Informing the Design and Refinement of Privacy and Security Controls

Thesis Proposal

Daniel Smullen

CMU-CS-XXX-XXX

April 8, 2021

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:
Norman Sadeh (Chair, Institute for Software Research)
Lorrie Faith Cranor (Institute for Software Research)
Alessandro Acquisti (Heinz College)
Rebecca Weiss (External, Mozilla)
Yaxing Yao (External, UMBC)

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

This research was supported in part by grants from DARPA and AFRL under the Brandeis project on Personalized Privacy Assistants For the Internet of Things (FA8750-15-2-0277) and by grants from the National Science Foundation Secure and Trustworthy Computing program (CNS-15-13957, CNS-1801316, CNS-1914486).

The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the NSF, DARPA, AFRL or the US Government. The US Government is authorized to reproduce and distribute reprints for Governmental purposes not withstanding any copyright notice.

Keywords: privacy, security, usability, settings, affordances, notice and choice, awareness, control, machine learning

Abstract

Amid increasing privacy and security risks, managing one's privacy and security choices is becoming ever more important. Yet, the proliferation of security and privacy controls is making this task overwhelmingly complex. Are so many controls actually needed? Are they the right controls, and are they effective? Ideally, the available controls should enable people to align system behaviors with their security and privacy preferences. These preferences typically reflect the flexibility users want systems to have, their tolerance for risk, and their confidence in their ability to mitigate these risks using available controls. This dissertation explores security and privacy settings in web browsers and mobile apps. Our overall objective is to determine how effective they are at giving users the awareness and control they need, and inform ways to improve these controls, if necessary. This includes determining whether users are able to identify risks, whether they are aware of what controls are available and what they do, and whether they are able to use available controls to restrict behaviors that they are uncomfortable with and mitigate risks. This dissertation comprises a series of three studies designed to help answer the questions identified above. Results from these three studies are the basis for some recommendations to improve privacy and security controls. We also discuss the public policy implications of some of our findings and how these findings extend beyond browsers and mobile app environments.

Our first study explores people's understanding of the privacy and security settings offered by today's five most popular browsers. We conduct an in-depth study of the extent to which people are able to use their browser to identify and mitigate privacy and security risks. This work is ongoing and details of what remains to be completed are included in this text.

Acknowledging the challenges people face in effectively using browser settings, in the next chapter we turn our attention to studying people's preferred settings in the context of a representative set of potentially intrusive online practices. We show that although people have diverse concerns and preferences, there are commonalities in the ways they prefer their browser to behave in different situations, such as restricting some practices (e.g., fingerprinting, behavioral profiling, targeted ads) across categories of websites by default. We use our findings to explore ways to simplify browser settings. Our results suggest that it is possible to better align settings with people's mental models and make them easier to configure. These findings also suggest that privacy and security controls could be specified in browsers, configured by users in one interface, and enforced by the browser across every website rather than requiring users to identify and configure settings for each individual website they visit. This would require the adoption of standards for presenting and capturing privacy and security preferences from users, which has been resisted by industry so far.

In our third study, we extend our exploration to mobile app permissions. Today's permission settings omit the purpose for which apps request permissions. For instance, there is no distinction between permissions required to support core functionality and those required for other purposes (e.g., advertising). Incidentally, new regulations such as GDPR require distinguishing between different purposes when requesting user consent. We study trade-offs between the user burden associated with configuring settings that are capable of distinguishing between different purposes, and the ability of such finer settings to better capture people's preferences. Unsurprisingly, we find that finer controls would increase the number of settings users would need to configure and thereby also increase user burden. We explore the extent to which this additional burden can be mitigated using machine learning. Our findings suggest that with finer permissions it is possible to build more accurate predictive models of people's preferences. In fact, our findings suggest that such models make it possible to get the benefits of finer controls that better capture people's concerns and preferences without increasing user burden.

In summary, expected contributions of this work include the identification of significant shortcomings in existing privacy and security controls available in popular browsers (currently underway) and for controlling mobile app permissions (already completed). Existing controls are misaligned with people's mental models, fall short when it comes to informing people about privacy and security risks, and do not adequately allow people to mitigate privacy and security risks they are concerned about. Approaches to address these shortcomings would benefit from more systematic analyses of people's privacy and security preferences, taking into account the limits of what people are realistically capable of doing. These approaches would also benefit from taking advantage of correlations in people's preferences and concerns to simplify privacy and security controls, and/or to leverage predictive machine learning models that could be used to recommend settings to users. Our findings also illustrate how greater levels of standardization in capturing and communicating privacy and security preferences could further simplify the task of configuring an otherwise unmanageable collection of privacy and security controls.

Except for one additional study, the work described in this proposal has already been completed, and the author believes that the remaining work can be submitted by the end of the August 2021.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This dissertation is about improving the management of privacy and security controls. What do we mean by privacy and security, and why are controls needed? Privacy has classically been characterized by a tension between the threat of pervasive and intrusive data collection, versus the preservation of individual liberties and autonomy [91]. Enabling personal choice is the most important guiding principle, since privacy preferences vary among individuals and are context-dependent [15]. In contrast, usable security experts have advocated to not offer any choices when there is a clearly optimal or "safe" default [22]. However, the reality of today's complex software systems is such that alternative security options are still necessary in many circumstances. People have different sets of concerns, different degrees of tolerance for risk, and different degrees of confidence in their ability to steer clear from threats or mitigate risks. There is no clearly optimal default that will satisfy every user in every possible situation. Thus, privacy and security controls are both based around individual risk tolerances as well as preferences, which can in turn be influenced by myriad human factors such as cost constraints, effort, time, expertise, and so on. It goes without saying that in cases where controls are needed, they should be offered. Thus, most software offers ways to configure at least some kind of privacy and security controls. Interfaces for configuring these controls should be designed to ensure that people can understand the practices they may encounter, the potential risks involved, and provide the ability to realign system behaviors based on their preferences and risk tolerance.

Evaluating Privacy and Security Controls

Software developers can evaluate the design of their privacy and security controls by considering two simple questions, which are central to this work: First, does the system do

a good enough job at informing users about the practices that they might potentially be concerned about? Second, do users have the ability to take control of the practices that specifically concern them? Unfortunately, privacy and security incidents are continuous reminders that what is offered to users has room for improvement. Therefore, in this work we ask, how effective are today's privacy and security controls, and can we improve on the current situation? In our first study, we begin by evaluating the state of the art, starting with the approaches taken by popular web browsers. This initial study sets the stage for a deeper exploration on how to improve what is offered online and elsewhere, informed by both our novel evaluation as well as issues seen in prior work.

Improving Privacy and Security Controls

Designing effective privacy and security controls is a hard problem, and our first chapter's evaluation of the state of the art is expected to uncover a variety of shortcomings that need to be addressed. Further research is needed to move beyond the state of the art, inform best practices, and promote better standards. Concretely, we must identify ways in which controls become overly burdensome or inaccurate as a result of sub-optimal design. We also plan to suggest possible ways to improve these designs.

In our second study, we explore rich user perspectives on controlling a variety of potentially intrusive (and insecure) practices in the browser such as fingerprinting, behavioral profiling, and crypto-mining. Specifically, we survey users' preferences and expectations for control as well as their desire to be notified about these practices. Using a mixed-methods approach, we analyze how people would prefer to be able to configure their browser in an attempt to find ways to improve what is currently offered. Based on our findings, we suggest concrete ways to address some of the problems that are associated with the ad-hoc solutions offered on individual websites. We observed that users are confused about the extent of their ability to restrict certain practices. Improved settings which offer the control they expect to have on individual websites would be overly burdensome. Many websites offer settings which are too complex, while many others offer only trivial settings that are too inaccurate; these controls often fail to achieve their goals due to users' unwillingness to make the effort to use them, various usability issues, or misconceptions about how and when to use them. Our recommendations aim to standardize settings along factors which we show to better align with users' mental models and preferences, such as categories of websites and individual practices. This better-aligned model, combined with restricting intrusive practices by default, can decrease burden without sacrificing accuracy. However, this proposed approach would also require standardized settings and APIs which do not currently exist. In principle, this approach would move browser settings to more

closely resemble mobile app permissions. In addition to these standards, we propose regulations which would potentially mitigate the possibility of websites intentionally breaking when users express stricter privacy and security preferences that conflict with websites' business goals. We have identified this as a possible dark pattern, where websites attempt to coerce users into relaxing their settings based on the knowledge that they may discard their privacy and security preferences in favor of completing their original task.

In our third study, we move beyond web browsers to study the potential for improving mobile app permissions. The design of privacy and security controls in web browsers are guided only by limited standards (many of which are voluntary or no longer supported [76]), resulting in settings which can vary from website to website. In contrast, mobile app permissions are standardized and are uniformly enforced by mobile operating systems. Yet despite app permissions already conforming to a well-defined standard, configuring permissions remains an overwhelming task given the explosive growth of apps and their increased use of sensitive permissions. Research has repeatedly shown that people express different preferences depending on the purpose for which a permission is being requested by an app [48, 51, 83]. Mobile app permissions currently do not capture this factor, but including settings subject to purpose would further increase burden. Mobile apps are in need of ways to provide more comprehensive settings without increasing the existing burden, which is already unrealistically high. Here, intelligent recommendations offer the promise of simplifying the management of complex and numerous settings. Using a combination of supervised and unsupervised machine learning techniques, we show that it is possible to shift the burden away from users. Decision support can provide recommendations based on more complex and nuanced factors. These factors would make permissions prohibitively burdensome if they were naively introduced into existing permissions. By sweeping the parameter space of this approach, we are able to demonstrate that it is capable of favoring accuracy, user burden, or balancing both objectives simultaneously. By quantifying the relationship between different parameters and the expected accuracy versus user burden, we provide concrete guidance that can enable developers to optimize their solutions according to their required levels of accuracy or burden.

Awareness and Control versus Notice and Choice

In this work, we focus on two key dimensions: *awareness and control*. These concepts refer to the need to both adequately inform users and provide them with adequate choices, whether in a privacy context (where this concept is often referred to as "notice and choice" [23]) or in a security context. As techniques such as machine learning, fingerprinting, profiling, and other forms of automated reasoning become increasingly pervasive and

capable [44, 5, 90, 85], people may experience such practices nearly constantly as they engage with software [73]. Unfortunately, in spite of their potential benefits, these practices also pose privacy and security risks; users may object to practices involving surveillance, and may feel violated if subjected to these practices without their consent. The associated data collection may also result in insecurity through unanticipated dissemination, breach, or secondary usage [6, 94, 19, 79]. The concepts of awareness and control are intended to provide ways of addressing and mitigating privacy and security risks.

In theory, frameworks such as awareness and control (or notice and choice) empower users to make themselves aware of the practices that they are subject to, understand their options, and take action to restrict those that they deem unacceptably intrusive or risky [23, 68]. However, in a practical sense, control is only possible within the design parameters of what developers provide to users. The basic assumption is that settings are provided to users with the expectation that they will understand and make full use of them. This results in the configuration burden being placed solely on users. However, users have limited attention, and configuring their settings is a secondary task [3]. This trade-off between increased burden and more comprehensive settings makes designing usable settings even more difficult, and evaluating whether a particular design has incorporated the right trade-off is critical. This also does not discount the importance of default settings (which should be conservative) to ensure that users are initially protected until or unless they decide to opt for less risk-averse settings. Designs which incorporate more expressive, granular settings can take into account different contexts, can enable settings to be more accurate, and can better account for users' various individual preferences – ideally, this is what developers should strive for in their designs. On the other hand, engaging with more complex settings comes at the cost of a higher cognitive and attentional burden [16]. If managing their settings becomes too burdensome, users will refuse to engage meaningfully with the settings altogether [5, 4, 64]. The trade-off between accuracy and user burden in privacy and security settings is an understudied phenomenon, despite being evident in the literature for many years [16, 50]. Each of these factors and more, in different contexts, is known to contribute to a measurable change in a particular individual's preferred settings [15]. Overly simplistic or one-size-fits-all solutions are therefore unlikely to be satisfying for many users [93]. Effective designs must maximize accuracy (the ability to correctly capture users' preferences) and minimize user burden (the amount of effort users must endure to enact their preferences) based on a principled approach.

