

DESIGNING COLLECTIVE ACTION SYSTEMS FOR USER PRIVACY

A Dissertation
Presented to
The Academic Faculty

By

Yuxi Wu

In Partial Fulfillment
of the Requirements for the Degree
in the
School of Interactive Computing
College of Computing

Georgia Institute of Technology

May 2024

© Yuxi Wu 2024

DESIGNING COLLECTIVE ACTION SYSTEMS FOR USER PRIVACY

Thesis committee:

Dr. Sauvik Das
Human-Computer Interaction Institute
Carnegie Mellon University

Dr. Christopher Le Dantec
College of Computer Sciences
Northeastern University

Dr. W. Keith Edwards
School of Interactive Computing
Georgia Institute of Technology

Dr. Richmond Wong
School of Literature, Media, and Communication
Georgia Institute of Technology

Dr. Joseph Calandrino
Bureau of Consumer Protection
Federal Trade Commission

Dr. Ellen Zegura
School of Computer Science
Georgia Institute of Technology

Date approved: April 11, 2024

ACKNOWLEDGMENTS

First and foremost, I am endlessly grateful to my advisors, Sauvik Das and Keith Edwards, for recognizing and believing in my potential to begin my PhD. Your patience, sensitivity, and deep knowledge have been indispensable in my journey toward forming and establishing my identity as an academic and a researcher.

I would also like to thank the members of my thesis committee for their invaluable help in preparation of this dissertation: Joe Calandrino for helping me understand how to frame my work to have real-world impacts; Chris Le Dantec for sustaining support for my work since early on in my PhD; Richmond Wong for encouraging me to pursue creative research methods outside of my comfort zone; and Ellen Zegura for sharing both your cross-domain insights and your lab space with me.

I am filled with gratitude for the friends and colleagues at Georgia Tech who made this work possible: Grace Guo for conversations about research and life that ranked high in both quality and quantity; Anh-Ton Tran for being a co-fish-out-of-water-in-computing early on in our PhDs and a critical research sounding board; Youngwook Do for being a bottomless well of not only novel technical modalities to apply to security and privacy, but also questionable puns; Beatriz Palacios Abad for exchanging fanciful dreams of thru-hiking long-distance trails and being my productivity accountability buddy; and Sachin Pendse for championing my work ever since our first CHI paper swap years ago. And, thank you to members and affiliates of the SPUD Lab—Willie Agnew, Ezra Awumey, Sydney Bice, Isadora Krsek, Hank Lee, Jacob Logas, Kyzyl Monteiro, Stephanie Yang—for being a consistent source of collaboration, cheerleading, and sanity-checking.

Special thanks to the faculty and staff throughout the years who have made my graduate experiences more enriching. At Georgia Tech, thank you to Amy Bruckman and Beki Grinter for being on my qualifying committee, and giving me research guidance early on in my PhD. Around campus, I also greatly appreciated Rosa Arriaga and Carl DiSalvo

for always being friendly faces at TSRB, and Annie Anton and Peter Swire for providing legal and policy insights, in research and in the classroom. Immense thanks also go to Jason Hong from Carnegie Mellon University for always giving thoughtful feedback on my research and advice for communicating my academic identity; to the research support staff at both Georgia Tech and Carnegie Mellon University for helping my research go smoothly; and to faculty at the University of Chicago, where I was first exposed to the world of computing—Anne Rogers for introducing me to Python, programming, and computational thinking, and Blase Ur for welcoming me into computer science research.

I also could not have done it without my dear friends from outside of computing—Ari Anisfeld, Jasmin Dial, Ratul Esrar, Jamie Kwon, and Yuqi Zhu—who deep-dove into the world of data science, policy, politics, and research, commenced PhD journeys of their own at other institutions, engaged in radical politics, shared unserious memes, and ran with me.

Finally, thank you to my parents, my brother Ethan, and my husband Benson for their unconditional love and support throughout my life.

TABLE OF CONTENTS

Acknowledgments	iii
List of Tables	xi
List of Figures	xii
List of Acronyms	xiii
Summary	xiv
Chapter 1: Introduction	1
1.1 Research Thesis Statement	6
1.2 Research Questions	6
1.3 Document Structure	6
1.4 Contributions	7
Chapter 2: Background and Related Work	8
2.1 Pain Points of Computing Systems Supporting (Privacy) Collective Action	8
2.1.1 Privacy Collective Action	9
2.1.2 Computer-Supported Collective Action	9
2.2 Design Probes and Preliminary Prototypes for Shepherding Collectives	10

Chapter 3: Creating a Unified Voice of Privacy Concerns	12
3.1 Introduction	12
3.2 Exploratory Interview Study: Identifying Inciting Incidents	15
3.2.1 Procedure	15
3.2.2 Data Analysis	16
3.2.3 Findings	17
3.3 The Representation RQ: Find-Fix-Verify	19
3.3.1 Recruitment and Overall Procedure	21
3.3.2 Phase Procedures	23
3.3.3 Find: Aggregating User Concerns	25
3.3.4 Fix: Proposing Demands to Concerns	28
3.3.5 Verify: Prioritizing Demands	32
3.4 The Stewardship RQ: Expert Panel	35
3.4.1 Procedure and Recruitment	35
3.4.2 Findings	36
3.5 Discussion	40
3.5.1 Towards Platforms to Facilitate Grassroots Privacy Collective Action	40
3.5.2 Reformist vs. Non-Reformist Reform	42
3.5.3 The Role of Experts	42
3.5.4 Limitations	44
3.6 Conclusion	45
Chapter 4: Interpreting the Unified Voice Through a Lens of Harm	46

4.1	Introduction	46
4.2	Related Work	49
4.2.1	Online Behavioral Advertising	49
4.2.2	Harms and Socio-Technical Systems	51
4.3	Methodology	52
4.3.1	Recruitment, Ethics, and Compensation	52
4.3.2	Survey	53
4.3.3	Analysis	54
4.4	Findings	55
4.4.1	Quantitative Breakdown	56
4.4.2	Psychological Distress	56
4.4.3	Loss of Autonomy	60
4.4.4	Constriction of User Behavior	64
4.4.5	Algorithmic Marginalization and Traumatization	65
4.5	Legal Recognition and Formal Measurement of Privacy Harms	69
4.5.1	Legally Recognizing Privacy Harms	71
4.5.2	Formally Measuring Harms	72
4.5.3	Limitations	73
4.6	Conclusion	74
Chapter 5: Imagining Formal Ways to Measure and Respond to Privacy Harms		75
5.1	Introduction	76
5.2	Related Work	78

5.3	Design Context and Approach	79
5.3.1	Design Context: GoTCHas	80
5.3.2	Design Approach	82
5.3.3	Comicboard Development	84
5.3.4	Comicboard Content	85
5.4	Study Design	87
5.4.1	Recruitment, Ethics, and Compensation	87
5.4.2	Survey Instrument	87
5.4.3	Analysis	88
5.5	Study Findings	89
5.5.1	UI and UX Elements of Tool	90
5.5.2	Post-Contribution Expectations	91
5.5.3	Costs of Contributing and Volunteering	93
5.5.4	Benefits of Contributing and Volunteering	95
5.5.5	Outlook on Future of Tool and OBA	96
5.5.6	Overall Findings	98
5.6	Design Principles for Instilling Trust in GoTCHas	100
5.6.1	Visible, Upfront Benefits	101
5.6.2	Timely, Useful Feedback	102
5.6.3	Contestability	103
5.6.4	Error Prevention Measures	103
5.6.5	Integration into Everyday Life	104
5.6.6	Consideration of Social Influence	105

