figure 5.12 : Results from the Likert questions on the feeling of privacy protection by the three logging concepts, (a) before the log data review and (b) after it.

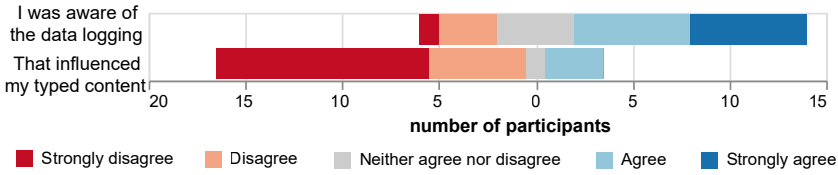


Figure 5.13 : Results on awareness of logging while typing and the reported influence of that awareness on the typed content.

abstraction" (Figure 5.12 top, 5=strongly agree). Here, participants rated Word Categorisation best (M=4), followed by Custom Regex Filtering (M=3.5) and Whitelist Counting (M=3.5).

We also asked people to order the text abstractions by how much these contribute to respecting their privacy: 13 of 20 people judged Word Categorisation to contribute the most to respecting privacy, followed by Whitelist Counting (4) and Custom Regex Filtering (3).

Influence of Log Data Review (Post-Study Q.) The post-study questionnaire presented users with a subset of their log data via a web UI, then asked again about privacy protection through the abstractions (Figure 5.12 bottom). We found no sig-

nificant changes comparing ratings before and after this review for the three text abstractions (Wilcoxon signed-rank tests, all $p > .05$). Hence, we found no evidence of an overall shift towards lower or higher ratings.

There were rather individual shifts, both ways: 16 of 20 people changed their opinion on at least one of the three concepts. Ratings for Whitelist Counting were changed by seven people, who all except for one lowered their ratings. In the interviews, two of those stated that they could relate the counting data more easily to themselves than expected. One participant noticed a high count for "I," making them wonder whether he was too egocentric. Another one recognised terms from band names and concluded that he might be identified by these otherwise rare words if his musical taste was known. For Word Categorisation we did not notice a trend (8 did not change, 6 increased, 6 decreased). For Custom Regex Filtering five people changed their rating (4 increases, 3 decreases).

Influence of Study Setting on Text Content (Post-Study Q.) Our post-study questionnaire had five-point Likert items on potential influences of the study situation and the keyboard functionality and UI. Figure 5.13 shows the results: Overall, people reported that they were aware of the data logging during the study (median 4). However, they indicated that this had little to no influence on the text they entered (median 1).

Influence of Study Keyboard on Usability (Post-Study Q.) Our post-study questionnaire also included five-point Likert items on potential differences of our app to people's usual keyboard apps. Most people found differences (median M=4) and felt that this influenced their interaction behaviour (M=4). More specifically, 16 people (rather) agreed that different auto-completion was an influence (M=5), while ten rated this way for visual UI differences (M=3). Free text questions and interviews showed that auto-completion was worse since the app had not learned users usual words (yet). Moreover, a key UI difference was how to access special characters and emojis.

5.2.5 Limitations

Our text abstraction methods and tool enable new studies of everyday mobile language use in the wild which have not been possible so far. Nevertheless, our approach comes with some limitations:

Our app only logs text entered by the participant. Hence, we can neither easily study full conversations, nor distinguish between receivers of text within one app. Extensions might, for example, extract chat partners from notifications (cf. [344]), although this raises issues of logging data from non-participants.

It is also difficult to use text abstractions that require word/category lists in open vocabulary work [362], which explores occurring terms instead of a defined set. This may also apply to slang terms. This could be partly addressed by defining a very comprehensive whitelist, or by informing word lists with a pre-study (e.g., to collect slang terms).

Furthermore, text abstractions inherently limit what kind of text representations are available. For instance, many computational methods in Natural Language Processing use word embeddings (i.e. words represented as high-dimensional vectors, e.g. [313]): In a privacy view, these embeddings are no different to logging words, since usually embeddings can be turned back into words. Hence, we do not support directly logging words as embeddings. However, the word data that is indeed logged with our tool (e.g., whitelisted words) could be converted to embeddings post-hoc (albeit not contextual ones [378]). Regarding further word metrics, a future version might be extended, for example, to count word co-occurrences for the whitelisted words.

Finally, introducing a keyboard app is unlikely to match people's own keyboards exactly. We used a common open-source keyboard from Google to optimise familiarity for many Android users. Nevertheless, people noticed differences for auto-correction and small UI design choices. While we do not expect these to considerably impact on language use, we plan to improve on this in future work (e.g., via options for UI customisation and importing existing user dictionaries).

5.2.6 Discussion

5.2.6.1 Logging Everyday Mobile Language Use

We have shown that it is feasible to log everyday mobile language use in a privacy-respectful way, using a keyboard app with integrated text abstractions. Researchers can further minimize the potential of reconstructing raw text from such abstracted data

by logging enough data per person. In our experiments and study, 50-100 sentences comes close to the minimum and can easily be achieved within a two-week study (for many people in our study much earlier, 1-2 days).

Data collected with our method and tool can meet many sought-after criteria, such as 1) covering all mobile text communications, 2) including personal ones (e.g., chats), 3) unobtrusive long-term measurements, and 4) varied natural everyday contexts. Very few people felt that the study influenced the content of their writing. People in our study also typed in their usual everyday apps, including chat apps and web browsers. These results are all positive with regard to achieving comprehensive and unbiased data collection. We thus conclude that our method and tool present a valuable addition to the toolset for research intersted in everyday mobile language use.

5.2.6.2 Remaining Privacy Risk

As in related work [71], our goal is to facilitate privacy-respectful studies – not to counter malicious attacks. Responsible study planning still comprises more than a keyboard (e.g., secure data storage). Besides attacks, it might be possible to gauge a user's interest in some topics from word counts (e.g., one interviewee mentioned rare words over-represented in his favourite band names). If this is to be avoided it can be addressed with a more selective whitelist for word counting. Overall, the choice of whitelist and category mappings influences what can be inferred from the abstracted data. Our app supports tailoring these settings to research questions, which can limit collected data and possibly unwanted inference opportunities. In all cases, our method avoids logging raw text.

5.2.6.3 Privacy Perception

The majority of participants found that the text abstraction methods contributed to protecting their privacy in our study. In more detail, perception of privacy was individual: In the interviews, some participants stated quite low concern and did not deem it necessary to use our provided opportunity for a log data review. Others were more sceptical and would have preferred more details about the logging system's inner workings. Fittingly, participants' scores also varied for the questionnaire on general

privacy concerns [70]. Individuality of perspectives is also indicated by the fact that the log data review influenced participants' perception of the abstractions in both directions (see next discussion point).

5.2.6.4 Log Data Review

In contrast to other methods, ours does not require participants to manually review data as text abstractions are applied automatically. We still tested the idea of data review, using a web UI after the study. Seeing their own abstracted data influenced some people's perceptions of the abstraction methods yet not in a systematic way overall. One exception was Whitelist Counting, which four people found less protective after the review. Crucially, motivations for rating protection more critically after seeing the data were not related to fears of raw text being reconstructed, indicating that our method succeeded with regard to this main concern. Instead, people noticed potential inferences of a more general kind (e.g., one interviewee wondered if a high count of "I" indicated egocentrism). Feedback generally showed that people liked this view on their data. Overall, we conclude that showing actual logged data could replace generic examples to inform participants of how the data is processed. Future work could also integrate such a view into the keyboard app directly.

5.2.6.5 Communicating Privacy-Aware Logging

Two people avoided entering passwords and locations. We do not record password fields and neither passwords nor specific location names are in the whitelist. Fittingly, one interviewee explained that she might not have behaved differently if she had better understood how the system worked exactly. We see a tradeoff between 1) providing extensive (technical) detail to facilitate trust and 2) not overwhelming people (cf. long terms of use). Future work could integrate some of these explanations into the app as part of a typical first-launch intro, which users know from many apps today.

5.2.6.6 Further Opportunities for Research and Applications

We illustrate the rich opportunities enabled by our method and tool with ideas for future studies, methods, and applications.

Analysing the Use of Non-Verbal Cues in Mobile Messaging An active line of research analyses non-verbal cues, such as emojis, for example to reveal misunderstandings [102, 284] and improve UIs [320, 321]. Recent work using questionnaires [409] concluded that real-world data collection is needed yet warned about privacy challenges due to logging private messages. Our tool enables such studies: For instance, researchers can set regular expressions that capture emojis in context (cf. [320]).

Long-Term Observations of Everyday Language Use Running silently in the background our app enables long-term observations without repeatedly asking people to do study tasks. For example, such studies could investigate changes in language use when a person moves to a new location, social circle, and/or starts to learn and use a new language. Related work also indicates changes with age [362].

Enriching Data Collected with Other Methods Our approach can be combined with other mobile study methods, such as experience sampling (ESM) or mobile (context) sensing: Studies could use ESM to ask users about subjective experiences (e.g., mood) and now relate this to in-situ language use collected with our tool at and around that moment. More generally, such combinations of different objective and subjective information channels enrich observations [72, 401].

Informing Intelligent Text Entry Systems Word suggestion and correction algorithms could use datasets on language use collected with our method, for example, to address the cold-start problem for new users of a keyboard. Moreover, intelligent reply systems (e.g., Google Smart Reply [212]) could consider a user's word usage to generate replies that are in line with personal style and language use.

Personalising Intelligent Mobile Applications Data recorded with our method could also inform intelligent systems. For instance, a chatbot might prefer words that fit the user's logged frequent vocabulary to avoid misunderstandings. Related, educational apps could be personalised to help users learn a foreign language, either by estimating progress from words used in that language or by teaching translations for the used words, increasing relevance for the learner (cf. [317]).

Integrating Logging Beyond the Keyboard LanguageLogger is implemented in a modular way. Thus it could be incorporated into other Android applications, dealing with other textual data than keyboard logs. For example, incoming notifications could be analyzed regarding their content. This could enhance previous studies which looked only at notification categories [344], e.g., to now also analyse content topics and sentiment.

Integration with Other Privacy Approaches There exist also other approaches to collect data in the wild in an anonymous way: For example, differential privacy [116] and randomized response techniques [138] add systematic noise to the data, without changing the full dataset's characteristics. However, such noise might be less suitable for studying language use of individual users, which is important for research in psychology. Such different approaches could also be combined: For example, differential privacy for a specific use case could be applied on top of our tool, e.g., as noise on our extracted word frequencies.

5.3 Fine-Grain, Continuous Smartphone Permissions

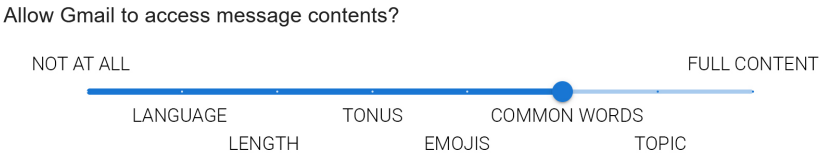


Figure 5.14 : Privacy Slider enables users to select which granularity of their data they want to give to smartphone apps.

> This section is based on the following publication:
>
> Upcoming: Florian Bemmann, Helena Stoll, and Sven Mayer. “Privacy Slider: Fine-Grain Privacy Control for Smartphones.” In: *Proc. ACM Hum.-Comput. Interact.* MobileHCI ’24 MHCI (2024). DOI: 10.1145/3676519

Permissions are nowadays most prominent approach to implement control on smartphone sensed data. However, only around 6% of users understand the scope of the permissions they agree to [368]. Shen et al. [368] argued that current mobile systems hardly convey to users what happens with their data and which specific data is used. Moreover, apps request permission, such as location, to display weather information and full device access to support accessibility. However, the weather could be forecasted by only knowing the current city, and the screen readers only need the screen content – not full access. Current permission systems do not allow fine-grain control but only some toggle switches for groups of data access [248]. In combination, users are often unaware of what they agree to and do not have the needed control when understanding the specific case. Thus, users need more transparency in the process coupled with better control of their data.

With smartphones becoming more intelligent, apps and services require increasingly more contextual user data to provide high-quality support, e.g., O’donoghue and Herbert [305]. To reduce potential privacy issues, the researcher has proposed a wide range of mechanisms to preserve the users’ privacy. The simplest solution is outsourcing the decision-making process to an algorithm [142, 299, 325]. However, it takes all control from the user. Thus, Gao et al. [162] opposed recommending decisions

to the user and not just taking them for them. Moreover, decisions might be context-depended [431]. This raises the question: Are all-or-nothing decisions, as they are implemented with the current grant or deny toggle switches, a good option in the first place? The current operating system addresses this by adding a frequency component to the permission, which the user can set for how long the permission is valid. However, the underlying decision as to which specific data the app has access remains the same. For this, Olejnik et al. [299] added an option "obfuscate" as an alternative to "deny" and "grant." Here, obfuscate helps to retain privacy which 73% of users found useful. An alternative approach is by Malviya et al. [270], who envisioned that fake data could be used to make a function work while protecting the users' privacy. While these are options to increase the privacy level of the users' data, they lack the ability to allow the users to determine the abstraction level of data to be shared. In other words, the user will still be given full device access and meter-precise location. In summary, the permission interface has to meet the tradeoff of giving users detailed control while sufficiently summarizing and reducing the options so that users are not overwhelmed.

In this work, we aim to address the limitations of today's all-or-nothing permission control systems. As such, we designed and developed a fine-grain permission system. In detail, we envision replacing today's toggle switch with a *Privacy Slider* to hand full data control back to the user. For this, we first conducted an online survey ($N = 123$) to understand what control users envision they need to gain control over their data. In detail, we asked how they could subdivide the all-or-nothing permissions; for instance, the location could be subdivided using the location precision. The results show a total of 135 potential steps for ten datatypes. Next, we investigated the user's concern concerning the steps using an online study ($N = 109$). This allowed us to rank and potentially group the steps to optimize usability; we supported this step by conducting a workshop ($N = 5$) before starting a multi-stage design and development process of our *Privacy Slider*. In a final evaluation ($N = 32$), we compared *Privacy Slider* to the classical toggle switches in two scenarios: a permission popup and the device permissions screen.

We make a set of contributions all leading up to the design and implementation of *Privacy Slider*. Our first two studies show how sliders can support control and transparency based on steps between today's all-or-nothing choices. The subsequent implementation showed that *Privacy Sliders* significantly outperformed toggle switches

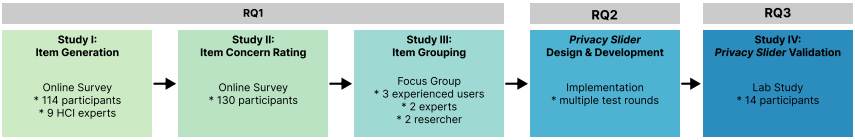


Figure 5.15 : The development process of *Privacy Slider*.

with respect to privacy, security, control, transparency, and understandability. Thus, *Privacy Sliders* have the potential to support users in making better decisions when controlling their devices' permissions.

5.3.1 Research Gap and Derived Concept of a Continuous Permission System

Only a few studies investigated the user perspective of control features (e.g., Pennekamp et al. [312] rated usability themselves). Finer-grained permissions (e.g., Scoccia et al. [363]) were rarely studied, and, if so, emphasize technical aspects rather than the user. Moreover, the choices given to the user by today's systems are typically binary, most prominently toggle switches. At the same time, it is clear that choices are often not binary. For instance, to get the weather forecast, apps do not need the user's precise location; the city the user is in would be enough to determine whether an umbrella is necessary today. However, current systems only allow an all-or-nothing choice, which causes many privacy concerns and allows for more potential to infringe on users' privacy than necessary.

As a result, we witness the need for more fine-grained mobile permission systems that enable users to configure their data logging on steps between granting access to all or nothing. For this, we pose three research questions:

RQ3f *What are helpful sub-steps for fine-grained data control?*

RQ3g *How do we deliver the additional control that is usable?*

RQ3h *How does it perform compared to the existing Android permission UI?*

We conducted four studies to address these questions; see Figure 5.15. In Study I ($N = 123$), we investigated the potential items that are of interest to users to be controllable. In Study II ($N = 109$), we asked users to rate the items concerning

their privacy concerns. We used these concerns in Study III ($N = 5$) to rank and group the items into semantically similar groups. Moreover, we investigated possible representation methods, such as how much abstraction and control users want. In combination with multiple rounds of testing and debugging, we developed the final look and feel of *Privacy Slider*. Lastly, we carried out an A/B testing ($N = 32$) to ensure the usability outperforms the industry standard using toggle switches to control users' data.

5.3.2 Study I: Item Gathering and Concern Rating (Online Survey - RQ3f)

In this study, we investigate what users would like to control beyond current binary options. Here, we intentionally ask users and experts to gain both perspectives: a) What do users understand to be important for them? and b) What could developers see as important to enable certain applications? Therefore, we conducted an online survey ($N = 123$) with smartphone users and HCI experts to reach a sample of diverse experience levels regarding UX and privacy.

We asked participants to envision the use of 10 different datatypes to explore new avenues for a future permission system. Here, we prompted them with the following ten datatypes: *app usage*, *camera usage*, *incoming message*, *notification*, *phone calls*, *screen content*, *text input*, *user activity*, *voice input*, and *volume & brightness*. The datatypes are rooted in a combination of typically tracked and collected data [344] and common activities [57]. Moreover, all of them can be tracked today and are typically controlled via phone permissions.

5.3.2.1 Procedure

First, we explained the procedure and content of the study and asked to give informed consent. Next, we asked participants' demographic data such as age, gender, education, and professional field. See the complete question in the supplementary materials. For each datatype, we collected participants' ideas using the question: *"Which intermediate stages would you find useful?"* Participants did this for ten specific datatypes. Additionally, we asked participants for the logging frequency, which is a property overarching over all datatypes. At the end of the survey, we rewarded participants with 9 GBP per hour.

5.3.2.2 Participants

In total, we recruited 123 participants. We recruited 28 participants from our institution and an additional 86 participants via Prolific to diversify the sample. Additionally, we supplemented the sample with 9 HCI experts whom we personally recruited to reflect expert opinions. We required all participants to use a mobile phone or tablet at last almost daily. Participants were between 19 and 74 years old ($M = 29.1, SD = 9.4$), and 64 identified as female, 58 as male, and one as diverse. The majority reported a university degree as their highest degree of education (85), 19 had a high school degree, 28 had a high school diploma, 5 had a completed apprenticeship, 4 had a secondary school degree, and one participant finished school without graduation. Their top 5 professional fields were IT, electrics and engineering (45), economy and logistics (14), social and pedagogy (12), services and sales (11), and arts and media (8). In total, our sample resided in 19 different countries, most from Germany (37), South Africa (33), Portugal (10), and the United Kingdom (9).

We determined the targeted sample size on the go via thematic saturation, i.e., when the recruitment of further participants did not reveal new steps [259]. We stopped recruitment when the saturation index reached a threshold of 90%, which, through the underlying information weighting model, expresses that the probability of mentions being shared between existing and new participants is 90%.

5.3.2.3 Results

We used Python, R, and Atlas.ti to analyze the data and ensure the validity of the responses. We received 1339 individual feedback statements for the different data types. We analyzed our participants' responses using Affinity Diagramming [183]. We formed code groups per datatype. Through this process, we sorted the 1339 statements into 135 distinct codes. On average, each participant contributed 9.9 codes ($SD = 11.9$). Figure 5.16 gives a visual overview of the final groups.We retrieved the most distinct codes for the datatype *text input* (21 distinct ideas), *phone calls* (16), and voice inputs (16).

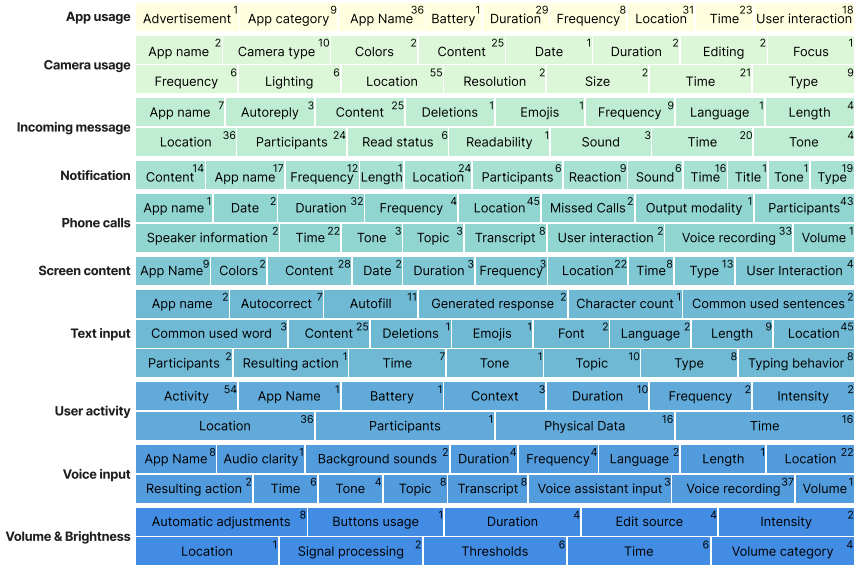


Figure 5.16 : The 135 codes for the ten datatypes that emerged from the 1339 participant statements.

The most common mention across many datatypes is the location at which a logging event/data item took place, i.e., the location of a physical activity, the location the user was at when receiving a text message, etc. It was mentioned as the most frequent step in 6 of 10 datatypes and was mentioned second frequently for the remaining 4.

For datatypes where it is appropriate, the name of the respective data item was mentioned often; i.e., for datatype *physical activity* the activity's name, and for *app usage* the app's name. The point of time was within the top 5 steps for all datatypes except *text input* and *voice input*. Also, the duration was mentioned often (e.g., on rank 4 for *activity* and *phone calls*, and rank 3 for *app usage*).

5.3.2.4 Summary

In total, our 123 participants came up with 135 distinct codes that are important for the ten scenarios. Many steps are common for multiple data types; for example, *App Name* was mentioned with all but one, or location was mentioned for every data type. The

frequency with which steps are mentioned varies. While *location* is among the top 3 most mentioned steps for most data types, it is among the two least mentioned steps for data type *Volume & Brightness*. However, we do not know for sure whether the mentioning frequency is an indicator of importance, relevance to the user, or privacy concern. It may rather depend on how present an aspect is in the users' minds. Existing research shows that people are initially rather unaware of privacy risks and do hard naming their concerns unless they are confronted with the topic (cf. [157]). The order of the collected steps is thus neither given by the survey participants nor naturally by the aspects' characteristics. To obtain a concern ranking, which is required in order to place them on a slider scale, we conducted another study.

5.3.3 Study II: Item Concern Rating (Online Survey - RQ3f)

With the results of Study I, we next investigate how the 135 codes (see Figure 5.16) are rated with respect to their privacy concern. It will inform the design of potential mechanisms to allow users to make fine-grain console adjustments. Thus, we conducted another online survey ($N = 109$) and let users rate their perceived privacy concerns for the codes.

5.3.3.1 Procedure

First, we explained the procedure and content of the study. Afterward, we asked them to consent to the data recording and storage. Next, we asked participants' demographic data such as age, gender, education, and professional field. For each step that resulted from the item gathering study (cf. Figure 5.16), participants had to rate their agreement with a statement worded "I am very concerned with my smartphone tracking *duration of the activity (e.g., 1h)*" on a continuous (101 point) slider item [156, 335]. We grouped the items on survey pages by datatype, presenting in total 135 items distributed over 10 pages to the user. Each page started with a short paragraph reminding the participants of their task and context. Steps for the overarching property *frequency* were not ranked, as they are all-time indications that have a natural order. We disclose all study instruments in the supplementary materials of this paper.

5.3.3.2 Participants

In total, we recruited 109 participants (42 from our institution and 67 via Prolific). As in the first study, participants had to be fluid in English or German and use a mobile phone or tablet at last almost daily. Their ages are between 18 and 60 years ($M = 29.7, SD = 9.5$), with 60 identifying as female, 47 as male, one as a diverse participant, and one who preferred not to disclose. The majority reported a university degree as their highest degree of education (65), 34 had a high school degree, 3 had a secondary school diploma, 5 had a completed apprenticeship, and two participants finished school without graduation. Their top 5 professional fields were IT, electrics and engineering (31), economy and logistics (13), health (11), social and pedagogy (9), unemployed (6), service and sale (5), and arts and media (2). 32 did not identify themselves with the given groups and specified other professions. In total, participants reside in 17 distinct countries. The most represented countries of residence were Germany (43), South Africa (21), Portugal (12), and Greece (4).

5.3.3.3 Results

The participant's general privacy concern ratings over all items are in the middle of the 1-100 scale ($M = 45.8$, $SD = 33.8$). The concern values are distributed in a rather bimodal distribution , i.e., the fewest values are in the middle range around 50, and most values are either very low or very high. A slight tendency to the left shows that people were more often rather less concerned than rather high.

Taking a look at the stated concern value grouped by datatype, we see that spoken/written contents were rated most concerning, with phone calls having the highest concern ratings ($Mdn = 59$, $SD = 35.1$) and voice input the second highest ($Mdn = 55$, $SD = 34.8$). Text inputs and screen contents are both on rank three ($Mdn = 46$, $SD = 33.8$). The lowest concern ratings were stated for volume and brightness ($Mdn = 24$, $SD = 31.4$). The values for all ten datatypes are in Figure 5.17.

5.3.3.4 Summary

As a result of Study I and II, we created a collection of steps for a set of common data types. We have not performed any reduction or grouping of steps yet. However, to not overload the UI, a reduction to a reasonable amount of steps has to be considered.

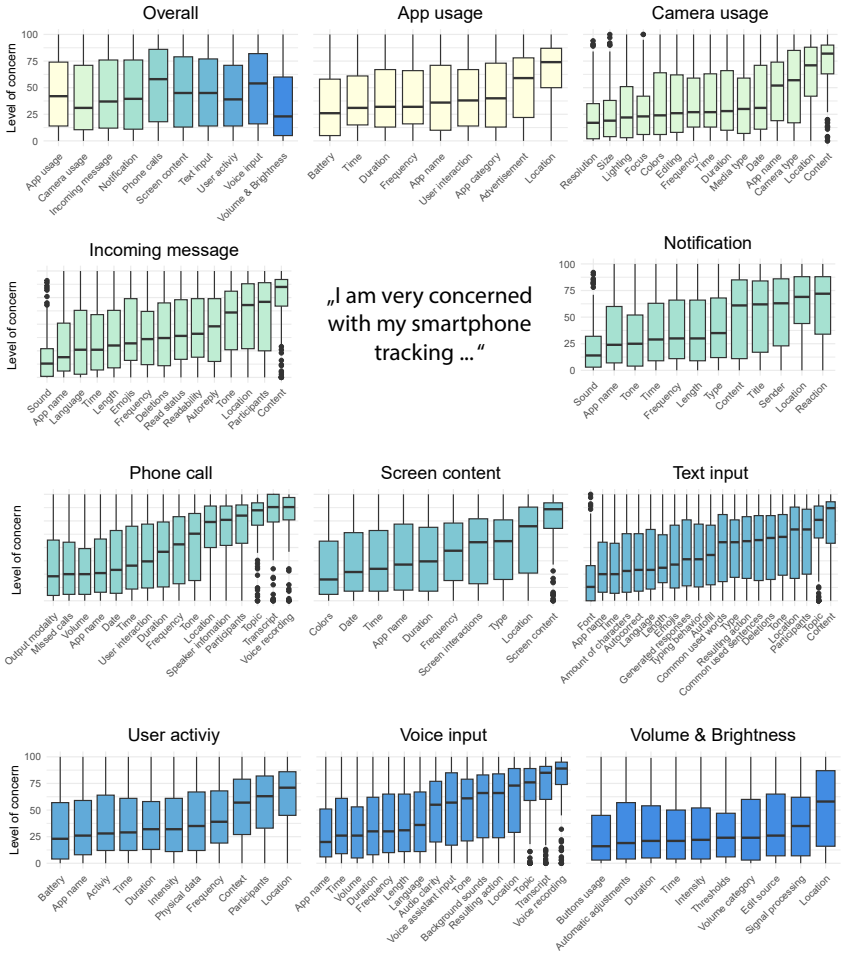


Figure 5.17 : Boxplot of the concern ratings in our item concern rating study. The black line indicates the median.

To group steps, we see two general options: (1) Numerically, i.e. merging steps with close-by concern rating into a group, or (2) semantically. To make this and other design decisions we will in the following conduct a focus group.

5.3.4 Study III: Focused Exploration (Focus Group - RQ3f)

In the previous two studies, we have collected steps that represent subaspects of logged data, accompanied by privacy concern ratings. The results of both studies suggest that it is possible to subdivide today's toggle switches to set up permissions. Study I gave a wide range of steps between all or nothing, and Study II ordered them according to the users' concern level. While this allows us to put all steps on a slider, ensuring full transparency and control; to the users; however, with so many steps, the usability might suffer. Thus, next, we investigate potential user-facing presentations of a novel permission system using a focus group. This investigation will inform the design.

5.3.4.1 Procedure

After explaining the focus group and answering any open questions, we asked participants to fill out a consent form and demographics questionnaire. Then the participants introduced themselves and were introduced to the topic. We introduced the general idea and discussed the differences, pros, and cons between toggles and sliders. As the next step, the focus group leader presented the collected steps from Study I augmented using the concerns of Study II for each datatype (see supplementary materials). With this information, we asked the participants to envision to group and potentially sort the steps with the goal of keeping transparency and control high but at the same time also the usability. Finally, the group discussed how the sliders could be best integrated into smartphone usage, i.e., when users would use them and what is needed for effective use.

5.3.4.2 Participants

The focus group was conducted with 5 participants (3 female, 2 male) and led by the two first authors. The participants were between 24 and 29 years old ($M = 26.6$, $SD = 2.4$). We were aiming for the perspective of both experts and the user side,

so we recruited two experts and three smartphone-experienced users. The experts were HCI researchers who are currently pursuing their Ph.D. The three users were two students and one public service employee.

5.3.4.3 Results

The focus group took just over 1 h. We did an audio recording and transcribed this into a text file, which we then coded using ATLAS.ti. This resulted in 49 distinct codes that were assigned in total 90 times. We grouped the codes into 7 code groups, which constitute to topics presented in the following.

Use Case and Target Audience of Privacy Sliders The focus group identified privacy sliders as **a way to simplify privacy settings for users who are not willing to spend time or do not have sufficient technology understanding** to do so. P1 mentioned *"I do not think the idea of the slider is a bad one because it would make it easier for a lot of people who do not deal with such things that much."* - [P1], and P2 came up with a concrete example of a family member *"[...] I can tell my mom, that once you are halfway through the slider, the app can do more but it also knows more about you."* The participants also pointed out that it is important not to curtail expert users in their control over their data, i.e., still **have detailed control options available on demand** *"But maybe it would be good, as [P2] says, that you would then still have the possibility to set individual things differently."* - [P1] and proposed separate toggles e.g., to turn off single aspects. P2 summarized the combination of privacy sliders and on-demand toggles as *"Good thing for people without much IT knowledge [...], but if I deal with it, then I can use the fine limb ticks."*

Privacy Slider Scope The focus group came up with the idea of having **one central, default privacy configuration that overarches all apps and datatypes**: *"I would expect to be able to set it in the system once, all abide by it and if they need something extra then I am asked."* - [P1]. When users are about to perform some action they are least willing to deal with privacy configuration and would benefit from a default configuration *"If you want to take a picture, it should be fast but first you have to adjust everything"* - [P3]; *"Therefore already before!"* - [P1].

Finer Granularity for Continuous Datatypes When the group discussed how much sense sliders would make for specific datatypes, P1 and P2 came up with examples of use cases where a reduced granularity of data would be sufficient. P1: *"I have an example for the location theme. If you could then just go in again, and change that again, then it's enough if I can set: City, State, Country. Or I would like to say more precisely that I am currently in XY street."* The group had the opinion that ordering steps for a slider specifically on location data is easy, concluding that **a slider to control location granularity** would be good. *"Although I would say with location, within locations I would find it exciting if it was a slider. From not at all, to city or urban area at 300m."* - [P1]. Similar ideas arose for content, such as texts, speech, and images. *"Kind of like the direct content, it is obvious what's on the far right and everything before that is what you can infer through the content. So like tone, emoji, language."* - [P2]. A level of *"content abstraction"* [P2] was proposed as a continuous scale that could be mapped onto a slider.