Chronology

This dissertation is structured around chapters which have been worked on in reverse chronological order. Traditionally, research problems are identified and the problem landscape studied in detail during early exploratory work. This initial stage is usually what motivates solutions to be developed after the problem has been better understood. After they are developed, solutions usually undergo further evaluation, and this is the culmination of the research. The order of the chapters in this dissertation have been arranged to reflect this more traditional, clearer narrative. This decision can be further justified by prior literature, which has repeatedly called attention to the issues that this work focuses on addressing. Privacy and security settings are a well studied area, but there are still many aspects which are in need of further research. Moreover, the rapidly evolving domains of web browsing and mobile apps routinely introduce new developments which reveal new research questions and create the possibility of new answers for previously answered questions.

Our first technical chapter proposes a deep qualitative study that seeks to systematically evaluate the status quo in browsers, and we offer ways to refine these controls further in subsequent chapters. We decided to present our work on web browsers before our work on mobile app permissions, because this ordering allows for a more natural progression. We propose to evaluate the extent to which browsers ensure that users are made aware of potential privacy and security risks, and provide ways to control these risks. This proposed study has the potential to uncover new problems as popular browsers revise their privacy and security interfaces, and it is worth providing a richer and more detailed evaluation of what is now offered. Presenting this exploratory study first better reflects a more traditional narrative.

1.1 Main Contributions

This dissertation evaluates the effectiveness of current browser interfaces and mobile app permission managers at communicating privacy and security risks, and providing meaningful controls. Part of this evaluation should determine and characterize the trade-offs that are inherent in the various settings and interfaces now being provided by different browsers. We want to do a thorough study of the way in which users are made aware of, interpret, understand, and make use of the controls they provide. The findings of this study are likely to lead naturally into further questions about users' preferences, and what they want to control versus what they are actually able to do in reality.

This dissertation identifies ways to improve software engineering practice. Our existing findings and the expected findings of our proposed study point to the need for ways to better educate and inform users, offer more rational controls, and incorporate more sensible defaults. If the state of practice is to move forward, we show that the following issues must be addressed in software generally. First, interfaces for managing privacy and security settings online are in need of improvement and require principled and systematic evaluation. People's preferences are often not well aligned with what is offered, and more effective standardized controls would be simpler and easier to configure. In the case of browsers, standardizing controls around factors such as website categories and intrusive practices can enable browsers to be a neutral platform for expressing people's preferences more broadly. Second, we must address public policy issues associated with standardization, such as the need for APIs, and misaligned incentives which make it difficult to improve upon the state of the art. Finally, in controls that already employ standardization, such as mobile apps, machine learning shows the potential to take advantage of complex information to make configuration easier.

This dissertation provides design guidance for practitioners which will enable them to maximize the benefits, accuracy, and acceptance of the controls they provide while minimizing user burden. Engineering is about making appropriate trade-offs, but we show that it may be possible to mitigate the trade-offs inherent in designing privacy and security controls. Our work using machine learning provides strong evidence that it can help to maximize the benefits of controls that are based on constrained permissions models such as those we propose for browsers, and which are already present in mobile apps. Machine learning can also minimize the drawbacks in terms of user burden by providing better recommendations. These recommendations make greater use of available information than existing models which rely solely on user input to manually configure permissions. Instead, we show that it is possible to use machine learning to accurately infer the preferred settings for most individuals. By characterizing the parameter space for a specific machine learning approach, we show that it is possible to tailor these recommendations to optimize for accuracy, user burden, or both.

1.2 Outline of Technical Chapters

Evaluating the Effectiveness of Browser Privacy and Security Settings The purpose of this proposed technical chapter is to evaluate a variety of popular browsers' privacy and security settings. We will explore the extent to which people understand the settings, how they interpret what they do, and gauge people's ability to make use of them to identify

and mitigate a variety of privacy and security risks. Today, settings are offered in web browsers through streamlined privacy and security dashboards, as well as related options available elsewhere in the browser. It is unclear whether these offerings can effectively provide awareness and control. People may also have a variety of assumptions about what the default settings are, and expectations about what the settings can control. It would be useful to better understand whether dashboards in particular can present information in a way that users can understand, particularly when these dashboards are often already simplified versions of more complex interfaces that exist in other places. Further, it would be important to identify whether the controls provided on these dashboards (and elsewhere) enable users to express their desired preferences, particularly subject to the limitations of the design choices implemented in different browsers.

As the first technical chapter, this study is intended to establish a deep, rich perspective of what is currently offered in browsers, and the problems associated with them. This chapter should also illustrate that there is more to study in terms of individual perspectives and preferences. This leads into the second technical chapter, which explores both individual and aggregate user perspectives in more detail.

Mismatches Between Desired and Available Privacy and Security Controls in Browsers
This technical chapter sheds light on the signals which users rely on when determining whether they are encountering or avoiding intrusive practices while browsing. We also uncover perspectives on the existing settings which people believe are associated with managing these practices, and survey how they would prefer to be able to configure their browser. This research was performed using a mixed-methods study consisting of qualitative and quantitative surveys. Confirming prior work, we find that people have very little insight into what is actually happening during browsing and have unrealistic expectations about controlling the types of intrusive practices they may encounter. We also find that people are unable to identify reliable ways of determining whether they are being subject to the practices that they would prefer to restrict, and whether their attempted interventions are effective. Often, they end up making inconsistent inferences based on the presence of ads. This opens the potential for unawareness and unmitigated risks when ad-blockers are used, further exacerbated by over-reliance on security tools which may not be intended to address the risks which people are primarily concerned about. Though they expect settings to be available to control most practices and express the desire to opt out, they are unable to take appropriate action. Users need better ways of objectively determining what is going on, and need to have the ability to take control, but what is offered ad-hoc across websites is falling short.

This chapter also identifies alternative controls which could potentially more accu-

rately capture people's preferences, given that most people want to opt out of most practices. Alternative settings which offer opt-out by default are likely to be more in line with people's expectations. We show that alternative settings would potentially be easier to configure, particularly when they incorporate factors which we show would better fit with people's mental models (e.g., website categories and intrusive practices). Importantly, this chapter establishes that people's preferences can be accurately captured by standardizing controls based on these factors, introducing the possibility to eliminate the need to reconfigure them on every website. These standards would need to be uniformly implemented through the use of APIs which do not currently exist. Improving the controls further also has the potential to be hampered by websites who may be incentivized not to accommodate people's preferences or even break intentionally. However, combined with regulation, new APIs could provide the standardization that is needed to enable browsers to act as a neutral platform for centrally managing users' preferences much in the way that mobile app permissions operate. We turn our attention to mobile app permissions in the final technical chapter.

Mitigating Trade-offs Between Accuracy and User Burden Using Machine Learning: A Study of Mobile App Permissions Mobile apps, unlike web browsers, employ well-defined standardized permissions that are configured centrally and enforced by mobile operating systems. In principle, similar standardization may help improve browser controls as well. However, standards are not a panacea. Mobile app permissions are simpler and more uniform, but also do not account for many important factors that are known to affect users' preferences to allow or deny them. One example is the purpose for an app requesting access to sensitive APIs. Unfortunately, the explosive growth of mobile apps makes configuring permissions too burdensome as they are. While the introduction of new factors such as purpose has the potential to make the controls more accurately express what people want, this would further exacerbate the burden. In this work, we began by measuring the impact of various factors on the ability to predict people's permission settings, and determined which factors would have an impact on people's likelihood to allow or deny. We approached this by collecting a large corpus of user preferences, and performed regression analysis to determine the impact of surveyed factors on likelihood to allow or deny app permissions across a broad range of different types of apps.

Unsurprisingly, we found that people's preferences are complex and can be influenced by a variety of factors. However, we found evidence that the addition of factors such as purpose, which would ordinarily increase user burden, had the potential to improve the predictive power of our models. These models could be used to build machine learning based recommendation systems, and in turn potentially alleviate user burden by accurately

inferring people's preferences. By experimenting with a combination of supervised and unsupervised machine learning approaches, we found that it was possible to leverage this predictive power and generate recommendations which make configuration easier. When incorporating additional factors which significantly differentiate users' preferences (such as the purpose), the machine learning models allow more comprehensive settings to be offered. As the accuracy of recommendations improves with this added complexity, configuration requires less manual decision-making. By sweeping a large portion of the parameter space for our approach, we found that it is possible to optimize for accuracy or user burden depending on the chosen parameters. Our results provide guidance for developers to tailor their implementation depending on their specifications for the minimum required accuracy, and the maximum tolerable burden. We also show that by optimizing for both objectives, it is possible to mitigate the trade-off between accuracy and user burden.

Chapter 2

Background and Related Work

In this chapter, we review prior literature and distinguish our work from prior studies. We begin by highlighting prior studies on design guidelines and evaluation principles. Moving on, we review several approaches which support awareness and control, revealing problems and trade-offs. We note recurring themes of promoting standardization, and the importance of user-centric design. Finally, acknowledging that standardization has limits, we outline prior work aimed at reducing user burden.

2.1 Design Guidelines and Evaluation Principles

Prior work has covered the parameters inherent in designing based on notice and control, and in particular sought ways to guide developers to make their designs effective. Our work applies these concepts to security settings, which is why we have adopted the use of the term awareness and control, which extends notice and choice. Schaub et al. describe the challenges, requirements, and best practices for designing privacy notices [71]. Other prior work also shows why many forms of awareness and control mechanisms have limited effectiveness and can lead to unnecessary burden in various domains, specifically browsing [70], mobile apps [13], and the Internet of Things [29]. This is the same phenomenon we see again and again in our work. We recognize browsers employ only variations of the "on demand" and "decoupled" archetypes when providing awareness [71]. In other words, browsers typically notify users about privacy and security risks as they occur (rather than in advance), but only if the user actively seeks out this information by opening the privacy and security interface at that time. We are motivated to study how these design choices may be working to create user burden, or otherwise limit the effectiveness of what is offered. In

the case of mobile app permissions, notifications are "just-in-time" and "blocking", which is part of what creates user burden – these notifications appear at the moment a permission is requested, and users cannot go back to their original task without making a decision to allow or deny at that moment [71, 13].

In this dissertation, we recognize the three key characteristics of effective methods of providing awareness and control identified in prior work [71, 23, 22]; they should be relevant, actionable, and understandable. We use these as guidelines for determining the effectiveness of methods to provide awareness and control in our research.

2.2 Mechanisms to Support Awareness and Control

Within the design space identified in the previous section, there are a multitude of solutions which have sought to improve privacy and security awareness and controls. One common thread going back many years is work which exposes disconnects between users' expectations and what is offered by different solutions [37, 21, 67]. Prior studies have evaluated a variety of solutions offered by browser add-ons and extensions [72, 84, 56], but we focus on what is offered in browsers' unmodified default configurations.