5.6.7	Commitment Diversity and Flexibility	106
5.7	Limitations and Future Work	106
5.7.1	Application to Non-Government Harm-Reporting Tools	107
5.7.2	Counter-Data Action	107
5.7.3	Design Fiction and Participant Outlook	108
5.7.4	Participant Sample and Study Context	108
5.8	Conclusion	109
Chapter 6: Discussion		110
6.1	Long-Standing Problematic Constants in Usable Privacy and Security . . .	111
6.1.1	P1: Expert-User Misalignment	111
6.1.2	P2: The Privacy Paradox	113
6.1.3	P3: Slow Violence	114
6.2	Remaining Challenges and Reflections	115
6.2.1	What Is Collective “Action”?	116
6.2.2	The Value of a Persistent Dataset of Privacy Harms	117
6.2.3	Co-Design, Field Deployment, and Real-World Evaluation	118
Chapter 7: Conclusion		119
Appendices		121
Appendix A:	OBA Harms Codebook	122
References		123

Vita 136

LIST OF TABLES

3.1	Demographics of Find-Fix-Verify participants. Three separate, non-overlapping sets of participants were recruited, one for each phase.	23
3.2	The final output of concerns and demands voted on by participants in the Verify phase.	34
3.3	Demographics of experts. A = academia, I = industry, G = government, L = law. *E4 did not wish to be identified.	36
4.1	Demographics of participants, broken down by whether they chose to share an account of their experiences with online behavioral advertising (OBA). .	53
4.2	Descriptions of the types of harms and their distinguishing characteristics. .	55
5.1	Demographics of Prolific participants in the main study.	89
5.2	Summary of qualitative findings as related to our research questions.	99
5.3	Initial evaluation of existing GoTCHas using our seven design principles. ✓: Satisfies this design principle. ✗: Does not address this principle at all. †: Partially satisfies this principle. *: Design principle is based on a Nielsen heuristic.	101
A.1	Codebook used in analysis of survey responses. The left column, “Initial Category”, refers to non-prescriptive categories we used as references in early discussions of the data. These groupings loosely formed a basis for the typology of harms.	122

LIST OF FIGURES

- 3.1 *The Find-Fix-Verify process for an inciting PVEI.* In the Find phase, a set of 200 participants identified concerns they had with a PVEI. In the Fix phase, another set of 100 participants picked the most pressing concerns and proposed demands to address them. In the Verify phase, a third set of 100 participants picked the most pressing remaining concerns, and ranked the corresponding demands from the previous phase. 24

- 4.1 A summary of the ways we found OBA harms manifesting in people’s lives. Aggregated, collective evidence of such experiences may help establish such harms as concrete injuries with legal standing. 70

- 5.1 The final two comicboards we developed. We refer to the first comicboard as the ”contribution” board, and the second as the “annotation” board. . . . 86

SUMMARY

People feel concerned, angry, and frustrated when subjected to data breaches, surveillance, and other privacy-violating experiences with large institutions. However, they also feel helpless to effect change. Collective action may empower groups of people affected by such experiences to jointly voice their stories of lived harm and demand redress.

In this thesis, I show that considering users' privacy concerns and lived harms on a collective level can empower users through allowing them to (1) understand they are not alone in their experiences; (2) recognize that their harms are significant and measurable; and (3) be equipped with the appropriate tools to regularly speak out about these harms. I do this through a series of work in which I create a unified collective voice of privacy concerns, interpret the unified voice in existing legal lenses of harm, and imagine formal ways to measure and respond to privacy harms. Reflecting upon my findings from this work, I discuss how the current lack of a collective action framing within the usable privacy and security field has led to the community not addressing multiple long-standing problems, and how my work can inform future directions of research in the field.

CHAPTER 1

INTRODUCTION

While many Internet users are concerned about how large institutions collect and handle their personal data, they may feel powerless to effect change. For example, prior work has shown that users express concern, anger and frustration when they encounter privacy-violating experiences with institutions (PVEIs)—be it through investigative exposés of surveillance, as in the Snowden revelations [1, 2, 3], or through personal exposure to data breaches, like the Equifax breach [4, 5]. Yet, a 2019 Pew study found that over 80% of adults in the U.S. believed that they had little or no control over the data that corporations and the government collected, and that it was impossible to go through daily life without having data about themselves collected [6]. This tension—between workaday people’s concerns over PVEIs and their perceived lack of agency to effect change—is indicative of a wider power chasm between data-aggregating institutions and the individual users whose data they collect and monetize.

How might we bridge this power chasm? One strategy that has been effective in other contexts is channeling the frustration of the dis-empowered masses into collective action—i.e., action taken by multiple people in pursuit of the same goal or collective good [7]—to demand redress. For example, in the Industrial Revolution, workers unionized, unilaterally agreeing to withhold labor from employers, tilting the balance of power toward workers and resulting in basic mainstays of modern society like minimum wages, the two-day weekend, and an 8-hour work day [8]. Importantly, prior to these worker victories, legal doctrines reinforced employer property rights over the ability of employees to organize [9]; regulatory efforts to support worker rights only came *after* sustained, collective effort. In short, history suggests that we cannot rely on existing legal structures alone to effect change in favor

of people and at the expense of powerful institutions; a sustained, united public pressure must come first.

In the context of privacy, there is some evidence that this sort of collective action can work. For example, a 2017 petition signed by California residents was the origin of today's California Consumer Protection Act (CCPA). However, the CCPA was heavily financed and driven by a small team of three individuals; the collective primarily contributed signatures necessary for a ballot measure rather than substantive policy recommendations [10]. More attempts at privacy collective action have, thus far, fallen short of effecting real change: for example, a Change.org petition responding to the Cambridge Analytica scandal garnered nearly 180,000 signatures [11], but did not result in any material redress. Other vectors for expressing collective frustrations similarly result in little material change, e.g., voicing concerns and sharing information about PVEIs on online forums. This discrepancy begs the following question that motivates my thesis: **What causes collective action efforts in privacy to fail, and how can we improve their likelihood of success?**

I argue that the usable privacy and security (UPS) research community's existing dominant approaches to tackling people's privacy problems are themselves a hindrance to successful privacy collective action. Currently, the UPS community primarily attempts to address people's individual privacy concerns through interaction nudges, educational interventions, and measuring people's (generally negative) responses to privacy violations. As a result, the community has historically failed to acknowledge and redress the massive scale of privacy *harms* that people have suffered in aggregate.

Introducing collective action as a new paradigm within UPS, where people can report on their privacy harms en masse and demand remedies or reparations, can open up promising paths forward to address these problems. If people can confirm that their privacy actions can be directly tied to real changes in policy or result in concrete reparations for their privacy harms, they might feel more motivated to act. If people can know that there are hundreds of others suffering the same psychological effects of invasive online behav-

ioral advertising as them, they might gain a sense of community and feel more encouraged to discuss their experiences publicly. And, if people are equipped with the appropriate resources and guidance to communicate effectively about their privacy harms—if they can tie specific terminology to their privacy violations beyond feeling “creeped out”—they might feel more empowered to call out these harms for what they are. If reporting privacy harms to a regulatory agency or third-party watchdog can become as mundane and ubiquitous as the privacy harms themselves, then people might actually do so.

To help explore these hypotheticals, I draw on the framework of computer-supported collective action (CSCA), proposed by Shaw et al. in 2014 [12]. In short, CSCA is collective action that is mediated by computing technology, include existing methods like using social media to spread the word and gain support, as well as other as yet unrealized designs. Shaw et al. introduced five key stages of CSCA that can help diagnose why CSCA efforts fail: many such efforts fail because they skip over requisite stages in the model. These stages include: (1) Identifying a problem; (2) Generating, debating and selecting solutions; (3) Coordinating and preparing to take action; (4) Taking action; and, (5) Following up, documenting and assessing action taken. (In the same vein, stalling and friction [13], similar to the transitions between stages of CSCA, are twin failures that keep tactical publics (i.e., collectives of users) from taking action.) In this thesis, I present a series of completed work that addresses the challenges of transitioning between these different stages of CSCA, which present in unique ways for user privacy.