Grouping of Steps Participants suggested to rather **group steps by topic** instead of strictly adhering to privacy concern levels. They proposed various groups that make sense to them, e.g., P3 suggesting that *"Time and Duration could be put together, so everything that has this time and duration aspect."* or P1 who distinguishes between personal and contextual data: *"I would try to separate it like this: personal data, the data that is more context and something like location or context plus data related to something like app name."* A general **desire for grouping** was expressed especially in cases where many steps exist *"It's just a lot. So you couldn't display it like that on the slider, you would have to group it in any case."* - [P3].

Ordering of the steps We discussed the order of steps in the focus group, which was derived from Study II (Item Concern Rating). Participants **overall agreed with the resulting order**, however, the difficulty of deciding on an appropriate order varied with the datatype. P1 and P2 stated that they **did hard ordering the steps of app usage and activity**, while they found that ordering text input went intuitively easy through the degree of content abstraction. The focus group participants could not comprehend why camera type and duration were rated relatively concerning, while physical data and emojis received a surprisingly low concern rating.

The Relation Between Data Privacy Concern and Importance When discussing the privacy concern rating in turn of the step order, the focus group also discussed whether the concern is the right ordering criterion in this case. P3 mentioned that she does not find the location very private if it is really necessary: *"I do not think that's so bad, because you need the location for many apps. Be it renting a car or scooter or Google Maps."* In line with this, the idea is to order by the ratio of privacy sensitivity and importance in a specific use case.

User Desires and Design Decisions When a system contains multiple privacy sliders (e.g., for various datatypes), it is important that **consistency of similar steps positions** is guaranteed. Otherwise, users might face unexpected behavior. *"The location should always be specified the same. Also, if you now have different sliders and I would now set everything to 60 or so, then I would also expect that it is somehow everywhere the same "safe." And if then suddenly a location is already at 50, then that would be super stupid, just because I do not want to read every time."* - [P1]. The exact position of steps on the slider was not deemed that important, P2 suggested mapping them with equal distances instead of trying to represent the exact concern values.

5.3.4.4 Summary

The focus group gave us a good understanding of people's opinions on how our insights from Study I and II could be fused into a privacy slider design. They agreed on the appropriateness of sliders, especially for datatypes that impose a natural order, such as location or content. A slider interface might especially be beneficial to non-expert users, but we also take away that it is important not to restrict expert users by removing detailed controls. While the idea of having a system-level slider to set a default privacy policy was liked, participants also pointed out the advantages of runtime permissions, aligning with findings in the literature (e.g., [363]). The privacy slider design should thus incorporate both concepts.

We noted that, in our focus group, the expert participants had higher speaking shares. We perceived that the non-expert users had hard imagining the slider concept in-depth. We, therefore, conclude that a study with an implemented prototype is necessary to get a sufficient user perspective.

5.3.5 Privacy Sliders: The Final Design (RQ3g)

Based on our two surveys, the focus group, and a review of related work, we propose a concept for privacy sliders – a novel user-centered mobile data permission system. Privacy sliders realize two central aspects: First, a simpler, easier user interface that enables quick and easy privacy setting-making. It targets users who are either novices or not willing to spend much time on their data privacy configuration. Second, privacy sliders enable users to choose a custom level of granularity, at which they want to allow to pass data to an app. Both are presented in detail in the following two subsections.

5.3.5.1 The System-Level Settings Slider - Sliders as Simplification for Fast and Consistent Privacy Configuration

We propose to implement one slider as a central, default privacy configuration, which overarches all apps and datatypes. A prototype is sketched in (Figure 5.18): The more to the right the user pushes the system-level slider, the more detail is granted for every datatype. To meet expert users' needs and individual needs on specific datatypes, the access level to a datatype can be overwritten. One slider per datatype allows overriding the system-level slider's setting (e.g., in Figure 5.18: For the location, the user has configured lower granularity data access).

Based on the results of our focus group we envision that this simplifies users indicating their privacy preferences. Especially novice users and those who do not want to spend much time on privacy configurations might benefit from the intuitive and fast UI of a slider.

5.3.5.2 Enhanced Permission Popups: Information Minimization of Continuous Datatypes

In the focus group, we found that for some data types, such as location and content (text, speech, and camera were mentioned), it makes sense to configure granularities. Steps for location data could, for example, be reduced to an accuracy of +/- 500 m, city, or country. For many use cases, that might be sufficient: For example, when using a weather forecast app it would be sufficient if the OS passes the city name to the app, instead of the user's precise location. For content, such as text messages, content abstraction procedures could similarly be applied.

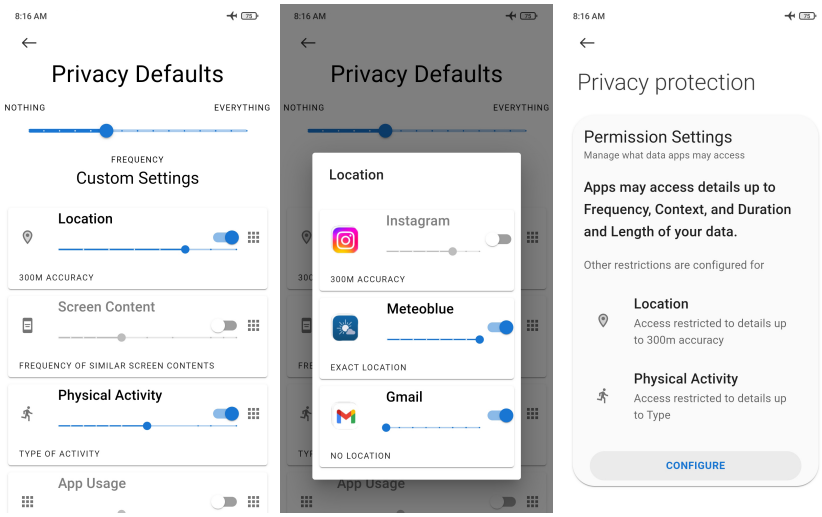


Figure 5.18 : The system-level slider (left) is used to configure the phone's general privacy settings, per default applying to all apps and overarching all datatypes. It is especially targeting users who do not want to spend much time with single privacy decisions. The sub sliders below allow to set overriding configurations for single datatypes. A popup (middle) allows to make settings per app, and the overview screen (right) summarizes the settings.

Supplementing the previously presented system-level slider, our privacy slider concept introduces such configuration options for location and content. These sliders (see Figure 5.19 (left) and (middle)) are meant to replace the current permission UIs, e.g., the *only one time, always, when using the app* single-choice radio button interface that Android currently uses for location access permission.

5.3.6 Study IV: Slider Validation (Lab Study - RQ3h)

To evaluate our privacy slider concept, we implemented it as a prototype and conducted a lab study. Participants were asked to use both a UI mockup of a traditional permission interface and a mockup of the privacy slider interface concept. We assessed both interfaces' effects in a mixed-method approach, using survey items and interview questions. The study consisted of 4 scenarios, with participants going through them two times, once using traditional Android permission UI and once more using the

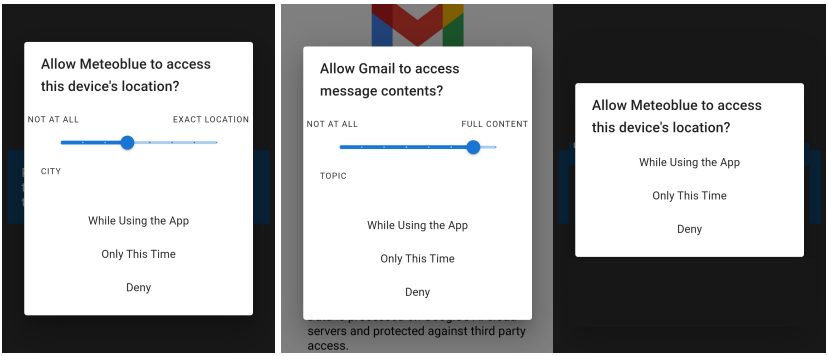


Figure 5.19 : We enhance permission popups with a slider that allows to choose a level of granularity. The screenshot on the left visualizes this on the example of the location permission, the middle one for text message contents. On the very right we show our control condition, consisting of a slider-less permission popup as it is implemented in the Android UI nowadays.

slider interface. Three of the scenarios were runtime permission popup situations (see Figure 5.20), and the remaining one was the general privacy settings menu deploying our system-level slider (see Figure 5.18). The order in which the four scenarios were presented was randomized.

5.3.6.1 Apparatus

We mocked the Android permission UI (runtime popups see Figure 5.19 (right), and the settings menu see Figure 5.18 (right)) with a Progressive Web App[1]. The runtime permission popup scenarios consisted mainly of a series of app screenshots that the users clicked through, augmented with button respectively slider UIs that mimic the permission interface. Launched in fullscreen mode and being tailored specifically to the study device, the UI experience was very close to real Android UI. To rule out the effects of the mock, we also mocked the traditional Android permission popup UI with this approach, instead of using the OS implementation.

Slider Steps: System-Level Slider In Section 5.3.2 we collected potential steps for privacy sliders for each datatype and ranked them by their level of privacy concern

[1] https://web.dev/progressive-web-apps/, last accessed 2024-12-06

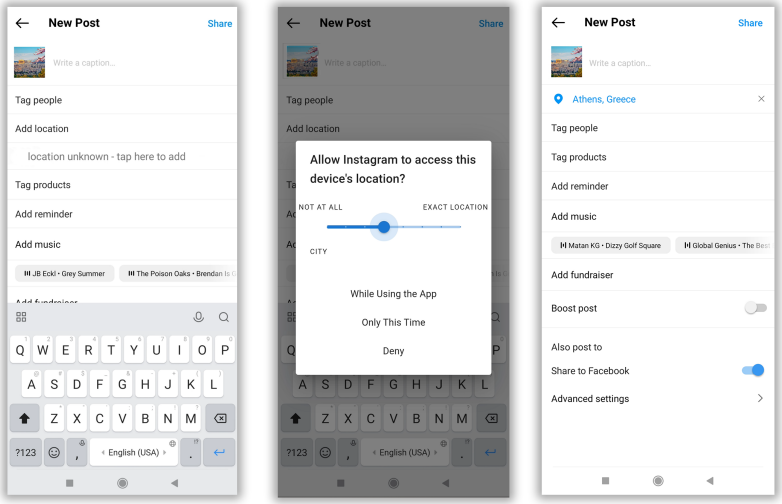


Figure 5.20 : A series of screenshots that shows one of the scenarios that we used in our studies from left to right: Here the participant is advised to craft an Instagram post, that is tagged with its location. This figure shows the privacy slider condition of the experiment.

in Section 5.3.3. However, in Section 5.3.4 we found that a strict order by privacy concern does not make sense to users, as the varying order of similar steps on different datatypes might lead to unexpected configurations. We thus follow the idea of the focus group to group the steps by topics *"that make sense."* These groups then constitute the steps of the system-level slider and its sub-sliders (except the continuous datatypes location and content). Thus the sliders for all datatypes are designed equal. We order the steps on their slider by the median concern value that our participants in Study II rated them.

Slider Steps: Location and Content As pointed out by the focus group, location and content pose an inherent granularity, which can be mapped to a continuous slider design. We chose the following steps, based on mentions from the focus group and proposed order in Section 5.3.3:

Location: not at all, country, state, city, urban area, 500m, 300m, street name, exact location

Content: not at all, language, length, tonus, emojis, common words/sentences, topic, raw content

5.3.6.2 Procedure

The study conductor met each participant in our lab in a separate room with a table. After explaining the study, the participants read and signed the consent form. We then started with a questionnaire on one's individual information privacy concern level using the IUIPC questionnaire [268]. Furthermore, we assessed affinity for technology interaction (ATI) [151] and demographics.

To get participants into the thinking mode and to affiliate with the scenario of giving permissions, the study conductor talked with them for a couple of minutes about their last contact with smartphone permissions, what they thought and felt in that situation, and what the decision was like. Then we introduced them to one of the two conditions *Classic* or *Slider* in randomized order. For both conditions, we ran participants through the same procedure: a demo of the condition, then they tested the three showcases, and finally, they tested the system-level settings application. After each but the demo, they filled in a system usability scale (SUS) [214] questionnaire and answered six items on perceived control, privacy, security, making sense, transparency, and understanding. These items were assessed on a continuous slider scale, we disclose the wording in the supplementary materials. At any time, participants could verbally or in writing articulate additional feedback.

In the end, we asked participants additionally if they had any further feedback. We audio-recorded the full procedure and rewarded participants with 10 EUR per hour or the respective amount of study credit points. Participation took approximately 30 minutes to complete the study.

5.3.6.3 Participants

We recruited 32 participants via our university mailing list, Slack channel, Instagram, and personal contacts. We required participants to be daily smartphone users and be fluent in English. Participants were between 21 and 70 years ($M = 28.0$, $SD = 8.5$),

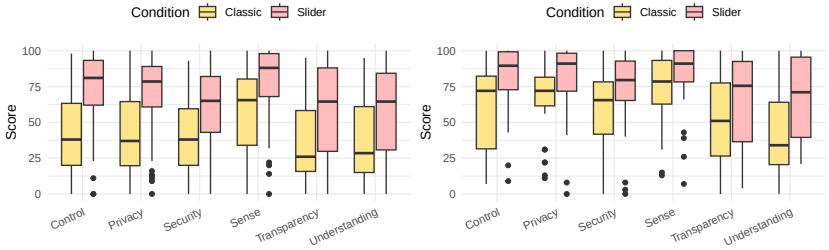


Figure 5.21 : Ratings on five aspects around privacy compared for the classic permission UI and the slider UI. On the left regrading permission popups on runtime, on right side for the device's settings menu.

with 18 female and 14 male participants. They reported having a Master's degree (16), a Bachelor's degree (9), a high school degree (4), a doctoral degree (1), not finished school (1), and a vocational education (1). All participants reside in Germany, besides one participant from the United States. To understand our sample's privacy perception, we assessed the IUIPC questionnaire [268]. Our participants rated their Awareness on average with 6.2 (SD = 1.3), Control with 5.7 (SD = 1.3), and Collection with 5.7 (SD = 1.2) (higher scores mean more privacy-affine). Their mean score of affinity for technology interaction (ATI) was 4.1 ($SD = 1.0$). This indicates a rather technology-affine sample. According to the classification of Franke et al. [151] the ATI of an average population is to be expected at around 3.5, with high ATI samples around 4.

5.3.6.4 Results

We run the statistical evaluation in Python and R. Moreover, we applied non-parametric tests when the normality was violated. We did the qualitative coding again in Atlas.ti.

Usability The users' rating on the system usability scores (SUS) was higher for the slider UI than for the classical UI, for both the runtime permission popup comparisons (see Table 5.8) and the system-level settings menu (see Table 5.9). According to the adjective classification of Brooke [67] all interfaces' usability can be regarded as *excellent*, only the classic device settings menu was rated as *good* only. However, only in the runtime comparison does the slider significantly outperform the classic approach.

Privacy Effects For the system-level settings, we can use the classical Wilcoxon signed-rank test after confirming the non-normality of the data. However, for the permission runtime popups, we perform an ART-ANOVA [436] with task and participant as random factors to account for the differences between the three showcases. We compared the effects of the classic and the slider UI on their users' privacy perception, see Figure 5.21. We show that *Privacy Slider* outperforms the classic approach significantly in nearly all measures. The only non-significant items are the measures for usability and making-sense in the system-level settings menu. However, descriptively the slider performs better also for these items. See Table 5.8 and Table 5.9 for all measures and test results.

Qualitative Feedback We coded the free text responses, transcribed audio recordings, and interview notes in Atlas.ti. We then organized the codes into code groups, which constitute the following topics.

Slider Interface is Preferred over Classic Button Interface

In general, comments on the slider interface were better than on the classic button-only interface. Many participants mentioned that they did prefer the slider interface, while none said that they'd rather like to stay with the classical version. P39 described it as *"a big improvement over the usual UI and would definitely prefer this in all cases."*. Similarly, P12 *"No, I think its a great addition."* and P21 *"liked it much better than the previous one."* In contrast, the classic condition was described as having *"not enough options for data privacy"* (P26) and being *"too general."* P53 expected a *"more*

	Runtime Popups							
	Classic		Slider		Normality		Wilcoxon*	
	M	SD	M	SD	W	p	W/F	p
SUS	81.9	9.2	88.2	11.1	.925	<.001	88.5	<.001
Control	39.7	26.2	73.7	24.8	.939	<.001	133.98	<.001
Privacy	40.6	25.9	70.5	25.0	.939	<.001	107.21	<.001
Security	39.8	24.4	61.9	26.1	.964	<.001	65.951	<.001
Sense	59.7	28.2	81.0	21.2	.885	<.001	52.23	<.001
Transparency	33.6	24.9	58.6	31.3	.93	<.001	54.58	<.001
Understanding	35.2	26.2	58.2	29.4	.935	<.001	54.183	<.001

Table 5.8 : The statistical results of Study IV, regarding the runtime permission popups. * we report F values for all but SUS using the ART-ANOVA.

detailed option display." For continuous datatypes, such as location, the slider interface was more intuitive to use for some participants. One mentioned that they'd prefer it for continuous datatypes only: *"Slider for me is only useful for continuous values like distance."* (P51). Especially in the weather app scenario the ability to configure a granularity makes sense to the participants, as e.g., P60 stated: *"The weather apps do not require my exact location so I like this slider feature here."*

System-Level Settings: Better Overview and Easier Getting-Into with Sliders

People liked the ability to use sliders for the device-wide privacy settings as well. Participants liked the overview that the sliders gave on data collection. P39 reported on the system-level slider: *"This was close to perfect, I think this is how it should be. The granularity of specific sliders also grants insight into all the different data that is being collected, which the normal UI completely lacks."* In contrast, the classical settings menu is often criticized as it is hard to get into it and lacks an overview. *"Especially at first glimpse, the system is not that easy to get an overview of."* (P62). The lack of transparency of the classical interfaces also leads to a lack of trust, as a further mention of P39 shows: *"Many options were hidden and you have to kind of guess what each thing does. Also, not sure if the options will be reverted after an update."*. P26 complained that *"you have to click through more"* with the classic interface. Besides clear benefits, our participants also pointed out some drawbacks of the new slider interface. Criticism mainly evolved around the system-level slider, which stood on top of the settings screen above all individual permission sliders. e.g., P22: *"I do not think I would use one slider on the top. Especially as the different categories only*

	Device Settings							
	Classic		Slider		Normality		Wilcoxon	
	M	SD	M	SD	W	p	W	p
SUS	78.0	17.0	80.0	18.1	.887	<.001	203	.382
Control	60.5	27.7	80.7	23.6	.876	<.001	63	<.001
Privacy	66.1	24.6	80.4	25.2	.86	<.001	83.5	<.001
Security	59.8	25.9	73.5	27.2	.9	<.001	84.5	<.003
Sense	73.2	25.0	82.2	23.3	.835	<.001	117	.086
Transparency	48.8	31.5	66.7	29.9	.922	<.001	114	<.009
Understanding	40.6	29.6	66.2	29.4	.92	<.001	58.5	<.001

Table 5.9 : The statistical results of Study IV, regarding the system level settings. * we report F values for all but SUS using the ART-ANOVA.

make sense for some apps. Instead, I would get rid of the slider on top and instead have a categorical setting like privacy level high/medium/low for example." P10 saw privacy risks for lazy, speeding users introduced by the system-level slider: *"On further reflection, this feature now strikes me as very risky to dangerous. For example, it could tempt me as an annoyed user to be happy to easily set the location permission for all my navigation and sports tracking apps to maximum, and thus unintentionally set e.g., the memory access permission for all apps ever downloaded to completely open as well."*

Besides some points of criticism, the participants overall liked the slider version. *"It is very sexy, please install it on every phone."* (P40). They saw the main benefit of the slider-based system-level settings menu in its intuitive understandability and overview. *"The permission settings were very clear and concise, it was easy to gain an impression of how the data would be used."* (P53). Less technology-experienced users would benefit: *"[The slider] controller [is] more intuitive for older people or people who do not have smartphone affinity."* (P25).

Sliders Improve Transparency

Besides improved control, participants also perceived higher transparency about the data collection. The sliders with their steps make transparent which aspects a permission encompasses *"The granularity of specific sliders also grants insight into all the different data that is being collected, which the normal UI completely lacks."* (P39), and even give the user more sense about how their data is used *"The permission settings were very clear and concise, and it was easy to gain an impression of how the data would be used."* (P53). Having an overview of the active steps of each permission slider, the user could quickly grasp what is collected *"They explained what exactly would be collected."* (P22). P28 further saw an explanatory effect and triggered reflection processes: *"Offers control but also explains the usage of the data, and by showing the different levels of data abstraction people get a feeling of how much the data can actually capture and gives an opportunity to realistically reflect on their own boundaries."*

More Perceived Control on How Data is Used

Participants mentioned that the system-level slider gives them control on *how* data is used: *"The permission settings were very clear and concise, and it was easy to gain an impression of how the data would be used."* (P53). However, we found that,

independently of the applied method, participants felt a general lack of control over what happens with their data after granting access to it. Especially regarding the classic interface, many participants mentioned that the given control options only give control on what data is *passed* to an app, but not at all what thereafter *happens* with their data; i.e. with whom it is shared, how it is processed, where it is stored. We observed a general lack of trust in all that happens behind/after the granting interface, independent of the method. It was mentioned that *"trust [is] missing, the method is not the problem"* (P13). Similarly, P38 expresses issues with control over the later stages in the processing pipeline: *"This Interface suggests some form of privacy control but unless one denies everything, there is limited control once data is in the app."* P28 expressed missing *"control over where the data is stored, with whom it is shared, and what information is drawn from it."* This issue is out of the scope of permission granting methods which we focus on in this paper, but nevertheless noteworthy for future work.

Slider Design A couple of participants expressed different preferences of the slider's step order, for example, P21 generally expressed that *"The order of some levels of privacy did not make much sense to me."* or P13 proposing based on their cultural background that *"emojis should go to the last. In India, different emojis are differently interpreted."* Customization of steps per app was suggested by P38 *"The slider might be adjusted by application, since for example weather is no more accurate than a couple of 100 meters anyway."* On the other hand, some participants were concerned about too many differences between the sliders. P10 said that inconsistencies in the slider steps could lead to unexpected behavior. A medium slider value should express a similar level of data and privacy across all sliders. In general, it was perceived as a *"very sufficient interface, and with a bit of background knowledge, it is easy to understand how it works and what it does."* (P52).

Detailed UI Comments As for the nature of a high-fidelity prototype study, participants also pointed out many detailed UX improvement suggestions and criticisms of our prototype. During the study, a couple of unclarities in the UI were pointed out, especially regarding the slider-based system-level settings (e.g., *"not clear what gradations mean."* (P51)). The behavior of the sub sliders and their toggles was unclear to some participants (*"unintuitive what you turn on with the toggle?"* (P10)), also explanations on the slider steps, for example with context menus, were desired. A few participants

generally misconceived the slider steps as selecting instead of summing up. However in general the participants made themselves familiar with the slider UIs quickly, and further explanations by the study conductor were necessary in individual cases only.

5.3.6.5 Summary

Summing up on the lab study where we tested our high-fidelity privacy slider prototype with 32 participants, we conclude that the concept of privacy sliders was perceived very positively. Participants saw benefits in several scenarios, the usability was rated better than with the traditional privacy settings UI, and we found positive effects on transparency and control. We collected points of criticism that can be improved in future iterations. For example, step positions on the sliders should be determined by their concern level instead of the percentage of the slider scale, and the cumulative behavior has to be better explained.

5.3.7 Discussion of Privacy Slider

5.3.7.1 Runtime Permission Sliders Outperform the Standard Permission UI

The feedback on runtime permission popup sliders was overall very good. Extending the current button UI with a continuous choice of data granularity made sense to our participants and was perceived as intuitive, regarding the system usability score it significantly outperformed Android's current UI. Especially for data types that impose a natural degree of granularity (such as location), it was liked, and participants envisioned situations where they see an advantage in continuous permissions. With its straightforward user flow, which is close to Android's current design, users got into it easily and there is not much that could trigger confusion. We argue for including this in future runtime permission popups. While fine-granular permission concepts have been published in the past occasionally, for example by Jeon et al. [204] and Scoccia et al. [363], the present study is, to the best of our knowledge, the first study that implements it as a well-usable slider UI and studies its usability and applicability with lay smartphone users.

5.3.7.2 System-Level Settings Slider

Also in the system-level settings menu sliders were preferred by our participants. We see a benefit, especially in the transparency and overview that they provide, what participants confirmed in their qualitative statements. The system-level slider on top was deemed a good feature for novice users, and the flat menu structure required users to click through less. However, the more complex nature of a whole settings menu in contrast to a single-case runtime popup makes designing challenging. This reflects in a couple of remaining usability issues that we found in our study, and need to be addressed before rolling this out in the wild.

Most importantly, user support should be included in helping users get into the principle of how the slider-based menu is working, such as a tour, as proposed by Carlèn [75]. Furthermore, it has to be ensured that users do not experience unexpected behavior across the sliders. In our study, we found that users perceived privacy concerns of specific steps differently, and thus would have expected a different order. However, with the system-level slider on top, it is essential that the subsliders that are moving alongside do not show unexpected configurations.

5.3.7.3 A Method-Independent Lack of Control of What Happens With the Data

In the qualitative feedback, participants mentioned that they desire more transparency and control over what happens with their data after a permission has been granted. As they also admit, this is out of the scope of our study on the permission granting interface, however, we think this finding is nevertheless important to note. Sliders could for this issue be part of the solution as well. In our case of granting data access, sliders enabled control and conveyed transparency to their users. Generalizing privacy sliders to a modality for configuring data transactions, they could also find application on other stages of the data pipeline. The setting options of how far data is passed on, or in what depth it is analyzed, pose a natural order (for example data not leaving the device, going to the app company's server only, being processed on a cloud server, being disclosed to third party companies, being made available publicly).

5.3.7.4 The Tradeoff Between Warning Fatigue and User Control

Permission interfaces have to deal with the tradeoff of warning fatigue, i.e. users' desire for control and on the other hand being mad about too much information, options, and time spent. Users tend to ignore privacy-enhancing technologies (also coined *the challenge of user ignorance*) [10, 25] and concepts that foster their usage have to be considered, such as nudging approaches (e.g., [394]). Control-providing concepts have to be designed with care, as sophisticated concepts may quickly annoy their users and thereby fail [312]. While this is not the case for our runtime permission slider, users do not have to deal with the system-level settings unless they actively look for them in the settings menu. Users could be triggered to use the system-level settings in appropriate situations. That could be once after the device has been set up, to choose the device's default privacy policy, after the installation of multiple new apps, or after context changes (for example when travels or holidays are detected).

5.3.7.5 Contextual Privacy and Personalization

The privacy slider configuration could also be context-dependent. Users prefer data disclosure differently depending on their context, as Wang et al. [417] show in the example of online behaviors. Including context in runtime permissions help users make their decision [363] and enables an even more fine-grained choice, also called *flexible permission* [363]. A system that is able to understand and extract a contextual difference given, that would even be possible without additional user burden. Contextual privacy has shown to be beneficial in various use cases, such as online privacy policies [433] or IoT [301]. Regarding smartphone permissions, research has shown that the incorporation of contextual cues can improve decisions [429] and studied machine-learning-based decision support [398]. We think that our fine-granular approach to permissions well integrates with such approaches. The non-binarity of continuous privacy configurations could be used to reflect model uncertainties, i.e. instead of a prediction model requiring to output a binary all-or-nothing decision, an insecure prediction could lead to a slider value somewhere in the middle. Furthermore, the context could be another dimension of configuration that could be controlled through a slider (e.g., rating private situations on weekends as more concerning and worthy of protection than behavior during office days).

5.3.7.6 Privacy Sliders Enable Novel Adaptive Use Cases

To avoid the unpleasant consequences of privacy issues, data usage by applications is restricted. Access to potentially sensitive resources, such as screen contents and detailed device activity, is in Android for example organized into the Android Accessibility services. Thereby access is highly restricted to a few purposes only. By going from the current binary approach to fine-grained configuration we envision that such resources could be opened to wider application purposes. Screen contents for example could be leveraged for adaptive application scenarios, such as predicting next-action sequences, if the data was abstracted to the smallest necessary level of detail. Exact text contents, like text messages, names, or login credentials could be abstracted to tokens like *textmessage*, *name*, and *logindata*. They would thereby still be useful for several application scenarios but way less privacy-invading. Continuous permissions, realized through privacy sliders, thus not only have a privacy-preserving effect on the users but also enable novel opportunities for application and system developers.

5.3.7.7 Future Work: Field Study

In the present study, we have focused on users' opinions on the privacy slider. We deliberately asked end-users because they are the major stakeholders regarding privacy; their data is worked with and they opt for buying and using their smartphone. In the next iteration of the privacy slider, developers should also be taken into account, to see which effects fine-granular permissions would have on them, what changes from the development perspective, and find solutions on how developers could deal with that. Developers would need to deal with data of different granularity. If they in the worst case simply reject all except the finest data levels and thus force users to push the privacy slider to the finest level, not much would be won for the user. They still would be in the dilemma of granting (full) data access or not using the application (cf. Stach and Mitschang [382]). Flexible data structures might be needed in mobile app frameworks, to make it easy to work with varying granularities of data. Furthermore, an in-the-wild study has to be conducted. Our lab study was appropriate to get first insights on privacy sliders, showed that it is promising, and yielded valuable insights for

the next iterations. However, to see how users actually use them in real situations and which effects that has, a field study, where privacy sliders are distributed to the users' own devices, has to be conducted.

5.4 Transparency Through Interactivity: Interactive Machine Learning

This section is based on the following publication:

Florian Bemmann, Daniel Buschek, and Hussmann Heinrich. "Interactive End-User Machine Learning to Boost Explainability and Transparency of Digital Footprint Data." In: Yokohama, Japan: HCXAI Workshop at ACM CHI 2021, 2021

Digital footprint data is increasingly used to fuel powerful artificial intelligence systems e.g., to predict future behaviour or characteristics of the user [117]. The purpose of predictions ranges from content personalization and recommendation [73], over adaptive user interfaces [369] to research purposes, e.g., in the fields of psychology [384]. With power comes responsibility: Tracking huge amounts of personal data demands for a good privacy protection concept to reach real transparency. The usual approach has been to inform the user about *who* collects and processes *what* data. However, with the possibilities which big data and psychometric modeling enable, that is not sufficient anymore. People should also be informed about *how* data is being used [179]. Therefore the full process of AI systems should be transparent [300].

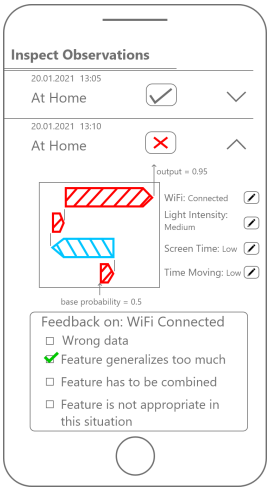


Figure 5.22 : Interactively giving feedback on a locally trained personalized predictor: In the middle, a local explanation gives feedback from the system to the user, explaining which mobile sensing features contributed to the model's decision. At the bottom, the user gives feedback to the system, communicating why the model falsely concluded that the user is at home. Here: User states that the feature *WiFi connected* is too general.

Digital footprint data contains a huge amount of (concealed) information about their users. Using methods of psychometrics and psychological targeting [275], the data can be exploited for unethical purposes. Targeted advertisements based on digital footprint data can influence societies and poses huge challenges to our democracies [122].

Unfortunately users do especially hard in understanding the capabilities of machine learning procedures. Related work and our studies in Chapter 4 have shown that users can hardly estimate what potential lies in digital footprint data, and thereby rate abstract-looking data as way less privacy-invasive. In this section, we study how an interactive machine learning interface can educate users on these procedures, to reach more transparency and sense for the inherited prediction potential. Therefore we put up the following research question for this section:

RQ3i *How can interactive machine learning yield transparency regarding prediction and inference potential of mobile sensing data?*

5.4.1 Explaining Through Interactivity for Better Transparency

If users would be more aware of and understand how their data is leveraged for psychometric modeling, they may less likely be susceptible to content targeted based on their personality. By bringing explainability to the process of digital footprint model building, we think one can increase transparency and an understanding about how psychometrics work.