Prior studies chronicling historical changes to browser privacy and security awareness mechanisms show that changes over time have been subtle, and many browsers incorporate similar approaches and interfaces with common design themes [26, 45]. Many of these have questionable effectiveness [7, 45]. Some researchers have proposed specialized dashboards [17] and alternative browsers [24] which are intended to more comprehensively reveal and control privacy- and security-relevant data flows. These examples, among others [16], illustrate design trade-offs favoring comprehensiveness or accuracy – but which can be too burdensome or technically involved for the average user [56, 39]. While the effectiveness of shorter and more targeted explanations of data practices can be further influenced by their framing [31], there is evidence that oversimplification worsens the likelihood of some risks [18]. Prior work repeatedly highlights the overwhelming amount of information that needs to be processed by users to understand risks and filter out less essential information [87, 80, 40]. This leads to users facing difficulty identifying and managing risks independently of the modality [59]. Many studies have shown that the ad-hoc approaches seen online make restricting intrusive data practices exceptionally difficult, necessitating tool support [14, 46]. However, the way that settings, browser features, and privacy and security tools are portrayed has also been shown to be misleading and may result in unmitigated risks, such as believing that security tools like anti-virus software also prevents online data collection, or that "private browsing" mode offers com-

prehensive privacy and security protection [81, 2, 1].

Standardized awareness and control mechanisms, such as nutrition labels [41, 43, 75] based on experts' advice [27], hold promise as a way of simplifying privacy and security awareness and making risks easier to understand. Such work is part of a broader theme which upholds the importance of user-centered design [1, 34, 30] and accompanying standards [42]. Nudges [3] and personalized notifications [36] based around a standard set of data practices, disclosure decisions, and security choices have also been shown to be effective ways of drawing attention to, mitigating, and even preventing [25] risks. Our work also shows that standardization has value for browsers, simplifying settings by emphasizing important factors that align with users' mental models. Our work explores the existing dashboards and streamlined privacy and security settings offered by many browsers, evaluates the effectiveness of their approaches, and broadly surveys what average users believe would be satisfactory for them to feel that they are in control. We show that the constraints offered by appropriately standardized settings would have the effect of reducing the burden of configuring settings, without compromising their ability to express what people want.

2.3 Approaches to Reducing User Burden

One of the biggest challenges in offering effective awareness and control mechanisms is to reduce the burden that they place on users. Strict standardization is already present in mobile app permissions, yet they face many of the same challenges of balancing comprehensiveness with user burden. This is due in part to the unique and emergent challenges associated with widespread data collection [74, 48], diverse preferences [49, 88], and the explosive growth of both apps [66, 47] and their associated permissions [50, 54, 9, 86]. Prior work proposes to reduce this burden through alternative interaction designs that involve users negotiating with systems to balance competing interests [12, 11], semantics [69], or dynamically granting permissions as the circumstances evolve [92].

The most promising approaches to mitigate user burden employ machine learning. These approaches ease the burden on users by either performing configuration automatically [63], or are capable of offering recommendations based on profiles [52] which limit the number of manual decisions required to configure settings [50, 51]. In our work, we show that machine learning has the capability of not only simplifying the configuration of existing mobile app permissions, but can further ease the burden of more complex permissions models that incorporate factors that the existing settings do not support.

Chapter 3

Evaluating the Effectiveness of Browser Privacy and Security Settings

This chapter consists of proposed work which has not yet been completed. Accordingly, we introduce this proposed study by describing the problem and motivate our research goals. Next, we describe the research questions, methodology, and expected results.

3.1 Introduction

Literature shows that people's preferences and risk tolerances towards intrusive data practices online vary [6, 19]; such practices can include fingerprinting, crypto-mining, tracking related to social media, and a whole host of others. Many people are intolerant of these practices, misunderstand them, and want to be able to take control over them. Most importantly, people want to set their preferences according to their own perceptions and understanding of what's going on [67, 94].

To deal with this problem, some browsers are implementing dashboards streamlined notices and controls pertaining to a variety of potentially objectionable practices that may be encountered while browsing [60]. Many browsers have extensive privacy and security dashboards built in, while others offer more limited settings. Some browsers offer no settings or mechanisms to support awareness towards privacy and security risks at all [32]. For those browsers that offer such interfaces, dashboards are streamlined shortcuts to settings and information that the browser can report on regarding privacy and security risks encountered by the user. The shortcuts to controls which limit these types of practices often draw attention to specific threats. It is worth noting that for many browsers, third

party add-ons and tools exist which are designed to assist users in much the same way as dashboards, or make up for the lack of privacy and security interfaces in some browsers. We have chosen not to include these add-ons in the scope of our study on the basis that we wish to limit the variability they introduce compared to the standard out of box experience offered by browsers. Many of these add-ons are directed at technically sophisticated users, or are designed to combat specific privacy/security risks [39, 56]. Many of the settings embedded in websites fail to achieve their goals due to users' unwillingness to make the effort to use them, various usability issues, or misconceptions about how and when to use them [46]. Instead, we focus on the controls built in to browsers in their default configuration.

Practical guidance for how to provide certain kinds of notices and controls to users is well studied in the literature [71]. For example, studies show that when users are made aware of potential threats, they are more likely to make protective decisions based on their individual privacy and security preferences [67, 3, 8]. This suggests that browser users should be made aware of privacy and security risks as they are encountered, and offer settings to manage them somehow. In other domains (e.g., mobile apps), notifications, privacy managers, and permissions deliver important privacy and security information to users and also offer relevant controls. Between various mobile platforms such as iOS and Android, there is evidence that the differences in design approaches to these interfaces and settings make them more or less accurate, more or less burdensome, and more or less effective [77]. However, there are very few studies on the effectiveness of comparable interfaces (such as dashboards) built in to web browsers. We aim to fill this gap.

In the past, browsers provided extensive arrays of privacy and security information and settings which were hidden away in specialized control panels that required several clicks to view and interact with. A classic example is the settings offered by Internet Explorer 6, which were comprehensive, supported machine-readable privacy policies, and security zones [46]. The interface also incorporated small icons which were intended to make users aware of privacy and security threats, as well as the presence of options to mitigate certain threats (in particular, associated with third party cookies). These interfaces were likely created in response to work on standards such as P3P [20] which were emerging at the time, however there is little evidence of systematic evaluations which reported the effectiveness of their approach. Rather, many users found the settings to be too complex and they were often ignored [46, 23].

While many of these older types of interfaces still exist today in modern browsers in some form, some privacy and security dashboards in browsers offer a smaller subset which are presumably thought to be the most relevant to their users, with more granular or advanced settings offered elsewhere. Unlike more detailed settings, most dashboards are

accessible with a single click during normal browsing activities. Ostensibly, these dashboards are designed to be streamlined and simple. How users feel about contemporary offerings is unknown. Though they may be far simpler, it is possible that even current dashboards are too burdensome to be useful for the average user, or provide interfaces which some people cannot understand. There are clear trade-offs which are being made between the ability to provide more comprehensive settings that are more in alignment with what users want, versus omitting or simplifying these settings in favor of simplicity (at the risk of ignoring users' need for awareness and control). Some of the most popular browsers do not have dashboards which offer any settings, though some offer some limited information [60, 58] – is this enough? Some popular browsers offer no settings or information, omitting dashboards entirely [32] – are users aware of the potential for unmitigated risks? In general, current browsers are trending towards less informative, less intrusive, less interactive interfaces with less text and fewer options. Our study offers the potential to challenge this emerging trend.

3.1.1 Research Goal

The goal of this work is to evaluate the current privacy and security dashboards offered by browsers, particularly with respect to their ability to provide adequate awareness and control over privacy and security risks. Our main research questions can be summarized as follows: how good are these browsers in ensuring that users are aware of relevant privacy and security threats, have the control necessary to mitigate these threats, and are able to effectively take advantage of these controls?

We will employ a contextual interview study methodology, with the goal of collecting rich qualitative information about users' experiences [38]. This methodology is well suited to providing detailed accounts of how users experience interacting with what is provided by their browser, and has the potential to reveal insights about where they are successful, and where they fall short. To minimize the potential for confounding factors, it is important that all participants have the same experience of the browser they are assigned to. Data should therefore be collected through remote screen-sharing sessions, an approach that enables users to interact with a browser controlled by the interviewer – this enables the interviewer to have full control over the browser, ensuring that it is always consistently configured according to the defaults. To maximize the ecological validity of the study, the browser which is presented through the screen-sharing session would be a real browser, and the websites which are displayed in the browser will be controlled by the interviewer.

Part of the research goal is to account for the familiarity of users with a particular browser or set of browsers, as most users may be familiar with only a subset of the most

popular browsers. Chrome in particular is an overwhelmingly popular browser [53]. Accordingly, this study will use a purposive sampling strategy, aimed at collecting accounts of users with strong familiarity with the browser we interview them about. Participants will be assigned to be interviewed about the browser they claim to have the strongest familiarity with during pre-screening surveys. To collect a diverse set of experiences for each browser, we will aim to capture at least 5 interviewees familiar with each of the 5 most popular browsers (Edge, Safari, Chrome, Firefox, and Brave). Interviews are expected to take 45 minutes, and participants will be paid at a rate of $12 per hour.

3.2 Proposed Methodology

The study will be divided into two parts. First, a pre-screening survey which will evaluate candidacy for the second part, consisting of contextual interviews. During the pre-screening, we will ask questions which measure the participants' self-professed familiarity with the browsers in our sample groups. Eligible participants who pass the pre-screening will be assigned to the browser they report as being their primary browser. In lieu of this, or in instances where the quota for less commonly used browsers are unmet, participants will be assigned to the browser they are most familiar with if it is not their primary browser.

The second part of the study is the interview. Prior to the interview itself, participants will complete a brief demographic questionnaire, and will schedule their interview session. Interviews will follow a contextual interview methodology [35], presenting the assigned browser and asking questions about it through a screen-sharing session. Rather than allowing participants to use their own browser, which may be subject to myriad differences in configuration, version, and so on, we wish to control what participants see and ensure that the browser which participants see is precisely the same between subjects. We also want to ensure that the participants are able to fully interact with this browser through the remote screen-sharing. The browser will be reset to the default settings, reflecting the out of box experience as though it were freshly installed. Users may ask questions or ask to interact with all parts of the browser, such as buttons and menus. It's important that the website within the browser does not influence the users' perceptions as we want responses concerning only the browser and its settings and other interfaces, so we will only include an extremely basic blank website with the various data practices embedded.

3.2.1 Sampling Strategy

This study will use purposive sampling to ensure that we interview participants about the browser they are most familiar with, out of the browsers we are studying. We will evaluate the 5 most popular browsers. Each participant will be assigned to only one browser, which we will present to them at the beginning of the interview. Candidates for the interview will be pre-selected and assigned based on their responses to a screening survey which will evaluate their familiarity with each of the 5 browsers. Participants who are unfamiliar with all 5 browsers will be excluded, as will participants who do not disclose that they regularly browse the internet using one of these browsers. We will also exclude mobile browsers and their associated users; we are only interested in desktop browsers as they offer a different user interface which includes dashboards. Mobile browsers do not include privacy and security dashboards.

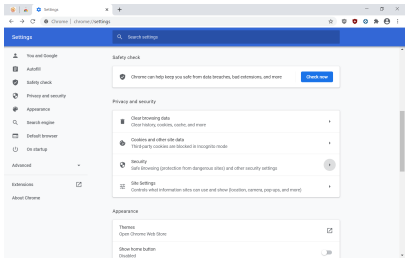
Brave browser (Figure 3.1(e) is the first browser we will study. This browser offers fine-grained information and extensive controls in both dashboards and additional settings. It has the smallest user base, and it may be difficult to find users of Brave among the general public. We may be required to recruit additional participants from the */r/bravebrowser* sub-reddit (which has approximately 3000 users) in order to fulfil our quota of interviewees for this browser.

Firefox (Figure 3.1(d) is the second browser we will study. It offers fine-grained information in dashboards, but with simplified controls. This browser has a moderately large user base, and it should not be too difficult to find users in the general public.