First, I explore how to **create a unified voice of privacy concerns** from a large collective of users via a three-part survey instrument sensitizing probe. I found that collectives of users can easily converge on high-priority concerns and demands for redress for PVEIs and that many of their demands indicated preferences for sweeping legislative reform and formal, legal recognition of the harms suffered. Participants also felt a strong sense of empathy and solidarity with each other, even though they never directly interacted in the process. However, I also found that security and privacy experts were uniformly dismissive

of these user demands, preferring incremental measures that cleanly mapped onto existing legal structures. I refer to this work throughout the thesis as the **unified voice** project.

To better operationalize the collective's demands of apologies for wrongdoing, reparations, and recognition of harm, I **interpret the unified voice in existing legal contexts** in my second completed work. Defining privacy harms is of increasing interest in legal scholarship: whereas financial losses and physical injury can be clearly identifiable as harms in a court of law, privacy harms are less well-understood and often unrecognized. I delve into one kind of PVEI—online behavioral advertising (OBA), which in a screener survey, participants deemed most personally-violating—to understand and taxonomize specific privacy harms that people endure. To provide evidence in support of a formal legal recognition of privacy harms related to OBA, I collected hundreds of user-reported accounts of violating experiences with OBA and analyzed the different resultant harms that people reported experiencing in their day-to-day lives. I found four key harms from OBA: psychological distress, loss of autonomy, constriction of user behavior, and algorithmic marginalization and traumatization. I refer to this work as the **harms taxonomy** project.

Such a categorization can be helpful with privacy harm measurement on a wider scale, perhaps via future systems that can programmatically help people attain recognition in a legal context. I thus **imagine formal ways to measure and respond to privacy harms** from online behavioral advertising in my third piece of completed work. Taking a design fiction approach, I explore the the design requirements and cultural ideals of a government-run system that empowers people to collectively report on and make sense of experiences of privacy harm from online behavioral advertising. Study participants had detailed conceptions of the user experience of such a tool, but wanted assurance that their labor and personal data would not be exploited further by the government if they contributed evidence of harm. I refer to this work as the **harm reporting design fiction** project.

Reflecting upon my findings from these three thrusts of work, I then discuss how the current *lack* of a collective action framing within the usable privacy and security field has

led to the community missing out on the opportunity to address several long-standing UPS problems often assumed to be constant:

P1: Expert-User Misalignment A mismatch between what experts think users should do and feel is realistic in terms of change, and what users actually want to address their problems and concerns. This mismatch might be due to technical feasibility, institutional entrenchment, the nature of “interaction” design, etc.;

P2: The Privacy Paradox A frequently-observed and historically-cited phenomenon where individuals appear to act against their privacy preferences, i.e., they express concern over their privacy but do not take action to protect it [14, 15, 16, 17]; and

P3: Slow Violence The extended temporal nature of certain kinds of privacy harms that feel unimportant or difficult to prioritize individually, but that can have significant effects collectively [18].

The **unified voice** work finds evidence of **P1** (expert-user misalignments) and introduces a potential path for breaking out of **P2** (the privacy paradox). The **harms taxonomy** work demonstrates evidence of **P3** (slow violence) and proposes a potential design space for **P1**. The **harm-reporting design fiction** work presents concrete design insights for all three.

Finally, on top of helping to address these three problems, a collective action framing can also encourage us think outside of the UPS research community and consider how such systems can be valuable to multiple stakeholders. To that end, I also extrapolate design insights from my third completed work to government-supported complaint-reporting platforms in other domains, finding multiple common design gaps that might disincentivize people to report experiences of harm, be they privacy-related or otherwise. Relatedly, I note that CSCA is not a terminal model, and my work can be viewed as an initial cycle through CSCA to prepare for other kinds of collective action for user privacy later on.

1.1 Research Thesis Statement

Considering users' privacy concerns and lived harms on a collective level can empower users through allowing them to (1) understand they are not alone in their experiences; (2) recognize that their harms are significant and measurable; and (3) equip them with the appropriate tools to regularly speak out about these harms.

1.2 Research Questions

In this thesis, I will answer the following research questions:

RQ1 How can we mobilize collectives to move from amorphous discontent and anger toward specific, representative privacy demands? How do security and privacy (S&P) experts help steward these demands?

RQ2 How can identifying and taxonomizing harms from specific privacy violations help collectives and experts better coordinate their efforts?

RQ3 How can we envision a world where evidence of privacy harms is systematically collected and used successfully?

1.3 Document Structure

In Chapter 2, I briefly provide related work on why and how privacy collective action has failed in the past, as well as other design probes to shepherd collectives toward meaningful creative outputs. In Chapter 3, I show that while it is possible for a crowd of people to distill their own views about PVEIs into a representative set of demands for redress, these views don't mesh well with the opinions of security and privacy experts, who find them unrealistic in existing regulatory environments. I also demonstrate that users want recognition of and apology for privacy harms by large institutions. In Chapter 4, I describe how analyzing and taxonomizing the privacy harms that people face in one specific context—online behavioral

advertising—can help situate such harms more cognizably in a legal context. I then argue that a systematic gathering of collective evidence of these harms can be a viable course of action for users to take against privacy-violating institutions. In Chapter 5, I use design fiction methods to illustrate a potential world where such evidence-gathering systems might be successful, and present how the design insights we gained from a fictional system for user privacy can be easily applicable to other domains as well. Finally, in Chapter 6, I discuss how the work I have completed has addressed the aforementioned long-standing issues in UPS and built a foundation for other types of collective action for user privacy in future work. I conclude my thesis in Chapter 7.

1.4 Contributions

In this thesis, I make the following contributions:

- A sensitizing cultural probe into how we might scaffold users into converging on a unified voice of privacy concerns and demands, and how security and privacy experts might help or hinder that unified voice;
- A taxonomy of privacy harms arising from online behavioral advertising, such that security and privacy experts can better interpret the unified voice;
- A speculative exploration of fictional futures, in which a government-supported tool might allow people to report their privacy harms; and
- A provocation targeted at the usable privacy and security research community to consider a broader scale of privacy concerns beyond individual user interactions.

CHAPTER 2

BACKGROUND AND RELATED WORK

There has been extensive prior work documenting users' reactions to a variety of PVEIs, from discovering their information was a part of a data breach [19, 20], to learning about new regulations in the news, to perceiving output from recommendation algorithms as being too specific or creepy [21]. A study of the aftermath of the 2017 Equifax data breach illustrated that while people were aware of the risks resulting from this breach, they tended not to take protective actions because they did not know enough about the breach or because it was cost-prohibitive to do so [5]. A theme in past work is that despite their worry, fear or anger in response to these PVEIs, users tend not to take further action to protect themselves or react to the institution behind the event. Similarly, people tend to reject security advice or adopting S&P behaviors because of stigmas of doing so as being "overkill" [22, 23], feeling that it was not their job to feel responsible, since they trusted corporations to protect them from outside threats [24]. This prior work does not speak to users' sense of responsibility for protecting themselves from corporations, however. And, even if people do not feel empowered to protect their own S&P, they feel responsible for others' [25, 26].

2.1 Pain Points of Computing Systems Supporting (Privacy) Collective Action

Online petitions, perhaps best exemplified by all-purpose, all-cause websites like Change.org, MoveOn.org and Care2.com, are commonly used for collective action online, but rarely translate to real-world action. Moreover, the types of people who sign these petitions might not be representative of the broader population: less than 5% of users on Change.org accounted for over half of all signatures, and more than 99% of petitions are never marked as "victorious" [27].