To do this we introduce participatory design and interactive explainable AI to data-collecting mobile apps. Instead of just being observed, users should be included in the full process of data collection and model building. Beyond showing what data is collected, it should be explained which features are extracted and what they are expressing, why the feature selection decides for certain features, and how a model can learn to predict a target variable from these inputs. Interactivity is therefore well suited, but unfortunately not very prominent in intelligent systems [8]. Research on intelligent systems calls for enabling rich feedback from the user towards the system [238], and interactive machine learning has demonstrated positive effects on learning [12] and explainability [130].

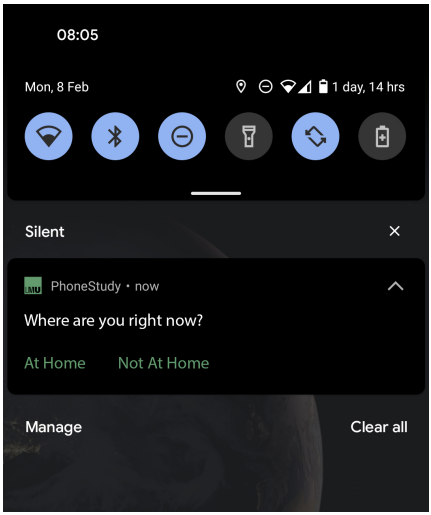


Figure 5.23 : To collect ground-truth data for mobile model building in our prototype, the user was asked to indicate when leaving from / returning to home via a permanent notification.

The concept of interactive user-involvement as explanations in XAI is contextually transferable and could be established as a general XAI technique. We envision more transparent and thus responsible intelligent systems, by letting users participate in and interact with data collection and model building.

5.4.2 Proof of Concept: Interactive Model Building Demonstrates the Hidden Information in Digital Footprint Data

To operationalize interactive explainability in practice, we incorporate it into a Mobile Sensing smartphone app. Our basis is the PhoneStudy app[1], which collects passive sensing data and self-reports in the wild to fuel offline model building for psychological research [384]. In a working prototype we brought the model building to the client device, allowing the user to train a personalized model on-device. Features are extracted locally from user's live sensing data, and self-reported data is used as prediction target.

In a first experiment we use the binary variable *being at home or not at home* as prediction target and basic device status features as input data. Although this is a trivial

[1] https://osf.io/ut42y/, last accessed 2024-12-09

example, it demonstrates our concept of mobile model building with live sensing data to show to users how “hidden” information in digital footprint data might be revealed: The model fitted nearly perfectly on features on the smartphone’s WiFi status.

We want users to better understand what can (and can’t) be inferred from digital footprint data, by making psychometric modeling explainable and interactive. Therefore we apply explainable artificial intelligence (XAI) techniques to a Mobile Sensing research project: We deploy a Machine Learning model that predicts personality from smartphone usage data to the user’s smartphone, incorporate XAI concepts, and study the interaction with and effects of local explanations. The functioning of psychological profiling is thereby explained to the user with their live data. Furthermore we introduce interactivity: In a working prototype the user selects a desired target variable, and trains a personalized model with their live data that tries to predict the target variable. We will investigate whether the usage of an explainable and interactive psychometrics app can lead to an understanding of the principles of psychological profiling, and raise awareness about the inference potential that lies in ones digital footprint data. While current work in the domain of explainable artificial intelligence (XAI) mostly explains decisions of models with the aim of interpretability:

Who? We target the non-professional, data-generating smartphone user, whose data is the basis of modern psychometric profiling models.

Why? We think that an understanding of what happens with users’ data and what can (and cannot) be inferred is important. Regarding research, our vision is to empower the user: Data collection and model building should be more a collaboration of user and researcher, than the researcher observing the user. After (1) having raised awareness for what data is recorded, we on the long term we aim for (2) creating an understanding of psychometric models and their capabilities (and limitations), (3) show what they can be used for in practice (esp. targeting and manipulation) to (4) create awareness against targeted (fake) content that manipulates people.

Where? A Mobile Sensing smartphone application is the optimal basis for our approach. It provides a plethora of data that is user-centric, making the explanations and

interactivity interesting for the user. The ability to deploy our system completely on the client-side avoids that privacy-sensitive data has to leave the user's device and makes it scaleable.

5.4.3 Discussion

5.4.3.1 Can Interactivity Boost Explainability?

We argue for interactivity – going beyond passive explanations. The literature supports this idea: For example, trying configurations of learning systems and observing effects is desired by users of ML systems and could result in a better understanding [130] and long-term learning effects. The effects of absence and presence of specific features on the model performance could convey their value to the user. Principles of correctability and rich feedback [238] should be incorporated in ML systems, and be used beyond improving models. To further inform and evaluate this approach, it should be studied how such interactivity and user-to-system feedback affects the understanding of a system: Which insights can be conveyed easier, and which cannot?

5.4.3.2 Interactive ML for Data Transparency

Transparency is limited in current systems that collect data for model building: Beyond explaining the user *who* collects *what* data, systems should convey *how* data can be used and what can be *inferred* [179] to reach real transparency. To implement this, we suggest to include the user in the model building process. We want to study whether interactively trying out inference techniques with their own data, thus experiencing which kinds of predictions can work and which are more difficult, support real transparency.

5.4.3.3 Participatory Model Design

In our concept the user is included in the model building process, rather than just observed [8]. To reach "Participatory Model Design," inspired by Participatory (Product) Design, new workflows for research working with user data should be studied: Can users be involved in the big challenges, e.g., feature design and selection? Users could create features they think are predictive for them, try them out in a local model, refine their features. Finally the researchers collect only a set of individual models and feature descriptions from their study participants.

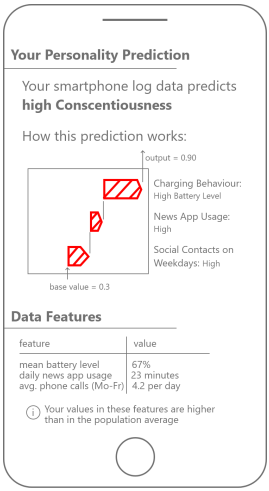


Figure 5.24 : Interactivity on personalized predictors is only suitable for intra-user variables. To explain inter-user variables like personality, we propose to deploy a pre-trained predictor on the client device and locally run and explain predictions using the data collected by the user's device.

5.4.3.4 Interactivity for Personalized vs. Universal Models

Personalized models can demonstrate predictions about intra-user variables, such as some status of the user (e.g., indoor/outdoor, mood, stress). To showcase models that compare users (inter-user variables) among certain characteristics (e.g., personality) a universal model is needed. While for a personalized model the full process of training, evaluation and prediction can be demonstrated live, for universal models it is only possible to show predictions using a pre-trained model. Thus, different concepts of interactivity have to be designed for both types of models, and their effects on the user may have to be studied separately.

5.4.3.5 Make a Difference: Inference Potential and (Unethical) Applications

Users being aware of what can be inferred from their digital footprints is an important first step. However, to make a difference with this work we encourage to think beyond: Building on the outlined interactive explainable mobile sensing app, it should be studied how it can further be used to "vaccine" people against (unethical) applications of personal data collection: McGuire's Inoculation Theory [278] proposes weakened pre-exposure to protect against persuasion. With *The Bad News Game*[1] the application

[1] https://www.getbadnews.com, last accessed 2024-12-09

of Inoculation Theory has shown a reduction of susceptibility in the domain of online misinformation [31]. Similar concepts seem promising to the domain of targeted content as well. Can we demonstrate the suggestibility of content that is targeted with personal data, to empower users to unmask and reluct against such in the future?

5.4.3.6 Motivating User Engagement

Our concept is not targeted to a specific user-group, instead any data generating smartphone user should be encouraged to use it. While a short term usage could be motivated by gamification techniques (e.g., little challenges or tasks), we assume that long-term use would require more sophisticated application concepts. Furthermore it should be discussed whether long-term use is even needed to yield the desired understanding.

5.5 Chapter Conclusion: Improving Privacy While Keeping Data Output

We started this chapter with motivating the need for transparency and control to improve user privacy. In a field study on privacy dashboards incorporating both features (cf. Section 5.1.1), we found that transparency initially decreased the user's perceived privacy. The offering of control features could then reverse this effect and lead to an improved privacy perception compared to a baseline condition. Interestingly, the control features were in fact rarely used, thus the sole availability of them gave users a sense of agency. This lets us conclude that transparency should always be accompanied by control.

Based on these findings we presented three approaches for privacy-enhancing technologies that are appropriate for mobile sensing. We explored on-device pre-processing of mobile language data as a means for data minimization (cf. Section 5.2), proposed and evaluated fine-grain control in Android permissions (cf. Section 5.3), and finally proposed the concept of interactive machine learning to inform and educate users about the hidden potential of data inference and predictions (cf. Section 5.4).

Chapter Take Away

Transparency and control are the two main components of privacy-enhancing interface concepts that can mitigate perceived privacy concerns. Transparency needs to be incorporated comprehensively, otherwise it can induce adverse effects. Privacy interface always act on the edge to user annoyance and warning fatigue, as users' motivation to spend effort on privacy belongings is mostly low. Their perception of an application's service privacy fit and obtained benefit by giving away their data mainly influence privacy decisions.

For future work, we would like to motivate more in-the-wild studies. Our lab study of the privacy slider was appropriate to get first insights, showed that it is promising, and yielded valuable insights for the next iterations. However, to see how users actually use them in real situations and which effects that has, a field study, where privacy sliders

are distributed to the users' own devices, has to be conducted. Similarly, we did not evaluate our proposed concept of interactive machine learning yet, but would like to do that in the future.

6

General Discussion

In this chapter, we regard the findings from the studies of this thesis in the big picture, discuss their implications, and draw conclusions for future research and industry. We therefore first **reflect on the overarching research questions** that guided the previous chapters. We especially **elaborate on the effect of transparency**, and explain its two-sided influence on app adoption behavior with the *valley of transparency*. We derive concepts to tackle the omnipresent lack of user motivation regarding privacy, and discuss various approaches of how to **integrate privacy interfaces best in apps' interaction flow**. Finally, we point out the specific challenges that mobile sensing brings privacy-wise, discuss to which scope our insights apply, and point out remaining issues and directions for future research.

6.1 RQ: How to Support User Quantification While Preserving Privacy?

Research on privacy-enhancing technologies in the context of mobile sensing is not new. Prior work's focus was especially on the user side, i.e., protecting user privacy

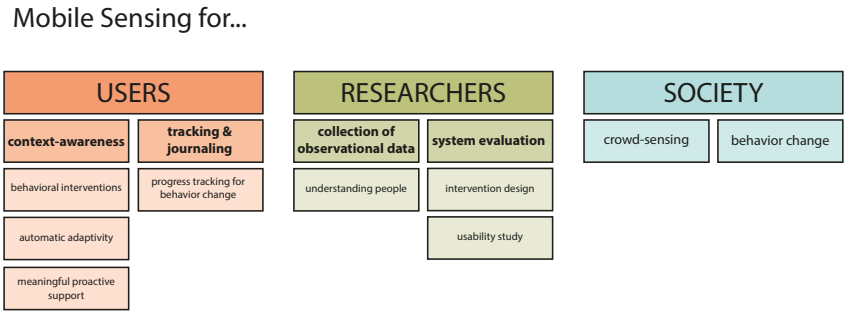


Figure 6.1 : Mobile Sensing applications can be to the benefit of three stakeholders. The user, researcher, and society.

as well as possible. Furthermore, proposed approaches are often rather technical and study the security side [89, 209, 239], which hardly keeps the user in the loop. However, in reality, multiple stakeholders are involved in the proliferation of mobile sensing systems. Besides the user, this includes especially the app publisher. The latter publish their apps mostly with an interest in mind, such as pursuing a business model through earning money or collecting data for further use. Approaches to improve privacy that focus only on the user thereby hardly show relevance in practice: App publishers will unlikely incorporate solutions that yield restrictions on the app publisher side. Furthermore, if apps become less intelligent and decline in user-friendliness, users also notice that and might, depending on their consideration of the service privacy fit, opt for a less privacy-friendly option. To have potential impact in-the-wild, both parties need to be regarded.

6.1.1 RQ1: Which Benefits Can Be Expected From Mobile User Quantification?

To give context on the mobile sensing landscape and its involved actors, we have shown in Chapter 3 who benefits from sensing data. We identified three major stakeholders that benefit from collecting and processing mobile sensing data.

Mobile sensing data collection methods enable **researchers** to collect data on people in the field. This data helps them understand human real-world behavior and how people use technology. It helps to explain psychological constructs such as

personality and can enable interventions that help users prevent or interrupt unwanted behaviors. Interdisciplinary research from multiple fields identified mobile sensing methodologies as data source that outperforms classical approaches such as manual journaling and observations in regard of data quantity, objectivity and scalability.

Detailed behavioral and contextual data furthermore allows to develop novel adaptive applications, of which **users** gain a direct benefit. By using passively sensed data, interactions are sped up, as users do not have to enter all information manually anymore. Contextual data can be auto-filled, most prominently location in navigation and location-aware search use cases. Furthermore, with systems becoming more proactive, the demand for contextual data rises. Without understanding the user's situation, recommendations and suggestions by interactive systems remain rather dumb and are low in relevance in regard to the user's daily life and real-world situations.

Furthermore, mobile sensing technologies can be deployed at scale, whereby they can influence whole social groups and our **society**. Crowd-sensing applications can be used to track environmental parameters, such as ambient noise. Mobile Sensing poses the opportunity to use data for the common good, e.g., to be one component of approaches that fight modern societal challenges such as polarization and climate change. All technology that comes with power also brings responsibility. When providing such data, it is essential to remember that it can also be used for unethical purposes, such as surveillance and control in authoritarian states.

In order to design adaptive, user-supporting applications well, a fine-grain and as comprehensive as possible view of the user's state and situation has to be available to the system. Without such contextual data, proposed actions and interventions may be irrelevant as they may occur in situations where they are not appropriate.

6.1.2 RQ2: What Concerns Arise From Mobile User Quantification?

However, mobile sensing applications face a low adoption rate. People are skeptical about apps that require much data from them, especially when it comes to passively collected aspects in the background. In mobile sensing research studies, where an app is deployed in the wild to collect user data, researchers face much lower consent rates than in traditional studies [221, 234, 338]. Conversion rates of studies can hardly be quantified with comparable numbers, there are many variables that influence the conversion rate, e.g., participation reward, advertisement, logged data,

other participation obligations, or the country in which a study was conducted. The participation rate in a mobile sensing study can be expected to be around 10%; for example in the IAB-SMART study of Kreuter et al. [234], 15,9% of all invited participants actual installed the app, of which 71% granted all requested permissions. Among mental health patients, Di Matteo et al. [114] measured that 41% were clearly willing to install an app that helps them detect potential disorders, with acceptance to requested permission ranging between 19% and 46% depending on the datatype. Although the app of Di Matteo et al. [114] deals with a more sensitive topic than the one of Kreuter et al. [234], the higher benefit for the user seems to lead to a higher adoption rate.

The use of rich and detailed data on users' behaviors and their smartphones' states also induces real security issues, such as stolen login credentials. Operating system developers are forced to protect their users and restrict access to such data. As currently no appropriate privacy-enhancing technology exists that allows to deal with highly sensitive data, they see access restriction as the only option.

We overall found that, from the user perspective, users are concerned about outcomes that affect their real-world lives. That are especially stolen wallet- and account information which may lead to financial loss or identity theft. Rich data types such as text messages and microphone data are deemed dangerous due to the large and varying amount of content. Concerns are mitigated if users expect a benefit alongside the data, especially benefits towards their productivity.

In general, users are initially unaware of many issues until they deal with the topic of data privacy. For rather abstract data types, they find it hard to estimate the contained information gain, and they can hardly judge which further information prediction models might infer from their data. They see the underlying issue mostly in data shared with or stolen by third-party actors such as companies or hackers. With these findings, we go beyond existing literature by investigating smartphone users' privacy concerns in depth rather than merely finding that privacy concerns are an important issue.

6.1.3 RQ3: How to Improve Privacy, Without Obstructing Usability?

To meet the demand for data unveiled by RQ1, while tackling the issues pointed out in RQ2, we, in this thesis, regard both the user and app publisher perspectives. We emphasized preserving privacy while keeping the data usable for its stakeholders.

Key Aspects of Privacy-Enhancing Technology Design Summarizing the findings from our studies and related work, we point out the following key aspects to consider with every privacy-enhancing technology.

Implement comprehensive transparency. Transparency is the first and essential building block to mitigate user privacy concerns. If users remain in doubt about what happens with their data, they can hardly put off their concerns. However, it is important to implement transparency comprehensively. If users are given only partial information, they may remain with more doubts and concerns than before. Furthermore, to avoid users feeling helpless, transparency should always be accompanied by control.

Offer Control. We found that control is the factor that mitigates users' privacy concerns. Interestingly, users rarely made use of the provided features. Solely providing control and giving them the ability to control data logging and processing happenings whenever they desire to do so made the difference. Control thereby offers itself as the most relevant feature of privacy-enhancing technologies. Control has been shown to make a difference in privacy concerns, while reductions in data quantity and quality have been shown to be negligible.

Do not annoy the user - the trade-off of warning fatigue and user control. Permission interfaces have to deal with the trade-off of warning fatigue, i.e., users' desire for control and, conversely, being mad about too much information, options, and time spent. Users tend to ignore privacy-enhancing technologies (also coined *the challenge of user ignorance*) [10, 25] and concepts that foster their usage have to be considered, such as nudging approaches (e.g., [394]). Control-providing concepts must be designed carefully, as sophisticated concepts may quickly annoy users and fail [312]. While this is not the case for our runtime permission slider, users do not have to deal with the system-level settings unless they actively look for them in the settings menu. Users could be triggered to use the system-level settings in an appropriate situation. That could be after the device has been set up, when choosing the device's default privacy policy, after installing multiple new apps, or after context changes (for example, when travels or holidays are detected).

Mind Service-Privacy fit and Privacy Calculus. Our research has identified the service-privacy fit as the most relevant decision model for privacy decisions. If users see a benefit in data usage, they tend to neglect privacy. For app designers, this

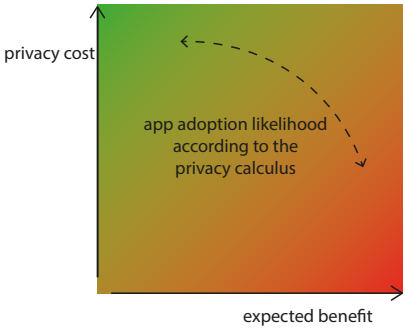


Figure 6.2 : We identified the privacy calculus and the related service-privacy fit as main decision models regarding app adoption decisions.

means that it is important to make the purpose of data collection and processing clear throughout the full pipeline. Thereby, an emphasis must be put on the user perspective, i.e., explaining why it is in the user's interest to leverage their mobile sensing data.

User-Centered Privacy. Things that users do not understand rather make them skeptical. We have also seen this effect in our studies on the privacy perception of data practices. Technical security measures that neglect to respond to the user are hardly perceived as privacy improvement. Technical measures for privacy and security should thus always be accompanied by a user-centered interface solution that conveys to the user what is happening.

Take Away

Offer control whenever possible: The ability to control data logging and processing practices mitigates users' privacy concerns. Although users actually use control features rarely, their availability and the option to do so anytime makes the difference.

Privacy-Perception of Three Methods In this thesis, we propose three approaches to improve privacy that keep the data output usable. These are namely on-device

preprocessing, fine-granular control, and interactive user involvement. We studied and presented them in Chapter 5, and the following shows what we found in our studies regarding the users' privacy perception of our methods.

Most participants felt protected by **on-device preprocessing procedures**. In more detail, perception of privacy was individual: In our interviews, some participants stated low concern and did not deem it necessary to use the opportunity we provided for a log data review. Others were more skeptical and preferred more details about the logging system's inner workings. Fittingly, participants' scores also varied for the questionnaire on general privacy concerns [70]. The log data review influenced participants' perceptions of the abstractions in both directions, indicating the individuality of perspectives (see next discussion point). Feedback generally showed that people liked this view of their data. Overall, we conclude that showing actual logged data could replace generic examples to inform participants of how the data is processed.

In our lab study, the feedback on more **fine-granular permission options** was overall excellent, especially for novices. Extending the current button UI with a continuous choice of data granularity made sense to our participants and was perceived as intuitive; regarding the system usability score, it significantly outperformed Android's current UI. Especially for data types that impose a natural degree of granularity (such as location), it was liked, and participants envisioned situations where they see an advantage in continuous permissions. With its straightforward user flow close to Android's current design, users got into it quickly. We argue for including this in future runtime permission popups. While fine-granular permission concepts have been published in the past occasionally, for example, by Jeon et al. [204] and Scoccia et al. [363], the present study is, to the best of our knowledge, the first study that implements it as a well-usable slider UI and studies its usability and applicability with lay smartphone users.

Our concept of **interactive machine learning**, presented in Section 5.4, is not evaluated yet. Its main contribution towards privacy is an increased understanding of machine learning procedures, thus *transparency*. We expect its influence on perceived privacy and user trust to be a two-sided sword. As outlined in section Section 6.2, where we discuss the effects of *transparency* in detail, interactive machine learning might require a careful and comprehensive design. If it partially explains inference and

prediction procedures and leaves its users with doubts and questions, its effects might be negative. Users could be scared by the presented potential and hidden opportunities that lie in such systems.

6.2 The Valley of Transparency

In our studies, we found that *transparency* in mobile sensing applications did not mitigate perceived privacy concerns. Instead, users became aware of privacy issues, which raised concerns and decreased app adoption. This finding aligns with past literature, which fits our results. For example, Tsai et al. [397] report that *transparency* tends to decrease service usage; however, if the lack of *transparency* means that users are not even aware of data usage practices, they are more likely to adopt a service. Regarded from a naive point of view, this questions whether we should implement *transparency* at all. While ethical and moral reasons require transparent data practices, the result-driven perspective could make developers ignore it completely. However, we think that this quite negative perspective on *transparency* is only half the truth.

To understand and solve this issue, a deeper understanding of the relationship between *transparency*, *service usage*, and other related factors is necessary. Notably, prior work found that *trust* is a mediating factor between various external and perceived factors and *system adoption intention* Choobineh and Kini [88]. *Trust* is known to positively influence *service adoption* [37, 279]. Janic et al. [203] also found that the effect of *privacy* on *trust* highly depends on other factors, especially the user's *understanding*. We know from other domains that *trust* is a mediating factor between system factors and *service adoption* [249].

Our findings and past work indicate that there may be a U-shaped relationship between both: Our field studies and surveys show that when *transparency* is introduced, users become aware of a system's data practices. As users are initially unaware of these, the *awareness* raises privacy concerns, and *trust* decreases due to the uncovered mismatch between initial expectations and actual practices. We hypothesize that these concerns can be mitigated when continuing to provide further *transparency* (trust is increasing again) unless any trust-inhibiting aspects are unraveled. We frame

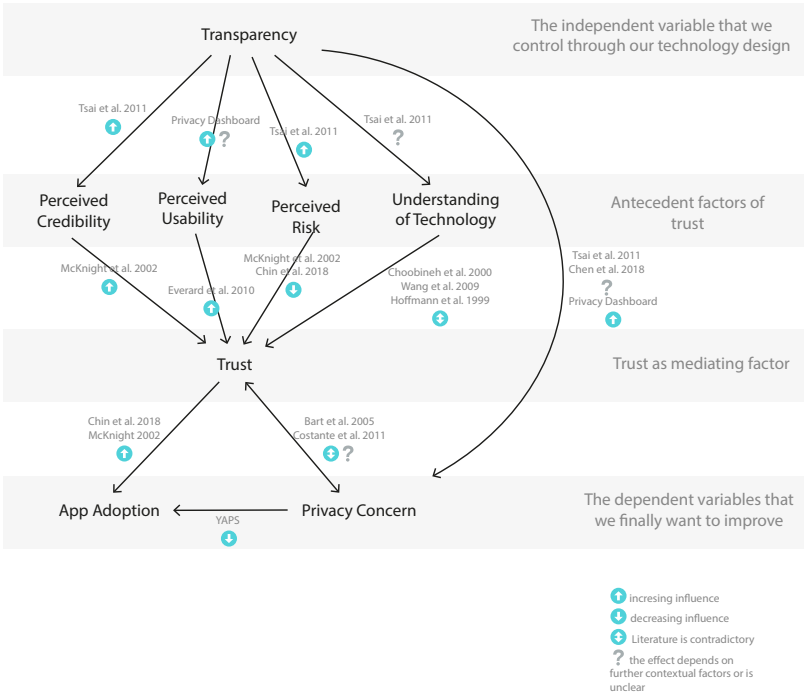


Figure 6.3 : Our hypothesized model on the relationship of *transparency* and *app adoption*. We argue that mediating factors, namely *trust* and its antecedents, connect *transparency* and *app adoption*. Further research is necessary, especially on the effect of *transparency* on the antecedent factors of *trust* and *privacy concerns*. The model builds in some parts on those of Corritore et al. [99] and Janic et al. [203].

this theory the *valley of transparency*. It explains the negative effects of a system's *transparency* by proposing that its *transparency* does not go far enough yet, i.e., awareness of data practices is raised, but yielded concerns could not be mitigated yet.

Influence of Mediating Factors on Trust Literature shows that *trust* has many antecedent factors, such as *perceived credibility* and *perceived usability* [139, 279], which are influenced by *transparency* [397]. I give an overview of the model of the mediating factors between *transparency* and *trust* in Figure 6.3. A comprehensive list

of mediating factors is hardly possible and has not yet been brought up by literature in the context of smartphone sensing. Based on literature from related areas, we assume *trust* to be, among others, yielded from *perceived credibility*, *usability*, and *risk* [88, 99, 139, 164, 231, 279], *understanding of the used technology* [88, 413] , *privacy concerns* [29, 100] , and *perceived security* [147] (list is non-exhaustive). The literature concludes that the influence of *credibility* on *trust* is positive [99, 279], however also shows that *credibility* itself has to be sub-distinguished into multiple factors (see [99]). We hypothesize that the effect of *credibility* can be stronger than that of *understanding the used technology*, as people can trust the app publisher unquestioningly and drop the need to understand what happens. While regarding the influence of *perceived risk* and *privacy concerns*, past work hypothesizes a negative correlation with *trust*, the evidence remains inconclusive. McKnight et al. [279] and Chin et al. [82] found no statistically significant effect, while Bart et al. [29] and Costante et al. [100] identified a weak negative influence. Overall literature is contradictory regarding the effect of *understanding of used technology* (e.g., [88, 413] vs. [189]). While *understanding*, on the one hand, is positive as it gives users a sense of agency and clarifies uncertainties, it can also raise awareness of negatively associated aspects. Thus, the effect of *understanding of used technology* on app adoption is complex and depends on what is achieved with it. As we did not find research stating the opposite, we assume these relationships are linear. They are, alongside the deducted overall relationship of these factors with trust, visualized in Figure 6.4.

Influence of Transparency on Mediating Factors *Transparency* influences many of these factors, likely increasing *perceived credibility*, but also increasing *perceived risk* [397]. The influence of *transparency* on privacy concerns and *understanding of technology* might depend on the level of *transparency* that is reached [397]. Chen and Sundar [78] further emphasized that if *transparency* is not applied holistically, i.e., not all *transparency* cues are understandable, it might have an adverse effect and rather reduce the user's perceived understanding of the system. Although we did not find evidence of the effect of *transparency* on *perceived credibility* in the literature, our studies indicate a similar behavior here (cf. Section 4.3): Initially, transparency can decrease a system's *credibility*; however, it can increase again when *transparency* reaches a certain level. It can be assumed that *transparency* is tightly related to

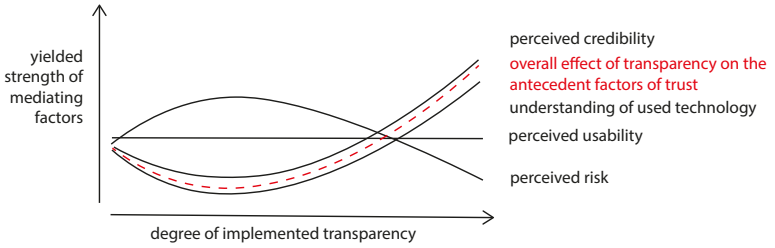


Figure 6.4 : We hypothesize the overall effect (dashed red line) of transparency on the antecedent factors of trust (black lines) to follow a U-shape. *Perceived credibility* and *understanding of used technology* initially decrease when users become aware of what happens to their data but can be increased again if transparency is provided comprehensively. *Perceived risk* initially increases as well and can decrease later on. However, its influence on trust opposes the two factors mentioned above. We did not find any indication regarding an effect of transparency on *perceived usability*, and thus assume no influence. The absolute positions of the visualized relationships are arbitrary, as we did not find comparative literature that provides sufficient evidence for a comparison.

awareness, i.e., being transparent can raise awareness of things that users did not have on their minds before. Based on our findings, we hypothesize a rather U-shaped, instead of linear, relationship for the effect of *transparency* on the aforementioned mediating factors (see Figure 6.5). *Transparency* initially raises awareness, which can be assumed to have a negative effect on the outlined mediating factors. For example, if users become aware of what data is logged, that likely has a negative impact on perceived privacy.

A Valley in the Relationship Between Transparency and App Adoption Joining the relationships of these two stages, i.e., the effects of *transparency* on the mediating effects *credibility*, *understanding of used technology*, *perceived risk and concerns*, and *usability*, and the effects of these mediating factors on *trust*, we obtain a U-shaped curve connecting *transparency* with *trust*. Such a relationship may explain the inconsistencies in current literature. Research, in its assessment methodology and statistical evaluations, usually assumes and thus tests for linear relationships, which does not allow for uncovering other kinds of relationships. Again, hypothesizing

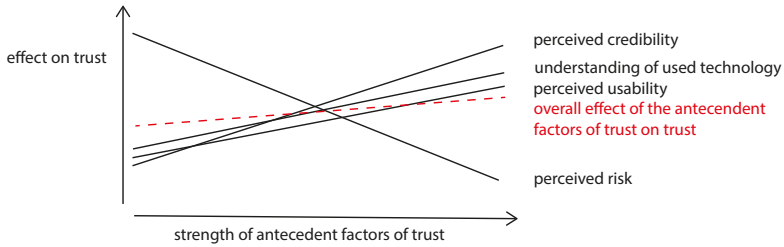


Figure 6.5 : Based on the literature and our studies, we hypothesize the effect of the antecedent factors of *trust* on *trust* to be slightly increasing. As we did not find any opposing evidence we assume their relationships to be linear.

a linear relationship between *trust* and *app adoption* based on research in related domains (e.g., [108]), we conclude: Based on our results and review of literature, we hypothesize (a) that *transparency* does not directly influence *app adoption*, instead this influence is mediated by the factor *trust*, and (b) that through this, *transparency* has a non-linear influence on *trust* and thus *app adoption*. We motivate future work to study *transparency* in more detail: Both should be regarded as a continuous variable rather than a binary exists or does not exist state. We hypothesize that those papers that found a negative influence of *transparency* on *app adoption* implemented a mediocre level of *transparency* that did not make it past the valley. Overall, this three-stage relationship is not sufficiently studied yet, with literature yielding inconsistent results (see the review of Janic et al. [203]), and thus needs further investigation. Work that we refer to in this section often evolved from the context of *trust* in web applications instead of smartphone apps. Also, the interplay with related constructs, such as *awareness*, needs to be studied. We visualize an overview of these dependencies in Figure 6.6.