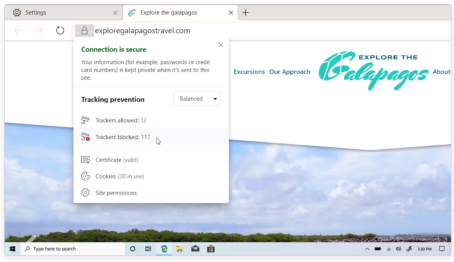
Edge (Figure 3.1(b) is the third browser we will study. This browser offers only simplified information and basic controls. It also has a small user base, and it may be somewhat difficult to find users in the general public. Like Brave, this browser may require us to recruit users from the */r/edge* sub-reddit, which has approximately 5200 users.

Safari (Figure 3.1(c) is the fourth browser we will study. It offers simplified information and no controls. It also has a very large user base, and we should have no problems finding users in the general public.

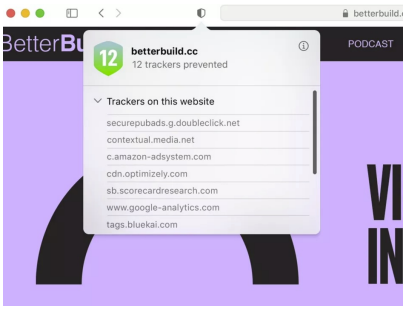
Chrome (Figure 3.1(a) is the fifth and final browser we will study. It offers no information and no controls, as it does not incorporate a privacy/security dashboard or any comprehensive settings. This browser will require a slight variation of the interview script since it does not offer dashboards for us to ask interviewees questions about. Chrome also has the largest user base, and we should have no problems finding users.



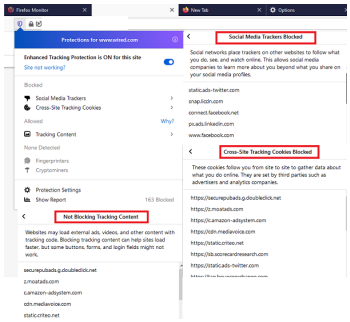
(a) Chrome does not offer any privacy or security dashboard, and only limited settings.



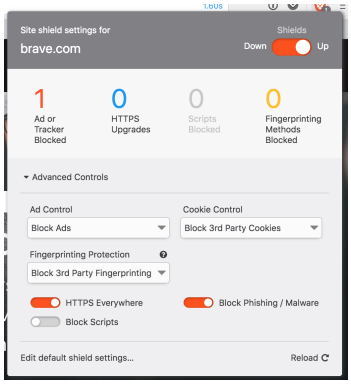
(b) Edge offers a dashboard with limited information and settings, hidden behind the HTTPS "lock" indicator.



(c) Safari offers a dashboard with no controls, but does provide reports in addition to information about blocked "trackers".



(d) Firefox offers a breakdown of data practices based on a detailed taxonomy, with simple controls.



(e) Brave has a large number of configurable options, paired with a detailed taxonomy of practices.

Figure 3.1: Screenshots of the 5 most popular browsers' privacy and security interfaces.

Recruiting

We will recruit participants primarily on CBDR, and if we wish to expand the number of participants we can send out small numbers of pre-screening surveys on MTurk and Prolific (we will ensure that these are all included on the IRB). Since it may be difficult for us to recruit users of browsers with small user bases (Brave, possibly others), we should attempt to recruit users from these browsers' sub-reddits in the event we cannot find enough users during our initial attempts. The goal is to collect data from 25 interviews, 5 per browser in total.

3.2.2 Interview Scenarios

The participant is shown a browser with a minimal fake example website. The contents of the website must not be visible (the website should not affect responses). It's crucial that the website itself does not influence responses; we only care about the browsers, not the websites. In each scenario, the participant is being subject to various practices, (e.g., fingerprinting, crypto-mining, etc. chosen at random) which are shown on the dashboard on each website. The list of practices has been reused from prior work [78], and is based on an expert-validated taxonomy that underwent several revisions through focus groups and pilot tests.

Participants are shown one website with some practices first, then another website with different practices, (all randomized, so the order of appearance of the different practices minimally affects responses) but covering the full extent of practices. Each scenario we show will incorporate half of the practices covered, so that we can have one scenario with some present and others not, and then the inverse. These scenarios will be created with full realism as we will be creating these websites and incorporating the various practices on them. The browsers will appear to users in the same way as they would if these websites were available online.

3.3 Expected Results

Prior research suggests it is possible that many users will not be able to fully understand or make use of the interfaces that browsers provide to manage privacy and security risks [78, 81], but it is unknown precisely what parts of this interaction users will have trouble with. This is in part because we expect to see that many users have never engaged with the dashboards in their browser, if they use a browser which offers one. Beyond this, it may be

reasonable to expect that the finer-grained information offered by some browsers (such as Firefox and Brave) may be more useful to some types of users, but it is difficult to predict whether this information will be overwhelming or more confusing than helpful to others. Most browsers do not offer fine-grained information about specific categories of privacy and security risks, but is more streamlined information easier to understand, or is it better to provide more detailed information? We expect that there is potentially an optimum which is somewhere in the middle of the spectrum that is currently offered. However, in contrast to browsers which offer no information, we expect to see that browsers which offer at least some details will be more effective.

When it comes to assumptions and expectations about settings to block or restrict data practices and their defaults, we expect to see a wide range of responses. This is motivated the data collected in our prior work on people's preferences and expectations while browsing [78]. However, this prior work did not study the actual interfaces provided by different browsers. Rather, users were surveyed about their preferences to use hypothetical controls which target a broad set of practices. We expect to see that when it comes to the specific behavior of the 5 browsers we study in this proposed work, users will have a variety of different mental models which differ from browser to browser.

Chapter 4

Mismatches Between Desired and Available Browser Privacy and Security Controls

4.1 Introduction

In the previous chapter, we propose an in-depth study of the effectiveness of today's browsers' privacy and security interfaces. In this chapter, we move beyond what is offered by current browsers and explore how people would ideally prefer to configure them when controlling a broad set of privacy and security risks. By analyzing people's preferences and expectations, we aim to uncover and suggest ways to reduce mismatches between what is desired and what is actually provided.

Reducing the mismatch between expectations and reality is important, because of the potential for harm. The likelihood of harm is also increasing – as techniques including machine learning, fingerprinting, profiling, and other forms of automated reasoning become increasingly pervasive, users may experience them nearly constantly during everyday Internet browsing [73]. However, the application of these techniques can often provide users with improved, safer, and more relevant online experiences [90, 85, 44, 5]. Therefore, users should be provided with controls which help them to restrict behaviors they are uncomfortable with in accordance with their preferences and tolerance for risk.

In this chapter, we refer to "potentially intrusive practices" (PIP) as common third-party tracking methods, as well as other types of malicious scripts that run in the browser to collect data, monitor activity, redirect users' attention, or operate in the background

to gather something of value. We focus specifically on 8 categories of practices that fit this definition: identity/sign-in services, targeted advertising, behavioral profiling, reporting and analytics, fingerprinting, nag screens, session replay, and crypto-mining. Each of these PIP can raise concerns associated with different dimensions of privacy captured by Solove's taxonomy [79], and have the potential to pose privacy and security risks. However, whether any of these practices are viewed as overly intrusive or risky is determined by the personal perspective of the individual – this user-centric aspect is the subject of interest in our work. In fact, these 8 PIP may be seen as valuable by some users. Generally, websites increasingly employ profiling, reporting and analytics, and session replay to improve their products and services, increase business intelligence, and capitalize upon data brokerage [33]. Many websites use nag screens, crypto-mining, or targeted advertising to highlight new features, generate revenue from monetization, or make ads more relevant [44, 5]. Sign-in services and fingerprinting are used ostensibly for user convenience and to increase security. However, PIP are increasingly ubiquitous [73], and lack transparency – many users experience annoyance, frustration, fear, and feelings of insecurity or being spied upon when they find out that they had been subjected to them (especially without their consent [6, 94, 19]).

Our work focuses on the awareness and control made available by the browser itself rather than the ever-increasing array of third-party add-ons and tools. Often, add-ons require technical expertise to install and use, and are not intended for the average user [56, 39]. Outside of this tool-centric perspective, few settings are available in browsers or on websites for users to manage PIP. Moreover, restricting PIP using mechanisms which are not explicitly supported by websites can be fragile. Websites are constantly updated, and breakage can occur when their contents are manipulated. As a result, rather than risking breakage and losing users, many browsers' default settings are limited and there is little that can be done to restrict or control PIP [28, 57]. Of the few controls which are supported explicitly on websites, many involve redirecting users through complex opt-out procedures requiring interaction with third-parties through labyrinthine external links [14].

4.1.1 Research Goal

Managing online PIP effectively is a significant problem and would benefit from user-centered research; PIP are complex, pervasive, and the extent to which users feel they have adequate awareness and control over them is unclear. However, a significant body of works has shown that users' privacy and security expectations are not currently fulfilled [67, 62, 55]. By modeling users' expectations, understanding, and preferences, we propose ways to improve the settings offered by browsers and shed light on some of the potential

implications of alternative designs.

We make the following main contributions:

1. Provide new insights into the understanding, preferences, and expectations of users toward PIP beyond the tool-centric approach seen in the prior art. Our user-centered approach should enable us to expose a variety of misunderstandings, misconceptions, and assumptions about practices on different websites (e.g., believing there is no PIP present if ads are not present).

2. Uncover ways to address participants' unfulfilled desire to be notified about and opt out of PIP across different contexts, and determine the extent to which their preferences can extend across categories of websites.

3. Find opportunities to revisit the settings that browsers make available, characterize their accuracy and user burden trade-offs, and highlight new research challenges that would need to be addressed for these settings to be better aligned with users' expectations.

4.1.2 Research Questions

This chapter is intended to address the following research questions. Each research question is focused on a particular aspect of designing interfaces for managing PIP.

RQ1 What are the signals that users rely on to determine whether they are being subjected to PIP or not during browsing? (Signals)

RQ2 What interfaces or settings do users associate with allowing or restricting PIP? (Interfaces)

RQ3 Are there PIP that users want to control (e.g. opt-in, or opt-out), and subject to what factors? (Controls)

RQ4 What are users' preferences to be notified about PIP on different types of websites? (Notifications)

RQ5 How well can the existing settings capture users' preferences, how often would they ideally need to be adjusted from the default, and what are the trade-offs associated with potential alternative settings? (Settings)

4.2 Methodology

Our study employed a mixed-methods approach, incorporating both qualitative and quantitative surveys which were administered to separate groups of participants. This way, we were able to gather a rich set of qualitative perspectives and a large quantitative dataset of preferences from participants. Both surveys are part of an IRB-approved protocol and incorporated attention check questions to ensure data quality. Our surveys were contextualized to 8 different website categories; News and Information, Entertainment and Games, Shopping, Travel, Finance, Adult, Health and Well-being, and Social Media and Blogging. We used high level categories from Alexa [10] which we believed were broadly representative, and selected the 1st (popular) and 500th (esoteric) examples from within each category. Each survey presented one category of PIP to each participant.

Qualitative surveys were intended to answer descriptive questions such as RQ1 and RQ2. These surveys underwent grounded analysis [89] to collect and categorize general themes, and use these findings to inform the design of a quantitative survey which could address statistical hypotheses. The grounded analysis results were used to discover trends in responses, find evidence of participants' assumptions, and determine their overall level of awareness and understanding of the surveyed concepts.

Quantitative surveys measured preferences to opt out of and be notified about PIP, intended to address RQ3 and RQ4. In order to determine which surveyed factors impacted participants' expressed likelihood to opt out of PIP, our quantitative survey results underwent regression analysis. Opt-out preferences from the quantitative survey were used to create simulations that tested alternative settings models, intended to address RQ5. These simulations characterized how accurately the settings could match with individuals' expressed preferences, and how many changes to the settings (within the constraints of the alternative models) would be required to bring the settings in alignment with individual preferences.