2.1.1 Privacy Collective Action

In the context of privacy, when people turn to platforms like Change.org to generate petitions to make demands of institutions after PVEIs, there is no guarantee of recourse. For example, at last count, 243,900 people had signed a petition titled “Don’t let EQUIFAX escape liability!”, addressed to the Federal Trade Commission, demanding that Equifax, in the wake of their 2017 breach [4], be forced to “pay for their greed, even if it drives them into dissolution” [28]. Four years after the event, people are still signing this petition, suggesting that their anger has not abated over time, and that there have still not been adequate amends made toward those affected.

There is also little productivity on discussion forums directly related to privacy. For example, on the r/privacy subreddit, Reddit users share privacy-related news and seek advice on privacy-enhancing behaviors. However, rarely do posts there result in more than a handful of replies, much less users organizing around their privacy grievances toward a particular institution. One promising movement arose in June 2020, where a Reddit user, fed up with being “watched by cops” in the wake of the protests after the murder of George Floyd, started scraping public records to monitor police officers’ on-the-job behaviors in retaliation [29, 30]. This user garnered thousands of upvotes and comments of support, launching a new subreddit (r/DataPolice) with 7000 members dedicated to the cause, as well as a Slack workplace with more than 2000 volunteers who wanted to contribute. However, just six months later, r/DataPolice was already bereft of new posts and comments. Those who signed up inquired about progress updates on the cause’s work, to little avail.

2.1.2 Computer-Supported Collective Action

Why do so many existing collective systems and platforms for enhanced privacy protections fail? A helpful framework for answering this question is Shaw et al.’s computer-supported collective action (CSCA) [12]. Shaw describes five key patterns of CSCA that computing systems need to address to successfully support collective action—viz., (1) Identifying a

problem, (2) Generating, debating, and selecting solutions, (3) Coordinating for action, (4) Taking action, and (5) Following up and assessing action—and claims that failures in CSCA occur not only within these patterns but also in the transitions *between* them. With my co-authors, I outlined a vision for an end-to-end system—*Privacy for the People* (PftP) [31]—that spans Shaw et al.’s CSCA framework for the context of privacy, in particular.

CSCA and PftP can help diagnose why online petitions, specifically those against large data aggregators, often fail. In the Change.org petition against Equifax, for example, after signing a petition and helping bring attention to a problem, signers had no way to collectively decide which of their concerns were most important to present as a united front. Nor could they debate on the exact mechanics of what solutions they wanted from Equifax, much less unilaterally move to a platform that would facilitate this debate or coordinate further action. In other words, the petition failed at the *second stage of CSCA and PftP*—the gathering and deliberation of ideas from the collective. However, neither CSCA nor PftP directly explore how such a structured gathering of ideas might successfully be collected, debated and filtered.

My work builds on CSCA and PftP by operationalizing and evaluating a specific structured process—that I set the stage for in the following section, and detail in the next chapter—to gather, debate and filter ideas for redress across a broad collective affected by a PVEI. In so doing, I aim to address the lack of buy-in or representation people have in privacy collective action. Additionally, I explore the external resources and scaffolding necessary to make such collective action more realizable.

2.2 Design Probes and Preliminary Prototypes for Shepherding Collectives

Prior art in shepherding collectives towards meaningful creative output informs my approach. ConsiderIt [32], for example, a platform for supporting public deliberation, surfaces and summarizes pro/con statements from individuals who are broadly for or against an issue up for public debate. However, ConsiderIt is meant to be used in a consultative

manner for policy makers and experts, rather than directly represent collectives. Zhang et al. designed WeDo [33], a prototype participatory, end-to-end collective action system built on top of Twitter designed to help transition collectives through Shaw et al.'s five phases of collective action. Through a preliminary deployment and evaluation, the authors found that to improve chances of a collective action campaign at succeeding, it was critical to identify and mobilize clear leaders—otherwise, the campaign could stall without clear direction on next steps. Salehi et al. [13] discovered two oppositional challenges—stalling and friction—when designing a collective action platform to congregate crowd workers into collectives. Stalling entails a loss of momentum: a collective would form around an issue but, without any tension or clarity in driving towards consensus, would quickly disassemble without acting. Friction entails an impasse in which two or more opposing ideas lead to a break down in civil discourse and progress. To overcome these challenges, the authors recommend design considerations that help structure the collective's labor, e.g., setting clear deadlines for consensus, allowing for decisions to move forward with space for undoing if necessary, encouraging reflection and producing hope. Based on this prior art, I concluded that a collective sense-making process must be carefully scaffolded in order to assure productive forward momentum [13]. In Chapter 3, I outline the specifics of one such process we employed and studied.

CHAPTER 3

CREATING A UNIFIED VOICE OF PRIVACY CONCERNS

In this chapter, I describe my **unified voice** project, published at CHI 2022. This work answers **RQ1**: “How can we mobilize collectives to move from amorphous discontent and anger toward specific, representative privacy demands? How do S&P experts help steward these demands?” It also explores stages 1 (identifying a problem) and 2 (generating, debating, and selection solutions) of CSCA.

In this work, I used a multi-part survey instrument prototype as a design probe to explore how to steward a large collective of users with shared privacy concerns and gather representative demands from that collective. I show that collectives of users can easily converge on high-priority concerns and demands for redress for PVEIs and that many of their demands indicated preferences for sweeping legislative reform and formal, legal recognition of the harms suffered. Participants also felt a strong sense of empathy and solidarity with each other, even though they never directly interacted in the process. However, I also find that security and privacy experts were uniformly dismissive of these user demands, preferring incremental measures that cleanly mapped onto existing legal structures. In the following sections, I will discuss the project’s research questions, methods, and findings.

3.1 Introduction

Adapting Shaw et al.’s model to the context of online privacy collective action [31], I and co-authors note that existing CSCA efforts for privacy often skip stage (2): after identifying a problem, e.g., the Cambridge-Analytica scandal (stage 1), typically one or a small group of individuals draft a petition and solicit signatures (stage 3). What’s missing is a structured gathering, debate and filtering of ideas from the collective that forms around an issue. Consequently, petitions are often not representative of the collective’s concerns, nor do

they necessarily represent the best ideas of that collective. In this work, thus, we build upon our prior vision of privacy collective action—what we call *Privacy for the People* [31]—by focusing on how we might design a process that facilitates this structured gathering of ideas (stage 2). Specifically, we explore the following two research questions:

RQ1 Representation. Viral petitions authored by only a few can demotivate those who might not know or trust the original author(s), and can overshadow other ideas that better represent the collective’s demands. *How can we mobilize collectives to move from amorphous discontent and anger toward specific, representative privacy demands?*

RQ2 Stewardship. Existing structures to interpret demands into actionable redress (e.g., filing class-action lawsuits, issuing FTC penalties) often require specialized expertise or access. *How do privacy experts view the privacy demands generated by a collective, and how might they be effective stewards for the collective in translating their demands into actionable recourse?*

First, to explore the *Representation* RQ, we designed a sensitizing concept based on Bernstein’s Find-Fix-Verify (FFV) crowd programming pattern [34]. Sensitizing concepts are exemplary artifacts intended to inspire other designers to new possibilities beyond the specific artifact that was created [35], and have been employed in HCI research as probes to explore and evaluate futuristic concepts to synthesize new design knowledge [36]. We implemented our FFV sensitizing concept as a series of questionnaires designed to guide a collective, presented with an emotionally-resonant user account of a PVEI, to (1) find specific privacy concerns demonstrated in the account, (2) propose concrete fixes for these concerns, and (3) verify that the proposed fixes address emergent concerns, and prioritize the most compelling fixes. We found compelling evidence that participants emotionally connected with strangers’ accounts of PVEIs and easily converged on concerns and demands; however, they had trouble articulating concrete demands for redress. Instead, the

collective proposed and voted for broad, systemic changes that would be difficult to formalize without significant access and expertise—e.g., “data protection laws” and “rethinking the algorithm”.