How to Sell Transparency to Companies? Regarding the raw, objective results showing that *transparency* leads to lower app adoption, one can doubt why app developers should incorporate *transparency* at all. Besides legal and ethical aspects, there seems to be less extrinsic motivation. However, on a deeper look, *transparency* can yield objective benefits for app publishers. The discussed relationships between *transparency*, *trust*, and *app adoption* show that, if incorporated consequently and

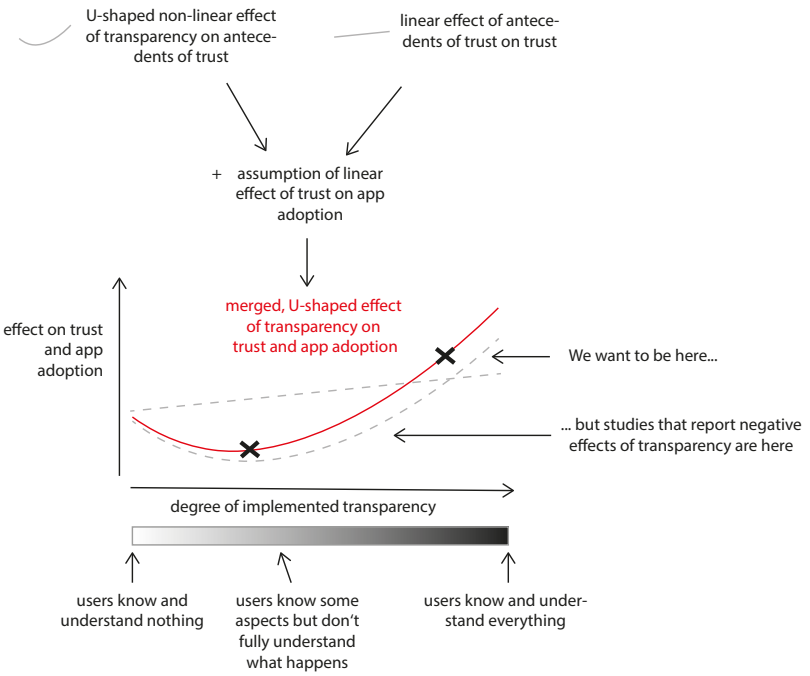


Figure 6.6 : We hypothesize that the relationship between *transparency* and *app adoption* follows a non-linear curve. We merged the relationships of *transparency* and antecedent factors of *trust* (cf. Figure 6.4), the relationship between these antecedents and *trust* Figure 6.5, and assume based on literature (e.g., [108]) a linear direct relationship between *trust* and *app adoption*.

understandable, *transparency* increases users' *trust* in an app publisher. From a service perspective, this is especially valuable as users then tend to continue using the service instead of switching to competitors.

Take Away

Mind to implement transparency holistically. If only awareness on data practices is raised, without explaining the details and how is justified by the service-privacy fit, users' concerns otherwise rather increase.

6.3 How to Approach a Lacking Motivation for Privacy Belongings?

Our study results have shown that the amount of time many users spend on privacy information and decisions is very low. In our interviews (cf. Section 4.3), users stated that comfort outweighs concerns, and questionnaire items on how much one familiarizes themselves with privacy aspects showed low efforts for a large share of our sample. Literature confirms that users lack the motivation to follow privacy and security practices [103, 160] and low usage of privacy-enhancing technologies [103, 237, 347]. Our study on privacy dashboards (see Section 5.1) also showed a very low usage rate of the provided transparency and control features. They are thereby caught in the triangular relationship of lack of motivation, high effort, and non-satisfying options. In order to yield privacy behaviors, some presence of all three factors is required. This aligns with the Protection Motivation Theory (PMT), which I map the three factors to and describe in the following.

The Protection Motivation Theory (PMT) describes that for a preventive behavior to happen, the probability of a threat occurrence and the effectiveness of coping behavior are essential [267]. In the PMT, this is subdivided into five factors: *perceived vulnerability* and *perceived severity* on the threat side, and *self-efficacy* and *response-efficacy* on the coping behavior side. Additionally, response cost describes the required effort to tackle the two ones mentioned above. Regarding privacy, *perceived vulnerability* (i.e., the likelihood of an incident - the *threat* side in the PMT) is rather low (although the severity is debatable and people indeed are afraid), *self-efficacy* and *response-efficacy* being rather low (i.e. the *coping behavior side*) and *response cost* being high (PMT's *cost* side). Thus, the resulting motivation for preventive behavior is low.

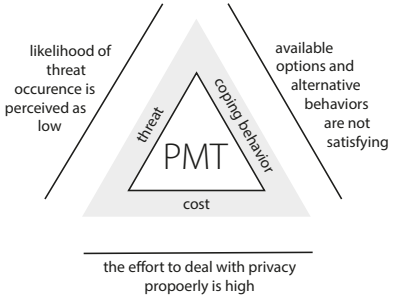


Figure 6.7 : The mapping of the three core aspects of the Protection Motivation Theory (PMT) (inner triangle) to issues of current smartphone privacy (outer triangle) explains low user motivation to spend time on privacy belongings.

6.3.1 Current User Experience Issues

The three aspects of PMT can be mapped to user experience issues in current privacy-enhancing technologies.

Lack of Motivation - Threat Side People lack motivation to actually follow privacy and security practices [103, 160]. Our studies show that users (1) do not see a direct benefit of spending time on privacy and (2) that expected damage seems very unlikely to them.

Unsatisfying Options - Coping Behavior Side People in our studies criticized the available options. While one theme of mentions evolves around doubts about whether the given options affect what they claim to do, users also claimed that some important aspects such as third-party behaviors are out of their reach of control. Due to missing satisfying options, users tend to adapt their behavior, for example, by entering false data or refusing service usage.

High Effort - Cost Side Grasping privacy information usually involves a significant time effort, for example as participants in Section 4.3 stated that information documents are long and hardly understandable.

6.3.2 Motivational Aspects Towards Privacy Behavior

According to the information-motivation-behavioral (IMB) skills model, people, besides being well-informed and having the behavioral skills to act, also need motivation [146]. While current research and privacy-enhancing technologies focus on the first two aspects, there is little on motivational factors. I.e., referring to the aforementioned triangular relationship, privacy-enhancing technologies provide better options with lower effort through transparency and control features but do not tackle motivational aspects.

From our studies and related literature, we conclude that, in addition to reducing user effort and improving available options, research should be conducted that targets user motivational factors. In the following, we propose ideas that should be followed up on in future work.

6.3.2.1 Gamification

Gamification is a major approach to support behavior change [32]. Especially with ubiquitous technologies that accompany us throughout the whole day, there are many options, and design spaces for gamification, such as Hallifax et al. [176], outline a wide set of possibilities on many dimensions. Mavroeidi et al. [276] also give an overview of game elements. While dedicated game applications to gamify privacy have been studied (e.g., [160]), approaches that apply gamification, i.e., integrate game elements into interfaces that are not games [112], are rare. We, in the following, propose which aspects of gamification are appropriate for privacy interfaces and critically discuss gamification.

Achievements, Awards, Goals Approaching behavior change by nature encompasses goals. These goals can either be a relative change in behavior or an absolute value that is targeted to be reached. Incorporating these goals is a common aspect of behavior change interfaces, done by various means such as quantifying the current status via points and progress indicators and collecting achievements and awards to honor reaching a goal. A prominent example are rings, which are initially open and fill up with progressing towards the goal, as used in the Apple Watch[1] and by

[1] https://www.apple.com/watch/close-your-rings/, last accessed 2024-12-09

Google Fit[1]. One could imagine a similar approach being used to visualize the user's status regarding privacy behaviors. A challenge, thereby, is to define goals, steps, and progress. As privacy is, by nature, not a quantifiable variable, research needs to develop ideas. Possible quantification metrics could include the risk of privacy issues and the gain of information from exposed data. For example, Alohaly and Takabi [11] propose an approach to quantify the amount of data an application collects. However, such metrics are rather the outcome of a privacy decision than a metric directly drawn from a behavior.

Social Privacy is a topic that affects every user, is relevant in many applications, and related behaviors are comparable across all application domains. This motivates the usage of social gamification elements, such as comparisons [2, 297]. Social elements have been shown effective in other domains such as physical activity promotion and weight control [247, 298]. People are especially receptive to social influences when they are overwhelmed or uncertain regarding configuration options [106], which, by our experience, applies to privacy decision-making. In the context of privacy decision-making, social aspects seem promising and hold future research opportunities, as outlined by Krsek et al. [237].

Ambient Ambient persuasion, as proposed already before the wide proliferation of smartphones in other contexts, helps to apply persuasive elements without requiring the user's full attention [177]. Murnane et al. [288] have shown design concepts to implement ambient gamification on smartphones, for example, through a lock screen that adapts depending on the user's behavior.

In the context of privacy behavior, ambient interfaces could be used to contextualize privacy. Omnipresent elements could convey the impact of a (former) privacy decision on a current situation. Contextual privacy has been shown promising in domains such as smart home or web services [302, 435], and should also find application in regard to mobile sensing data.

[1] https://blog.google/products/android/introducing-new-google-fit/, last accessed 2024-12-09

6.3.2.2 Reflection

Literature on gamification comes with contrasting results. In some cases, gamification elements do not outperform interfaces that enable reflection. For example, Zuckerman and Gal-Oz [458] observed no benefit of a gamified interface supporting physical activity compared to a quantified interface. Hanus and Fox [178] even find negative effects of gamification in the context of learning in classrooms. The authors of these papers thus argue that gamification has to be treated with care, and sole quantification of behaviors to support reflection should be considered first. Krsek et al. [237], for example, propose reflective writing, where users write about potential outcomes of privacy behaviors to foster self-reflection and yield behavior change.

6.4 In-Situ vs. In-Context

It is important to contextualize privacy matters. Context helps privacy decision-making and is in the scope of smartphone permissions realized through just-in-time permissions [15, 25, 59, 347]. Especially for sensitive and extracted data, which might also, to some extent, be unexpected by the user, the context has shown to be relevant [271]. However, just-in-time permission methods and at-setup permissions both have pros and cons.

For example, in-situ permission methods can raise mental overload [281]. Prior research tried to accommodate this with semi-automatic approaches that offload the user, e.g., automatic granting of subsequent requests [281], and prediction of privacy decisions [299, 430].

However, users are quickly feeling annoyed by in-situ privacy notices [10, 257], and all immediate privacy feedback is disrupting, especially when tied with actionability [306]. To not annoy the user too much and cause habituation and ignorance to the prompts, Patil et al. [306] suggest that they should only be used for important belongings. Schaub et al. [347] furthermore argue that privacy prompts tend to meet inappropriate moments, as they likely **conflict with a user's primary task**. Inglesant and Sasse [200] also reported this effect in the context of passwords, where effort spent on creating a strong password also conflicts with the user following a task.

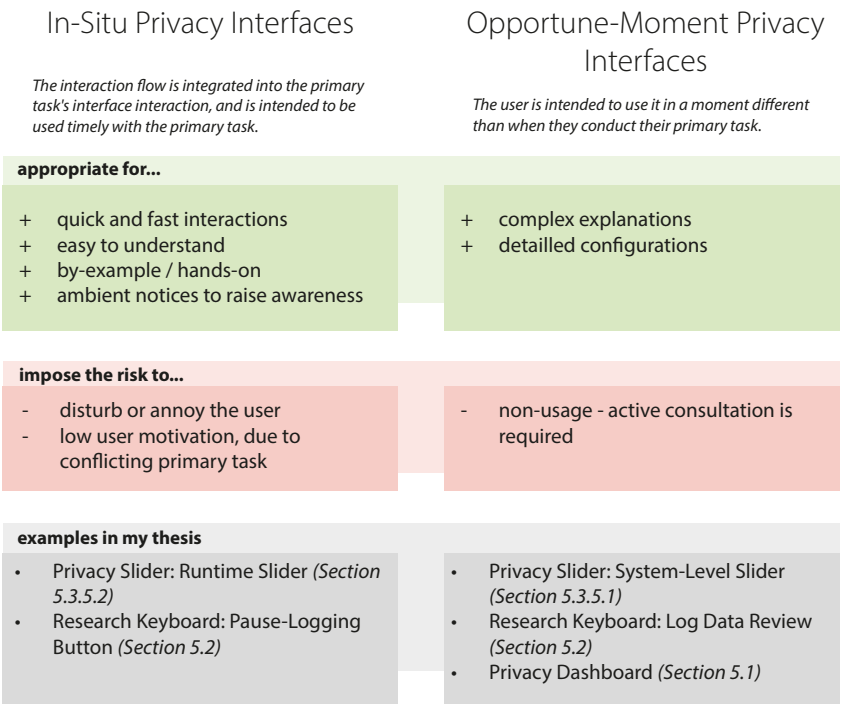


Figure 6.8 : We distinguish between privacy interfaces integrated into the user's primary task's interaction flow and those intended to be used in an independent, opportune moment. This figure shows the advantages, risks, and examples for both.

6.4.1 In-Situ- vs. Opportune-Moment Privacy Interfaces

We propose to distinguish between *in-situ privacy interfaces* and *opportune-moment privacy interfaces*: The first refers to interfaces whose interaction flow is integrated into the primary task's interface interaction and is intended to be used timely with the primary task. The latter are interface concepts that the user is intended to use in a moment different than when they conduct their primary task. How *opportune* is defined precisely is up for discussion and future research. We visualize the advantages and appropriate interface examples for both in Figure 6.8 and describe them in the following.

While our studies have confirmed some clear advantages of in-situ privacy information and action, we also found differences in the appropriateness of different interface purposes. In situ, we found that information mechanisms that were especially fast and easy to understand were perceived well by study participants. Ideally, they should be presented with examples of data that emerged in that situation instead of static contents. Information should be quickly perceptible and included in the primary-task interface ambiently so that the user is delivered with the information without having to leave their primary task proactively or is interrupted by the system. One example is the runtime permission slider that we have designed in Section 5.3.5.2, which participants in our evaluation found quick and easy to use in situ, especially for novices and people who do not want to spend much effort in privacy belongings. Similarly appropriate was judged the privacy button we implemented in our data-logging smartphone keyboard, which allowed users to turn on and off data logging at any time with just one tap Section 5.2. As inappropriate for an in situ deployment, we have proven privacy dashboards and the system-level settings slider. Participants valued the benefit of a privacy dashboard that we tested in an in-the-wild study Section 5.1.4.1; however, we also saw low proactive usage of it. In a dedicated usage session, the log data review applied in Section 5.2 instead received positive feedback. Also, the system level settings slider, besides receiving positive feedback, was mentioned by participants as something to be done once to set up a default configuration rather than a recurrently visited interface.

6.4.2 Untangling Context from Situation

While context is beneficial and important for people regarding privacy decisions, implementing it through in-situ approaches brings a couple of disadvantages. Concluding from the depicted literature and experiences from our studies, we propose to try to **untangle context from in-situ**. Context is currently mostly realized through in-situ methods, and both terms are even used equivalently in the literature (e.g., in [281, 347]).

Schaub et al. [347] in their design space for privacy notices define *Timing* as one dimension and there distinguish in-situ from context-dependent, and also propose at-setup, periodic, persistent and on-demand. As their design space is on privacy notices instead of consent mechanisms, not all options apply in regard to permissions, however, based on this, research has come up with other ideas to contextualize privacy.

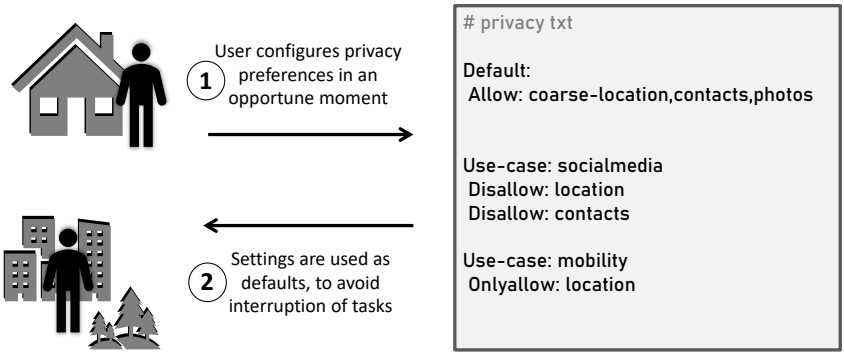


Figure 6.9 : Privacy in-situ brings context into privacy decisions; however, it also interrupts users in their tasks. In-situ users thus show an especially low motivation to spend time on privacy belongings. We propose to use other ways to contextualize privacy and allow users to define privacy rules at an opportune moment instead of in-situ.

Almuhimedi et al. [10] and Elbitar et al. [133] propose periodic nudges, where users configure the point of time. Patil et al. [306] found that information prompts can be delayed a bit to avoid interruption without yielding negative side effects. Further ideas to make privacy decision prompts context-based are proposed by Schaub et al. [347] themselves in their design space, for example, tying them to new locations, proximity to smart home sensors, or, more generally, detecting special situations. Another approach we have come up with is configuring a default privacy policy in advance. We have included this principle in Section 5.3 through the system level slider. One could develop this further by thinking of standardized formats to configure, save, and apply privacy decisions. For example, users could once configure their preferences and save them in a text file *privacy.txt*, inspired by the robots.txt[1] file used to configure search engine behavior. Whenever a new application is used, it reads this privacy.txt file and pre-configures all privacy decisions using the rules defined therein.

Overall, we motivate more studies on context-dependent privacy information and consent mechanisms. Emphasis should thereby be placed on opportune situations, i.e., not trying to interrupt users right when they are about to pursue a task. Based on Nissenbaum [291]'s theory of privacy as contextual integrity, we should regard

[1] https://developers.google.com/search/docs/crawling-indexing/robots/intro, last accessed 2025-01-08

privacy more in the light of the users' expectations. Wijesekera et al. [428] argue that smartphone permissions would be more efficient if the user is prompted only when data access to sensitive data likely defies the user's expectation. This is difficult to realize in practice and has hardly been implemented yet.

6.5 Personalizing Privacy Interfaces to Individual Attitudes

Privacy-related behavior is very individual. In our interviews and surveys, we saw both poles - users who did really care about their privacy and others who judged it unimportant for them or had resigned from the topic. When assessing users' engagement with privacy, what we did, for example, in Section 4.2, we found rather symmetrical distributions with two peaks and a valley in-between. Novice users want different things than experts (cf. Section 5.3), and the effects of privacy features accordingly vary across user groups (cf. Section 5.2).

Literature has yet came up with typologies that cluster users by their privacy behaviors, such as Schomakers et al. [360] who distinguish between *Privacy Guardians*, *Privacy Cynics*, and *Privacy Pragmatists*. Privacy guardians are highly concerned about their privacy and mind to protect it. Privacy cynics are also concerned but feel powerless and, therefore, show less protection motivation and behavior. In contrast, the privacy pragmatists are not actually concerned, and value provided benefits higher than their privacy. Most classifications evolve around the concern level, for example the *Privacy Segmentation Index* of Westin [367] which distinguishes between *Fundamentalists*, *Pragmatists* and *Unconcerned*, or Smit et al. [375] who simply segment into *highly concerned* and *low concern* users. The second prominent aspect in most models (e.g., those of Schomakers et al. [360] and Westin [367]) is user frustration and helplessness. In our studies, a distinction between users with high and low *expertise* and knowledge of technology was present. Different approaches to privacy interfaces for both user groups were deemed appropriate, for example, in the context of the Privacy Slider (cf. Section 5.3). Literature also backs this aspect of user segmentation with studies investigating differences in protective behavior between high and low-knowledgeable users (e.g., [233]).

Privacy-enhancing technologies should be designed with their target user group in mind. Depending on who an app targets and which types of users are mostly expected to use it, privacy-enhancing technologies should be designed differently. We propose to regard especially the two outpointed key dimensions of (1) level of expertise and (2) level of resignation and helplessness to adjust privacy-enhancing interfaces: **A high level of expertise should follow a high level of offered detailedness.** Respectively, novice users should be started with overviews, general configuration options, and hands-on aspects that they can directly relate to (e.g., practical examples of their data to turn on and off instead of configuring abstract concepts of data types). In contrast, for more experienced people, more detailed privacy interfaces are appropriate. Our interviews have shown this, and prior work also backs that by finding that people with high privacy and security knowledge show more protection behavior [233]. Second, **the prominence of privacy-enhancing features, especially control, should increase with the level of resignation and helplessness**. Users who resigned, thus giving up on coping with the topic of privacy, do not spend the effort to look for privacy information and intervention options actively. They have to be made offers unobtrusively that they can stumble upon. Offers should be prominent, ambiently integrated into the interaction flow, and show a low-effort affordability. Thereby, we see a chance to catch these users again.

Take Away

Adjust privacy-enhancing interfaces to the level of expertise that your target user group brings.

As a method to personalize the privacy interface appearance, reactive or proactive personalization approaches could be used [450]. Reactive personalization thereby relies on active user choice, i.e., the user could be asked for what they prefer. Directly asking the user might yield the best-fitting result, but it may be perceived as annoying and ignored. Proactive personalization instead automatically infers an appropriate choice for a user. Past user behavior could, therefore, be taken into account: **Resignation and unwillingness to deal with privacy could be detected by past interaction behavior** with privacy interfaces. Thereby, ideally, data across multiple apps, i.e., observations on the operating system level, should be taken into account. This ap-

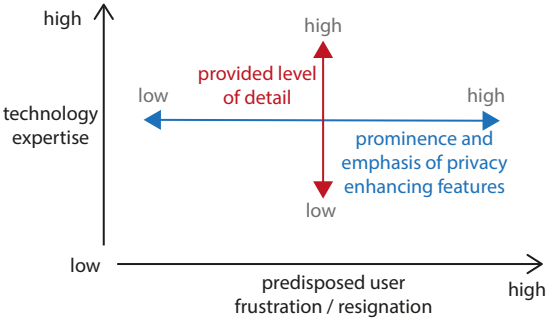


Figure 6.10 : Privacy-enhancing interfaces should adapt to (1) their user's technology expertise and (2) the user's inclined privacy resignation.

proach could be combined with the proposed idea of global privacy configurations (the *privacy.txt*, see Figure 6.9), i.e., the assessed user preferences stored on system-level and reused as default for future decisions. Technology expertise and knowledge could also be sensed from past interaction behaviors, such as leveraging interaction speed, error rate, and feature choices.

Take Away

Motivation to deal with privacy belongings and technology expertise could be sensed from past interaction behavior and used to proactively personalize privacy interfaces.

6.6 The Challenges of Rich, Detailed Data

Whereas most prominent smartphone permissions mostly deal with rather simple, manageable data types (i.e., location, contacts, physical activity), I in my thesis regard also rich, contentful datatypes such as UI tree data and mobile typing behavior. For the location or physical activity permission, to take these as examples, it is easy to grasp what information is contained therein. In contrast, UI tree data, for example, can contain a plethora of (hidden) information. Users are thus confronted with the **uncertainty of**

not knowing what is contained in rich data comprehensively. By examples and reasoning, one can list things that are possible, however such thoughts and created lists are never exhaustive. Rich data brings a **wider space of possibilities of what can be done with it**, especially concerning knowledge discovery and prediction procedures. As such algorithms advance over time, the capabilities of data inference may also increase in the future. Thus, the obtained information gain may be higher when revisiting data in a couple of years. This makes rich and contentful data special privacy-wise and brings challenges for privacy-enhancing technologies.

To deal with such rich data, we propose to describe their information gain and inherent potential through **boundaries instead of listings**. While recent approaches to inform users about the inherent information in and potential risks of data formulate an exhaustive list, this is hardly possible for rich data as outlined above. Boundaries instead describe a space in which information inference and risks evolve. Instead of providing users with a list of what is possible, this aims to help them understand and feel the hidden information in data. Users learn to adapt to changing environments and become less dependent on long and time-consuming information materials. They are also updated frequently on changing conditions.

6.7 Implications, Applicability, and Limitations of Privacy Studies

Many of our studies were conducted in the context of smartphone sensing projects for research purposes, such as Section 3.1, Section 5.2, and Section 5.1. We argue that choosing the domain of mobile sensing research studies is advantageous for our types of studies, in contrast to other options. Mobile sensing research questions in HCI and personality psychology have few requirements to their audience, as they by nature try to attract a wide i.e. representative sample of people. Other alternative application are mostly designed for a target audience, and thereby are biased in who they attract.

According to the concept of the service-privacy fit, app adoption behaviors are largely decided by a balance between the service benefit of an application and its privacy implications [192, 197]. In order to study the effect of the *privacy* component, it is thus important to keep the *service* component as controlled as possible. We

achieve this by choosing mobile sensing research studies as a domain. Mobile sensing research methods generally score low on the service side, as users do not get any direct benefit from it. Participants are usually compensated for their effort with money (cf. Haas et al. [174]), which is in the control of the researcher and thus does not yield many uncontrolled effects on user motivation. Distinguishing the service-induced motivation into extrinsic and intrinsic motivation, one can say that mobile sensing studies impose a low uncontrolled extrinsic motivation for users. Intrinsic motivation, however, varies as some people show intrinsic motivation to participate in studies in order to contribute to research. However, as in our studies and scenarios, participants obtained a monetary compensation, the effects of uncontrolled, intrinsic motivation can be assumed to be rather outperformed by the extrinsic motivation to obtain money [174]. With all other application choices, motivational factors are harder to control. Users have a predisposition to some apps from related apps that they have yet been using, from familiarity with specific use cases, and general affinity to a topic (such as maps, social media, or fitness tracking). We noticed the difficulty of controlling these effects in Section 5.3. Here, an application use case that users are motivated for was needed to yield natural interaction behavior with different privacy interfaces that we studied. To balance out these effects, we used four different scenarios in that study as additional independent variables.

As we did not find any indicators speaking against it, we argue that our findings apply to smartphone applications using mobile sensing in general. While our studies are limited to smartphones, we hypothesize that they transfer to other omnipresent and always-connected technologies, such as smartwatches. Crossler and Bélanger [103] argue that omnipresence and connectedness are the specific aspects of mobile technology due to which privacy-related behaviors differ from those of other technologies, such as desktop computers.

Take Away

Studying smartphone sensing privacy, using the example of mobile sensing research projects, avoids motivational biases, while insights to the best of our knowledge transfer well to other application contexts.

The transferability of the results of our studies on alternative data practices on-device (Section 5.3, Section 5.2) is unknown. In the study context, it is people's task to deal with privacy and use our proposed system; for example, people in Section 5.3 were instructed to make privacy decisions and were observed while doing so. We know from literature (cf. [103, 160]) and our surveys and interviews (cf. Section 4.3) that people's motivation to invest time and effort into privacy varies and often is low. While the study on on-device preprocessing (Section 5.2) was conducted in the wild, the lab setting of the study on privacy slider (Section 5.3) can not accommodate this issue. A follow-up study in the wild is necessary here. However, in-the-wild studies of privacy interfaces are technically hard to realize. Privacy interfaces such as permission popups are implemented on the operating system level and can thereby not be changed that easily. To integrate alternative privacy interfaces in the wild in-depth, the participants' smartphones would either have to be rooted (i.e., their operating system being changed) or they would need to be given alternative devices. Both are, in our opinion, not appropriate for a user study. Rooting a device brings irreversible changes and security issues to people's personal devices. Providing them alternative devices, on the other hand, does again not yield natural behavior. Research is needed to find solutions to this issue.

Take Away

Studies of smartphone privacy interfaces in-the-wild are difficult to conduct. Changes to privacy interfaces are hardly possible for researchers, as they would need to be implemented on operating system level.

6.8 Future Directions

Based on the insights of our work that we have presented in this thesis, some points of unclarity remain, and new gaps for future work arise.

Studies on non-linear effects of transparency and app adoption behavior In prior work (e.g., [78, 397]) and our studies, we found that the effect of transparency on perceived privacy concerns and app adoption behavior can be both of a positive

or adverse nature. We hypothesize a U-shaped relationship between transparency and app adoption (see Section 6.2) to explain this effect. Further research is needed to consolidate this: Studies on transparency and app adoption behavior should test for non-linear effects and regard more contextual and mediating factors, such as *understanding of the used technology* and *awareness of data practices*.

Comprehensive Transparency - How to Overcome the Valley of Transparency? Following up on the abovementioned future studies that investigate the effects of transparency in-depth, concepts need to be derived that make technology overcome the valley of transparency. Interaction components could thereby encompass feedback mechanisms that assess the user's understanding and make the system aware of raised doubts and concerns. If our technology became aware of the effect that it has on its users, opportunities would be opened up to avoid negative effects, such as getting stuck in the valley of transparency. Reactivity and adaptivity to the user are then needed to accommodate their demands. LLMs could, therefore, be leveraged to reach a satisfying level of general adaptivity and reactivity to the user.

Motivation Towards Privacy Behavior: Social Gamification In our survey and interview studies on users' privacy perceptions, we found that there is a general lack of motivation to spend effort on smartphone privacy. We identified three contributing factors: (1) Users do not perceive a high likelihood of privacy incidents actually occurring, (2) users have ineffective coping behaviors, and (3) the effort required to mitigate privacy issues is high. These three factors, which we explain in more detail regarding the Protection Motivation Theory (PMT) in Section 6.3.2, essentially would spark user motivation, but none is fulfilled. We propose to inspire privacy interface concepts with ideas from the field of persuasive technologies and gamification. Especially social components show promising results in other contexts, as social influences are known to be one of the strongest [149].

Untangling Privacy Decisions from Users' Primary Tasks Privacy information decisions in-situ face a high risk of being ignored by users. In-situ, i.e. when following a primary task, user motivation to deal with privacy aspects is especially low, as they

rather want to pursue the aim of their primary task. Studies on other concepts to contextualize privacy are needed, such as interaction concepts to make users configure privacy defaults and general rules in an opportune moment.

Control Beyond the Local Device The perception of being in control, and the Users' ability to enact control on their data anytime, has shown to be most beneficial to privacy perception and app adoption behaviors in our studies (see esp. Section 5.1). However, realizing and conveying effective control becomes difficult once data leaves the originating user's device. This was also apparent in our surveys and interviews on underlying reasons for privacy concerns. Users were most afraid of third parties getting access to their data. Concepts of differential privacy (e.g., [345, 414]), personal data stores [408], and platforms that manage data on users' behalves as trusted third parties [236], are all promising approaches to exert control outside of their sphere of influence. Procedures such as federated learning [193] have also shown that crowd-sensing applications are possible without exposing individual data points to remote servers. These solutions rely on algorithmic and mathematical procedures that technically ensure what is promised. However, as we have seen in our interviews on privacy concerns, users have a hard time trusting and accepting systems that they do not understand. HCI research needs to come up with solutions to keep the user in the loop with such systems. If ways are found to communicate these approaches to users sufficiently, they could enable users to exert control of their data throughout the full mobile sensing data pipeline.

Preventing Misuse of Data Besides opening opportunities for use cases that bring benefits to their users, mobile sensing technology may also be used for purposes that can hardly be regarded as *good* or are disputable. In general, misuse of data, i.e., its use for other purposes than its originating user consented to or assumes it is used for, should be avoided. Future research should critically discuss the risks that its technology brings besides use cases for the good.

7

Conclusion

In this thesis, we studied how we can improve the privacy of smartphone users while keeping the data usable for application purposes. We found that mobile sensing-based applications are to the benefit of three stakeholders, namely users, researchers, and society. Rich, contentful data seems thereby especially promising to realize novel adaptive, intelligent interaction concepts. However, we found that such data induces privacy concerns among its potential users. Privacy concerns are a major factor that hinders app adoption. Especially with application use cases where users get few or no direct benefits, and the service privacy fit is weak, users are reluctant to adopt such apps. While prior work improves privacy by simply minimizing data use, we study privacy-enhancing approaches while aiming to keep data usable. Our findings evolve around an overall lack of user-centeredness in current privacy-enhancing technologies and security measures. Users lack an understanding of what happens with their data, and they do not feel in control of data collection and processing. Technical security measures hardly mitigate concerns, users often do not trust the implementing companies and cannot retrace their inner workings. We identify transparency and control features as key elements towards an improved perception of privacy. While we found that offering control features directly reduces privacy concerns, transparency

initially worsens the situation unless it is applied comprehensively. We furthermore discuss different approaches to contextual privacy, the (in)appropriateness of privacy interfaces in situ, and the issue of a lack of user motivation regarding privacy belongings.