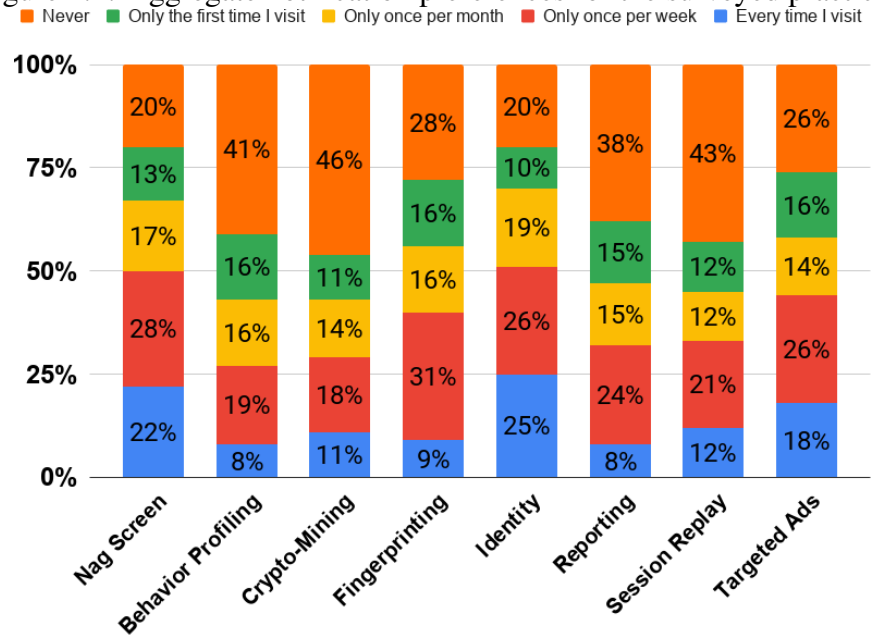
4.3 Summary of Results

Prior research focusing on specific practices and mitigation tools is consistent with our observations that most people are unaware of how to effectively identify or restrict the practices we surveyed [81, 2]. We show that people are generally unaware of the presence of intrusive practices, and don't seem to know how to detect these practices (independently of the browsers they use). In contrast to our proposed study on today's 5 most popular browsers, in this work we studied preferences which were browser-agnostic. This included

Table 4.1: Mean opt-out preferences for the surveyed practices across all website categories.

	Opt-In	**Opt-Out**
Session Replay	18%	82%
Targeted Ads	19%	81%
Behavioral Profiling	20%	80%
Crypto-Mining	14%	86%
"Nag" Screens	17%	83%
Fingerprinting	23%	77%
Identity and Sign-In Services	18%	82%
Reporting and Analytics	20%	80%

Figure 4.1: Aggregate notification preferences for the surveyed practices.



identifying controls and interfaces which people believed were associated with restricting PIP; many participants had unrealistic expectations about how their browsers and different websites gave them control over these practices. There was a clear mismatch between people's expectations and reality, particularly evident because people expressed the desire to opt out of PIP but could not. However, our user-centric approach also revealed diverse views about the perceived risks and benefits of PIP. Though a majority wished to opt out in general, there were those that had slightly different preferences depending on the category of website. Still, as can be seen in Table 4.1 and Figure 4.1, the majority of participants wished to restrict and be explicitly notified about the surveyed practices.

Table 4.2: Accuracy of the various alternative setting models.

	No Toggle	**No Toggle**	**Category Toggles**	**Category Toggles**	**Website Toggles**	**Website Toggles**
Default Setting	**Opt In**	**Opt Out**	**Opt In**	**Opt Out**	**Opt In**	**Opt Out**
Profiling	25.8%	74.2%	92.3%	92.3%	100.0%	100.0%
Reporting	27.9%	72.1%	91.9%	91.9%	100.0%	100.0%
Session Replay	24.5%	75.5%	92.7%	92.7%	100.0%	100.0%
Targeted Ads	24.6%	75.4%	90.0%	90.0%	100.0%	100.0%
Crypto-Mining	19.6%	80.4%	95.7%	95.7%	100.0%	100.0%
Identity Services	25.6%	74.4%	90.7%	90.7%	100.0%	100.0%
Fingerprinting	33.6%	66.4%	89.9%	89.9%	100.0%	100.0%
"Nag" Screens	24.5%	75.5%	92.1%	92.1%	100.0%	100.0%
Mean	25.8%	74.2%	91.9%	91.9%	100.0%	100.0%

Table 4.3: Aggregate user burden (average number of setting changes per user, per practice) associated with configuring the various alternative settings.

	No Toggle	**No Toggle**	**Category Toggles**	**Category Toggles**	**Website Toggles**	**Website Toggles**
Default Setting	**Opt In**	**Opt Out**	**Opt In**	**Opt Out**	**Opt In**	**Opt Out**
Profiling	0.00	1.00	5.32	1.45	11.87	4.13
Reporting	0.00	1.00	5.12	1.58	11.54	4.46
Session Replay	0.00	1.00	5.45	1.37	12.08	3.92
Targeted Ads	0.00	1.00	5.23	1.17	12.06	3.94
Crypto-Mining	0.00	1.00	6.09	1.23	12.86	3.14
Identity Services	0.00	1.00	5.21	1.30	11.91	4.09
Fingerprinting	0.00	1.00	4.50	1.88	10.63	5.37
"Nag" Screens	0.00	1.00	5.40	1.33	12.08	3.92
Mean	0.00	1.00	5.29	1.41	11.88	4.12

Though prior research shows that meaningful controls are rarely available [46, 14], we found that many participants mistakenly assume appropriate opt-out settings are common but are just too difficult to find. However, even if they were hypothetically available on every website, our findings suggest that pervasive opt-out settings may be more burdensome than alternative models which are contextualized to website categories instead. People's preferences are not completely uniform across categories of websites, but our analysis suggests that introducing settings which can distinguish between website categories may

help as this is a factor which was consistently found to be significant. This can be seen

Table 4.4: Regression model factors shown to significantly impact likelihood to opt out of the surveyed practices.

	Age Range	Education Level	Education Field	City Size	Marital Status	Employment Status	Employment Field	Looked at Settings	Changed Settings	Browser Used	Recent Surveys	At Risk	Website Category
Profiling	ns	ns	ns	ns	ns	ns	ns	⋆	⋆	ns	ns	ns	⋆
Reporting	ns	ns	ns	ns	ns	⋆	ns	ns	ns	⋆	ns	⋆	⋆
Session Replay	ns	ns	ns	⋆	ns	ns	ns	ns	ns	ns	⋆	⋆	⋆
Targeted Ads	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	⋆	ns	⋆
Crypto-Mining	ns	ns	ns	ns	ns	ns	⋆	ns	ns	ns	ns	ns	⋆
Identity Services	⋆	ns	ns	ns	⋆	⋆	ns	ns	⋆	ns	ns	ns	⋆
Fingerprinting	ns	⋆	⋆	ns	⋆	ns	ns	ns	ns	ns	ns	ns	⋆
"Nag" Screens	ns	ns	ns	ns	ns	ns	ns	⋆	ns	ns	ns	ns	⋆

in Table 4.4, where significant factors ($\alpha = 0.05$, $p < .05$) are starred ($\star$). Factors which were not found to be significant were marked as well (ns). Browser, gender and SA-6 were never found to significantly influence opt-out likelihood and were therefore excluded from the table.

Our simulations suggest that these settings have the potential to be far less burdensome than website-specific settings (as seen in Table 4.3), yet still retain a high degree of accuracy (as seen in Table 4.2). In these tables, the settings used in Chrome (No Toggle, opted in to all practices) are considered the default. This was the most popular browser among our respondents, as well as among the overall population. Our results show that the settings which are currently offered (block everything/nothing, or individual per-website settings) are either too coarse and do not offer the control people want, or are overly burdensome. Our qualitative study in particular shows that many people are often unaware of where the settings are available or how to use them. Together, this may explain why very few people use these settings when they are offered.

Our findings may not fully generalize across all browsers, as our random sampling method did not enable us to capture a large enough sample of each browser to maximize the likelihood of finding significant differences among this factor. However, we do not see any statistically significant differences between the different browsers, when we consider likelihood to opt out (this factor was excluded from Table 4.4 because of this).

Overall, our results argue for contextually-aware settings which can distinguish among categories where certain practices are seen as permissible, proactively notify users about their presence, and otherwise restrict intrusive practices by default. Standardizing these settings in the browser rather than on websites would have the advantage of providing a uniform interface to support awareness and control, which would be easier to configure in one place. This would have the effect of eliminating the need for users to reconfigure their settings across individual websites as they browse.

Chapter 5

Mitigating Trade-offs Between Accuracy and User Burden Using Machine Learning: A Study of Mobile App Permissions

5.1 Introduction

Not everyone feels the same way about the collection and use of their data, hence the need to provide users with privacy and security options that enable them to control data flows and ensure that these flows are aligned with their individual privacy preferences. Regulations such as the EU General Data Protection Regulation (GDPR) mandate that users be given proper control over the collection and use of their data such as securing informed consent [65]. Effectively, as data continues to be collected and used in ever more diverse ways, users are also expected to make an increasingly unrealistic number of decisions about whether to allow or restrict these practices.

A rather prominent example of the trade-off between accuracy and user burden is found in the context of mobile app privacy and security settings (referred to as *app permissions*), which allow users to control the sensitive APIs an app can access. Prompts appear when apps first request access to sensitive data categories (e.g., contacts, location, audio, calendar, camera, etc.). A permission is the ability for an app to access a specific category of data on the smartphone. Many apps ask users to grant them access to multiple permissions. On average, Android users would have to make over a hundred decisions to configure the

permission settings associated with their apps [8, 51]. It is no surprise that the vast majority of users do not take the time to configure many of these settings, even though research shows that they truly care about many of them. Indeed, many users express both surprise and discomfort when asked to take a look at what their permission settings actually allow [48, 8, 51].

Recent research has shown that, using machine learning techniques, it is often possible to predict many of people's preferences based on a relatively small number of factors such as prior permissions, decisions or answers to related questions [61, 50, 52, 51]. This approach offers the promise of helping reduce the number of decisions users have to make by possibly giving users individual recommendations on how they might want to configure their permission settings, or by possibly combining multiple closely correlated decisions for individual users. While research on how to best take advantage of these findings is still ongoing, early results involving the deployment of personalized assistants that use these models to recommend settings to users suggest that such an approach can make a big difference [51]. The question that no one has attempted to answer yet is to what extent more expressive mobile app permissions might lend themselves to the construction of preference models with greater predictive power. To what extent might these stronger predictive models help mitigate the greater user burden that would otherwise be associated with the configuration of more expressive privacy and security settings? Specifically, we focus on answering this question in the context of mobile app permissions, comparing models with permission settings that take the purpose of permissions into account versus models that do not. We present quantitative results aimed at evaluating this trade-off between accuracy and user burden across a number of parameter configurations.

5.1.1 Research Goal

Our first goal was to create a large corpus of user preferences about a variety of app permissions, for a variety of Android apps. The purpose of this corpus was to perform data mining which could potentially reveal insights into common factors along which user preferences would align. These patterns are indicative of the potential for improving predictive power. Our study sampled 5964 observations of preferences toward three sensitive Android app permissions (calendar, location, contacts), with user preferences across 108 apps, from a large sample of Android users ($n = 994$) in the United States. Having analyzed this corpus to find statistically significant factors, the next goal was to determine how to use this predictive power to improve recommendation models for preferences, leveraging profiles that incorporate a combination of supervised and unsupervised machine learning (agglomerative hierarchical clusters and conditional inference trees).