To address the *Stewardship* RQ, we asked a panel of security and privacy (S&P) experts to interpret the collective’s concerns and demands. Importantly, our goal here was not to prove or disprove that our sensitizing concept “worked” but to uncover insights into how we might solicit expert stewards to help translate the collective’s high-level demands into concrete compensatory or punitive requests for recourse. Unexpectedly, we found a strong tension between what the panel deemed to be appropriate responses and what the collective desired. The panel, while sympathetic to the collective’s *frustrations*, tended to dismiss their *demands* altogether. They either preferred highly-specific, one-off penalties unrelated to the collective’s demands—e.g., fines and FTC consent decrees—or expressed that the collective’s demands were altogether unrealistic. In short, the collective wanted broad, systemic change but did not have the expertise to translate these desires into specific, actionable demands. The panel expressed little desire to steward the collective, and wanted instead to work within existing legal structures to provide one-off relief and/or punishment that was often unrelated to the collective’s demands. This disconnect suggests that S&P experts, however well-meaning, may play a role in upholding the power divide between end-users and institutions, and that there is ample room for future work in aligning the desires of non-expert collectives with the knowledge of experts.

From this work, I contribute the following:

1. A sensitizing exploration, modeled after Bernstein et al.’s Find-Fix-Verify [34], of how non-expert collectives might generate representative concerns and demands for redress in response to a PVEI.
2. An assessment of how collective-constructed concerns and demands map onto existing existing mechanisms for redress by a panel of S&P experts spanning industry, academia, government, and law.

3. A discussion of how existing mechanisms for privacy reform can be misaligned with end-user desires, along with recommendations for designing privacy collective action platforms that help alleviate this misalignment.

3.2 Exploratory Interview Study: Identifying Inciting Incidents

Our primary goal was to explore how we might design a system that facilitates the second phase of Shaw et al.’s framework for CSCA in the context of privacy collective action—the systematic gathering, refining and selection of compensatory and punitive demands in the wake of an “inciting” PVEI. We leapfrog Phase 1—the proactive identification of a problem and finding others who care—because responding to PVEIs is inherently reactive: collectives, like those who signed the petition demanding Equifax be held accountable for their data breach, naturally form *after*, e.g., the publicizing of a data breach or a media exposé. Nevertheless, a necessary precursor to Phase 2 is to find an emotionally resonant inciting PVEI that might motivate a collective to act. To find candidate “inciting” PVEIs for our sensitizing design probe, we started with an exploratory interview study to identify emotionally-resonant PVEIs, hypothesizing that PVEI accounts where individuals articulated specific emotional impact and concrete demands for redress would be good candidates for collective action.

3.2.1 Procedure

We ran semi-structured interviews with 10 participants over the BlueJeans video conferencing tool due to the COVID-19 pandemic. Our procedure was approved by an IRB. Participants were: located in the United States; active users of Internet-based services and devices; and aged between 18 to 44 years old. Half were women, and half were men. Seven had a formal computer science education or career.

We first asked participants about experiences where a large institution had handled their personal data in a way that was unexpected or violating. Then, we asked about how the

institution could remedy the situation. Many participants struggled to envision themselves making demands of the institutions, so we asked them to imagine themselves in a variety of “power roles” relative to the offending institution: (1) The participant was assigned to design a new competitor institution that had all the features, benefits, and social reach of the initial institution, but with the privacy practices that the participant wanted; (2) the participant was part of a class-action lawsuit by a third-party law firm and was asked to make a statement to the institution about their experiences and demands; and, (3) the participant wielded enough power to create legislation or regulation that would force the violating institution to comply. Then, to unpack the emotional resonance of these PVEIs, we asked participants to describe how they felt about the PVEIs by asking them to select five emotional adjectives adapted from the expanded Positive and Negative Affect Schedule (PANAS-X) [37].

Interviews lasted 30 to 60 minutes, depending on how many experiences participants felt comfortable sharing. Participants were paid 10 USD in gift cards. We promoted the study on our personal Facebook and Twitter accounts. Participants were notified in consent forms that their responses would be used to inform future design opportunities relating to user agency over personal data.

3.2.2 Data Analysis

One member of the research team transcribed the contents of the interviews. Then, through repeated discussions involving the entire research team, we performed reflexive thematic analysis [38] on the interview contents to categorize the types of PVEIs participants reported and the emotional resonance of those experiences. We also noted patterns in the types of power roles that led participants to come up with specific demands against privacy-violating institutions.

We further evaluated the types of experiences that people reported, as well as the differences in emotional reaction that these experiences elicited. We considered the types

of language participants used to talk about themselves in relation to the corporations, as well as what role they felt they could play in the situations. We organized the PANAS-X adjectives they chose to describe their experiences based on the categories outlined in the original PANAS-X manual: fear, hostility, guilt, sadness, joviality, self-assurance, attentiveness, shyness, fatigue, serenity, and surprise [37].

3.2.3 Findings

Participants reported three broad categories of PVEIs: targeting and personalization, data breaches, and surveillance. *Targeting and personalization* experiences included feeling harmed by perceived reductive profiling, or being “creeped out” by high levels of specificity in targeting. Participants who brought up *data breaches* had been victims of a wide range of breaches: e.g., Yahoo [39], Equifax [4] and OPM [40]

Participants mentioned having their personal contact information, passwords, credit card data, and financial history all violated. Participant accounts of *surveillance* included being concerned that Amazon Alexas or Google Homes were listening, or worried that Facebook was surreptitiously using the microphone on their smartphones to record their conversations and target them with ads. We took participants’ concerns at face value, and did not attempt to fact check their concerns; our goal, in this phase of our work, was only to derive a sense of which PVEIs could be emotionally-resonant enough to serve as the basis for a collective action effort.

Emotional Impact of Different PVEIs

The PVEIs that participants shared with us spanned a wide, primarily negative emotional range. Notably, data breaches incited particularly high negative emotions in participants; participants reporting these incidents frequently chose words corresponding with fear (e.g., “nervous”, “scared”) and hostility (e.g., “disgusted”, “angry”). One participant, a victim of

multiple simultaneous breaches involving the same sensitive information about him, could not identify where threats were coming from, and felt scared for his friends and family.

Instances of targeting and personalization that had materially, negatively impacted participants' lives and influenced their Internet usage also elicited strong reactions, e.g., “disgusted” and “frightened”. One participant noted that her mother was concerned about getting targeted with inappropriate ads for dating websites even though she was happily married; the participant was at a loss for how to explain the situation to her mother, who was less technically-educated. On the other hand, those who generally disliked the level of data collection needed to enable targeted advertisements and personalization conceded that they reaped small benefits from it, mirroring previous work [21]. They chose words like “amazed” or “surprised” to describe their experiences. Even when they chose negative words, participants felt the consequences of the targeting were not severe enough for them to warrant redress, and blamed themselves for not doing or knowing more. One participant admitted, *“I would like to be more informed about it. . . I just really haven't been able to do that yet, and I've already given them so much information.”*

Similarly, those who mentioned surveillance were concerned that their smart home device or smartphone was recording audio, but also felt that they could simply disable microphone permissions, turn off the devices, or get rid of the devices. Several participants felt simultaneously “afraid” of the surveillance and “amazed” at how pervasive it was. One participant noted possible malicious intent behind Google offering free Google Home devices—to collect as much home audio data as possible—and said she warned her friends not to redeem the free devices.