With more transparency and control being given to the users, we envision improved user privacy while allowing detailed contextual data to be used for adaptive applications and research purposes. Developers and designers need to mind the effect that privacy-enhancing technologies have on their users, especially transparency and in-situ interfaces that need careful consideration. Although conducted in the domain of smartphones, we argue that our research applies to other omnipresent and always-connected mobile systems as well.

Bibliography

[1] Bentolhoda Abdollahbeigi, Farhang Salehi. “The Role of Information and Communication Industry (ICT) in the Reduction of Greenhouse Gas Emissions in Canada.” In: *International Research Journal of Business Studies* 13.3 (2020). DOI: 0.21632/irjbs.

[2] Charles Abraham, Susan Michie. “A taxonomy of behavior change techniques used in interventions.” In: *Health psychology* 27.3 (2008), p. 379. DOI: 10.1037/0278-6133.27.3.379.

[3] Silvia Abrahão, Emilio Insfran, Arthur Sluÿters, Jean Vanderdonckt. “Model-based intelligent user interface adaptation: challenges and future directions.” In: *Software and Systems Modeling* 20.5 (2021), pp. 1335–1349. DOI: 10.1007/s10270-021-00909-7.

[4] Mark S Ackerman, Scott D Mainwaring. “Privacy issues and human-computer interaction.” In: *Computer* 27.5 (2005), pp. 19–26.

[5] Alessandro Acquisti, Jens Grossklags. “Privacy and rationality in individual decision making.” In: *IEEE Security Privacy* 3.1 (2005), pp. 26–33. DOI: 10.1109/MSP.2005.22.

[6] Tanisha Afnan, Yixin Zou, Maryam Mustafa, Mustafa Naseem, Florian Schaub. “Aunties, Strangers, and the FBI: Online Privacy Concerns and Experiences of Muslim-American Women.” In: *Eighteenth Symposium on Usable Privacy and Security*. SOUPS’22. Usenix, 2022, pp. 387–406. DOI: 10.5555/3563609.3563630.

[7] Divine Q Agozie, Tugberk Kaya. “Discerning the effect of privacy information transparency on privacy fatigue in e-government.” In: *Government Information Quarterly* 38.4 (2021), p. 101601. DOI: 10.1016/j.giq.2021.101601.

[8] Nitin Agrawal, Reuben Binns, Max Van Kleek, Kim Laine, Nigel Shadbolt. "Exploring Design and Governance Challenges in the Development of Privacy-Preserving Computation." en. In: *arXiv:2101.08048 [cs]* (Jan. 2021). arXiv: 2101.08048. DOI: 10.1145/3411764.3445677.

[9] Icek Ajzen. "From intentions to actions: A theory of planned behavior." In: *Action control*. Springer, 1985, pp. 11–39. DOI: 10.1007/978-3-642-69746-3_2.

[10] Hazim Almuhimedi, Florian Schaub, Norman Sadeh, Idris Adjerid, Alessandro Acquisti, Joshua Gluck, Lorrie Faith Cranor, Yuvraj Agarwal. "Your location has been shared 5,398 times! A field study on mobile app privacy nudging." In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 2015, pp. 787–796. DOI: 10.1145/2702123.2702210.

[11] Manar Alohaly, Hassan Takabi. "Better Privacy Indicators: A New Approach to Quantification of Privacy Policies." In: *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*. Denver, CO: USENIX Association, June 2016.

[12] Saleema Amershi, Maya Cakmak, William Bradley Knox, Todd Kulesza. "Power to the People: The Role of Humans in Interactive Machine Learning." en. In: *AI Magazine* 35.4 (Dec. 2014). Number: 4, pp. 105–120. DOI: 10.1609/aimag.v35i4.2513.

[13] Catherine L Anderson, Ritu Agarwal. "The digitization of healthcare: boundary risks, emotion, and consumer willingness to disclose personal health information." In: *Information Systems Research* 22.3 (2011), pp. 469–490. DOI: 10.1287/isre.1100.0335.

[14] Benjamin Andow, Akhil Acharya, Dengfeng Li, William Enck, Kapil Singh, Tao Xie. "Uiref: analysis of sensitive user inputs in android applications." In: *Proceedings of the 10th ACM Conference on Security and Privacy in Wireless and Mobile Networks*. 2017, pp. 23–34. DOI: 10.1145/3098243.3098247.

[15] Panagiotis Andriotis, Gianluca Stringhini, Martina Angela Sasse. "Studying users' adaptation to Android's run-time fine-grained access control system." In: *Journal of Information Security and Applications* 40 (2018), pp. 31–43. DOI: 10.1016/j.jisa.2018.02.004.

[16] Sarah Aragon Bartsch, Christina Schneegass, Florian Bemmann, Daniel Buschek. "A day in the life: Exploring the use of scheduled mobile chat messages for career guidance." In: *Proceedings of the 20th International Conference on Mobile and Ubiquitous Multimedia*. 2021, pp. 24–34. DOI: 10.1145/3490632.3490637.

[17] Jessica Aschemann-Witzel, Emilie Marie Niebuhr Aagaard. “Elaborating on the attitude–behaviour gap regarding organic products: young D anish consumers and in-store food choice.” In: *International Journal of Consumer Studies* 38.5 (2014), pp. 550–558. DOI: 10.1111/ijcs.12115.

[18] William Ascher. “Long-term strategy for sustainable development: strategies to promote far-sighted action.” In: *Sustainability Science* 1 (2006), pp. 15–22. DOI: 10.1007/s11625-006-0001-x.

[19] Kathy Wain Yee Au, Yi Fan Zhou, Zhen Huang, Phillipa Gill, David Lie. “Short paper: a look at smartphone permission models.” In: *Proceedings of the 1st ACM workshop on Security and privacy in smartphones and mobile devices*. 2011, pp. 63–68. DOI: 10.1145/2046614.2046626.

[20] Naveen Farag Awad, M. S. Krishnan. “The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to Be Profiled Online for Personalization.” In: *MIS Quarterly* 30.1 (2006), pp. 13–28. DOI: 10.2307/25148715.

[21] Oshrat Ayalon, Eran Toch. “Evaluating Users’ Perceptions about a System’s Privacy: Differentiating Social and Institutional Aspects.” In: *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. 2019, pp. 41–59. DOI: 10.5555/3361476.3361480.

[22] AB Karthick Anand Babu, R Sivakumar. “Multi agents for context awareness in ambient intelligence: a survey.” In: *Int J Eng Res Technol* 4 (2015), pp. 983–991.

[23] Michael Backes, Sebastian Gerling, Christian Hammer, Matteo Maffei, Philipp von Styp-Rekowsky. “AppGuard-Enforcing User Requirements on Android Apps.” In: *TACAS*. Vol. 13. Springer, 2013, pp. 543–548. DOI: 10.1007/978-3-642-36742-7_39.

[24] Christopher A Bail, Lisa P Argyle, Taylor W Brown, John P Bumpus, Haohan Chen, MB Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, Alexander Volfovsky. “Exposure to opposing views on social media can increase political polarization.” In: *Proceedings of the National Academy of Sciences* 115.37 (2018), pp. 9216–9221. DOI: 10.1073/pnas.1804840115.

[25] Rebecca Balebako, Jaeyeon Jung, Wei Lu, Lorrie Faith Cranor, Carolyn Nguyen. “"Little Brothers Watching You": Raising Awareness of Data Leaks on Smart-

phones." In: *Proceedings of the Ninth Symposium on Usable Privacy and Security*. SOUPS '13. Newcastle, United Kingdom: Association for Computing Machinery, 2013. DOI: 10.1145/2501604.2501616.

[26] Syagnik Banerjee, Ruby Roy Dholakia. "Location-based mobile advertisements and gender targeting." In: *Journal of Research in Interactive Marketing* 6.3 (2012), pp. 198–214. DOI: 10.1108/17505931211274679.

[27] Natã M Barbosa, Joon S Park, Yaxing Yao, Yang Wang. ""What if?" Predicting Individual Users' Smart Home Privacy Preferences and Their Changes." In: *Proceedings on Privacy Enhancing Technologies* 2019.4 (2019), pp. 211–231. DOI: 10.2478/popets-2019-0066.

[28] David Barrera, H Güneş Kayacik, Paul C Van Oorschot, Anil Somayaji. "A methodology for empirical analysis of permission-based security models and its application to android." In: *Proceedings of the 17th ACM conference on Computer and communications security*. 2010, pp. 73–84. DOI: 10.1145/1866307.1866317.

[29] Yakov Bart, Venkatesh Shankar, Fareena Sultan, Glen L Urban. "Are the drivers and role of online trust the same for all web sites and consumers? A large-scale exploratory empirical study." In: *Journal of marketing* 69.4 (2005), pp. 133–152.

[30] Susanne Barth, Menno DT de Jong, Marianne Junger, Pieter H Hartel, Janina C Roppelt. "Putting the privacy paradox to the test: Online privacy and security behaviors among users with technical knowledge, privacy awareness, and financial resources." In: *Telematics and informatics* 41 (2019), pp. 55–69. DOI: 10.1016/j.tele.2019.03.003.

[31] Melisa Basol, Jon Roozenbeek, Sander van der Linden. "Good news about bad news: Gamified inoculation boosts confidence and cognitive immunity against fake news." In: *Journal of cognition* 3.1 (2020). DOI: 10.5334/joc.91.

[32] Simone Bassanelli, Nicola Vasta, Antonio Bucchiarone, Annapaola Marconi. "Gamification for behavior change: A scientometric review." In: *Acta Psychologica* 228 (2022), p. 103657. DOI: 10.1016/j.actpsy.2022.103657.

[33] Roy F Baumeister, Kathleen D Vohs, David C Funder. "Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior?" In: *Perspectives on psychological science* 2.4 (2007), pp. 396–403.

[34] Carolin Baur. “Nachhaltigkeit in der Wertschöpfungskette: Das Problem des eingeschränkten moralischen Bewusstseins.” In: *Psychologie und Nachhaltigkeit: Konzeptionelle Grundlagen, Anwendungsbeispiele und Zukunftsperspektiven* (2018), pp. 149–163. DOI: 10.1007/978-3-658-19965-4_13.

[35] Morten Bay. “Social media ethics: A Rawlsian approach to hypertargeting and psychometrics in political and commercial campaigns.” In: *ACM Transactions on Social Computing* 1.4 (2018), pp. 1–14. DOI: 10.1145/3281450.

[36] Felix Beierle, Sandra C Matz, Mathias Allemand. “Mobile sensing in personality science.” In: *Mobile sensing in psychology: Methods and applications* (2023), p. 479.

[37] France Belanger, Janine S Hiller, Wanda J Smith. “Trustworthiness in electronic commerce: the role of privacy, security, and site attributes.” In: *The journal of strategic Information Systems* 11.3-4 (2002), pp. 245–270. DOI: 10.1016/S0963-8687(02)00018-5.

[38] Florian Bemmann, Daniel Buschek, Hussmann Heinrich. “Interactive End-User Machine Learning to Boost Explainability and Transparency of Digital Footprint Data.” In: Yokohama, Japan: HCXAI Workshop at ACM CHI 2021, 2021.

[39] Florian Bemmann, Daniel Buschek. “Interaction Challenges for N-Of-One Experiments based on Mobile Sensing Data.” In: Online: CHI’22 Workshop: Grand Challenges in Personal Informatics and AI, 2022.

[40] Florian Bemmann, Daniel Buschek. “LanguageLogger: A Mobile Keyboard Application for Studying Language Use in Everyday Text Communication in the Wild.” In: *Proc. ACM Hum.-Comput. Interact.* 4.EICS (June 2020). DOI: 10.1145/3397872.

[41] Florian Bemmann, Heinrich Hussmann. “Self-Reflection as a Tool to Foster Profound Sustainable Consumption Decisions.” In: *International Conference on ICT for Sustainability* (2020).

[42] Florian Bemmann, Timo Koch, Maximilian Bergmann, Clemens Stachl, Daniel Buschek, Ramona Schoedel, Sven Mayer. “Putting Language into Context Using Smartphone-Based Keyboard Logging.” In: *arXiv preprint arXiv:2403.05180* (2024).

[43] Florian Bemmann, Carmen Mayer, Sven Mayer. “Leveraging mobile sensing technology for societal change towards more sustainable behavior.” In: *arXiv preprint arXiv: 2303.12426* (2023).

[44] Florian Bemmann, Sven Mayer. “The Impact of Data Privacy on Users’ Smartphone App Adoption Decisions.” In: *Proc. ACM Hum.-Comput. Interact.* MobileHCI ’24 MHCI (2024). DOI: 10.1145/3676525.

[45] Florian Bemmann, Sven Mayer. “User-Centered Sustainable Technology Design: A Reflection on Human-Computer Interaction Research for a Sustainable Society.” In: *International Conference on ICT for Sustainability* (2023).

[46] Florian Bemmann, Helena Stoll, Sven Mayer. “Privacy Slider: Fine-Grain Privacy Control for Smartphones.” In: *Proc. ACM Hum.-Comput. Interact.* MobileHCI ’24 MHCI (2024). DOI: 10.1145/3676519.

[47] Florian Bemmann, Maximiliane Windl, Jonas Erbe, Sven Mayer, Heinrich Hussmann. “The Influence of Transparency and Control on the Willingness of Data Sharing in Adaptive Mobile Apps.” In: *Proc. ACM Hum.-Comput. Interact.* 6.MHCI (2022). DOI: 10.1145/3546724.

[48] Florian Bemmann, Maximiliane Windl, Tobias Knobloch, Sven Mayer. *Users’ In-Depth Privacy Concerns with Smartphone Data Collection*.

[49] Florian Bemmanna, Ramona Schoedela, Niels Van Berkel, Daniel Buschek. “Chatbots for experience sampling-initial opportunities and challenges.” In: *CEUR Workshop Proceedings*. Vol. 2903. CEUR Workshop Proceedings. 2021.

[50] Colin J Bennett, Jesse Gordon. “Understanding the “Micro” in Political Micro-Targeting: An Analysis of Facebook Digital Advertising in the 2019 Federal Canadian Election.” In: *Canadian Journal of Communication* 46.3 (2021), pp. 431–459. DOI: 10.22230/cjc.2021v46n3a3815.

[51] Marit Bentvelzen, Jasmin Niess, Mikołaj P Woźniak, Paweł W Woźniak. “The Development and Validation of the Technology-Supported Reflection Inventory.” In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021, pp. 1–8. DOI: 10.1145/3411764.3445673.

[52] Alastair R Beresford, Andrew Rice, Nicholas Skehin, Ripduman Sohan. “Mockdroid: trading privacy for application functionality on smartphones.” In: *Proceedings of the 12th workshop on mobile computing systems and applications*. 2011, pp. 49–54.

[53] Jan Hendrik Betzing, Matthias Tietz, Jan vom Brocke, Jörg Becker. “The impact of transparency on mobile privacy decision making.” In: *Electronic Markets* 30.3 (2020), pp. 607–625. DOI: 10.1007/s12525-019-00332-3.

[54] Christoph Bier, Kay Kühne, Jürgen Beyerer. "PrivacyInsight: the next generation privacy dashboard." In: *Annual Privacy Forum*. Springer. 2016, pp. 135–152. DOI: 10.1007/978-3-319-44760-5_9.

[55] Efsun Birtwistle, Ramona Schoedel, Florian Bemmann, Astrid Wirth, Christoph Sürig, Clemens Stachl, Markus Bühner, Frank Niklas. "Mobile sensing in psychological and educational research: Examples from two application fields." In: *International Journal of Testing* 22.3-4 (2022), pp. 264–288. DOI: 10.1080/15305058.2022.2036160.

[56] Jan Blom, Daniel Gatica-Perez, Niko Kiukkonen. "People-Centric Mobile Sensing with a Pragmatic Twist: From Behavioral Data Points to Active User Involvement." In: *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*. MobileHCI '11. Stockholm, Sweden: Association for Computing Machinery, 2011, 381–384. DOI: 10.1145/2037373.2037431.

[57] Matthias Böhmer, Brent Hecht, Johannes Schöning, Antonio Krüger, Gernot Bauer. "Falling Asleep with Angry Birds, Facebook and Kindle: A Large Scale Study on Mobile Application Usage." In: *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*. MobileHCI '11. Stockholm, Sweden: Association for Computing Machinery, 2011, 47–56. DOI: 10.1145/2037373.2037383.

[58] Kallista Bonawitz, Peter Kairouz, Brendan Mcmahan, Daniel Ramage. "Federated learning and privacy." In: *Communications of the ACM* 65.4 (2022), pp. 90–97.

[59] Bram Bonné, Sai Teja Peddinti, Igor Bilogrevic, Nina Taft. "Exploring decision making with Android's runtime permission dialogs using in-context surveys." In: *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. 2017, pp. 195–210.

[60] Katrin Borcea-Pfitzmann, Andreas Pfitzmann, Manuela Berg. "Privacy 3.0: Data minimization, user control, contextual integrity." In: *it-Information Technology* 53.1 (2011), pp. 34–40. DOI: 10.1524/itit.2011.0622.

[61] JL Boyles, A Smith, M Madden. *Apps and privacy: More than half of app users have uninstalled or decided to not install an app due to concerns about their personal information*. 2015.

[62] Jan Lauren Boyles, Aaron Smith, Mary Madden. "Privacy and data management on mobile devices." In: *Pew Internet & American Life Project* 4 (2012), pp. 1–19.

[63] Virginia Braun, Victoria Clarke. “Using thematic analysis in psychology.” In: *Qualitative research in psychology* 3.2 (2006), pp. 77–101. DOI: 10.1191/1478088706QP063OA.

[64] Taylor A Braund, Bridianne O’Dea, Debopriyo Bal, Kate Maston, Mark Larsen, Aliza Werner-Seidler, Gabriel Tillman, Helen Christensen. “Associations Between Smartphone Keystroke Metadata and Mental Health Symptoms in Adolescents: Findings From the Future Proofing Study.” In: *JMIR Mental Health* 10 (2023), e44986. DOI: 10.2196/44986.

[65] Saša Brdnik, Tjaša Heričko, Boštjan Šumak. “Intelligent user interfaces and their evaluation: a systematic mapping study.” In: *Sensors* 22.15 (2022), p. 5830. DOI: 10.3390/s22155830.

[66] Christina Bremer, Bran Knowles, Adrian Friday. “Have We Taken On Too Much?: A Critical Review of the Sustainable HCI Landscape.” In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 2022, pp. 1–11. DOI: 10.1145/3491102.3517609.

[67] John Brooke. “Sus: a “quick and dirty’usability.” In: *Usability evaluation in industry* 189.3 (1996), pp. 189–194.

[68] Lucas Brutschy, Pietro Ferrara, Omer Tripp, Marco Pistoia. “Shamdroid: gracefully degrading functionality in the presence of limited resource access.” In: *ACM SIGPLAN Notices* 50.10 (2015), pp. 316–331. DOI: 10.1145/2858965.2814296.

[69] Hronn Brynjarsdottir, Maria Håkansson, James Pierce, Eric Baumer, Carl DiSalvo, Phoebe Sengers. “Sustainably unpersuaded: how persuasion narrows our vision of sustainability.” In: *Proceedings of the sigchi conference on human factors in computing systems*. 2012, pp. 947–956. DOI: 10.1145/2207676.2208539.

[70] Tom Buchanan, Carina Paine, Adam N. Joinson, Ulf-Dietrich Reips. “Development of Measures of Online Privacy Concern and Protection for Use on the Internet.” In: *J. Am. Soc. Inf. Sci. Technol.* 58.2 (Jan. 2007), pp. 157–165. DOI: 10.1002/asi.v58:2.

[71] Daniel Buschek, Benjamin Bisinger, Florian Alt. “ResearchIME: A Mobile Keyboard Application for Studying Free Typing Behaviour in the Wild.” In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI ’18. Montreal QC, Canada: ACM, 2018, 255:1–255:14. DOI: 10.1145/3173574.3173829.

[72] Daniel Buschek, Sarah Völkel, Clemens Stachl, Lukas Mecke, Sarah Prange, Ken Pfeuffer. “Experience Sampling As Information Transmission: Perspective and Implications.” In: *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. UbiComp ’18. Singapore, Singapore: ACM, 2018, pp. 606–611. DOI: 10.1145/3267305.3267543.

[73] Hong Cao, Miao Lin. “Mining smartphone data for app usage prediction and recommendations: A survey.” en. In: *Pervasive and Mobile Computing* 37 (June 2017), pp. 1–22. DOI: 10.1016/j.pmcj.2017.01.007.

[74] Andrea Capponi, Claudio Fiandrino, Burak Kantarci, Luca Foschini, Dzmitry Kliazovich, Pascal Bouvry. “A survey on mobile crowdsensing systems: Challenges, solutions, and opportunities.” In: *IEEE communications surveys & tutorials* 21.3 (2019), pp. 2419–2465. DOI: 10.1109/COMST.2019.2914030.

[75] Filip Carlèn. “User Onboarding An investigation in how to increase the activation of new customers using design.” In: (2017).

[76] Ann Cavoukian, Scott Taylor, Martin E Abrams. “Privacy by Design: essential for organizational accountability and strong business practices.” In: *Identity in the Information Society* 3 (2010), pp. 405–413.

[77] Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, W Philip Kegelmeyer. “SMOTE: synthetic minority over-sampling technique.” In: *Journal of artificial intelligence research* 16 (2002), pp. 321–357. DOI: 10.1613/jair.953.

[78] Tsai-Wei Chen, S Shyam Sundar. “This app would like to use your current location to better serve you: Importance of user assent and system transparency in personalized mobile services.” In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018, pp. 1–13.

[79] Paula Glenda Ferrer Cheng, Roann Munoz Ramos, Jó Ágila Bitsch, Stephan Michael Jonas, Tim Ix, Portia Lynn Quetulio See, Klaus Wehrle. “Psychologist in a pocket: lexicon development and content validation of a mobile-based app for depression screening.” In: *JMIR mHealth and uHealth* 4.3 (2016), e5284. DOI: 10.2196/mhealth.5284.

[80] Mauro Cherubini, Rodrigo de Oliveira, Anna Hiltunen, Nuria Oliver. “Barriers and Bridges in the Adoption of Today’s Mobile Phone Contextual Services.” In: *Proceedings of the 13th International Conference on Human Computer Inter-*

action with Mobile Devices and Services. MobileHCI '11. Stockholm, Sweden: Association for Computing Machinery, 2011, 167–176. DOI: 10.1145/2037373.2037400.

[81] Bachir Chihani, Emmanuel Bertin, Noël Crespi. "A user-centric context-aware mobile assistant." In: *2013 17th International Conference on Intelligence in Next Generation Networks (ICIN)*. IEEE. 2013, pp. 110–117. DOI: 10.1109/ICIN.2013.6670901.

[82] Amita Goyal Chin, Mark A Harris, Robert Brookshire. "A bidirectional perspective of trust and risk in determining factors that influence mobile app installation." In: *International Journal of Information Management* 39 (2018), pp. 49–59. DOI: 10.1016/j.ijinfomgt.2017.11.010.

[83] Erika Chin, Adrienne Porter Felt, Vyas Sekar, David Wagner. "Measuring User Confidence in Smartphone Security and Privacy." In: *Proceedings of the Eighth Symposium on Usable Privacy and Security*. SOUPS '12. Washington, D.C.: Association for Computing Machinery, 2012. DOI: 10.1145/2335356.2335358.

[84] Hyunsung Cho, DaEun Choi, Donghwi Kim, Wan Ju Kang, Eun Kyoung Choe, Sung-Ju Lee. "Reflect, Not Regret: Understanding Regretful Smartphone Use with App Feature-Level Analysis." In: *Proc. ACM Hum.-Comput. Interact.* 5.CSCW2 (2021). DOI: 10.1145/3479600.

[85] Hyuntae Cho. "Walking Speed Estimation and Gait Classification Using Plantar Pressure and On-Device Deep Learning." In: *IEEE Sensors Journal* (2023). DOI: 10.1109/JSEN.2023.3305024.

[86] Hanbyul Choi, Jonghwa Park, Yoonhyuk Jung. "The role of privacy fatigue in online privacy behavior." In: *Computers in Human Behavior* 81 (2018), pp. 42–51. DOI: 10.1016/j.chb.2017.12.001.

[87] Noam Chomsky. "Explaining language use." In: *Philosophical topics* 20.1 (1992), pp. 205–231. DOI: 10.5840/philtopics19922017.

[88] Joobin Choobineh, Anil D Kini. "An Empirical Evaluation of the Factors Affecting Trust in Web Bankng Systems." In: *AMCIS 2000 Proceedings* (2000), p. 169.

[89] Delphine Christin, Andreas Reinhardt, Salil S Kanhere, Matthias Hollick. "A survey on privacy in mobile participatory sensing applications." In: *Journal of systems and software* 84.11 (2011), pp. 1928–1946. DOI: 10.1016/j.jss.2011.06.073.

[90] Kenneth Ward Church. “Word2Vec.” In: *Natural Language Engineering* 23.1 (2017), pp. 155–162. DOI: 10.1017/S1351324916000334.

[91] Adrian K Clear, Adrian Friday, Mark Rouncefield, Alan Chamberlain. “Supporting sustainable food shopping.” In: *IEEE Pervasive Computing* 14.4 (2015), pp. 28–36. DOI: 10.1109/MPRV.2015.78.

[92] Adrian K Clear, Kirstie O’neill, Adrian Friday, Mike Hazas. “Bearing an Open “Pandora’s Box”: HCI for Reconciling Everyday Food and Sustainability.” In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 23.5 (2016), p. 28. DOI: 10.1145/2970817.

[93] Roger Collier. “Mental health in the smartphone era.” In: *CMAJ : Canadian Medical Association journal = journal de l’Association medicale canadienne* 188.16 (2016), pp. 1141–1142. DOI: 10.1503/cmaj.109-5336.

[94] Jessica Colnago, Lorrie Faith Cranor, Alessandro Acquisti, Kate Hazel Stanton. “Is it a concern or a preference? An investigation into the ability of privacy scales to capture and distinguish granular privacy constructs.” In: *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*. 2022, pp. 331–346. DOI: 10.5555/3563609.3563627.

[95] Mauro Conti, Vu Thien Nga Nguyen, Bruno Crispo. “Crepe: Context-related policy enforcement for android.” In: *Information Security: 13th International Conference, ISC 2010, Boca Raton, FL, USA, October 25-28, 2010, Revised Selected Papers 13*. Springer. 2011, pp. 331–345. DOI: 10.5555/1949317.1949355.

[96] Kovila PL Coopamootoo. “Usage patterns of privacy-enhancing technologies.” In: *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*. 2020, pp. 1371–1390. DOI: 10.1145/3372297.3423347.

[97] Kovila PL Coopamootoo, Thomas Groß. “Mental models: an approach to identify privacy concern and behavior.” In: *Symposium on Usable Privacy and Security (SOUPS)*. 2014, pp. 9–11.

[98] Victor P Cornet, Richard J Holden. “Systematic review of smartphone-based passive sensing for health and wellbeing.” In: *Journal of biomedical informatics* 77 (2018), pp. 120–132. DOI: 10.1016/j.jbi.2017.12.008.

[99] Cynthia L Corritore, Beverly Kracher, Susan Wiedenbeck. “On-line trust: concepts, evolving themes, a model.” In: *International journal of human-computer studies* 58.6 (2003), pp. 737–758. DOI: 10.1016/S1071-5819(03)00041-7.

[100] Elisa Costante, Jerry Den Hartog, Milan Petkovic. “On-line trust perception: What really matters.” In: *2011 1st Workshop on Socio-Technical Aspects in Security and Trust (STAST)*. IEEE. 2011, pp. 52–59. DOI: 10.1109/STAST.2011.6059256.

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

[102] Henriette Cramer, Paloma de Juan, Joel Tetreault. “Sender-intended Functions of Emojis in US Messaging.” In: *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI ’16. Florence, Italy: ACM, 2016, pp. 504–509. DOI: 10.1145/2935334.2935370.

[103] Robert E Crossler, France Bélanger. “Why would I use location-protective settings on my smartphone? Motivating protective behaviors and the existence of the privacy knowledge–belief gap.” In: *Information Systems Research* 30.3 (2019), pp. 995–1006. DOI: 10.1287/isre.2019.0846.

[104] Douglas P Crowne, David Marlowe. “The approval motive: Studies in evaluative dependence.” In: (1964).

[105] Mary J Culnan, Pamela K Armstrong. “Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation.” In: *Organization science* 10.1 (1999), pp. 104–115. DOI: 10.1287/orsc.10.1.104.

[106] Sauvik Das, Tiffany Hyun-Jin Kim, Laura A. Dabbish, Jason I. Hong. “The Effect of Social Influence on Security Sensitivity.” In: *10th Symposium On Usable Privacy and Security (SOUPS 2014)*. Menlo Park, CA: USENIX Association, July 2014, pp. 143–157.

[107] Mark de Reuver, Harry Bouwman. “Dealing with self-report bias in mobile Internet acceptance and usage studies.” In: *Information & Management* 52.3 (2015), pp. 287–294. DOI: 10.1016/j.im.2014.12.002.

[108] Mark de Reuver, Shahrokh Nikou, Harry Bouwman. “The interplay of costs, trust and loyalty in a service industry in transition: The moderating effect of smartphone adoption.” In: *Telematics and Informatics* 32.4 (2015), pp. 694–700.

[109] Robert F DeVellis, Carolyn T Thorpe. *Scale development: Theory and applications*. Sage publications, 2021.

[110] Kenan Degirmenci. “Mobile users’ information privacy concerns and the role of app permission requests.” In: *International Journal of Information Management* 50 (2020), pp. 261–272. DOI: 10.1016/j.ijinfomgt.2019.05.010.

[111] Paula Delgado-Santos, Giuseppe Stragapede, Ruben Tolosana, Richard Guest, Farzin Deravi, Ruben Vera-Rodriguez. “A Survey of Privacy Vulnerabilities of Mobile Device Sensors.” In: *ACM Comput. Surv.* 54.11s (2022). DOI: 10.1145/3510579.

[112] Sebastian Deterding, Dan Dixon, Rilla Khaled, Lennart Nacke. “From game design elements to gamefulness: defining "gamification".” In: *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*. MindTrek ’11. Tampere, Finland: Association for Computing Machinery, 2011, 9–15. DOI: 10.1145/2181037.2181040.

[113] Saurabh Dhawan, Simon Hegelich. “From outside in: profiling, persuasion and political opinion in the age of big data.” In: *Digital Phenotyping and Mobile Sensing: New Developments in Psychoinformatics*. Springer, 2022, pp. 151–169. DOI: 10.1007/978-3-030-98546-2_10.

[114] Daniel Di Matteo, Alexa Fine, Kathryn Fotinos, Jonathan Rose, Martin Katzman, et al. “Patient willingness to consent to mobile phone data collection for mental health apps: structured questionnaire.” In: *JMIR mental health* 5.3 (2018), e9539. DOI: 10.2196/mental.9539.

[115] Tamara Dinev, Heng Xu, Jeff H Smith, Paul Hart. “Information privacy and correlates: an empirical attempt to bridge and distinguish privacy-related concepts.” In: *European Journal of Information Systems* 22.3 (2013), pp. 295–316. DOI: 10.1057/ejis.2012.23.

[116] Irit Dinur, Kobbi Nissim. “Revealing Information While Preserving Privacy.” In: *Proceedings of the Twenty-Second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*. PODS ’03. San Diego, California: Association for Computing Machinery, 2003, 202–210. DOI: 10.1145/773153.773173.

[117] Trinh Minh Tri Do, Daniel Gatica-Perez. “Where and what: Using smartphones to predict next locations and applications in daily life.” In: *Pervasive and Mobile Computing* 12 (2014), pp. 79–91. DOI: 10.1016/j.pmcj.2013.03.006.

[118] Leyla Dogruel, Sven Joeckel, Nicholas D Bowman. “Choosing the right app: An exploratory perspective on heuristic decision processes for smartphone app selection.” In: *Mobile Media & Communication* 3.1 (2015), pp. 125–144. DOI: 10.1177/2050157914557509.

[119] Stewart I Donaldson, Elisa J Grant-Vallone. “Understanding self-report bias in organizational behavior research.” In: *Journal of business and Psychology* 17 (2002), pp. 245–260. DOI: 10.1023/a:1019637632584.