We also sought to create empirically derived guidance for the design of systems which employ these machine learning techniques to improve permissions management systems. We empirically determined the number of questions required to successfully profile users and count the instances where additional user input is required to make strong predictions. We measured the differences in efficiency and accuracy between the models which consider purpose and those which do not. We find that models which incorporate purpose make more accurate predictions, and can also reduce the overall user burden, even when compared to other similar state of the art approaches [51]. Using machine learning, our approach demonstrates that it is possible to improve the expressiveness of mobile app permissions model without trading off accuracy for user burden and vice versa.

5.1.2 Research Questions

This chapter is intended to address the following research questions. Each of these questions surrounds an aspect of improving Android permissions through the use of machine learning.

RQ1 What is the impact of purpose (and other contextual factors) on the predictive power of machine learning models for Android permission preferences?

RQ2 What effect does this predictive power have on the accuracy of recommendations made by profiles?

RQ3 Can we make better predictions without increasing user burden?

5.2 Methodology

In this work, we administered a survey which collected participants' Android permissions preferences for a variety of apps under two conditions: one with purpose-specific permissions and another with permissions that extend across all possible purposes. We analyzed responses using logistic regression and machine learning. Our aim was to discover whether machine learning could help mitigate the trade-off between accuracy and user burden when it comes to configuring Android app permissions.

Next, we applied machine learning techniques to evaluate if profile-based models could improve app permission management in terms of accuracy and user burden. We generated agglomerative hierarchical clusters for similar individuals in our data set, aggregating

their preferences into profiles. A *profile* is a model of either $(app\ category \times permission)$ recommendations or, $(app\ category \times permission \times purpose)$ recommendations. All machine learning models we tested either contain purpose-specific or purpose-independent permissions, but not both. Once an unknown individual has been matched to a profile (referred to as *profiling*), the profile can be queried for recommendations across all permissions and app categories. Conditional inference decision trees are used to perform profiling and to evaluate the number of questions needed to profile. Profiles and the decision trees used in profiling are static models. Once trained, they do not continue to learn from profiling or queries.

One way that profiles differ from traditional classifiers and recommendation systems is that in some cases profiles cannot make a recommendation for a particular permission. This can be due to sparse data or lack of consensus. Where the clusters of individuals that make up a profile have greater than a specified threshold for consensus about a preference, the profile makes a recommendation. In our study, we tested multiple thresholds between 70% and 90%. Where recommendations cannot not be made (known as *null recommendations*), we default to the original prompt where the user is directly asked whether to allow or deny a permission instead. Traditional measures of classifier performance (such as precision and recall) are limited to evaluate our techniques, since they cannot account for null recommendations. We employ two alternative measures of performance. Our measure of accuracy is the cases where recommendations are made that coincide with participants' surveyed preferences, divided by the total number of recommendations made. User burden in contrast, is the measure of individual user interactions required to both perform profiling plus the number of traditional preference elicitation prompts users would encounter in cases of null recommendations.

Profiling uses conditional inference decision trees to re-estimate the regression relationship between clusters and individual preferences. Trees are composed of unidirectional connected decision nodes based on the most statistically significant model features: app categories, demographic factors, and permissions from the design matrix used in the logistic regression analysis. The permutation tests used in the tree generation are based on Strasser and Weber's method, using Bonferroni-corrected p-values [82]. Significance is the same as the original logistic regression models ($\alpha = 0.05$). The length of the tree path traversals from root to leaf nodes are used to characterize the number of questions required to profile an unknown individual from the test data set. The decision nodes in the trees are questions that must be answered by the participant which determine which profile they should be assigned to. The answers are known *a priori* from their survey responses. The leaf nodes represent a probabilistic conclusion about which profile that the individual ought to be assigned to. By counting the number of decision nodes required to arrive at a

leaf node, we can directly observe the number of user interactions required to profile an individual.

Regardless of the number of recommendations a profile is queried about, profiling need only occur once per individual user. For any given individual's profile, the ability to make a recommendation does not change based on the number of queries it undergoes or the number of recommendations it makes over time. Profiles and the decision trees used in profiling are static. Therefore, with respect to user burden, no additional assumptions are required for our analysis or evaluation. Profiles can be queried for recommendations *ad infinitum*, and can be asked to make recommendations for an unlimited number of new apps without the need to profile individuals more than once. As such, the number of interactions required for profiling is always constant for any given individual, and user burden can only increase proportionally to the number of instances where no recommendation is made. Querying a profile about additional apps for a particular individual introduces opportunities to make more recommendations and possibly null recommendations, worsening burden.

Measuring accuracy is based on the proportion of correct recommendations, and is not sensitive to the number of user interactions. Profile accuracy (A), is given as $A = (C + null)/Q$ where C is the number of correct recommendations, $null$ is the number of instances where recommendations were not made, and Q is the number of queries for recommendations. Based on this formula, in an instance where no recommendations can be made, the accuracy is assumed to be 100%, as we must assume that the interactions that would take place in lieu of a recommendation always elicit user preferences accurately.

To simulate the accuracy of a profile when queried about an arbitrary number of apps, we must make an additional conservative assumption; that the expected accuracy of the profiles' recommendations for an arbitrary number of apps lies within the Bootstrap distribution of accuracy for our 6-app data set. We use the mean of this distribution for 6 apps when simulating querying profiles for 36 apps, for all values of k in our hyperparameter sweep. This is a reasonable assumption given that the profiles that are being queried in our simulation are the same static profiles that were trained and evaluated with 6 apps, subjected to additional queries.

Because of our limited assumptions, our analysis, simulation, and evaluation are conservative. Our results show that profiles can help mitigate the need for additional interactions by users to elicit their preferences as more apps are installed, in many circumstances. In contrast, the current permissions model in Android always requires the maximum number of additional interactions to elicit preferences for new apps, in all circumstances.

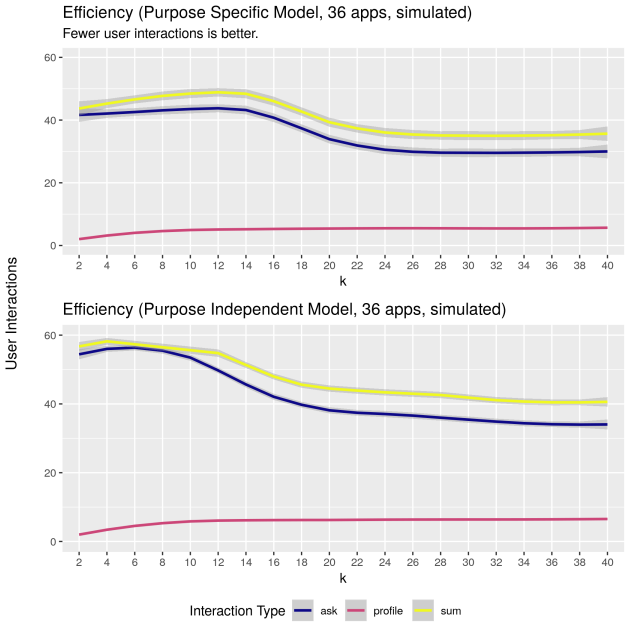
Table 5.1: ANOVA of purpose-independent regression models (Null) versus purpose-specific models.

	Contacts			
	Df	χ^2	χDf	$pr(>\chi^2)$
Null vs. Internal	57	≈ 0	0	≈ 1
Null vs. Advertisement*	57	1039.1	0	$\leq 2.2 \times 10^{-16}$
Null vs. Other/Unspec.*	57	1577.6	0	$\leq 2.2 \times 10^{-16}$
	Calendar			
Null vs. Internal	57	≈ 0	0	≈ 1
Null vs. Advertisement*	57	1292.1	0	$\leq 2.2 \times 10^{-16}$
Null vs. Other/Unspec.*	57	2025	0	$\leq 2.2 \times 10^{-16}$
	Location			
Null vs. Internal	57	≈ 0	0	≈ 1
Null vs. Advertisement*	57	1180.7	0	$\leq 2.2 \times 10^{-16}$
Null vs. Other/Unspec.*	57	1952.1	0	$\leq 2.2 \times 10^{-16}$

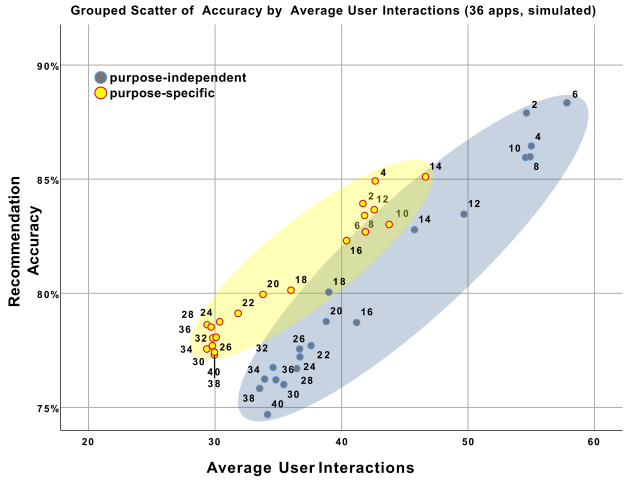
5.3 Summary of Results

Android privacy and security permissions have advantages over browser settings because they are standard, incorporating factors which allow permissions to be expressed for categories of APIs and their associated data practices, such as location access. However, adding additional factors to the permissions model to increase accuracy also increases the amount of user burden. For example, mobile app permissions could offer the ability to moderate permissions subject to purpose, but this would multiply the configuration burden by the number of specified purposes. Our results suggest that machine learning can indeed help mitigate trade-offs between accuracy and user burden. In the context of models that take the purpose of permissions into account, our study suggests that it is possible to get the "best of both worlds", namely doing a better job at accurately capturing people's preferences while simultaneously reducing the number of decisions they have to make. This is accomplished using machine learning to assign users to privacy profiles and using these profiles to infer many permissions for each user. Our results show that greater expressiveness in the settings does not have to necessarily translate into greater user burden and that machine learning can help mitigate trade-offs between user burden and accuracy. This is evident in Figure 5.1(a) and 5.1(b), which show that permissions subject to purpose are consistently less burdensome, yet are able to achieve higher accuracy.

In general, we would have expected to see similar results with other machine learning techniques (e.g. collaborative filtering techniques or techniques such as those discussed in prior work [52]). However, in examining the studied contextual factors, we found that preferences change significantly subject to the more expressive permissions which incorporate purpose. There is also evidence that participants cannot distinguish between cases where purpose is unspecified and cases where the purpose is "internal" (for the app to



(a) The number of interactions required to profile a user is static. The model which does not include purpose makes fewer recommendations, and must ask additional questions more often.



(b) The model incorporating purpose is less burdensome while providing higher accuracy.

Figure 5.1: Hyperparameter sweep for 36 apps, simulated using the Bootstrap distribution from 6 apps.

provide basic functionality). This is shown in Table 5.1, where factors which significantly differ from the null hypothesis are marked with an asterisk. These findings are consistent with prior research which shows the purpose for granting a permission has an effect on the likelihood to allow or deny it [83].

In addition, our results also strongly argue for the introduction of purpose-specific permissions in mobile operating systems such as Android and iOS. As our results show, people's privacy preferences are strongly influenced by the purpose for which permissions are requested (see Table 5.1). Regulations such as the EU GDPR further mandate obtaining consent from users for the collection of their data for specific purposes [65]. Our results further suggest that, using machine learning, interfaces could be built to mitigate the increase in user burden that would otherwise result from the introduction of purpose-specific mobile app permissions.

Chapter 6

Proposed Timeline

The proposed timeline is seen in Table 6.1. This incorporates approximately 6 weeks of slack time, in order to allow for some extra flexibility to account for unexpected issues.

For the remaining study proposed in the first technical chapter, the amount of time allocated to data collection has been based on an expected rate of 8 interviews per week, at an average of 2 interviews per day.

Table 6.1: Proposed timeline for the completion of remaining tasks.