Conflicting Senses of Agency

Regardless of the “power role” we asked participants to imagine themselves in, they struggled to make specific demands of institutions. Participants had deeply-ingrained notions that they were themselves to blame, and were concerned about their lack of alternatives.

One participant wondered if they were really allowed to feel violated if they had consented: *“I guess I am giving my consent when I sign up for a Google account. Like it’s all in the fine print. So I almost feel like I shouldn’t feel violated because I’m consenting to these things by using a phone.”* Others felt they should have simply known better and taken better precautions or not consented, but sometimes didn’t do so anyway. In response to this phenomenon, one participant remarked, *“That cognitive dissonance makes me sad.”*

This is not to say participants felt completely defeated: they often felt responsible for protecting others, and also found solace in sharing their experiences. However, this pressure of social responsibility made them feel both more frustrated and more powerless about the privacy-violating experiences, because they could not help those with less S&P knowledge navigate these experiences. One participant admitted that even though she was not completely knowledgeable about best S&P practices, she still had to take care of her family: *“My family kind of relies on me to keep them safe, so I feel a lot of pressure that way. I know that I’m really not the best at this, but I would like to help keep them safe.”*

So, while participants often felt helpless to effect change as *individuals*, their senses of solidarity with and accountability to others could present an opportunity for collective action that is triggered from a personal narrative.

3.3 The Representation RQ: Find-Fix-Verify

The exploratory interview study helped us learn about the types of PVEIs that might inspire privacy collective action. Given this set of inciting PVEIs, to address the *Representation RQ*—How can we mobilize collectives to move from amorphous discontent and anger toward specific, representative privacy demands?—we next employed methods from concept-driven design in HCI research [41]. Specifically, we designed a collective sense-making process as a “sensitizing concept” that envisions one way privacy collectives might transition from an individual’s PVEI to collective-synthesized demands for redress. Sensitizing concepts are exemplary artifacts meant to probe a design space and inspire designers

to new possibilities in the space [35]. Importantly, the utility of a sensitizing concept is *not* the artifact itself, but the synthesis of new design knowledge from the creation and evaluation of that artifact [36].

Our sensitizing concept provides scaffolding for a collective sense-making process via three online questionnaires presented to users on Prolific, a platform where people can perform Internet-based tasks for monetary compensation. The questionnaires were modeled after Bernstein et al.’s Find-Fix-Verify (FFV) crowd programming pattern [34], which presents a distributed crowd with a high-level open-ended task (e.g., shortening a block of text), and funnels individual crowdworkers’ attention to smaller sub-tasks that, in aggregate, accomplish the high-level task. Here, the high-level open-ended task was to take an emotionally resonant account of a PVEI and have a distributed collective converge on a set of core privacy concerns and demands for redress. To that end, our three surveys presented participants with an emotionally resonant account of an inciting PVEI, and then participants had to: (1) **find** concerns they had with the account, (2) propose concrete **fixes** for these concerns, and (3) **verify** that the proposed fixes addressed emergent concerns, and prioritize the most compelling of the proposed fixes, respectively.

FFV simplifies the complex task of asynchronously distilling a unified set of demands from a distributed collective. The three stages are easily translatable to design requirements, each offering different insights about how to support the collective: whether a collective even cares about others’ concerns (**Find**), whether it can understand and provide support for itself (**Fix**), and whether it can unite (**Verify**). While other methods like focus groups provide explanatory insights about consensus-making at a small scale, our goal was to uncover insights that would more directly inform the design of a larger scale system. FFV has been shown to be effective at shepherding groups to produce high-quality outputs for open-ended creative tasks [34].

At the same time, no amount of guided scaffolding can be successful without an unifying motivation. One failure point of privacy petitions previously mentioned is that they

are ad-hoc, or only responsive to immediate events like a singular data breach; such actions tap into a collective identity of all being part of the same breach. However, a given user has been affected by multiple breaches, violating feelings from targeted advertising, and experiences of unwanted surveillance, and they connect to and support each other by sharing these experiences with each other [42, 43]. Users experience unique combinations of PVEIs, the overlaps of which create an even bigger base for collective action. Thus, in the case of end-user privacy, we believe that tapping into collective empathy with other users' personal related experiences, via Bennett and Segerberg's connective action [44], is more imperative as a motivation than membership in a single data breach's class of victims.

In this section, we will examine the results from each phase of our FFV concept, which will help address **RQ1**. Specifically, we split the problem into two sub-questions:

- **RQ1.1:** *How do people empathize with and relate to others' PVEIs?*
- **RQ1.2:** *What external stewardship is necessary to guide collectives toward concrete demands for redress?*

Answering **RQ1.1 (Empathy)** will shed light on whether an emotionally-resonant inciting PVEI can motivate someone to substantively contribute to a collective effort. Answering **RQ1.2 (Scaffolding)** will unpack the potential design requirements of an effective platform for orchestrating collectives into jointly composing privacy concerns and demands in response to an inciting PVEI.

3.3.1 Recruitment and Overall Procedure

For our FFV sensitizing concept, we picked three PVEI accounts from our exploratory interviews that had the strongest emotional resonance with our participants, hypothesizing that these accounts would be more likely to elicit empathic reactions from the collective. The accounts we selected reflected reactions to specific violations, rather than general opinions on phenomena like the “creepiness” of audio recording or aggregate data collection.

The interview participants who provided these three accounts gave high levels of detail in their answers about what specifically made them uncomfortable, and how they wished institutions had responded. Each scenario was comprised of direct quotes from respective individual participant accounts, edited for brevity.

At each phase, FFV participants were assigned to read one of the three PVEI accounts (henceforth known as either the *Equifax data breach scenario*, the *Instagram profiling scenario*, or *OPM data breach scenario*, respectively). The author of the *Equifax data breach scenario* described his experience of finding out he was affected by the breach [4] by checking Equifax’s online tool, relaying his frustration at the one year of credit monitoring and small settlement payment Equifax offered in response. The author of the *Instagram profiling scenario* described seeing recommended posts on Instagram based on perceived profiling of his identity as a gay man, and expressed discomfort with this reductionist targeting. The author of the *OPM data breach scenario* described how he felt confused and alone when he found out he was affected by the breach [45, 40], and how he wished there had been more direct followup from OPM (the U.S. Office of Personnel Management, which manages the employment data of all federal government employees). The full text of the three accounts that we chose can be found in the appendices.

Following the original Find-Fix-Verify design [34], participants were split into three independent groups across the phases. As with the crowd-powered text-processing system, Soylent, for which the FFV design pattern was originally developed, we felt that differing effort levels in participants could yield incomplete coverage of the concerns in the accounts. Thus, recruiting independent groups for each of the three phases allowed us to limit the amount of effort required from any individual contributor. (For example, if Find and Fix were combined, participants might simply choose concerns that they felt were easiest to address, rather than ones they felt were most concerning.) Recruiting three independent groups could also ensure participation from a greater number and diversity of voices

from the collective, and the separate Verify participant pool also provides an independent verification of the best fixes.

Each phase consisted of a short online questionnaire, each of which we will detail in the following subsection. A diagram showing the overall flow of the FFV process is found in Fig. 3.1. We recruited three sets of participants (200 for the Find phase, 100 for Fix, and 100 for Verify) on Prolific. Participants were over the age of 18, located in the United States, fluent in English, and active users of Internet-based services like social media, a smartphone, or a smart home device. For each phase, participants took part in a short questionnaire hosted on Qualtrics, taking on average 5 minutes, for which they were compensated 1.50 USD on the Prolific platform (the equivalent of 18 USD per hour). Participants had a mean Prolific score of 99.7. Demographics of the participants across the three phases can be found in Table 3.1.