[120] Lauren Donis. “How filter bubbles and echo chambers reinforce negative beliefs and spread misinformation through social media.” In: *Debating Communities and Networks Conference XII*. 2021.

[121] Nick Doty, Mohit Gupta. “Privacy design patterns and anti-patterns patterns misapplied and unintended consequences.” In: (2013).

[122] Gibbs A Doward J. *Did Cambridge Analytica Influence the Brexit Vote and the US Election?* 2017. URL: https://www.theguardian.com/politics/2017/mar/04/nigel-oakes-cambridge-analytica-what-role-brexit-trump.

[123] Nora A Draper, Joseph Turow. “The corporate cultivation of digital resignation.” In: *New media & society* 21.8 (2019), pp. 1824–1839. DOI: 10.1177/1461444819833331.

[124] Michelle Drouin, Claire Davis. “R u txting? Is the Use of Text Speak Hurting Your Literacy?” In: *Journal of Literacy Research* 41.1 (2009), pp. 46–67. DOI: 10.1080/10862960802695131.

[125] Sophia Xiaoxia Duan, Hepu Deng. “Exploring privacy paradox in contact tracing apps adoption.” In: *Internet Research* 32.5 (2022), pp. 1725–1750. DOI: 10.1108/INTR-03-2021-0160.

[126] Joy Dutta, Chandreyee Chowdhury, Sarbani Roy, Asif Iqbal Middya, Firoj Gazi. “Towards smart city: sensing air quality in city based on opportunistic crowd-sensing.” In: *Proceedings of the 18th international conference on distributed computing and networking*. 2017, pp. 1–6. DOI: 10.1145/3007748.3018286.

[127] Aarthi Easwara Moorthy, Kim-Phuong L Vu. “Privacy concerns for use of voice activated personal assistant in the public space.” In: *International Journal of Human-Computer Interaction* 31.4 (2015), pp. 307–335. DOI: 10.1080/10447318.2014.986642.

[128] Christos Efstratiou, Ilias Leontiadis, Marco Picone, Kiran K Rachuri, Cecilia Mascolo, Jon Crowcroft. “Sense and sensibility in a pervasive world.” In: *International Conference on Pervasive Computing*. Cham: Springer, 2012, pp. 406–424. DOI: 10.1007/978-3-642-31205-2_25.

[129] Serge Egelman, Adrienne Porter Felt, David Wagner. “Choice architecture and smartphone privacy: There’sa price for that.” In: *The economics of information security and privacy* (2013), pp. 211–236. DOI: 10.1007/978-3-642-39498-0_10.

[130] Malin Eiband, Sarah Theres Völkel, Daniel Buschek, Sophia Cook, Heinrich Hussmann. “When People and Algorithms Meet: User-reported Problems in Intelligent Everyday Applications.” en. In: (2019), p. 11. DOI: 10.1145/3301275.3302262.

[131] Johannes C. Eichstaedt, Robert J. Smith, Raina M. Merchant, Lyle H. Ungar, Patrick Crutchley, Daniel Preoţiuc-Pietro, David A. Asch, H. Andrew Schwartz. “Facebook language predicts depression in medical records.” In: *Proceedings of the National Academy of Sciences* 115.44 (2018), pp. 11203–11208. DOI: 10.1073/pnas.1802331115.

[132] Gudrun Eisele, Hugo Vachon, Ginette Lafit, Peter Kuppens, Marlies Houben, Inez Myin-Germeys, Wolfgang Viechtbauer. “The effects of sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population.” In: *Assessment* 29.2 (2022), pp. 136–151.

[133] Yusra Elbitar, Michael Schilling, Trung Tin Nguyen, Michael Backes, Sven Bugiel. “Explanation beats context: The effect of timing & rationales on users’ runtime permission decisions.” In: *30th USENIX Security Symposium (USENIX Security 21)*. 2021, pp. 785–802.

[134] Anne Elevelt, Peter Lugtig, Vera Toepoel. “Doing a time use survey on smartphones only: What factors predict nonresponse at different stages of the survey process?” In: *Survey Research Methods*. Vol. 13. 2. 2019, pp. 195–213. DOI: 10.18148/srm/2019.v13i2.7385.

[135] Chris Elsden, Abigail C Durrant, David S Kirk. “It’s just my history isn’t it? Understanding smart journaling practices.” In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2016, pp. 2819–2831.

[136] William Enck, Machigar Ongtang, Patrick McDaniel. “On Lightweight Mobile Phone Application Certification.” In: *Proceedings of the 16th ACM Conference on Computer and Communications Security*. CCS ’09. Chicago, Illinois, USA: Association for Computing Machinery, 2009, 235–245. DOI: 10.1145/1653662.1653691.

[137] Anita Engels, Jochem Marotzke, Eduardo Gresse, Andrés López-Rivera, Anna Pagnone, Jan Wilkens. *Hamburg Climate Futures Outlook: The plausibility of a 1.5°C limit to global warming - social drivers and physical processes*. Feb. 2023. DOI: 10.25592/uhhfdm.11230.

[138] Úlfar Erlingsson, Vasyl Pihur, Aleksandra Korolova. "RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response." In: *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. CCS '14. Scottsdale, Arizona, USA: Association for Computing Machinery, 2014, 1054–1067. DOI: 10.1145/2660267.2660348.

[139] Andrea Everard, Scott McCoy. "Effect of Presentation Flaw Attribution on Website Quality, Trust, and Abandonment." In: *Australasian Journal of Information Systems* 16.2 (2010). DOI: 10.3127/ajis.v16i2.516.

[140] Zheran Fang, Weili Han, Yingjiu Li. "Permission based Android security: Issues and countermeasures." In: *computers & security* 43 (2014), pp. 205–218.

[141] Florian M Farke, David G Balash, Maximilian Golla, Markus Dürmuth, Adam J Aviv. "Are Privacy Dashboards Good for End Users? Evaluating User Perceptions and Reactions to Google's My Activity." In: *30th USENIX Security Symposium (USENIX Security 21)*. 2021, pp. 483–500. DOI: 10.48550/arXiv.2105.14066.

[142] Johannes Feichtner, Stefan Gruber. "Understanding Privacy Awareness in Android App Descriptions Using Deep Learning." In: *Proceedings of the Tenth ACM Conference on Data and Application Security and Privacy*. CODASPY '20. New Orleans, LA, USA: Association for Computing Machinery, 2020, 203–214. DOI: 10.1145/3374664.3375730.

[143] Dan Feldman, Ashwin Rao, Zihao He, Kristina Lerman. "Affective polarization in social networks." In: *arXiv preprint arXiv:2310.18553* (2023).

[144] Adrienne Porter Felt, Elizabeth Ha, Serge Egelman, Ariel Haney, Erika Chin, David Wagner. "Android permissions: User attention, comprehension, and behavior." In: *Proceedings of the eighth symposium on usable privacy and security*. 2012, pp. 1–14. DOI: 10.1145/2335356.2335360.

[145] Yuanyuan Feng, Yaxing Yao, Norman Sadeh. "A Design Space for Privacy Choices: Towards Meaningful Privacy Control in the Internet of Things." In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. Yokohama, Japan: Association for Computing Machinery, 2021. DOI: 10.1145/3411764.3445148.

[146] Jeffrey D Fisher, William A Fisher. “Changing AIDS-risk behavior.” In: *Psychological bulletin* 111.3 (1992), p. 455.

[147] Carlos Flavián, Miguel Guinalíu. “Consumer trust, perceived security and privacy policy: three basic elements of loyalty to a web site.” In: *Industrial management & data Systems* 106.5 (2006), pp. 601–620.

[148] Daniel Fleischhauer, Benjamin Engelstätter, Omid Tafreschi. “The Privacy Paradox in Smartphone Users.” In: *Proceedings of the 21st International Conference on Mobile and Ubiquitous Multimedia*. MUM ’22. Lisbon, Portugal: Association for Computing Machinery, 2022, 62–70. DOI: 10.1145/3568444.3568467.

[149] Brain J Fogg. “Creating persuasive technologies: an eight-step design process.” In: *Proceedings of the 4th international conference on persuasive technology*. 2009, pp. 1–6. DOI: 10.1145/1541948.1542005.

[150] Claes Fornell, David F Larcker. “Evaluating structural equation models with unobservable variables and measurement error.” In: *Journal of marketing research* 18.1 (1981), pp. 39–50. DOI: 10.1177/002224378101800104.

[151] Thomas Franke, Christiane Attig, Daniel Wessel. “A personal resource for technology interaction: development and validation of the affinity for technology interaction (ATI) scale.” In: *International Journal of Human–Computer Interaction* 35.6 (2019), pp. 456–467. DOI: 10.1080/10447318.2018.1456150.

[152] Alisa Frik, Juliann Kim, Joshua Rafael Sanchez, Joanne Ma. “Users’ expectations about and use of smartphone privacy and security settings.” In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 2022, pp. 1–24. DOI: 10.1145/3491102.3517504.

[153] Jon Froehlich, Tawanna Dillahunt, Predrag Klasnja, Jennifer Mankoff, Sunny Consolvo, Beverly Harrison, James A Landay. “UbiGreen: investigating a mobile tool for tracking and supporting green transportation habits.” In: *Proceedings of the sigchi conference on human factors in computing systems*. 2009, pp. 1043–1052. DOI: 10.1145/1518701.1518861.

[154] Chris Fullwood, Lisa J. Orchard, Sarah A. Floyd. “Emoticon convergence in Internet chat rooms.” In: *Social Semiotics* 23.5 (2013), pp. 648–662. DOI: 10.1080/10350330.2012.739000.

[155] Carol Fung, Vivian Motti, Katie Zhang, Yanjun Qian. “A Study of User Concerns about Smartphone Privacy.” In: *2022 6th Cyber Security in Networking Conference (CSNet)*. IEEE. 2022, pp. 1–8. DOI: 10.1109/CSNet56116.2022.9955623.

[156] Frederik Funke, Ulf-Dietrich Reips. “Why semantic differentials in web-based research should be made from visual analogue scales and not from 5-point scales.” In: *Field methods* 24.3 (2012), pp. 310–327.

[157] Marco Furini, Silvia Mirri, Manuela Montangero, Catia Prandi. “Privacy perception when using smartphone applications.” In: *Mobile Networks and Applications* 25.3 (2020), pp. 1055–1061. DOI: 10.1007/s11036-020-01529-z.

[158] R Michael Furr. “Personality psychology as a truly behavioural science.” In: *European Journal of Personality* 23.5 (2009), pp. 369–401. DOI: 10.1002/per.724.

[159] Sandra Gabriele, Sonia Chiasson. “Understanding Fitness Tracker Users’ Security and Privacy Knowledge, Attitudes and Behaviours.” In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI ’20. Honolulu, HI, USA: Association for Computing Machinery, 2020, 1–12. DOI: 10.1145/3313831.3376651.

[160] Anirudh Ganesh, Chinenye Ndulue, Rita Orji. “Smartphone security and privacy–a gamified persuasive approach with protection motivation theory.” In: *International Conference on Persuasive Technology*. Springer. 2022, pp. 89–100.

[161] Raghu K Ganti, Fan Ye, Hui Lei. “Mobile crowdsensing: current state and future challenges.” In: *IEEE communications Magazine* 49.11 (2011), pp. 32–39. DOI: 10.1109/MCOM.2011.6069707.

[162] Hongcan Gao, Chenkai Guo, Yanfeng Wu, Naipeng Dong, Xiaolei Hou, Sihan Xu, Jing Xu. “AutoPer: Automatic Recommender for Runtime-Permission in Android Applications.” In: *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*. Vol. 1. IEEE, 2019, pp. 107–116. DOI: 10.1109/COMPSAC.2019.00024.

[163] Roxana Geambasu, Tadayoshi Kohno, Amit A Levy, Henry M Levy. “Vanish: Increasing Data Privacy with Self-Destructing Data.” In: *USENIX security symposium*. Vol. 316. 2009, pp. 10–5555.

[164] David Gefen, Elena Karahanna, Detmar W Straub. “Trust and TAM in online shopping: An integrated model.” In: *MIS quarterly* (2003), pp. 51–90.

[165] Surjya Ghosh, Niloy Ganguly, Bivas Mitra, Pradipta De. “Tapsense: Combining self-report patterns and typing characteristics for smartphone based emotion detection.” In: *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*. 2017, pp. 1–12.

[166] J. Golbeck, C. Robles, M. Edmondson, K. Turner. “Predicting Personality from Twitter.” In: *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*. 2011, pp. 149–156. DOI: 10.1109/PASSAT/SocialCom.2011.33.

[167] Alejandra Gomez Ortega, Jacky Bourgeois, Gerd Kortuem. “Towards designerly data donation.” In: *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*. 2021, pp. 496–501. DOI: 10.1145/3460418.3479362.

[168] Dale L Goodhue. “Understanding user evaluations of information systems.” In: *Management science* 41.12 (1995), pp. 1827–1844. DOI: 10.1287/mnsc.41.12.1827.

[169] Samuel D Gosling, Oliver P John, Kenneth H Craik, Richard W Robins. “Do people know how they behave? Self-reported act frequencies compared with online codings by observers.” In: *Journal of personality and social psychology* 74.5 (1998), p. 1337. DOI: 10.1037/0022-3514.74.5.1337.

[170] Fenne große Deters, Ramona Schoedel. “Keep on scrolling? Using intensive longitudinal smartphone sensing data to assess how everyday smartphone usage behaviors are related to well-being.” In: *Computers in Human Behavior* 150 (2024), p. 107977. DOI: 10.1016/j.chb.2023.107977.

[171] Jie Gu, Yunjie Calvin Xu, Heng Xu, Cheng Zhang, Hong Ling. “Privacy concerns for mobile app download: An elaboration likelihood model perspective.” In: *Decision Support Systems* 94 (2017), pp. 19–28. DOI: 10.1016/j.dss.2016.10.002.

[172] Anshita Gupta, Sudip Misra, Nidhi Pathak. “StressAlly: A Smartphone-Based Stress Companion Recommender System for Students.” In: *GLOBECOM 2023-2023 IEEE Global Communications Conference*. IEEE. 2023, pp. 504–509. DOI: 10.1109/GLOBECOM54140.2023.10437258.

[173] Mattia Gustarini, Katarzyna Wac, Anind K Dey. “Anonymous smartphone data collection: factors influencing the users’ acceptance in mobile crowd sensing.” In: *Personal and Ubiquitous Computing* 20 (2016), pp. 65–82. DOI: 10.1007/s00779-015-0898-0.

[174] Georg-Christoph Haas, Frauke Kreuter, Florian Keusch, Mark Trappmann, Sebastian Bähr. “Effects of incentives in smartphone data collection.” In: *Big Data Meets Survey Science: A Collection of Innovative Methods* (2020), pp. 387–414.

[175] Cecilia Håkansson, Göran Finnveden. “Indirect rebound and reverse rebound effects in the ICT-sector and emissions of CO2.” In: *EnviroInfo and ICT for Sustainability 2015*. Atlantis Press. 2015, pp. 66–73.

[176] Stuart Hallifax, Audrey Serna, Jean-Charles Marty, Elise Lavoué. “A Design Space For Meaningful Structural Gamification.” In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI EA ’18. Montreal, QC, Canada: Association for Computing Machinery, 2018, 1–6. DOI: 10.1145/3170427.3188442.

[177] Jaap Ham, Cees Midden. “Ambient persuasive technology needs little cognitive effort: the differential effects of cognitive load on lighting feedback versus factual feedback.” In: *Persuasive Technology: 5th International Conference, PERSUASIVE 2010, Copenhagen, Denmark, June 7-10, 2010. Proceedings 5*. Springer. 2010, pp. 132–142. DOI: 10.1007/978-3-642-13226-1_14.

[178] Michael D. Hanus, Jesse Fox. “Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance.” In: *Computers & Education* 80 (2015), pp. 152–161. DOI: 10.1016/j.compedu.2014.08.019.

[179] Gabriella M Harari. “A process-oriented approach to respecting privacy in the context of mobile phone tracking.” In: *Current opinion in psychology* 31 (2020), pp. 141–147. DOI: 10.1016/j.copsyc.2019.09.007.

[180] Gabriella M Harari, Samuel D Gosling. “Understanding behaviours in context using mobile sensing.” In: *Nature Reviews Psychology* (2023), pp. 1–13.

[181] Gabriella M Harari, Sandrine R Müller, Min SH Aung, Peter J Rentfrow. “Smartphone sensing methods for studying behavior in everyday life.” In: *Current opinion in behavioral sciences* 18 (2017), pp. 83–90. DOI: 10.1016/j.cobeha.2017.07.018.

[182] Gabriella M Harari, Sandrine R Müller, Clemens Stachl, Rui Wang, Weichen Wang, Markus Bühner, Peter J Rentfrow, Andrew T Campbell, Samuel D Gosling. “Sensing sociability: Individual differences in young adults’ conversation, calling, texting, and app use behaviors in daily life.” In: *Journal of personality and social psychology* 119.1 (2020), p. 204. DOI: 10.1037/pspp0000245.

[183] Gunnar Harboe, Elaine M. Huang. “Real-World Affinity Diagramming Practices: Bridging the Paper-Digital Gap.” In: *Proc. 33rd Annual ACM Conf. Human Factors in Computing Systems*. New York, NY, USA: ACM, 2015, pp. 95–104. DOI: 10.1145/2702123.2702561.

[184] Andrew Hard, Kanishka Rao, Rajiv Mathews, Swaroop Ramaswamy, Françoise Beaufays, Sean Augenstein, Hubert Eichner, Chloé Kiddon, Daniel Ramage. “Federated learning for mobile keyboard prediction.” In: *arXiv preprint arXiv:1811.03604* (2018).

[185] Mike Hazas, AJ Bernheim Brush, James Scott. “Sustainability does not begin with the individual.” In: *Interactions* 19.5 (2012), pp. 14–17. DOI: https://psycnet.apa.org/doi/10.1037/pspp0000245.

[186] Cornelia Helfferich. *Die Qualität qualitativer Daten*. Vol. 4. Springer, 2011.

[187] Eelco Herder, Olaf van Maaren. “Privacy Dashboards: The Impact of the Type of Personal Data and User Control on Trust and Perceived Risk.” In: *Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization*. UMAP ’20 Adjunct. Genoa, Italy: Association for Computing Machinery, 2020, 169–174. DOI: 10.1145/3386392.3399557.

[188] Susan C. Herring, John C. Paolillo. “Gender and genre variation in weblogs.” In: *Journal of Sociolinguistics* 10.4 (2006), pp. 439–459. DOI: 10.1111/j.1467-9841.2006.00287.x.

[189] Donna L Hoffman, Thomas P Novak, Marcos Peralta. “Building consumer trust online.” In: *Communications of the ACM* 42.4 (1999), pp. 80–85.

[190] Jason I Hong, Yuvraj Agarwal, Matt Fredrikson, Mike Czapik, Shawn Hanna, Swarup Sahoo, Judy Chun, Won-Woo Chung, Aniruddh Iyer, Ally Liu, et al. *Designing Privacy-Enhanced Android’s User Interfaces*. Tech. rep. Carnegie Mellon University, 2021.

[191] Christel Hopf. “5.2 Qualitative Interviews–ein Überblick.” In: *Qualitative Forschung. Ein Handbuch* 9 (2012), pp. 349–360.

[192] Jung-Kuei Hsieh, Hsiang-Tzu Li. “Exploring the fit between mobile application service and application privacy.” In: *Journal of Services Marketing* 36.2 (2022), pp. 264–282. DOI: 10.1108/JSM-01-2021-0023.

[193] Rui Hu, Yuanxiong Guo, Hongning Li, Qingqi Pei, Yanmin Gong. “Personalized federated learning with differential privacy.” In: *IEEE Internet of Things Journal* 7.10 (2020), pp. 9530–9539. DOI: 10.1109/JIOT.2020.2991416.

[194] Jianjun Huang, Zhichun Li, Xusheng Xiao, Zhenyu Wu, Kangjie Lu, Xiangyu Zhang, Guofei Jiang. "{SUPOR}: Precise and scalable sensitive user input detection for android apps." In: *24th USENIX Security Symposium (USENIX Security 15)*. 2015, pp. 977–992. DOI: 10.5555/2831143.2831205.

[195] Martina Z Huber, Lorenz M Hilty. "Gamification and sustainable consumption: overcoming the limitations of persuasive technologies." In: *ICT innovations for sustainability*. Springer. 2015, pp. 367–385. DOI: 10.1007/978-3-319-09228-7_22.

[196] Galen Chin-Lun Hung, Pei-Ching Yang, Chen-Yi Wang, Jung-Hsien Chiang. "A smartphone-based personalized activity recommender system for patients with depression." In: *Proceedings of the 5th EAI international conference on wireless mobile communication and healthcare*. 2015, pp. 253–257.

[197] Joshua B Hurwitz. "User choice, privacy sensitivity, and acceptance of personal information collection." In: *European data protection: Coming of age*. Springer, 2012, pp. 295–312. DOI: 10.1007/978-94-007-5170-5_13.

[198] Giovanni Iachello, Jason Hong, et al. "End-user privacy in human–computer interaction." In: *Foundations and Trends® in Human–Computer Interaction* 1.1 (2007), pp. 1–137.

[199] Ozlem Durmaz Incel, Sevda Özge Bursa. "On-device deep learning for mobile and wearable sensing applications: A review." In: *IEEE Sensors Journal* 23.6 (2023), pp. 5501–5512.

[200] Philip G. Inglesant, M. Angela Sasse. "The true cost of unusable password policies: password use in the wild." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '10. Atlanta, Georgia, USA: Association for Computing Machinery, 2010, 383–392. DOI: 10.1145/1753326.1753384.

[201] Annette Jäckle, Jonathan Burton, Mick P Couper, Carli Lessof. "Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel: Coverage and participation rates and biases." In: *Survey Research Methods*. Vol. 13. 1. 2019, pp. 23–44. DOI: 10.18148/srm/2019.v1i1.7297.

[202] Tun-Min Catherine Jai, Nancy J King. "Privacy versus reward: Do loyalty programs increase consumers' willingness to share personal information with third-party advertisers and data brokers?" In: *Journal of Retailing and Consumer Services* 28 (2016), pp. 296–303. DOI: 10.1016/j.jretconser.2015.01.005.

[203] Milena Janic, Jan Pieter Wijbenga, Thijs Veugen. “Transparency enhancing tools (TETs): an overview.” In: *2013 Third Workshop on Socio-Technical Aspects in Security and Trust*. IEEE. 2013, pp. 18–25. DOI: 10.1109/STAST.2013.11.

[204] Jinseong Jeon, Kristopher K Micinski, Jeffrey A Vaughan, Ari Fogel, Nikhilesh Reddy, Jeffrey S Foster, Todd Millstein. “Dr. android and mr. hide: fine-grained permissions in android applications.” In: *Proceedings of the second ACM workshop on Security and privacy in smartphones and mobile devices*. 2012, pp. 3–14.

[205] Haojian Jin. “Modular Privacy Flows: A Design Pattern for Data Minimization.” PhD thesis. PhD thesis, Carnegie Mellon University, 2022. 6.2, 2022.

[206] Leslie K John, Alessandro Acquisti, George Loewenstein. “Strangers on a plane: Context-dependent willingness to divulge sensitive information.” In: *Journal of consumer research* 37.5 (2011), pp. 858–873. DOI: 10.1086/656423.

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

[208] Yoonhyuk Jung. “What a smartphone is to me: understanding user values in using smartphones.” In: *Information Systems Journal* 24.4 (2014), pp. 299–321. DOI: 10.1111/isj.12031.

[209] Nesrine Kaaniche, Maryline Laurent, Sana Belguith. “Privacy enhancing technologies for solving the privacy-personalization paradox: Taxonomy and survey.” In: *Journal of Network and Computer Applications* 171 (2020), p. 102807.

[210] Vaiva Kalnikaite, Yvonne Rogers, Jon Bird, Nicolas Villar, Khaled Bachour, Stephen Payne, Peter M Todd, Johannes Schöning, Antonio Krüger, Stefan Kreitmayer. “How to nudge in Situ: designing lambent devices to deliver salient information in supermarkets.” In: *Proceedings of the 13th international conference on Ubiquitous computing*. ACM. 2011, pp. 11–20. DOI: 10.1145/2030112.2030115.

[211] Yufan Kang, Mohammad Saiedur Rahaman, Yongli Ren, Mark Sanderson, Ryen W White, Flora D Salim. “App usage on-the-move: Context-and commute-aware next app prediction.” In: *Pervasive and Mobile Computing* 87 (2022), p. 101704.

[212] Anjuli Kannan, Karol Kurach, Sujith Ravi, Tobias Kaufmann, Andrew Tomkins, Balint Miklos, Greg Corrado, Laszlo Lukacs, Marina Ganea, Peter Young, Vivek Ramavajjala. “Smart Reply: Automated Response Suggestion for Email.” In: *Pro-*

ceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '16. San Francisco, California, USA: ACM, 2016, pp. 955–964. DOI: 10.1145/2939672.2939801.

[213] Sabrina Karwatzki, Olga Dytynko, Manuel Trenz, Daniel Veit. "Beyond the personalization–privacy paradox: Privacy valuation, transparency features, and service personalization." In: *Journal of Management Information Systems* 34.2 (2017), pp. 369–400. DOI: 10.1080/07421222.2017.1334467.

[214] Aycan Kaya, Reha Ozturk, Cigdem Altin Gumussoy. "Usability measurement of mobile applications with system usability scale (SUS)." In: *Industrial Engineering in the Big Data Era: Selected Papers from the Global Joint Conference on Industrial Engineering and Its Application Areas, GJCIE 2018, June 21–22, 2018, Nevsehir, Turkey*. Springer. 2019, pp. 389–400.

[215] Flavius Kehr, Tobias Kowatsch, Daniel Wentzel, Elgar Fleisch. "Blissfully ignorant: the effects of general privacy concerns, general institutional trust, and affect in the privacy calculus." In: *Information Systems Journal* 25.6 (2015), pp. 607–635. DOI: 10.1111/isj.12062.

[216] Mark J Keith, Samuel C Thompson, Joanne Hale, Paul Benjamin Lowry, Chapman Greer. "Information disclosure on mobile devices: Re-examining privacy calculus with actual user behavior." In: *International journal of human-computer studies* 71.12 (2013), pp. 1163–1173. DOI: http://dx.doi.org/10.1016/j.ijhcs.2013.08.016.

[217] Patrick Gage Kelley, Sunny Consolvo, Lorrie Faith Cranor, Jaeyeon Jung, Norman Sadeh, David Wetherall. "A conundrum of permissions: installing applications on an android smartphone." In: *Financial Cryptography and Data Security: FC 2012 Workshops, USEC and WECSR 2012, Kralendijk, Bonaire, March 2, 2012, Revised Selected Papers 16*. Springer. 2012, pp. 68–79. DOI: 10.1007/978-3-642-34638-5_6.

[218] Shawna Kelly, Bonnie Nardi. "Playing with sustainability: Using video games to simulate futures of scarcity." In: *First Monday* (2014).

[219] Paul E Ketelaar, Mark Van Balen. "The smartphone as your follower: The role of smartphone literacy in the relation between privacy concerns, attitude and behaviour towards phone-embedded tracking." In: *Computers in Human Behavior* 78 (2018), pp. 174–182. DOI: 10.1016/j.chb.2017.09.034.

[220] Florian Keusch, Sebastian Bähr, Georg-Christoph Haas, Frauke Kreuter, Mark Trappmann. “Coverage error in data collection combining mobile surveys with passive measurement using apps: Data from a German national survey.” In: *Sociological Methods & Research* 52.2 (2023), pp. 841–878.

[221] Florian Keusch, Bella Struminskaya, Christopher Antoun, Mick P Couper, Frauke Kreuter. “Willingness to participate in passive mobile data collection.” In: *Public opinion quarterly* 83.S1 (2019), pp. 210–235. DOI: 10.1093/poq/nfz007.

[222] Florian Keusch, Alexander Wenz, Frederick Conrad. “Do you have your smartphone with you? Behavioral barriers for measuring everyday activities with smartphone sensors.” In: *Comput. Hum. Behav.* 127.C (2022). DOI: 10.1016/j.chb.2021.107054.

[223] Wazir Zada Khan, Yang Xiang, Mohammed Y Aalsalem, Quratulain Arshad. “Mobile phone sensing systems: A survey.” In: *IEEE Communications Surveys & Tutorials* 15.1 (2012), pp. 402–427.

[224] Aparup Khatua, Apalak Khatua, Erik Cambria. “Predicting political sentiments of voters from Twitter in multi-party contexts.” In: *Applied Soft Computing* 97 (2020), p. 106743. DOI: 10.1016/j.asoc.2020.106743.

[225] Hee-Woong Kim, Hock Chuan Chan, Sumeet Gupta. “Value-based adoption of mobile internet: an empirical investigation.” In: *Decision support systems* 43.1 (2007), pp. 111–126. DOI: 10.1016/j.dss.2005.05.009.

[226] Michael M Klammer. “Videogames for Future: How digital games can sensitize for climate change and promote sustainability.” PhD thesis. Master thesis. Fachhochschule Oberösterreich., 2020.

[227] Predrag Klasnja, Sunny Consolvo, Tanzeem Choudhury, Richard Beckwith, Jeffrey Hightower. “Exploring privacy concerns about personal sensing.” In: *International Conference on Pervasive Computing*. Springer, 2009, pp. 176–183. DOI: 10.1007/978-3-642-01516-8_13.

[228] Timo K. Koch, Peter Romero, Clemens Stachl. “Age and gender in language, emoji, and emoticon usage in instant messages.” In: *Computers in Human Behavior* (2021), p. 106990. DOI: 10.1016/j.chb.2021.106990.

[229] Timo Koch, Johannes Eichstädt, Clemens Stachl. “Affect Experience in Everyday Language Logged with Smartphones.” In: (2022). DOI: 10.23668/psycharchives.5399.

[230] *Korpusbasierte Wortgrundformenliste DEREWO, v-ww-bll-320000g-2012-12-31-1.0, mit Benutzerdokumentation*. Germany, 2013.

[231] Marios Koufaris, William Hampton-Sosa. "The development of initial trust in an online company by new customers." In: *Information & management* 41.3 (2004), pp. 377–397.

[232] Michael D Krämer, Yannick Roos, Ramona Schoedel, Cornelia Wrzus, David Richter. "Social dynamics and affect: Investigating within-person associations in daily life using experience sampling and mobile sensing." In: *Emotion* (2023).

[233] Lydia Kraus, Ina Wechsung, Sebastian Möller. "A comparison of privacy and security knowledge and privacy concern as influencing factors for mobile protection behavior." In: *Workshop on Privacy Personas and Segmentation*. 2014, p. 2014.

[234] Frauke Kreuter, Georg-Christoph Haas, Florian Keusch, Sebastian Bähr, Mark Trappmann. "Collecting survey and smartphone sensor data with an app: Opportunities and challenges around privacy and informed consent." In: *Social Science Computer Review* 38.5 (2020), pp. 533–549. DOI: 10.1177/0894439318816389.

[235] Jacob Kröger. "Unexpected inferences from sensor data: a hidden privacy threat in the internet of things." In: *Internet of Things. Information Processing in an Increasingly Connected World: First IFIP International Cross-Domain Conference, IFIPIoT 2018, Held at the 24th IFIP World Computer Congress, WCC 2018, Poznan, Poland, September 18-19, 2018, Revised Selected Papers 1*. Springer. 2019, pp. 147–159. DOI: 10.1007/978-3-030-15651-0_13.

[236] Ioannis Krontiris, Tassos Dimitriou. "A platform for privacy protection of data requesters and data providers in mobile sensing." In: *Computer Communications* 65 (2015), pp. 43–54.

[237] Isadora Krsek, Kimi Wenzel, Sauvik Das, Jason I. Hong, Laura Dabbish. "To Self-Persuade or be Persuaded: Examining Interventions for Users' Privacy Setting Selection." In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. New Orleans, LA, USA: Association for Computing Machinery, 2022. DOI: 10.1145/3491102.3502009.