Milestone	Expected Duration	Begin	End
Propose			April 8th
IRB for Browser Dashboard Study		March 26th	April 2nd - April 8th
PETS Rebuttals		April 7th	April 9th
Pilot for Browser Dashboard Study	2 weeks	April 2nd - April 8th	April 23rd
Data Collection for Browser Dashboard Study	3 weeks	April 23rd	May 14th
Data Analysis for Browser Dashboard Study	2 weeks	May 14th	May 28th
Reporting Results, Incorporating Into Dissertation	4-8 weeks	May 28th	July 23rd
Dissertation Slack Time	2-3 weeks	July 23rd	August 13th
Defend			August 20th
Final Revision Slack Time	2-4 weeks		September 17th

Bibliography

[1] Ruba Abu-Salma. *Designing User-Centered Privacy-Enhancing Technologies*. PhD thesis, UCL (University College London), 2020. 2.2

[2] Ruba Abu-Salma and B. Livshits. Evaluating the end-user experience of private browsing mode. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020. 2.2, 4.3

[3] Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, Norman Sadeh, Florian Schaub, Manya Sleeper, et al. Nudges for privacy and security: Understanding and assisting users' choices online. *ACM Computing Surveys (CSUR)*, 50(3):1–41, 2017. 1, 2.2, 3.1

[4] Alessandro Acquisti and Jens Grossklags. Privacy and rationality in individual decision making. *IEEE Security Privacy*, 3(1):26–33, Jan 2005. 1

[5] Alessandro Acquisti, Curtis Taylor, and Liad Wagman. The economics of privacy. *Journal of Economic Literature*, 54(2):442–92, June 2016. 1, 4.1

[6] Lalit Agarwal, Nisheeth Shrivastava, Sharad Jaiswal, and Saurabh Panjwani. Do not embarrass: Re-examining user concerns for online tracking and advertising. In *Proceedings of the Ninth Symposium on Usable Privacy and Security*, SOUPS '13, New York, NY, USA, 2013. Association for Computing Machinery. 1, 3.1, 4.1

[7] Devdatta Akhawe and Adrienne Porter Felt. Alice in warningland: A large-scale field study of browser security warning effectiveness. In *22nd USENIX Security Symposium (USENIX Security 13)*, pages 257–272, 2013. 2.2

[8] Hazim Almuhimedi, Florian Schaub, Norman Sadeh, Idris Adjerid, Alessandro Acquisti, Joshua Gluck, Lorrie Faith Cranor, and 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*, CHI '15, pages 787–796, New York, NY, USA, 2015. ACM. 3.1, 5.1

[9] Ali Alshehri, Pawel Marcinek, Abdulrahman Alzahrani, Hani Alshahrani, and Huirong Fu. Puredroid: Permission usage and risk estimation for android applications. In *Proceedings of the 2019 3rd International Conference on Information System and Data Mining*, pages 179–184, 2019. 2.3

[10] Amazon. Alexa top sites. `https://www.alexa.com/topsites`, 2020. 4.2

[11] Reyhan Aydoğan, Pinar Øzturk, and Yousef Razeghi. Negotiation for incentive driven privacy-preserving information sharing. In *International Conference on Principles and Practice of Multi-Agent Systems*, pages 486–494. Springer, 2017. 2.3

[12] Tim Baarslag, Alan Alper, Richard Gomer, Muddasser Alam, Perera Charith, Enrico Gerding, and m.c. schraefel. An automated negotiation agent for permission management. In *AAMAS 2017: Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, pages 380–390. ACM, May 2017. 2.3

[13] Rebecca Balebako, Florian Schaub, Idris Adjerid, Alessandro Acquisti, and Lorrie Cranor. The impact of timing on the salience of smartphone app privacy notices. In *Proceedings of the 5th Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices*, pages 63–74, 2015. 2.1

[14] Vinayshekhar Bannihatti Kumar, Roger Iyengar, Namita Nisal, Yuanyuan Feng, Hana Habib, Peter Story, Sushain Cherivirala, Margaret Hagan, Lorrie Cranor, Shomir Wilson, Florian Schaub, and Norman Sadeh. Finding a choice in a haystack: Automatic extraction of opt-out statements from privacy policy text. In *Proceedings of The Web Conference 2020*, WWW '20, page 1943–1954, New York, NY, USA, 2020. Association for Computing Machinery. 2.2, 4.1, 4.3

[15] Adam Barth, Anupam Datta, John C Mitchell, and Helen Nissenbaum. Privacy and contextual integrity: Framework and applications. In *2006 IEEE symposium on security and privacy (S&P'06)*, pages 15–pp. IEEE, 2006. 1, 1

[16] Michael Benisch, Patrick Gage Kelley, Norman Sadeh, and Lorrie Faith Cranor. Capturing location-privacy preferences: Quantifying accuracy and user-burden tradeoffs. *Personal Ubiquitous Comput.*, 15(7):679–694, October 2011. 1, 2.2

[17] Christoph Bier, Kay Kühne, and Jürgen Beyerer. Privacyinsight: the next generation privacy dashboard. In *Annual Privacy Forum*, pages 135–152. Springer, 2016. 2.2

[18] Johana Cabinakova, Christian Zimmermann, and Guenter Mueller. An empirical analysis of privacy dashboard acceptance: the google case. *Proceedings of the 2016 European Conference on Information Systems*, 2016. 2.2

[19] Hongliang Chen, Christopher E. Beaudoin, and Traci Hong. Securing online privacy: An empirical test on internet scam victimization, online privacy concerns, and privacy protection behaviors. *Computers in Human Behavior*, 70:291 – 302, 2017. 1, 3.1, 4.1

[20] Lorrie Cranor and Rigo Wenning. Platform for privacy preferences (p3p) project. `https://www.w3.org/P3P/`, Feb 2018. 3.1

[21] Lorrie Faith Cranor. What do they "indicate?" evaluating security and privacy indicators. *Interactions*, 13(3):45–47, 2006. 2.2

[22] Lorrie Faith Cranor. A framework for reasoning about the human in the loop. In *Proceedings of the 1st Conference on Usability, Psychology, and Security*, UPSEC'08, USA, 2008. USENIX Association. 1, 2.1

[23] Lorrie Faith Cranor. Necessary but not sufficient: Standardized mechanisms for privacy notice and choice. *JTHTL*, 10:273–308, 2012. 1, 2.1, 3.1

[24] Willem De Groef, Dominique Devriese, Nick Nikiforakis, and Frank Piessens. Flowfox: A web browser with flexible and precise information flow control. In *Proceedings of the 2012 ACM Conference on Computer and Communications Security*, CCS '12, page 748–759. Association for Computing Machinery, 2012. 2.2

[25] Nicolás E Díaz Ferreyra, Tobias Kroll, Esma Aïmeur, Stefan Stieglitz, and Maritta Heisel. Preventative nudges: Introducing risk cues for supporting online self-disclosure decisions. *Information*, 11(8):399, 2020. 2.2

[26] Joseph Dickinson. Tracking changes in browser security indicators. In *The best of ECE undergraduate research*. Illinois Digital Environment for Access to Learning and Scholarship, 2018. 2.2

[27] Pardis Emami-Naeini, Yuvraj Agarwal, Lorrie Faith Cranor, and Hanan Hibshi. Ask the experts: What should be on an iot privacy and security label? In *2020 IEEE Symposium on Security and Privacy (SP)*, pages 447–464. IEEE, 2020. 2.2

[28] Steven Englehardt and Arvind Narayanan. Online tracking: A 1-million-site measurement and analysis. In *Proceedings of the 2016 ACM SIGSAC Conference on*

Computer and Communications Security, CCS '16, page 1388–1401, New York, NY, USA, 2016. Association for Computing Machinery. 4.1

[29] Yuanyuan Feng, Yaxing Yao, and 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, page 1–15, New York, NY, USA, 2021. Association for Computing Machinery. 2.1

[30] Denis Feth, Andreas Maier, and Svenja Polst. A user-centered model for usable security and privacy. In *International Conference on Human Aspects of Information Security, Privacy, and Trust*, pages 74–89. Springer, 2017. 2.2

[31] Joshua Gluck, Florian Schaub, Amy Friedman, Hana Habib, Norman Sadeh, Lorrie Faith Cranor, and Yuvraj Agarwal. How short is too short? implications of length and framing on the effectiveness of privacy notices. In *Twelfth Symposium on Usable Privacy and Security ({SOUPS} 2016)*, pages 321–340, 2016. 2.2

[32] Google. Choose your privacy settings. `https://support.google.com/chrome/answer/114836`, 2021. 3.1

[33] Peiqing Guan and Wei Zhou. Business analytics generated data brokerage: Law, ethical and social issues. In Robin Doss, Selwyn Piramuthu, and Wei Zhou, editors, *Future Network Systems and Security*, pages 167–175, Cham, 2017. Springer International Publishing. 4.1

[34] Margaret Hagen. User-centered privacy communication design. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, 2016. 2.2

[35] Karen Holtzblatt and Sandra Jones. Conducting and analyzing a contextual interview (excerpt). In *Readings in Human–Computer Interaction*, pages 241–253. Elsevier, 1995. 3.2

[36] Corey Brian Jackson and Yang Wang. Addressing the privacy paradox through personalized privacy notifications. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, 2(2):1–25, 2018. 2.2

[37] Carlos Jensen, Colin Potts, and Christian Jensen. Privacy practices of internet users: Self-reports versus observed behavior. *International Journal of Human-Computer Studies*, 63(1-2):203–227, 2005. 2.2

[38] Laurie Kantner, Deborah Hinderer Sova, and Stephanie Rosenbaum. Alternative methods for field usability research. In *Proceedings of the 21st Annual International Conference on Documentation*, SIGDOC '03, page 68–72, New York, NY, USA, 2003. Association for Computing Machinery. 3.1.1

[39] Soroush Karami, Panagiotis Ilia, Konstantinos Solomos, and Jason Polakis. Carnus: Exploring the privacy threats of browser extension fingerprinting. In *Proceedings of the Symposium on Network and Distributed System Security (NDSS)*, 2020. 2.2, 3.1, 4.1

[40] Mark J Keith, Courtenay Maynes, Paul Benjamin Lowry, and Jeffry Babb. Privacy fatigue: The effect of privacy control complexity on consumer electronic information disclosure. In *International Conference on Information Systems (ICIS 2014), Auckland, New Zealand, December*, pages 14–17, 2014. 2.2

[41] Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W Reeder. A "nutrition label" for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, pages 1–12, 2009. 2.2

[42] Patrick Gage Kelley, Lucian Cesca, Joanna Bresee, and Lorrie Faith Cranor. Standardizing privacy notices: an online study of the nutrition label approach. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems*, pages 1573–1582, 2010. 2.2

[43] Patrick Gage Kelley, Lorrie Faith Cranor, and Norman Sadeh. Privacy as part of the app decision-making process. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 3393–3402, 2013. 2.2

[44] Eunjin Kim and Byungtae Lee. E-service quality competition through personalization under consumer privacy concerns. *Electronic Commerce Research and Applications*, 8(4):182 – 190, 2009. Special Issue: Economics and Electronic Commerce. 1, 4.1

[45] Kat Krol and Sören Preibusch. Control versus effort in privacy warnings for webforms. In *Proceedings of the 2016 ACM on Workshop on Privacy in the Electronic Society*, pages 13–23, 2016. 2.2

[46] Pedro Leon, Blase Ur, Richard Shay, Yang Wang, Rebecca Balebako, and Lorrie Cranor. Why johnny can't opt out: A usability evaluation of tools to limit online behavioral advertising. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, page 589–598, New York, NY, USA, 2012. Association for Computing Machinery. 2.2, 3.1, 4.3