Table 3.1: Demographics of Find-Fix-Verify participants. Three separate, non-overlapping sets of participants were recruited, one for each phase.

Phase	% Women	% Men	% Other	Mean Age	Std.Dev.
Find	47.0	52.5	0.5	29.96	9.99
Fix	61.0	38.0	1.0	29.87	9.97
Verify	48.0	52.0	0.0	30.62	10.06
All Phases	50.8	48.8	0.4	30.10	9.98

We parsed the text responses from each phase manually. To minimize researcher guidance and best mimic the hands-off nature of such a platform in the real world, we did not edit or paraphrase any concern or solution statements. We uncovered emergent patterns in each of the phases through thematic analysis [46].

3.3.2 Phase Procedures

In the **Find** phase, participants were asked to first express their emotional reaction to reading their assigned scenario by selecting two of Ekman’s six emotions [47]—afraid, angry, disgusted, sad, happy, surprised—and explaining why they had chosen those emotions. We

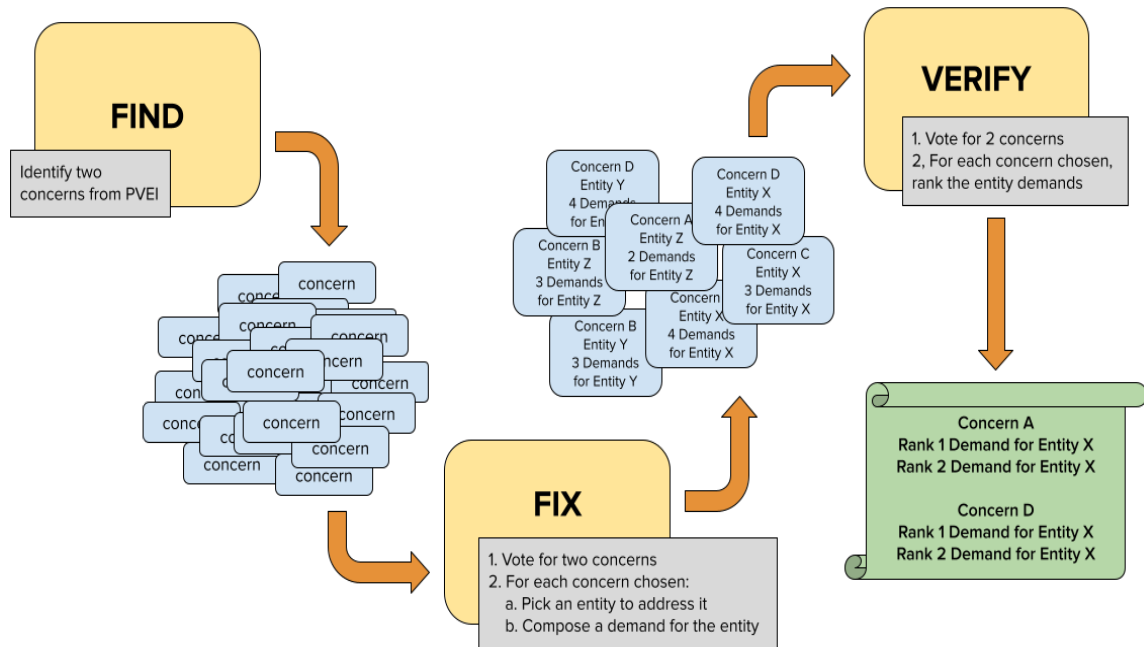


Figure 3.1: *The Find-Fix-Verify process for an inciting PVEI.*

In the Find phase, a set of 200 participants identified concerns they had with a PVEI. In the Fix phase, another set of 100 participants picked the most pressing concerns and proposed demands to address them. In the Verify phase, a third set of 100 participants picked the most pressing remaining concerns, and ranked the corresponding demands from the previous phase.

presented this question as a space for participants to first vent any feelings or biases directed toward the institutions themselves. Afterwards, we asked participants to identify two specific concerns from the passage that they found most concerning, and explain why.

In the **Fix** phase, after reading their assigned scenarios, participants were randomly presented with five concern statements (parsed from the previous stage and respective to their assigned scenario). They were then asked to pick two that they felt should be prioritized. In pilot studies, we found that participants interpreted “prioritize” to mean whatever resonated most with them. Then, for each of the two concerns they chose, we asked participants to pick one entity—out of (a) the offending institution, (b) other people who had been affected by a similar experience, (c) a government agency or regulatory body, or (d) another third party like a law firm or advocacy group—that they felt could take action to address the concern. We then asked them to detail what actions they believed this entity should take. This

simplified categorizing participants' responses, and also guided them into thinking about parties outside of only those in the passage they read. Since we also wanted to know about how users viewed their agency in these scenarios, we also asked participants how they felt after completing the survey.

Finally, in the **Verify** phase, we carried forward only concerns from the fix phase that had actions where participants had chosen the three entities with the most votes to take action. To further streamline convergence on a few concerns and recommended actions, we restricted the concerns (and corresponding actions) to only those that had gotten at least two votes. This resulted in five to 10 concerns and 10 to 20 actions per scenario. After reading their assigned scenario, participants were, similar to the fix phase, randomly presented with five concerns and asked to choose the two they felt should be prioritized. For each of the concerns they picked, we showed them all of the respective actions that fix-phase participants had authored and asked our verify-stage participants to rank-order the proposed fixes. We asked participants to keep in mind three criteria for ranking: (1) specificity, i.e., whether the action included specific tasks or steps for the entity to take; (2) effectiveness, i.e., how well the action addresses the concern; and (3) desirability, i.e., how much the participant personally wanted the action to be implemented.

3.3.3 Find: Aggregating User Concerns

Collective output

Participants indicated negative emotional reactions across the board in reaction to reading the accounts. 69% of Equifax-assigned participants felt angry or disgusted; 56% of Instagram-assigned participants felt sad or disgusted; and 45% of OPM-assigned participants felt angry or afraid.

Equifax Data Breach Concerns. Participants expressed strong disappointment and dismay at the lack of adequate compensation, remediation, and general effort on Equifax's part. Many strongly empathized with the original author's perceptions of being neglected

and brushed aside. Participants felt that a year of free credit monitoring offered by Equifax to those affected by the leak was inadequate, as was the class action lump sum payment, and also made value judgments about Equifax's perceived priorities. One participant said, for example: *"Nobody really cared about what happened; [Equifax] couldn't be bothered to actually take the time to properly structure payout and examine the issue. It was just a large chunk of money which, honestly, doesn't even really hurt these large corporations all that much."*

Instagram Profiling Concerns. Participants expressed confusion about Instagram's recommendation algorithm, as well as a lack of control over the posts they saw. Some tried to guess at the mechanisms of the recommendation algorithm, remarking on their perceived problems with how the algorithm works: *"The user is not putting every post using the hashtag 'gay' or anything related. Yet the recommended posts do not reflect what hashtags the account primarily uses."* This is not to say that participant conceptions of the Instagram recommendation algorithm were accurate, but rather that the perceived outcome drew negative reactions. Outside of algorithmic concerns, participants also worried about the possibility of context collapse—i.e., how different social contexts like the personal the professional blur together in online environments [48]—when accidentally displaying not-safe-for-work photos in a work environment. They noted that this experience could restrict how the original author uses Instagram: *"This limits when the user can use the application because it may show a NSFW post that he doesn't want someone to think he is looking at [in the workplace]."*

OPM Data Breach Concerns. While the OPM and Equifax scenarios were similar in that participants felt the punishment for OPM was insufficient, several participants directly cited OPM's status as a governmental agency, holding OPM to a higher standard of security and transparency: *"You would think that a government agency would be more equipped and transparent about things that could compromise safety and livelihood."* They also worried that the government could not be held accountable: *"The lack of options given by a govern-*

ment entity always makes me uncomfortable, because ultimately they are unaccountable to anyone and their actions are hardly scrutinized with any consequence.”