[238] Todd Kulesza, Margaret Burnett, Weng-Keen Wong, Simone Stumpf. "Principles of Explanatory Debugging to Personalize Interactive Machine Learning." en. In: *Proceedings of the 20th International Conference on Intelligent User Interfaces*. Atlanta Georgia USA: ACM, Mar. 2015, pp. 126–137. DOI: 10.1145/2678025.2701399.

[239] Florian Künzler. “Context-aware notification management systems for just-in-time adaptive interventions.” In: *2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE, 2019, pp. 435–436. DOI: 10.1109/PERCOMW.2019.8730874.

[240] Ivan Ka Wai Lai, Guicheng Shi. “The impact of privacy concerns on the intention for continued use of an integrated mobile instant messaging and social network platform.” In: *International Journal of Mobile Communications* 13.6 (2015), pp. 641–669. DOI: 10.1504/IJMC.2015.072086.

[241] J Richard Landis, Gary G Koch. “The measurement of observer agreement for categorical data.” In: *biometrics* (1977), pp. 159–174.

[242] Nicholas D Lane, Emiliano Miluzzo, Hong Lu, Daniel Peebles, Tanzeem Choudhury, Andrew T Campbell. “A survey of mobile phone sensing.” In: *IEEE Communications magazine* 48.9 (2010), pp. 140–150. DOI: 10.1109/MCOM.2010.5560598.

[243] Marc Langheinrich. “Privacy by design—principles of privacy-aware ubiquitous systems.” In: *International conference on ubiquitous computing*. Springer. 2001, pp. 273–291.

[244] Francisco Laport-López, Emilio Serrano, Javier Bajo, Andrew T Campbell. “A review of mobile sensing systems, applications, and opportunities.” In: *Knowledge and Information Systems* 62.1 (2020), pp. 145–174. DOI: 10.1007/s10115-019-01346-1.

[245] Reed Larson, Mihaly Csikszentmihalyi. “The experience sampling method.” In: *New directions for methodology of social & behavioral science* (1983).

[246] John Laugesen, Khaled Hassanein. “Adoption of Personal Health Records by Chronic Disease Patients.” In: *Comput. Hum. Behav.* 66.C (2017), 256–272. DOI: 10.1016/j.chb.2016.09.054.

[247] Tricia M. Leahey, Melissa M. Crane, Angela Marinilli Pinto, Brad Weinberg, Rajiv Kumar, Rena R. Wing. “Effect of teammates on changes in physical activity in a statewide campaign.” In: *Preventive Medicine* 51.1 (2010), pp. 45–49. DOI: 10.1016/j.ypmed.2010.04.004.

[248] Hansoo Lee, Joonyoung Park, Uichin Lee. “A systematic survey on android api usage for data-driven analytics with smartphones.” In: *ACM Computing Surveys* 55.5 (2022), pp. 1–38. DOI: 10.1145/3530814.

[249] Lorraine S Lee, William D Brink. “Trust in cloud-based services: A framework for consumer adoption of software as a service.” In: *Journal of Information Systems* 34.2 (2020), pp. 65–85.

[250] Seokjun Lee, Rhan Ha, Hojung Cha. “Click sequence prediction in Android mobile applications.” In: *IEEE Transactions on Human-Machine Systems* 49.3 (2018), pp. 278–289.

[251] Toby Jia-Jun Li, Lindsay Popowski, Tom Mitchell, Brad A Myers. “Screen2vec: Semantic embedding of gui screens and gui components.” In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021, pp. 1–15.

[252] Weijian Li, Yuxiao Chen, Tianran Hu, Jiebo Luo. “Mining the Relationship between Emoji Usage Patterns and Personality.” In: *International AAAI Conference on Web and Social Media*. Palo Alto, CA, USA: AAAI Publications, 2018.

[253] Ya-Cheng Li, Shin-Ming Cheng. “Privacy preserved mobile sensing using region-based group signature.” In: *IEEE Access* 6 (2018), pp. 61556–61568. DOI: 10.1109/ACCESS.2018.2868502.

[254] Yuanchun Li, Hao Wen, Weijun Wang, Xiangyu Li, Yizhen Yuan, Guohong Liu, Jiacheng Liu, Wenxing Xu, Xiang Wang, Yi Sun, et al. “Personal llm agents: Insights and survey about the capability, efficiency and security.” In: *arXiv preprint arXiv:2401.05459* (2024).

[255] Jialiu Lin, Shahriyar Amini, Jason I Hong, Norman Sadeh, Janne Lindqvist, Joy Zhang. “Expectation and purpose: understanding users’ mental models of mobile app privacy through crowdsourcing.” In: *Proceedings of the 2012 ACM conference on ubiquitous computing*. 2012, pp. 501–510. DOI: 10.1145/2370216.2370290.

[256] Jialiu Lin, Bin Liu, Norman Sadeh, Jason I Hong. “Modeling Users’ Mobile App Privacy Preferences: Restoring Usability in a Sea of Permission Settings.” In: *10th Symposium On Usable Privacy and Security (SOUPS 2014)*. USENIX Association, 2014, pp. 199–212. DOI: 10.5555/3235838.3235856.

[257] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhimedi, Shikun Aerin Zhang, Norman Sadeh, Yuvraj Agarwal, Alessandro Acquisti. “Follow my recommendations: A personalized privacy assistant for mobile app permissions.” In: *Twelfth symposium on usable privacy and security (SOUPS 2016)*. 2016, pp. 27–41.

[258] Changchang Liu, Supriyo Chakraborty, Prateek Mittal. *DEEProtect: Enabling Inference-based Access Control on Mobile Sensing Applications*. 2017. DOI: 10.48550/arXiv.1702.06159.

[259] Andrew Lowe, Anthony C Norris, A Jane Farris, Duncan R Babbage. "Quantifying thematic saturation in qualitative data analysis." In: *Field methods* 30.3 (2018), pp. 191–207.

[260] Hong Lu, Denise Frauendorfer, Mashfiqui Rabbi, Marianne Schmid Mast, Gokul T Chittaranjan, Andrew T Campbell, Daniel Gatica-Perez, Tanzeem Choudhury. "Stresssense: Detecting stress in unconstrained acoustic environments using smartphones." In: *Proceedings of the 2012 ACM conference on ubiquitous computing*. 2012, pp. 351–360.

[261] Hong Lu, Wei Pan, Nicholas D Lane, Tanzeem Choudhury, Andrew T Campbell. "Soundsense: scalable sound sensing for people-centric applications on mobile phones." In: *Proceedings of the 7th international conference on Mobile systems, applications, and services*. 2009, pp. 165–178.

[262] Adrienn Lukács. "What is privacy? The history and definition of privacy." In: (2016).

[263] Kai Lukoff, Ulrik Lyngs, Himanshu Zade, J. Vera Liao, James Choi, Kaiyue Fan, Sean A. Munson, Alexis Hiniker. "How the Design of YouTube Influences User Sense of Agency." In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. Yokohama, Japan: Association for Computing Machinery, 2021. DOI: 10.1145/3411764.3445467.

[264] Kai Lukoff, Cissy Yu, Julie Kientz, Alexis Hiniker. "What Makes Smartphone Use Meaningful or Meaningless?" In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2.1 (Mar. 2018), 22:1–22:26. DOI: 10.1145/3191754.

[265] Scott M Lundberg, Su-In Lee. "A Unified Approach to Interpreting Model Predictions." In: *Advances in Neural Information Processing Systems*. Ed. by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett. Curran Associates, Inc., 2017, pp. 4765–4774.

[266] Elsa Macias, Alvaro Suarez, Jaime Lloret. "Mobile sensing systems." In: *Sensors* 13.12 (2013), pp. 17292–17321. DOI: 10.3390/s131217292.

[267] James E Maddux, Ronald W Rogers. “Protection motivation and self-efficacy: A revised theory of fear appeals and attitude change.” In: *Journal of experimental social psychology* 19.5 (1983), pp. 469–479.

[268] Naresh K Malhotra, Sung S Kim, James Agarwal. “Internet users’ information privacy concerns (IUIPC): The construct, the scale, and a causal model.” In: *Information systems research* 15.4 (2004), pp. 336–355. DOI: 10.1287/isre.1040.0032.

[269] Anshul Malik, S Suresh, Swati Sharma. “Factors influencing consumers’ attitude towards adoption and continuous use of mobile applications: a conceptual model.” In: *Procedia computer science* 122 (2017), pp. 106–113. DOI: 10.1016/j.procs.2017.11.348.

[270] Vikas Kumar Malviya, Chee Wei Leow, Ashok Kasthuri, Yan Naing Tun, Lwin Khin Shar, Lingxiao Jiang. “Right to Know, Right to Refuse: Towards UI Perception-Based Automated Fine-Grained Permission Controls for Android Apps.” In: *Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering*. ASE ’22. Rochester, MI, USA: Association for Computing Machinery, 2023. DOI: 10.1145/3551349.3559556.

[271] Marie-Helen Maras. “Internet of Things: security and privacy implications.” In: *Int’l Data Priv. L.* 5 (2015), p. 99.

[272] Justin Matejka, Michael Glueck, Tovi Grossman, George Fitzmaurice. “The effect of visual appearance on the performance of continuous sliders and visual analogue scales.” In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2016, pp. 5421–5432. DOI: 10.1145/2858036.2858063.

[273] Anijo Punnen Mathew. “Using the environment as an interactive interface to motivate positive behavior change in a subway station.” In: *CHI’05 Extended Abstracts on Human Factors in Computing Systems*. 2005, pp. 1637–1640. DOI: 10.1145/1056808.1056985.

[274] Sandra C Matz, Ruth E Appel, Michal Kosinski. “Privacy in the age of psychological targeting.” In: *Current opinion in psychology* 31 (2020), pp. 116–121.

[275] Sandra C Matz, Michal Kosinski, Gideon Nave, David J Stillwell. “Psychological targeting as an effective approach to digital mass persuasion.” In: *Proceedings of the national academy of sciences* 114.48 (2017), pp. 12714–12719. DOI: 10.1073/pnas.1710966114.

[276] Aikaterini-Georgia Mavroeidi, Angeliki Kitsiou, Christos Kalloniatis. “The Role of Gamification in Privacy Protection and User Engagement.” In: *Security and Privacy From a Legal, Ethical, and Technical Perspective*. Ed. by Christos Kalloniatis, Carlos Travieso-Gonzalez. Rijeka: IntechOpen, 2020. Chap. 5. DOI: 10.5772/intechopen.91159.

[277] Aleecia M McDonald, Lorrie Faith Cranor. “The cost of reading privacy policies.” In: *I/S: A Journal of Law and Policy for the Information Society* 4 (2008), pp. 543–568.

[278] William J McGuire. “Inducing resistance to persuasion. Some contemporary approaches.” In: 1 (1964), pp. 191–229.

[279] D Harrison McKnight, Vivek Choudhury, Charles Kacmar. “The impact of initial consumer trust on intentions to transact with a web site: a trust building model.” In: *The journal of strategic information systems* 11.3-4 (2002), pp. 297–323.

[280] Lakmal Meegahapola, Daniel Gatica-Perez. “Smartphone sensing for the well-being of young adults: A review.” In: *IEEE Access* 9 (2020), pp. 3374–3399.

[281] Ricardo Mendes, André Brandão, João P. Vilela, Alastair R. Beresford. “Effect of User Expectation on Mobile App Privacy: A Field Study.” In: *2022 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. 2022, pp. 207–214. DOI: 10.1109/PerCom53586.2022.9762379.

[282] Lemay Michael. *Understanding the Mechanism of Panel Attrition*. 2009.

[283] Kristopher Micinski, Daniel Votipka, Rock Stevens, Nikolaos Kofinas, Michelle L. Mazurek, Jeffrey S. Foster. “User Interactions and Permission Use on Android.” In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI ’17. Denver, Colorado, USA: Association for Computing Machinery, 2017, 362–373. DOI: 10.1145/3025453.3025706.

[284] Hannah Jean Miller, Daniel Kluver, Jacob Thebault-Spieker, Loren G Terveen, Brent J Hecht. “Understanding Emoji Ambiguity in Context: The Role of Text in Emoji-Related Miscommunication.” In: *International AAAI Conference on Web and Social Media*. Palo Alto, CA, USA: AAAI Publications, 2017, pp. 152–161.

[285] David C Mohr, Mi Zhang, Stephen M Schueller. “Personal sensing: understanding mental health using ubiquitous sensors and machine learning.” In: *Annual review of clinical psychology* 13 (2017), pp. 23–47.

[286] Ivan Montiel, Javier Delgado-Ceballos, Natalia Ortiz-de Mandojana. *Mobile Apps for Sustainability Management Education: The Example of GoodGuide,(http://www. goodguide. com/about/mobile)*. 2017.

[287] Matteo Moschelli. "Using Feedback to Promote Meaningful App-Switching Suggestions." PhD thesis. Politecnico di Torino, 2022.

[288] Elizabeth L Murnane, Xin Jiang, Anna Kong, Michelle Park, Weili Shi, Connor Soohoo, Luke Vink, Iris Xia, Xin Yu, John Yang-Sammataro, et al. "Designing ambient narrative-based interfaces to reflect and motivate physical activity." In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020, pp. 1–14. DOI: 10.1145/3313831.3376478.

[289] Mohammad Nauman, Sohail Khan, Xinwen Zhang. "Apex: extending android permission model and enforcement with user-defined runtime constraints." In: *Proceedings of the 5th ACM symposium on information, computer and communications security*. 2010, pp. 328–332.

[290] Elisabet M Nilsson, Anders Jakobsson. "Simulated sustainable societies: Students' reflections on creating future cities in computer games." In: *Journal of Science Education and Technology* 20 (2011), pp. 33–50.

[291] Helen Nissenbaum. "Privacy as contextual integrity." In: *Wash. L. Rev.* 79 (2004), p. 119.

[292] Patricia A Norberg, Daniel R Horne, David A Horne. "The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors." In: *Journal of Consumer Affairs* 41.1 (2007), pp. 100–126. DOI: 10.1111/j.1745-6606.2006.00070.x.

[293] Midas Nouwens, Ilaria Liccardi, Michael Veale, David Karger, Lalana Kagal. "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, pp. 1–13. DOI: 10.1145/3313831.3376321.

[294] Jonathan A Obar, Anne Oeldorf-Hirsch. "The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services." In: *Information, Communication & Society* 23.1 (2020), pp. 128–147. DOI: 10.1080/1369118X.2018.1486870.

[295] Obi Ogbanufe, Robert Pavur. “Going through the emotions of regret and fear: Revisiting protection motivation for identity theft protection.” In: *International Journal of Information Management* 62 (2022), p. 102432. DOI: 10.1016/j.ijinfomgt.2021.102432.

[296] Derek H. Ogle, Jason C. Doll, Powell Wheeler, Alexis Dinno. *FSA: Fisheries Stock Analysis*. R package version 0.9.3. 2022.

[297] Harri Oinas-Kukkonen, Marja Harjumaa. “Persuasive systems design: Key issues, process model, and system features.” In: *Communications of the association for Information Systems* 24.1 (2009), p. 28. DOI: 10.17705/1CAIS.02428.

[298] Ellinor K Olander, Helen Fletcher, Stefanie Williams, Lou Atkinson, Andrew Turner, David P French. “What are the most effective techniques in changing obese individuals’ physical activity self-efficacy and behaviour: a systematic review and meta-analysis.” In: *International Journal of Behavioral Nutrition and Physical Activity* 10 (2013), pp. 1–15.

[299] Katarzyna Olejnik, Italo Dacosta, Joana Soares Machado, Kévin Huguenin, Mohammad Emtiyaz Khan, Jean-Pierre Hubaux. “SmarPer: Context-Aware and Automatic Runtime-Permissions for Mobile Devices.” In: *2017 IEEE Symposium on Security and Privacy (SP)*. 2017, pp. 1058–1076. DOI: 10.1109/SP.2017.25.

[300] Independent High-Level Expert Group on Artificial Intelligence Set Up by the European Commission. *Ethics Guidelines for Trustworthy AI*. European Union, 2019.

[301] Emmanuel Onu, Michael Mireku Kwakye, Ken Barker. “Contextual privacy policy modeling in iot.” In: *2020 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*. IEEE. 2020, pp. 94–102.

[302] Anna-Marie Ortloff, Maximiliane Windl, Valentin Schwind, Niels Henze. “Implementation and In Situ Assessment of Contextual Privacy Policies.” In: *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. DIS ’20. Eindhoven, Netherlands: Association for Computing Machinery, 2020, 1765–1778. DOI: 10.1145/3357236.3395549.

[303] Gene Ouellette, Melissa Michaud. “Generation text: Relations among undergraduates’ use of text messaging, textese, and language and literacy skills.” In: *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement* 48.3 (2016), pp. 217–221. DOI: 10.1037/cbs0000046.

[304] Cliodhna O'Connor, Helene Joffe. "Intercoder reliability in qualitative research: debates and practical guidelines." In: *International journal of qualitative methods* 19 (2020).

[305] John O'donoghue, John Herbert. "Data management within mHealth environments: Patient sensors, mobile devices, and databases." In: *Journal of Data and Information Quality (JDIQ)* 4.1 (2012), pp. 1–20.

[306] Sameer Patil, Roberto Hoyle, Roman Schlegel, Apu Kapadia, Adam J. Lee. "Interrupt Now or Inform Later? Comparing Immediate and Delayed Privacy Feedback." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. Seoul, Republic of Korea: Association for Computing Machinery, 2015, 1415–1418. DOI: 10.1145/2702123.2702165.

[307] Trista Patterson, Sam Barratt. "Playing for the planet: How video games can deliver for people and the environment." In: (2019).

[308] Paul Pearce, Adrienne Porter Felt, Gabriel Nunez, David Wagner. "Addroid: Privilege separation for applications and advertisers in android." In: *Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security*. 2012, pp. 71–72.

[309] James W Pennebaker. "Mind mapping: Using everyday language to explore social & psychological processes." In: *Procedia computer science* 118 (2017), pp. 100–107.

[310] James W. Pennebaker, Ryan L. Boyd, Kayla Jordan, Kate Blackburn. "The Development and Psychometric Properties of LIWC2015." In: (2015).

[311] James W Pennebaker, Martha E. Francis, Roger J. Booth. "Linguistic inquiry and word count: LIWC 2001." In: *Mahway: Lawrence Erlbaum Associates* 71.2001 (2001), p. 2001.

[312] Jan Pennekamp, Martin Henze, Klaus Wehrle. "A survey on the evolution of privacy enforcement on smartphones and the road ahead." In: *Pervasive and Mobile Computing* 42 (2017), pp. 58–76.

[313] Jeffrey Pennington, Richard Socher, Christopher Manning. "Glove: Global Vectors for Word Representation." en. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, 2014, pp. 1532–1543. DOI: 10.3115/v1/D14-1162.

[314] Iryna Pentina, Lixuan Zhang, Hatem Bata, Ying Chen. “Exploring privacy paradox in information-sensitive mobile app adoption: A cross-cultural comparison.” In: *Computers in Human Behavior* 65 (2016), pp. 409–419. DOI: 10.1016/j.chb.2016.09.005.

[315] Nataniel Pereira Borges Junior. “Learning the language of apps.” In: (2020).

[316] Andreas Pfitzmann, Marit Hansen. *A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management*. 2010.

[317] Paul R. Pintrich. “A Motivational Science Perspective on the Role of Student Motivation in Learning and Teaching Contexts.” In: *Journal of Educational Psychology* 95.4 (2003), pp. 667–686. DOI: 10.1037/0022-0663.95.4.667.

[318] Johanna Pirker, Christian Gütl. “Educational gamified science simulations.” In: *Gamification in education and business* (2015), pp. 253–275. DOI: 10.1007/978-3-319-10208-5_13.

[319] Isabella Poggi, Catherine Pelachaud, Fiorella de Rosis, Valeria Carofiglio, Berardina De Carolis. “Greta. A Believable Embodied Conversational Agent.” In: *Multimodal Intelligent Information Presentation*. Springer, 2005. DOI: 10.1007/1-4020-3051-7_1.

[320] Henning Pohl, Christian Domin, Michael Rohs. “Beyond Just Text: Semantic Emoji Similarity Modeling to Support Expressive Communication.” In: *ACM Trans. Comput.-Hum. Interact.* 24.1 (Mar. 2017), 6:1–6:42. DOI: 10.1145/3039685.

[321] Henning Pohl, Dennis Stanke, Michael Rohs. “EmojiZoom: Emoji Entry via Large Overview Maps.” In: *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI ’16. Florence, Italy: ACM, 2016, pp. 510–517. DOI: 10.1145/2935334.2935382.

[322] Barbara Prainsack. “Data donation: How to resist the iLeviathan.” In: *The ethics of medical data donation* (2019), pp. 9–22.

[323] Sarah Prange, Sven Mayer, Maria-Lena Bittl, Mariam Hassib, Florian Alt. “Investigating User Perceptions Towards Wearable Mobile Electromyography.” In: *Human-Computer Interaction – INTERACT 2021*. Cham: Springer International Publishing, 2021, pp. 339–360. DOI: 10.1007/978-3-030-85610-6_20.

[324] Ismini Psychoula, Deepika Singh, Liming Chen, Feng Chen, Andreas Holzinger, Huansheng Ning. "Users' privacy concerns in IoT based applications." In: *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation*. IEEE, 2018, pp. 1887–1894. DOI: 10.1109/SmartWorld.2018.00317.

[325] Yiting Qu, Suguo Du, Shaofeng Li, Yan Meng, Le Zhang, Haojin Zhu. "Automatic permission optimization framework for privacy enhancement of mobile applications." In: *IEEE Internet of Things Journal* 8.9 (2020), pp. 7394–7406. DOI: 10.1109/JIOT.2020.3039472.

[326] Taoufik Rachad, Ali Idri. "Intelligent mobile applications: A systematic mapping study." In: *Mobile information systems* 2020 (2020), pp. 1–17.

[327] Emilee Rader. "Normative and Non-Social Beliefs about Sensor Data: Implications for Collective Privacy Management." In: *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*. 2022, pp. 653–670.

[328] Mika Raento, Antti Oulasvirta, Nathan Eagle. "Smartphones: An emerging tool for social scientists." In: *Sociological methods & research* 37.3 (2009), pp. 426–454.

[329] Philip Raschke, Axel Küpper, Olha Drozd, Sabrina Kirrane. "Designing a GDPR-compliant and usable privacy dashboard." In: *IFIP international summer school on privacy and identity management*. Springer, 2017, pp. 221–236. DOI: 10.1007/978-3-319-92925-5_14.

[330] Gerhard Reese. "Common human identity and the path to global climate justice." In: *Climatic Change* 134 (2016), pp. 521–531. DOI: 10.1007/s10584-015-1548-2.

[331] Gerhard Reese, Karen RS Hamann, Claudia Menzel, Stefan Drews. "Soziale Identität und nachhaltiges Verhalten." In: *Psychologie und Nachhaltigkeit: Konzeptionelle Grundlagen, Anwendungsbeispiele und Zukunftsperspektiven* (2018), pp. 47–54.

[332] Byron Reeves, Nilam Ram, Thomas N Robinson, James J Cummings, C Lee Giles, Jennifer Pan, Agnese Chiatti, MJ Cho, Katie Roehrick, Xiao Yang, et al. "Screenomics: A framework to capture and analyze personal life experiences and the ways that technology shapes them." In: *Human–Computer Interaction* 36.2 (2021), pp. 150–201.

[333] Nils Reimers, Iryna Gurevych. “Sentence-bert: Sentence embeddings using siamese bert-networks.” In: *arXiv preprint arXiv:1908.10084* (2019).

[334] Andreas Reinhardt, Frank Englert, Delphine Christin. “Averting the privacy risks of smart metering by local data preprocessing.” In: *Pervasive and Mobile Computing* 16 (2015), pp. 171–183.

[335] Ulf-Dietrich Reips, Frederik Funke. “Interval-level measurement with visual analogue scales in Internet-based research: VAS Generator.” In: *Behavior research methods* 40.3 (2008), pp. 699–704. DOI: 10.3758/BRM.40.3.699.

[336] Thomas Reiter, Ramona Schoedel. “Never miss a beep: Using mobile sensing to investigate (non-) compliance in experience sampling studies.” In: *Behavior Research Methods* (2023), pp. 1–23.

[337] Robert Remus, Uwe Quasthoff, Gerhard Heyer. “Sentiws-a publicly available german-language resource for sentiment analysis.” In: *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC’10)*. 2010.

[338] Melanie Revilla, Mick P Couper, Carlos Ochoa. “Willingness of online panelists to perform additional tasks.” In: *Methods, data, analyses: a journal for quantitative methods and survey methodology (mda)* 13.2 (2019), pp. 223–252. DOI: 10.12758/mda.2018.01.

[339] Christian P Robert, Nicolas Chopin, Judith Rousseau. “Harold Jeffreys’s theory of probability revisited.” In: *Statistical Science* 24.2 (2009), pp. 141–172. DOI: 10.1214/09-STS284.

[340] Alberto Roffarello, Luigi De Russis. “The Race Towards Digital Wellbeing: Issues and Opportunities.” In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Glasgow, Scotland Uk: ACM, Apr. 2019, pp. 1–14. DOI: 10.1145/3290605.3300616.

[341] Yannick Roos, Michael D Krämer, David Richter, Ramona Schoedel, Cornelia Wrzus. “Does your smartphone “know” your social life? A methodological comparison of day reconstruction, experience sampling, and mobile sensing.” In: *Advances in Methods and Practices in Psychological Science* 6.3 (2023), p. 25152459231178738.

[342] Avi Rosenfeld, Sigal Sina, David Sarne, Or Avidov, Sarit Kraus. “A Study of WhatsApp Usage Patterns and Prediction Models without Message Content.” In: *CoRR* abs/1802.03393 (2018).

[343] Anna Małgorzata Rudnicka. “Disclosure of personal data in citizen science settings.” PhD thesis. UCL (University College London), 2020.

[344] Alireza Sahami Shirazi, Niels Henze, Tilman Dingler, Martin Pielot, Dominik Weber, Albrecht Schmidt. “Large-scale Assessment of Mobile Notifications.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’14. Toronto, Ontario, Canada: ACM, 2014, pp. 3055–3064. DOI: 10.1145/2556288.2557189.

[345] Nazir Saleheen, Supriyo Chakraborty, Nasir Ali, Md Mahbubur Rahman, Syed Monowar Hossain, Rummana Bari, Eugene Buder, Mani Srivastava, Santosh Kumar. “mSieve: differential behavioral privacy in time series of mobile sensor data.” In: *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 2016, pp. 706–717. DOI: 10.1145/2971648.2971753.

[346] Kanthashree Mysore Sathyendra, Shomir Wilson, Florian Schaub, Sebastian Zimmeck, Norman Sadeh. “Identifying the provision of choices in privacy policy text.” In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. 2017, pp. 2774–2779. DOI: 10.18653/v1/D17-1294.

[347] Florian Schaub, Rebecca Balebako, Adam L Durity, Lorrie Faith Cranor. “A design space for effective privacy notices.” In: *Eleventh Symposium On Usable Privacy and Security ({SOUPS} 2015)*. 2015, pp. 1–17. DOI: 10.5555/3235866.3235868.

[348] Maxim Schessler, Eva Gerlitz, Maximilian Häring, Matthew Smith. “Replication: Measuring User Perceptions in Smartphone Security and Privacy in Germany.” In: *European Symposium on Usable Security 2021*. 2021, pp. 165–179. DOI: 10.1145/3481357.3481511.

[349] H. Schmid. “Improvements in Part-of-Speech Tagging with an Application to German.” In: *Natural Language Processing Using Very Large Corpora*. Ed. by Susan Armstrong, Kenneth Church, Pierre Isabelle, Sandra Manzi, Evelyne Tzoukermann, David Yarowsky. Dordrecht: Springer Netherlands, 1999, pp. 13–25. DOI: 10.1007/978-94-017-2390-9_2.

[350] Albrecht Schmidt. “Interactive context-aware systems interacting with ambient intelligence.” In: *Ambient intelligence* 159 (2005).

[351] Christina Schneegass, Diana Irmscher, Florian Bemmann, Daniel Buschek. “LYLO–Exploring Disclosed Configurations for Inter-Personal Location Sharing.” In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021, pp. 1–6. DOI: 10.1145/3411763.3451652.

[352] Stefan Schneegass, Romina Poguntke, Tonja Machulla. “Understanding the Impact of Information Representation on Willingness to Share Information.” In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland Uk: Association for Computing Machinery, 2019, 1–6. DOI: 10.1145/3290605.3300753.

[353] Sebastian Schnorf, Martin Ortlieb, Nikhil Sharma. “Trust, transparency & control in inferred user interest models.” In: *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. 2014, pp. 2449–2454. DOI: 10.1145/2559206.2581141.

[354] Patrick Schober, Christa Boer, Lothar A Schwarte. “Correlation coefficients: appropriate use and interpretation.” In: *Anesthesia & analgesia* 126.5 (2018), pp. 1763–1768. DOI: 10.1213/ANE.0000000000002864.

[355] Ramona Schoedel, Fiona Kunz, Maximilian Bergmann, Florian Bemmann, Markus Bühner, Larissa Sust. “Snapshots of daily life: Situations investigated through the lens of smartphone sensing.” In: *Journal of Personality and Social Psychology* (2023). DOI: 10.1037/pspp0000469.

[356] Ramona Schoedel, Michelle Oldemeier. “Basic Protocol: Smartphone Sensing Panel Study.” In: (2020). DOI: 10.23668/PSYCHARCHIVES.2901.

[357] Ramona Schoedel, Michelle Oldemeier, Léonie Bonauer, Larissa Sust. *Dataset for: Systematic categorisation of 3091 smartphone applications from a large-scale smartphone sensing dataset*. 2022. DOI: 10.23668/psycharchives.5362.

[358] Ramona Schoedel, Florian Pargent, Quay Au, Sarah Theres Völkel, Tobias Schuwerk, Markus Bühner, Clemens Stachl. “To challenge the morning lark and the night owl: Using smartphone sensing data to investigate day–night behaviour patterns.” In: *European Journal of Personality* 34.5 (2020), pp. 733–752.

[359] Ramona Schoedel, Larissa Sust, Timo Koch, Florian Bemmann, Bernd Bischl, Heinrich Hußmann, Markus Bühner, Clemens Stachl, Holger Steinmetz, Stefanie Müller, et al. “Preregistration: Smartphone Sensing Panel Study.” In: (2020).

[360] Eva-Maria Schomakers, Chantal Lidynia, Martina Ziefle. “A typology of online privacy personalities: Exploring and segmenting users’ diverse privacy attitudes and behaviors.” In: *Journal of Grid Computing* 17.4 (2019), pp. 727–747. DOI: 10.1007/s10723-019-09500-3.

[361] Martin Schrepp, Andreas Hinderks, Jörg Thomaschewski. “Applying the user experience questionnaire (UEQ) in different evaluation scenarios.” In: *International Conference of Design, User Experience, and Usability*. Springer. 2014, pp. 383–392. DOI: 10.1007/978-3-319-07668-3_37.

[362] H. Andrew Schwartz, Johannes C. Eichstaedt, Margaret L. Kern, Lukasz Dziurzynski, Stephanie M. Ramones, Megha Agrawal, Achal Shah, Michal Kosinski, David Stillwell, Martin E. P. Seligman, Lyle H. Ungar. “Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach.” In: *PLOS ONE* 8.9 (Sept. 2013), pp. 1–16. DOI: 10.1371/journal.pone.0073791.

[363] Gian Luca Scoccia, Ivano Malavolta, Marco Autili, Amleto Di Salle, Paola Inverardi. “Enhancing trustability of android applications via user-centric flexible permissions.” In: *IEEE Transactions on Software Engineering* 47.10 (2019), pp. 2032–2051. DOI: 10.1109/TSE.2019.2941936.

[364] John S Seberger, Marissel Llavore, Nicholas Nye Wyant, Irina Shklovski, Sameer Patil. “Empowering resignation: There’s an app for that.” In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021, pp. 1–18. DOI: 10.1145/3411764.3445293.