[47] Tong Li, Mingyang Zhang, Hancheng Cao, Yong Li, Sasu Tarkoma, and Pan Hui. "what apps did you use?": Understanding the long-term evolution of mobile app usage. In *Proceedings of The Web Conference 2020*, pages 66–76, 2020. 2.3

[48] Jialiu Lin, Shahriyar Amini, Jason I. Hong, Norman Sadeh, Janne Lindqvist, and 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*, UbiComp '12, pages 501–510, New York, NY, USA, 2012. ACM. 1, 2.3, 5.1

[49] Jialiu Lin, Michael Benisch, Norman Sadeh, Jianwei Niu, Jason Hong, Banghui Lu, and Shaohui Guo. A comparative study of location-sharing privacy preferences in the united states and china. *Personal Ubiquitous Comput.*, 17(4):697–711, April 2013. 2.3

[50] Jialiu Lin, Bin Liu, Norman Sadeh, and 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)*, pages 199–212, Menlo Park, CA, 2014. USENIX Association. 1, 2.3, 5.1

[51] Bin Liu, Mads Schaarup Andersen, Florian Schaub, Hazim Almuhimedi, Shikun (Aerin) Zhang, Norman Sadeh, Yuvraj Agarwal, and Alessandro Acquisti. Follow my recommendations: A personalized privacy assistant for mobile app permissions. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, pages 27–41, Denver, CO, 2016. USENIX Association. 1, 2.3, 5.1, 5.1.1

[52] Bin Liu, Jialiu Lin, and Norman Sadeh. Reconciling mobile app privacy and usability on smartphones: Could user privacy profiles help? In *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14, pages 201–212, New York, NY, USA, 2014. ACM. 2.3, 5.1, 5.3

[53] Awio Web Services LLC. Web browser market share. `http://www.w3counter.com/globalstats.php?year=2021&month=1`, Jan 2021. 3.1.1

[54] Yemian Lu, Qi Li, Purui Su, Juan Pan, Jia Yan, Pengyi Zhan, and Wei Guo. A comprehensive study of permission usage on android. In *International Conference on Network and System Security*, pages 64–79. Springer, 2018. 2.3

[55] Kirsten Martin and Katie Shilton. Putting mobile application privacy in context: An empirical study of user privacy expectations for mobile devices. *The Information Society*, 32(3):200–216, 2016. 4.1.1

[56] Arunesh Mathur, Jessica Vitak, Arvind Narayanan, and Marshini Chetty. Characterizing the use of browser-based blocking extensions to prevent online tracking. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*, pages 103–116, Baltimore, MD, August 2018. USENIX Association. 2.2, 3.1, 4.1

[57] G. Merzdovnik, M. Huber, D. Buhov, N. Nikiforakis, S. Neuner, M. Schmiedecker, and E. Weippl. Block me if you can: A large-scale study of tracker-blocking tools. In *2017 IEEE European Symposium on Security and Privacy (EuroS P)*, pages 319–333, 2017. 4.1

[58] Microsoft. Microsoft edge, browsing data, and privacy. `https://support.microsoft.com/en-us/windows/microsoft-edge-browsing-data-and-privacy-bb8174ba-9d73-dcf2-9b4a-c582b4e640dd`. 3.1

[59] Heather Molyneaux, Elizabeth Stobert, Irina Kondratova, and Manon Gaudet. Security matters... until something else matters more: Security notifications on different form factors. In *International Conference on Human-Computer Interaction*, pages 189–205. Springer, 2020. 2.2

[60] Mozilla. Enhanced tracking protection in firefox. `https://support.mozilla.org/en-US/kb/enhanced-tracking-protection-firefox-desktop`, 2021. 3.1

[61] Jonathan Mugan, Tarun Sharma, and Norman Sadeh. Understandable learning of privacy preferences through default personas and suggestions. `http://reports-archive.adm.cs.cmu.edu/anon/isr2011/abstracts/11-112.html`, Aug 2011. 5.1

[62] Pardis Emami Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Faith Cranor, and Norman Sadeh. Privacy expectations and preferences in an iot world. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*, pages 399–412, 2017. 4.1.1

[63] Toru Nakamura, Shinsaku Kiyomoto, Welderufael B Tesfay, and Jetzabel Serna. Easing the burden of setting privacy preferences: A machine learning approach. In *International Conference on Information Systems Security and Privacy*, pages 44–63. Springer, 2016. 2.3

[64] Patricia A Norberg, Daniel R Horne, and David A Horne. The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of consumer affairs*, 41(1):100–126, 2007. 1

[65] Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union*, L119:1–88, May 2016. 5.1, 5.3

[66] Thanasis Petsas, Antonis Papadogiannakis, Michalis Polychronakis, Evangelos P Markatos, and Thomas Karagiannis. Rise of the planet of the apps: A systematic study of the mobile app ecosystem. In *Proceedings of the 2013 conference on Internet measurement conference*, pages 277–290, 2013. 2.3

[67] Ashwini Rao, Florian Schaub, Norman Sadeh, Alessandro Acquisti, and Ruogu Kang. Expecting the unexpected: Understanding mismatched privacy expectations online. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, pages 77–96, Denver, CO, June 2016. USENIX Association. 2.2, 3.1, 4.1.1

[68] Joel R Reidenberg, N Cameron Russell, Alexander J Callen, Sophia Qasir, and Thomas B Norton. Privacy harms and the effectiveness of the notice and choice framework. *ISJLP*, 11:485, 2015. 1

[69] Odnan Ref Sanchez, Ilaria Torre, and Bart P Knijnenburg. Semantic-based privacy settings negotiation and management. *Future Generation Computer Systems*, 111:879–898, 2020. 2.3

[70] Florian Schaub, Rebecca Balebako, and Lorrie Faith Cranor. Designing effective privacy notices and controls. *IEEE Internet Computing*, 2017. 2.1

[71] Florian Schaub, Rebecca Balebako, Adam L. Durity, and Lorrie Faith Cranor. A design space for effective privacy notices. In *Proceedings of the Eleventh USENIX Conference on Usable Privacy and Security*, SOUPS'15, pages 1–17, Berkeley, CA, USA, 2015. USENIX Association. 2.1, 3.1

[72] Florian Schaub, Aditya Marella, Pranshu Kalvani, Blase Ur, Chao Pan, Emily Forney, and Lorrie Faith Cranor. Watching them watching me: Browser extensions' impact on user privacy awareness and concern. In *NDSS workshop on usable security*, pages 1–10, 2016. 2.2

[73] Sebastian Schelter and Jérôme Kunegis. On the ubiquity of web tracking: Insights from a billion-page web crawl. *Journal of Web Science*, 4:53–66, 2018. 1, 4.1

[74] Tanusree Sharma, Hunter A Dyer, and Masooda Bashir. Enabling user-centered privacy controls for mobile applications: Covid-19 perspective. *ACM Transactions on Internet Technology (TOIT)*, 21(1):1–24, 2021. 2.3

[75] Yun Shen and Pierre-Antoine Vervier. Iot security and privacy labels. In *Privacy Technologies and Policy*, pages 136–147. Springer International Publishing, 2019. 2.2

[76] Michael Simon. Apple is removing the do not track toggle from safari, but for a good reason. `https://www.macworld.com/article/3338152/apple-safari-removing-do-not-track.html`, Feb 2019. 1

[77] Daniel Smullen, Yuanyuan Feng, Shikun Aerin Zhang, and Norman Sadeh. The best of both worlds: Mitigating trade-offs between accuracy and user burden in capturing mobile app privacy preferences. *Proceedings on Privacy Enhancing Technologies*, 2020(1):195 – 215, 01 Jan. 2020. 3.1

[78] Daniel Smullen, Yaxing Yao, Yuanyuan Feng, Norman Sadeh, Arthur Edelstein, and Rebecca Weiss. Managing potentially intrusive practices in the browser: A user-centered perspective. In *Under Review*, 2021. 3.2.2, 3.3

[79] Daniel Solove. A taxonomy of privacy. *University of Pennsylvania Law Review*, 154:477, 2005-2006. 1, 4.1

[80] B. Stanton, M. F. Theofanos, S. S. Prettyman, and S. Furman. Security fatigue. *IT Professional*, 18(5):26–32, 2016. 2.2

[81] Peter Story, Daniel Smullen, Yaxing Yao, Alessandro Acquisti, Lorrie Faith Cranor, Norman Sadeh, and Florian Schaub. Awareness, adoption, and misconceptions of web privacy tools. In *Proceedings on Privacy Enhancing Technologies Symposium (PoPETS 2021)*, volume 3, July 2021. 2.2, 3.3, 4.3

[82] Helmut Strasser and Christian Weber. On the asymptotic theory of permutation statistics. *Mathematical Methods of Statistics*, 8, 02 1970. 5.2

[83] Joshua Tan, Khanh Nguyen, Michael Theodorides, Heidi Negrón-Arroyo, Christopher Thompson, Serge Egelman, and David Wagner. The effect of developer-specified explanations for permission requests on smartphone user behavior. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 91–100, New York, NY, USA, 2014. ACM. 1, 5.3

[84] Nikolaos Tsalis, Alexios Mylonas, and Dimitris Gritzalis. An intensive analysis of security and privacy browser add-ons. In *International Conference on Risks and Security of Internet and Systems*, pages 258–273. Springer, 2015. 2.2

[85] R. Upathilake, Y. Li, and A. Matrawy. A classification of web browser fingerprinting techniques. In *7th International Conference on New Technologies, Mobility and Security (NTMS)*, pages 1–5, 2015. 1, 4.1

[86] Max Van Kleek, Ilaria Liccardi, Reuben Binns, Jun Zhao, Daniel J. Weitzner, and 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*, CHI '17, pages 5208–5220, New York, NY, USA, 2017. ACM. 2.3

[87] Anthony Vance, David Eargle, Jeffrey L Jenkins, C Brock Kirwan, and Bonnie Brinton Anderson. The fog of warnings: how non-essential notifications blur with security warnings. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, 2019. 2.2

[88] Daniel Votipka, Seth M. Rabin, Kristopher Micinski, Thomas Gilray, Michelle L. Mazurek, and Jeffrey S. Foster. User comfort with android background resource accesses in different contexts. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*, pages 235–250, Baltimore, MD, August 2018. USENIX Association. 2.3

[89] Diane Walker and Florence Myrick. Grounded theory: An exploration of process and procedure. *Qualitative Health Research*, 16(4):547–559, 2006. PMID: 16513996. 4.2

[90] R. Wang, S. Chen, and X. Wang. Signing me onto your accounts through facebook and google: A traffic-guided security study of commercially deployed single-sign-on web services. In *2012 IEEE Symposium on Security and Privacy*, pages 365–379, 2012. 1, 4.1

[91] Alan F. Westin. *Privacy and freedom*. Atheneum, 1970. 1

[92] Primal Wijesekera, Arjun Baokar, Lynn Tsai, Joel Reardon, Serge Egelman, David Wagner, and Konstantin Beznosov. The feasibility of dynamically granted permissions: Aligning mobile privacy with user preferences. In *2017 IEEE Symposium on Security and Privacy*, pages 1077–1093, May 2017. 2.3

[93] Daricia Wilkinson, Moses Namara, Karla Badillo-Urquiola, Pamela J. Wisniewski, Bart P. Knijnenburg, Xinru Page, Eran Toch, and Jen Romano-Bergstrom. Moving beyond a "one-size fits all": Exploring individual differences in privacy. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, page 1–8, New York, NY, USA, 2018. Association for Computing Machinery. 1

[94] Yixin Zou, Kevin Roundy, Acar Tamersoy, Saurabh Shintre, Johann Roturier, and Florian Schaub. Examining the adoption and abandonment of security, privacy, and identity theft protection practices. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–15, New York, NY, USA, 2020. Association for Computing Machinery. 1, 3.1, 4.1