[365] Sandra Servia-Rodríguez, Kiran K Rachuri, Cecilia Mascolo, Peter J Rentfrow, Neal Lathia, Gillian M Sandstrom. “Mobile sensing at the service of mental well-being: a large-scale longitudinal study.” In: *Proceedings of the 26th international conference on world wide web*. 2017, pp. 103–112. DOI: 10.1145/3038912.3052618.

[366] Heather Shaw, David A Ellis, Fenja V Ziegler. “The Technology Integration Model (TIM). Predicting the continued use of technology.” In: *Computers in Human Behavior* 83 (2018), pp. 204–214. DOI: 10.1016/j.chb.2018.02.001.

[367] Kim Bartel Sheehan. “Toward a typology of Internet users and online privacy concerns.” In: *The information society* 18.1 (2002), pp. 21–32. DOI: 10.1080/01972240252818207.

[368] Bingyu Shen, Lili Wei, Chengcheng Xiang, Yudong Wu, Mingyao Shen, Yuanyuan Zhou, Xinxin Jin. “Can Systems Explain Permissions Better? Understanding Users’ Misperceptions under Smartphone Runtime Permission Model.” In: *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, Aug. 2021, pp. 751–768.

[369] Choonsung Shin, Jin-Hyuk Hong, Anind K. Dey. “Understanding and Prediction of Mobile Application Usage for Smart Phones.” In: *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. UbiComp ’12. Pittsburgh, Pennsylvania: Association for Computing Machinery, 2012, 173–182. DOI: 10.1145/2370216.2370243.

[370] Gulshan Shrivastava, Prabhat Kumar, Deepak Gupta, Joel JPC Rodrigues. “Privacy issues of android application permissions: A literature review.” In: *Transactions on Emerging Telecommunications Technologies* 31.12 (2020), e3773. DOI: 10.1002/ett.3773.

[371] Pekka Siirtola, Juha Röning. “User-independent human activity recognition using a mobile phone: Offline recognition vs. real-time on device recognition.” In: *Distributed Computing and Artificial Intelligence: 9th International Conference*. Springer. 2012, pp. 617–627. DOI: 10.1007/978-3-642-28765-7_75.

[372] Johanneke Siljee. “Privacy Transparency Patterns.” In: *Proceedings of the 20th European Conference on Pattern Languages of Programs*. EuroPLoP ’15. Kaufbeuren, Germany: Association for Computing Machinery, 2015. DOI: 10.1145/2855321.2855374.

[373] Janice C Sipior, Burke T Ward, Linda Volonino. “Privacy concerns associated with smartphone use.” In: *Journal of Internet Commerce* 13.3-4 (2014), pp. 177–193. DOI: 10.1080/15332861.2014.947902.

[374] Anya Skatova, James Goulding. “Psychology of personal data donation.” In: *PloS one* 14.11 (2019), e0224240. DOI: 10.1371/journal.pone.0224240.

[375] Edith G Smit, Guda Van Noort, Hilde AM Voorveld. “Understanding online behavioural advertising: User knowledge, privacy concerns and online coping behaviour in Europe.” In: *Computers in human behavior* 32 (2014), pp. 15–22. DOI: 10.1016/j.chb.2013.11.008.

[376] H. Jeff Smith, Tamara Dinev, Heng Xu. “Information Privacy Research: An Interdisciplinary Review.” In: *MIS Quarterly* 35.4 (2011), pp. 989–1015.

[377] H Jeff Smith, Sandra J Milberg, Sandra J Burke. “Information privacy: Measuring individuals’ concerns about organizational practices.” In: *MIS quarterly* (1996), pp. 167–196. DOI: 10.2307/249477.

[378] Noah A. Smith. “Contextual Word Representations: A Contextual Introduction.” In: *arXiv:1902.06006 [cs]* (Feb. 2019). arXiv: 1902.06006.

[379] Daniel J Solove. “A taxonomy of privacy.” In: *University of Pennsylvania law review* (2006), pp. 477–564. DOI: 10.2307/40041279.

[380] Zhiyi Song, Stephanie Strassel, Haejoong Lee, Kevin Walker, Jonathan Wright, Jennifer Garland, Dana Fore, Brian Gainor, Preston Cabe, Thomas Thomas, Brendan Callahan, Ann Sawyer. “Collecting Natural SMS and Chat Conversations in Multiple Languages: The BOLT Phase 2 Corpus.” In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*. Reykjavik, Iceland: European Languages Resources Association (ELRA), May 2014, pp. 1699–1704.

[381] Dimitris Spathis, Sandra Servia-Rodriguez, Katayoun Farrahi, Cecilia Mascolo, Jason Rentfrow. “Passive mobile sensing and psychological traits for large scale mood prediction.” In: *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 2019, pp. 272–281. DOI: 10.1145/3329189.3329213.

[382] Christoph Stach, Bernhard Mitschang. “Privacy management for mobile platforms–a review of concepts and approaches.” In: *2013 IEEE 14th International Conference on Mobile Data Management*. Vol. 1. IEEE. 2013, pp. 305–313. DOI: 10.1109/MDM.2013.45.

[383] Clemens Stachl, Quay Au, Ramona Schoedel, Daniel Buschek, Sarah Völkel, Tobias Schuwerk, Michelle Oldemeier, Theresa Ullmann, Heinrich Hussmann, Bernd Bischl, et al. “Behavioral patterns in smartphone usage predict big five personality traits.” In: (2019). DOI: 10.1073/pnas.1920484117.

[384] Clemens Stachl, Quay Au, Ramona Schoedel, Samuel D Gosling, Gabriella M Harari, Daniel Buschek, Sarah Theres Völkel, Tobias Schuwerk, Michelle Oldemeier, Theresa Ullmann, et al. “Predicting personality from patterns of behavior collected with smartphones.” In: *Proceedings of the National Academy of Sciences* 117.30 (2020), pp. 17680–17687.

[385] Emily Stark, Ryan Sleevi, Rijad Muminovic, Devon O'Brien, Eran Messeri, Adrienne Porter Felt, Brendan McMillion, Parisa Tabriz. "Does Certificate Transparency Break the Web? Measuring Adoption and Error Rate." In: *2019 IEEE Symposium on Security and Privacy (SP)*. 2019, pp. 211–226. DOI: 10.1109/SP.2019.00027.

[386] Katarzyna Stawarz, Anna L Cox, Ann Blandford. "Beyond self-tracking and reminders: designing smartphone apps that support habit formation." In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 2015, pp. 2653–2662. DOI: 10.1145/2702123.2702230.

[387] Kathy A Stewart, Albert H Segars. "An empirical examination of the concern for information privacy instrument." In: *Information systems research* 13.1 (2002), pp. 36–49. DOI: 10.1287/isre.13.1.36.97.

[388] Eric Struse, Julian Seifert, Sebastian Üllenbeck, Enrico Rukzio, Christopher Wolf. "PermissionWatcher: Creating user awareness of application permissions in mobile systems." In: *International Joint Conference on Ambient Intelligence*. Springer. 2012, pp. 65–80. DOI: 10.1007/978-3-642-34898-3_5.

[389] Andrew Sullivan. Website. Retrieved January 24, 2023 from https://nymag.com/intelligencer/2016/09/andrew-sullivan-my-distraction-sickness-and-yours.html. Sept. 2016.

[390] Mohammad Tahaei, Ruba Abu-Salma, Awais Rashid. "Stuck in the Permissions With You: Developer & End-User Perspectives on App Permissions & Their Privacy Ramifications." In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Hamburg, Germany: Association for Computing Machinery, 2023. DOI: 10.1145/3544548.3581060.

[391] Nađa Terzimehić, Mohamed Khamis, Florian Bemmann, Heinrich Hussmann. "Lunchocracy: Improving Eating Dynamics in the Workplace Using a Bot-Based Anonymous Voting System." In: *Proceedings of the 36th Annual ACM Conference on Human Factors in Computing Systems*. CHI EA '18. MontrÃ©al, QuÃ©bec, Canada: ACM, 2018. DOI: 10.1145/3170427.3188626.

[392] Nađa Terzimehić, Svenja Yvonne Schött, Florian Bemmann, Daniel Buschek. "MEMEories: internet memes as means for daily journaling." In: *Proceedings of the 2021 ACM Designing Interactive Systems Conference*. 2021, pp. 538–548. DOI: 10.1145/3461778.3462080.

[393] Nada Terzimehic, Florian Bemmann, Miriam Halsner, Sven Mayer. “A Mixed-Method Exploration into the Mobile Phone Rabbit Hole.” In: *Proc. ACM Hum.-Comput. Interact.* 7.MHCI (2023). DOI: 10.1145/3604241.

[394] Christopher Thompson, Maritza Johnson, Serge Egelman, David Wagner, Jennifer King. “When It’s Better to Ask Forgiveness than Get Permission: Attribution Mechanisms for Smartphone Resources.” In: *Proceedings of the Ninth Symposium on Usable Privacy and Security*. SOUPS ’13. Newcastle, United Kingdom: Association for Computing Machinery, 2013. DOI: 10.1145/2501604.2501605.

[395] Sabina Tomkins, Steven Isley, Ben London, Lise Getoor. “Sustainability at Scale: Towards Bridging the Intention-Behavior Gap with Sustainable Recommendations.” In: *Proceedings of the 12th ACM Conference on Recommender Systems*. RecSys ’18. Vancouver, British Columbia, Canada: Association for Computing Machinery, 2018, 214–218. DOI: 10.1145/3240323.3240411.

[396] Bill Tomlinson. “Prototyping a community-generated, mobile device-enabled database of environmental impact reviews of consumer products.” In: *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*. IEEE. 2008, pp. 154–154. DOI: 10.1109/HICSS.2008.365.

[397] Janice Y Tsai, Serge Egelman, Lorrie Cranor, Alessandro Acquisti. “The effect of online privacy information on purchasing behavior: An experimental study.” In: *Information systems research* 22.2 (2011), pp. 254–268. DOI: 10.1287/isre.1090.0260.

[398] Lynn Tsai, Primal Wijesekera, Joel Reardon, Irwin Reyes, Serge Egelman, David Wagner, Nathan Good, Jung-Wei Chen. “Turtle guard: Helping android users apply contextual privacy preferences.” In: *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. 2017, pp. 145–162.

[399] Katerina Tzafilkou, Anastasios A Economides, Nikolaos Protogeros. “Mobile sensing for emotion recognition in smartphones: a literature review on non-intrusive methodologies.” In: *International Journal of Human–Computer Interaction* 38.11 (2022), pp. 1037–1051. DOI: 10.1080/10447318.2021.1979290.

[400] Simone Ueberwasser, Elisabeth Stark. “What’s up, Switzerland? A corpus-based research project in a multilingual country.” In: *Linguistik online* 84.5 (2017).

[401] Niels Van Berkel, Denzil Ferreira, Vassilis Kostakos. “The experience sampling method on mobile devices.” In: *ACM Computing Surveys* 50.6 (2017). DOI: 10.1145/3123988.

[402] Max Van Kleek, Ilaria Liccardi, Reuben Binns, Jun Zhao, Daniel J Weitzner, Nigel Shadbolt. “Better the devil you know: Exposing the data sharing practices of smartphone apps.” In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2017, pp. 5208–5220. DOI: 10.1145/3025453.3025556.

[403] Thea F Van de Mortel. “Faking it: social desirability response bias in self-report research.” In: *Australian Journal of Advanced Nursing, The* 25.4 (2008), pp. 40–48.

[404] Mariek M P Vanden Abeele. “Digital Wellbeing as a Dynamic Construct.” In: *Communication Theory* 31.4 (Oct. 2020), pp. 932–955. DOI: 10.1093/ct/qtaa024.

[405] Lev Velykoivanenko, Kavous Salehzadeh Niksirat, Noé Zufferey, Mathias Humbert, Kévin Huguenin, Mauro Cherubini. “Are those steps worth your privacy? Fitness-tracker users’ perceptions of privacy and utility.” In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5.4 (2021), pp. 1–41. DOI: 10.1145/3494960.

[406] Viswanath Venkatesh, Michael G Morris, Gordon B Davis, Fred D Davis. “User acceptance of information technology: Toward a unified view.” In: *MIS quarterly* (2003), pp. 425–478. DOI: 10.2307/30036540.

[407] Lieke Verheijen, Wessel Stoop. “Collecting Facebook Posts and WhatsApp Chats: Corpus Compilation of Private Social Media Messages.” In: *Text, Speech, and Dialogue*. Cham: Springer International Publishing, 2016, pp. 249–258. DOI: 10.1007/978-3-319-45510-5_29.

[408] Michele Vescovi, Christos Perentis, Chiara Leonardi, Bruno Lepri, Corrado Moiso. “My data store: toward user awareness and control on personal data.” In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. 2014, pp. 179–182. DOI: 10.1145/2638728.2638745.

[409] Sarah Theres Völkel, Daniel Buschek, Jelena Pranjic, Heinrich Hussmann. “Understanding Emoji Interpretation through User Personality and Message Context.” In: *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI ’19. Taipeh, Taiwan: ACM, 2019. DOI: 10.1145/3338286.3340114.

[410] Daniel Votipka, Seth M Rabin, Kristopher Micinski, Thomas Gilray, Michelle L Mazurek, Jeffrey S Foster. "User comfort with android background resource accesses in different contexts." In: *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*. 2018, pp. 235–250. DOI: 10.5555/3291228.3291247.

[411] Minh Duc Vu, Han Wang, Zhuang Li, Jieshan Chen, Shengdong Zhao, Zhenchang Xing, Chunyang Chen. "GPTVoiceTasker: LLM-Powered Virtual Assistant for Smartphone." In: *arXiv preprint arXiv:2401.14268* (2024).

[412] Rafael Wampfler, Severin Klingler, Barbara Solenthaler, Victor R. Schinazi, Markus Gross, Christian Holz. "Affective State Prediction from Smartphone Touch and Sensor Data in the Wild." In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. New Orleans, LA, USA: Association for Computing Machinery, 2022. DOI: 10.1145/3491102.3501835.

[413] Chih-Chien Wang, Chun-An Chen, Jui-Chin Jiang. "The Impact of Knowledge and Trust on E-Consumers' Online Shopping Activities: An Empirical Study." In: *J. Comput.* 4.1 (2009), pp. 11–18.

[414] Leye Wang, Daqing Zhang, Dingqi Yang, Brian Y Lim, Xiaojuan Ma. "Differential location privacy for sparse mobile crowdsensing." In: *2016 IEEE 16th International Conference on Data Mining (ICDM)*. IEEE. 2016, pp. 1257–1262. DOI: 10.1109/ICDM.2016.0169.

[415] Na Wang, Bo Zhang, Bin Liu, Hongxia Jin. "Investigating effects of control and ads awareness on android users' privacy behaviors and perceptions." In: *Proceedings of the 17th international conference on human-computer interaction with mobile devices and services*. 2015, pp. 373–382. DOI: 10.1145/2785830.2785845.

[416] Tianyu Wang, Giuseppe Cardone, Antonio Corradi, Lorenzo Torresani, Andrew T Campbell. "Walksafe: a pedestrian safety app for mobile phone users who walk and talk while crossing roads." In: *Proceedings of the twelfth workshop on mobile computing systems & applications*. 2012, pp. 1–6. DOI: 10.1145/2162081.2162089.

[417] Yang Wang, Huichuan Xia, Yun Huang. "Examining American and Chinese internet users' contextual privacy preferences of behavioral advertising." In: *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*. 2016, pp. 539–552. DOI: 10.1145/2818048.2819941.

[418] Yongfeng Wang, Zheng Yan, Wei Feng, Shushu Liu. “Privacy protection in mobile crowd sensing: a survey.” In: *World Wide Web* 23.1 (2020), pp. 421–452. DOI: 10.1007/s11280-019-00745-2.

[419] Alicia Wanless. *TED Talk: We Are All Alice Now: Falling Down a Digital Rabbit Hole*. 2016.

[420] Tanapuch Wanwarang, Nataniel P. Borges, Leon Bettscheider, Andreas Zeller. “Testing Apps With Real-World Inputs.” In: *Proceedings of the IEEE/ACM 1st International Conference on Automation of Software Test*. AST ’20. Seoul, Republic of Korea: Association for Computing Machinery, 2020, 1–10. DOI: 10.1145/3387903.3389310.

[421] Rick Wash. “Folk models of home computer security.” In: *Proceedings of the Sixth Symposium on Usable Privacy and Security*. 2010, pp. 1–16. DOI: 10.1145/1837110.1837125.

[422] Sunil Wattal, Rahul Telang, Tridas Mukhopadhyay, Peter Boatwright. “What’s in a “name”? Impact of use of customer information in e-mail advertisements.” In: *Information Systems Research* 23.3-part-1 (2012), pp. 679–697. DOI: 10.1287/isre.1110.0384.

[423] Clara Weber, Birgitta Gatersleben. “Office relocation: changes in privacy fit, satisfaction and fatigue.” In: *Journal of Corporate Real Estate* 24.1 (2022), pp. 21–39. DOI: 10.1108/JCRE-12-2020-0066.

[424] Paul Weiser, Simon Scheider, Dominik Bucher, Peter Kiefer, Martin Raubal. “Towards sustainable mobility behavior: Research challenges for location-aware information and communication technology.” In: *GeoInformatica* 20.2 (2016), pp. 213–239. DOI: 10.1007/s10707-015-0242-x.

[425] Hao Wen, Yuanchun Li, Guohong Liu, Shanhui Zhao, Tao Yu, Toby Jia-Jun Li, Shiqi Jiang, Yunhao Liu, Yaqin Zhang, Yunxin Liu. “Empowering llm to use smartphone for intelligent task automation.” In: *arXiv preprint arXiv:2308.15272* (2023).

[426] Alan F Westin. “Privacy and freedom.” In: *Washington and Lee Law Review* 25.1 (1968), p. 166.

[427] Angela Wichmann. “Quantitative und qualitative Forschung im Vergleich.” In: *Denkweisen, Zielsetzungen und Arbeitsprozesse, Berlin* (2019).

[428] Primal Wijesekera, Arjun Baokar, Ashkan Hosseini, Serge Egelman, David Wagner, Konstantin Beznosov. "Android Permissions Remystified: A Field Study on Contextual Integrity." In: *24th USENIX Security Symposium (USENIX Security 15)*. Washington, D.C.: USENIX Association, Aug. 2015, pp. 499–514.

[429] Primal Wijesekera, Arjun Baokar, Lynn Tsai, Joel Reardon, Serge Egelman, David Wagner, Konstantin Beznosov. "Dynamically regulating mobile application permissions." In: *IEEE Security & Privacy* 16.1 (2018), pp. 64–71. DOI: 10.1109/MSP.2018.1331031.

[430] Primal Wijesekera, Arjun Baokar, Lynn Tsai, Joel Reardon, Serge Egelman, David Wagner, Konstantin Beznosov. "The Feasibility of Dynamically Granted Permissions: Aligning Mobile Privacy with User Preferences." In: *2017 IEEE Symposium on Security and Privacy (SP)*. 2017, pp. 1077–1093. DOI: 10.1109/SP.2017.51.

[431] Primal Wijesekera, Joel Reardon, Irwin Reyes, Lynn Tsai, Jung-Wei Chen, Nathan Good, David Wagner, Konstantin Beznosov, Serge Egelman. "Contextualizing privacy decisions for better prediction (and protection)." In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018, pp. 1–13. DOI: 10.1145/3173574.3173842.

[432] Michael Wilson. "MRC psycholinguistic database: Machine-usable dictionary, version 2.00." In: *Behavior Research Methods, Instruments, & Computers* 20.1 (1988), pp. 6–10. DOI: 10.3758/BF03202594.

[433] Maximiliane Windl, Niels Henze, Albrecht Schmidt, Sebastian S Feger. "Automating contextual privacy policies: Design and evaluation of a production tool for digital consumer privacy awareness." In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 2022, pp. 1–18. DOI: 10.1145/3491102.3517688.

[434] Maximiliane Windl, Sven Mayer. "The Skewed Privacy Concerns of Bystanders in Smart Environments." In: *Proc. ACM Hum.-Comput. Interact.* MobileHCI (6 Sept. 28, 2022). DOI: 10.1145/3546719.

[435] Maximiliane Windl, Albrecht Schmidt, Sebastian S. Feger. "Investigating Tangible Privacy-Preserving Mechanisms for Future Smart Homes." In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Hamburg, Germany: Association for Computing Machinery, 2023. DOI: 10.1145/3544548.3581167.

[436] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, James J. Higgins. “The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’11. Vancouver, BC, Canada: Association for Computing Machinery, 2011, 143–146. DOI: 10.1145/1978942.1978963.

[437] Verena M Wottrich, Eva A van Reijmersdal, Edith G Smit. “The privacy trade-off for mobile app downloads: The roles of app value, intrusiveness, and privacy concerns.” In: *Decision support systems* 106 (2018), pp. 44–52. DOI: 10.1016/j.dss.2017.12.003.

[438] Kuang-Wen Wu, Shaio Yan Huang, David C Yen, Irina Popova. “The effect of online privacy policy on consumer privacy concern and trust.” In: *Computers in human behavior* 28.3 (2012), pp. 889–897. DOI: 10.1016/j.chb.2011.12.008.

[439] Ossy Dwi Endah Wulansari, Johanna Pirker, Johannes Kopf, Christian Guetl. “Video Games and Their Correlation to Empathy.” In: *The Impact of the 4th Industrial Revolution on Engineering Education*. Springer International Publishing, 2020. DOI: 10.1007/978-3-030-40274-7_16.

[440] Stephen Xia, Xiaofan Jiang. “Pams: Improving privacy in audio-based mobile systems.” In: *Proceedings of the 2nd International Workshop on Challenges in Artificial Intelligence and Machine Learning for Internet of Things*. 2020, pp. 41–47. DOI: 10.1145/3417313.3429383.

[441] Heng Xu, Sumeet Gupta, Mary Beth Rosson, John Carroll. “Measuring mobile users’ concerns for information privacy.” In: (2012), pp. 2278–2293. DOI: 10.1.1.668.3794.

[442] Heng Xu, Hock-Hai Teo, Bernard CY Tan, Ritu Agarwal. “Research note—effects of individual self-protection, industry self-regulation, and government regulation on privacy concerns: a study of location-based services.” In: *Information systems research* 23.4 (2012), pp. 1342–1363.

[443] Mengwei Xu, Feng Qian, Qiaozhu Mei, Kang Huang, Xuanzhe Liu. “Deeptype: On-device deep learning for input personalization service with minimal privacy concern.” In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2.4 (2018), pp. 1–26. DOI: 10.1145/3287075.

[444] Shuiqing Yang, Yaobin Lu, Yuangao Chen, Sumeet Gupta. “Understanding Consumers’ Mobile Channel Continuance: An Empirical Investigation of Two Fitness Mechanisms.” In: *Behav. Inf. Technol.* 34.12 (2015), 1135–1146. DOI: 10.1080/0144929X.2014.988176.

[445] Tevfik Sukru Yaprakli, Musa Unalan. “Consumer Privacy in the Era of Big Data: A Survey of Smartphone Users’ Concerns.” In: *PressAcademia Procedia* 4.1 (2017), pp. 1–10. DOI: 10.17261/Pressacademia.2017.509.

[446] Tal Yarkoni. “Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers.” In: *Journal of Research in Personality* 44.3 (2010), pp. 363–373. DOI: 10.1016/j.jrp.2010.04.001.

[447] Özgür Yürür, Chi Harold Liu, Zhengguo Sheng, Victor CM Leung, Wilfrido Moreno, Kin K Leung. “Context-awareness for mobile sensing: A survey and future directions.” In: *IEEE Communications Surveys & Tutorials* 18.1 (2014), pp. 68–93. DOI: 10.1109/COMST.2014.2381246.

[448] Marco Zappatore, Antonella Longo, Mario A Bochicchio. “Using mobile crowd sensing for noise monitoring in smart cities.” In: *2016 international multidisciplinary conference on computer and energy science (Splitech)*. IEEE. 2016, pp. 1–6. DOI: 10.1109/SpliTech.2016.7555950.

[449] Eric Zeng, Shrirang Mare, Franziska Roesner. “End user security and privacy concerns with smart homes.” In: *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. 2017, pp. 65–80. DOI: 10.5555/3235924.3235931.

[450] Bo Zhang, S Shyam Sundar. “Proactive vs. reactive personalization: Can customization of privacy enhance user experience?” In: *International journal of human-computer studies* 128 (2019), pp. 86–99. DOI: 10.1016/j.ijhcs.2019.03.002.

[451] Heng Zhang, Ahmed Ibrahim, Bijan Parsia, Ellen Poliakoff, Simon Harper. “Passive social sensing with smartphones: a systematic review.” In: *Computing* 105.1 (2023), pp. 29–51. DOI: 10.1007/s00607-022-01112-2.

[452] Shaokun Zhang, Hanwen Lei, Yuanpeng Wang, Ding Li, Yao Guo, Xiangqun Chen. “How Android Apps Break the Data Minimization Principle: An Empirical Study.” In: *2023 38th IEEE/ACM International Conference on Automated Software Engineering (ASE)*. IEEE. 2023, pp. 1238–1250. DOI: 10.1109/ASE56229.2023.00141.

[453] Ying Zhao, Jinjun Chen. “A survey on differential privacy for unstructured data content.” In: *ACM Computing Surveys (CSUR)* 54.10s (2022), pp. 1–28. DOI: 10.1145/3490237.

[454] Xin Zhou, Yang Li. “Large-scale modeling of mobile user click behaviors using deep learning.” In: *Proceedings of the 15th ACM Conference on Recommender Systems*. 2021, pp. 473–483. DOI: 10.1145/3460231.3474264.

[455] Yajin Zhou, Xinwen Zhang, Xuxian Jiang, Vincent W Freeh. “Taming information-stealing smartphone applications (on android).” In: *Trust and Trustworthy Computing: 4th International Conference, TRUST 2011, Pittsburgh, PA, USA, June 22-24, 2011. Proceedings 4*. Springer. 2011, pp. 93–107. DOI: 10.1007/978-3-642-21599-5_7.

[456] J Christopher Zimmer, Riza Arsal, Mohammad Al-Marzouq, Dewayne Moore, Varun Grover. “Knowing your customers: Using a reciprocal relationship to enhance voluntary information disclosure.” In: *Decision Support Systems* 48.2 (2010), pp. 395–406. DOI: 10.1016/j.dss.2009.10.003.

[457] Christian Zimmermann, Rafael Accorsi, Günter Müller. “Privacy dashboards: reconciling data-driven business models and privacy.” In: *2014 Ninth International Conference on Availability, Reliability and Security*. IEEE. 2014, pp. 152–157. DOI: 10.1109/ARES.2014.27.

[458] Oren Zuckerman, Ayelet Gal-Oz. “Deconstructing gamification: evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity.” In: *Personal and ubiquitous computing* 18 (2014), pp. 1705–1719. DOI: 10.1007/s00779-014-0783-2.

[459] Niels van Berkel, Jorge Goncalves, Simo Hosio, Zhanna Sarsenbayeva, Eduardo Velloso, Vassilis Kostakos. “Overcoming compliance bias in self-report studies: A cross-study analysis.” In: *International Journal of Human-Computer Studies* 134 (2020), pp. 1–12. DOI: 10.1016/j.ijhcs.2019.10.003.

All URLs cited were checked in December 2024.

List of Figures

List of Tables

Appendix

8.1 Declaration of Writing Aids

This PhD thesis is composed by myself and my co-authors original thoughts an comments. As writing aids, I used DeepL[1] to translate German text into English. I used ChatGPT[2] for ideation. To create some of the figures, such as the front matter, I used the image generation AI Adobe Firefly[3]. The figures in this thesis consist of contents created by myself, except contents of screenshots which are covered by the fair use principles, icons from the Adobe font EmojiOne[4], and graphics obtained through Microsoft Office[5].

[1] https://www.deepl.com/
[2] https://chatgpt.com/
[3] https://firefly.adobe.com/
[4] https://github.com/adobe-fonts/emojione-color
[5] https://www.microsoft.com/de-de/microsoft-365

8.2 Clarification of Contributions

	My Contribution	Contribution of Co-authors
[42]	I was the project lead and first author of the resulting publication. I developed the research idea and implemented it both on app- and data evaluation side. I came up with the field hint text categorization approach, proposed the initial coding structure, and implemented it in our large-scale field study.	Timo Koch helped with the framing and writing, keyword coding, and parts of the analysis code. Maximilian Bergmann participated in the keyword coding process. Ramona Schödel supervised the research direction, worked on data evaluations, and their reporting and general paper writing. Clemens Stachl, Daniel Buschek and Sven Mayer provided feedback on the research concept and final paper during all stages of the research project. This publication uses data that was collected in a large-scale, interdisciplinary field study where many further people contributed to [359].
[393]	Together with Nađa Terzimehić I share the first authorship. Besides being involved in all conceptualization and writing duties, I was especially responsible for the quantitative parts of this paper.	The project is based on the Bachelor's thesis of Miriam Halsner, who implemented the Android app and conducted the user study. Nađa Terzimehić primarily conceptualized this research. During analysis and writing her focus was on the qualitative parts. Sven Mayer provided feedback on the data analysis and final paper during all stages of the research project.
[43]	I was the project lead and first author of the resulting publication. I developed the research idea, conducted the underlying research, and developed the presented concepts.	Sven Mayer provided feedback on the final paper during all stages of the research project.
[45]	I was the project lead and first author of the resulting publication. I developed the research idea, conducted the underlying research, and developed the presented concepts.	Sven Mayer provided feedback on the final paper during all stages of the research project.
[44]	I was the project lead and first author of the resulting publication. I developed the research idea, conducted the underlying research, and developed the presented concepts.	Sven Mayer provided feedback on the final paper during all stages of the research project.
[48]	I was the project lead and first author of the resulting publication. I conceptualized the underlying studies, supervised the data evaluation, and primarily implemented all quantitative analyses.	Maximiliane Windl participated in all stages of the project. We especially shared the qualitative coding and she engaged in writing contents. The publication is based on the Bachelor's thesis of Tobias Knobloch. He conducted both user studies, helped in the coding process, and contributed ideas. Sven Mayer provided feedback on the final paper during all stages of the research project. He was furthermore especially engaged in creating visuals and optimizing wordings.

Table 8.1 : This table provides an overview of the contributions of my own and my collaborators to the publications incorporated in this thesis.

	My Contribution	Contribution of Co-authors
[47]	I was the project lead and first author of the resulting publication. I supervised the underlying bachelor's thesis, that laid the basis for this work. I conceptualized, implemented, and conducted the main user study that the publication consists of.	The research project was based on the Bachelor's thesis of Jonas Erbe. He proposed the general research idea, implemented the first prototype, and conducted and evaluated a pre-study. Maximiliane Windl advised the writing of the final paper with her expertise in the privacy domain, and engaged in optimizing the final paper. Heinrich Hußmann supervised the project in its early stages. Sven Mayer supervised the research project in its later stages. He provided feedback on the final paper during all publication stages.
[40]	I was the project lead and first author of the resulting publication. I developed the research idea and software artifact. I planned and conducted the study and experiments.	Daniel Buschek supervised the research project during all stages. He steered it into the right direction and provided feedback for the research concept, study, as well as evaluation and paper writing.
[46]	I was the project lead and first author of the resulting publication. I came up with the research idea, concept, and developed the prototype. I conducted and supervised all prestudies and conducted the lab study.	The design studies were conducted in the context of the Master's thesis of Helena Stoll. She conducted and evaluated the three preliminary studies. Sven Mayer supervised the research project in all stages, especially regarding the study-based concept development. He actively participated in writing, coding analyses and creating visualizations. He provided feedback for the final publication in all stages.
[38]	I was the project lead and first author of the resulting publication. I developed the research idea, conducted the underlying research, and developed the presented concepts.	Carmen Mayer contributed with thoughts on the presented concepts and created all visualizations. Sven Mayer provided feedback on the final paper during all stages of the research project.

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbständig und nur mit den angegebenen Hilfsmitteln verfasst habe. Alle Passagen, die ich aus der Literatur oder aus anderen Quellen übernommen habe, habe ich deutlich als Zitat mit Angabe der Quelle kenntlich gemacht.

München, 08.01.2025 Florian Bemmann

Ort, Datum, Unterschrift