Findings

Concerns raised in this phase were abundant but similar, suggesting that a collective could empathize with an inciting PVEI and rally around common concerns. We found initial answers to both RQ1 sub-questions.

RQ1.1 Empathy. The strong emotional language participants used to detail their concerns, plus their relation of the accounts to their own experiences, suggests that participants could strongly empathize with accounts of inciting PVEIs.

For example, participants assigned the Equifax scenario felt angry for the original author and were reminded of their own anger towards Equifax as a result of the breach. Many strongly empathized with the original author’s perceptions of being neglected: *“Just a year? Credit monitoring? This is a company that does not care about people.”* Some also worried of becoming habituated to learning about data breaches like the Equifax one. For example, one participant said, *“Honestly, I had almost forgotten about the Equifax thing until I read this, which sort of scares me too. Large institutions are able to get away with near criminal behavior and we as a public just tend to brush it off after a few years.”*

Participants similarly related the Instagram scenario to their own lives, e.g., feeling *“frustrated on behalf of a gay friend who uses Instagram”*. Several directly mentioned the harmful nature of certain stereotypes about gay men: *“As someone who’s well acquainted with multiple members of the LGBT community, I fully recognize the stereotype that all gay people are driven by sexual desires, and that’s exactly what Instagram is perpetuating here.... Instagram’s algorithm simply labeled him as ‘GAY’ and shoved pictures of shirtless men at him.”*

Those assigned the OPM scenario empathized more broadly with the author’s feelings of loneliness and confusion in the middle of interpreting legal documents without help,

generalizing the author’s experiences to other data breaches. One participant related, *“I’ve gotten so many notices over the years about my info being stolen from companies, and there’s never any follow-up. Literally never. Is my data being sold on the dark web? Did they track down the thief? Did they at least improve their security practices? Who the hell knows.”* Participants also opined about the burden of reacting to a data breach, echoing past work [49]: *“Individuals are left in these situations where they are given a kind of impossible task. Sure, you can theoretically protect yourself, if you have hours and hours of your private time to spend reading fine print, and understanding complex legal things, and sitting on hold and getting transferred, and still not really getting answers.”*

RQ1.2 Scaffolding. There were hints of an answer to RQ1.2 in this phase. Participants often generalized specific concerns into broader statements about the power that institutions have to come out unscathed after negative events. No additional researcher input was necessary to guide participants into these responses, because the scenarios were emotionally resonant in and of themselves. However, due to the number of concerns brought up—two per participant, 400 total—the next challenge the collective would need to overcome would be converging on a shortlist of actionable priorities.

3.3.4 Fix: Proposing Demands to Concerns

Collective Output

Free-text responses from this phase were noticeably less impassioned and detailed than the previous phase. The level of detail in the actions that participants authored were also dependent on both which scenario they were assigned and which entity they chose to take action. We also found the entities that most participants believed should take action for each scenario. 51.5% of participants assigned the Equifax scenario believed a governmental agency or regulatory body should take action to address the concern; 63.2% assigned the Instagram scenario believed Instagram, itself, should do something; and 51.6% of participants assigned the OPM scenario believed OPM, itself, should act.

Equifax Data Breach Scenario. Even though a majority of participants wanted a governmental agency or regulatory body to take action, the level of specificity in the actions they wrote was low. A large number of people who chose a government action came up with general phrases like “Policies, fines”, “Data protection laws”, or “Anything that would punish Equifax for what was done”; very few described what kind of regulation specifically. On the other hand, participants who chose Equifax to act were very detailed about what they wanted from Equifax. For example, one participant wanted Equifax *“to be responsible for any and all identity thefts for the next 5 years as well as total cooperation and customer service and complete ownership of the problem. oh and any money made from my information goes directly to me as well as a promise that any information sold equals a fine (paid also to the victim) of 50,000 USD.”*

Instagram Profiling Scenario. Since a large majority of participants wanted action from Instagram itself, the specificity of participants’ recommended actions varied depending on the concern rather than the entity. Those who prioritized the aforementioned context collapse problem overwhelmingly wanted Instagram to implement an end-user toggle for displaying NSFW photos. Meanwhile, participants concerned with reductive inferences or a confusing algorithm tended to provide less specific responses. Several participants wanted Instagram to “fix/update/reconfigure/revamp their algorithm”. A significant minority of participants chose anyone who’d been affected by a similar experience as the entity that should take action, urging those affected to *“petition other parties and the service being used to not be reduced to a gay person, as their combined lived experiences may have more sway than a bunch of individual complaints.”*

OPM Data Breach Scenario. We defined OPM itself and the entity of “governmental agencies or regulatory bodies” to be distinct, even though OPM is a governmental agency itself. One participant wrote extensively that they wanted another governmental agency *“To fully investigate the breach, publish the investigation in easily understood prose, and follow up with every individual who had personal information compromised. The follow up should*

include information...on how the breach occurred and how it's being corrected, as well as very specific information on how the breach could affect the individual and steps to take if it did (such as fraudulent charges, sold social security number, etc) so that the affected individuals have a starting place to fix their lives and livelihoods." Broadly, participants kept in mind the long term effects of the OPM breach. One participant mused, *"If they put someone in this position, it's their responsibility to make sure it doesn't have catastrophic effects. The leak may be brief, but people can easily hold onto the info for later use."*

Findings

RQ1.1 Empathy. We uncovered three themes in how participants felt after proposing fixes to emergent concerns: heightened frustration, greater alertness about the concerns raised, and resignation at the chance to voice their demands.

Several participants felt frustrated that their contributions would be fruitless. For example, one participant said they felt *"like nothing is going to change...because the problem is bigger than one company and some data mining. It's difficult to say there should be government oversight when we all know the government is doing the exact same thing and we're most likely vulnerable there, too."* Similarly, another participant said they felt *"more outraged? I feel like, I just realized that they #1 made money off me (they should give it to me), #2 they don't have to help with the fact that they ruined my life (they should take care of me), and they still didn't come clean (if they do it again there should be a fine.)"*

Others felt that their participation reintroduced previous concerns highlighted in the inciting PVEI. For example, one participant said, *"I feel a bit better [about contributing a demand], but the guy that had his information leaked was right. I totally forgot about the equifax thing until I read that [the inciting PVEI] again."* Another noted after completing the fix phase, *"I feel like I got the chance to better organize my opinion on these concerns and have a more solid stance on the issue."* Others felt more empowered: *"I feel that I ought to get more involved in local politics to express my views more freely."* Participants

were not completely satisfied, though: *“I feel a little bit better that I got [my opinion] out, but I would prefer to see actions related to this being taken/being held accountable before I let this issue go to rest.”*

Even without being asked about the likelihood that their demands would be realized, though, participants felt pessimistic about their prospects. As one participant said, *“I feel like my opinions were valid and easily executable but will not bear any weight on how companies actually treat breaches.”* In other words, some participants felt that: (i) powerful institutions that have the ability to effect change will not listen; and, (ii) that they, themselves, have no power against these institutions. This belief, in turn, could have a negative motivational effect on their participation.

RQ1.2 Scaffolding. In a naive reading of the question, our strict rules for advancing concerns and demands to specific entities easily drilled down on numerically fewer concerns and demands. And there were indeed specific entities that participants wanted to take action more than others: for the Equifax data breach, a government agency or regulatory body; for the Instagram profiling scenario, Instagram itself; and for the OPM data breach, OPM itself. Carrying forward only the concerns and demands that (1) involved these entities respectively and (2) had been picked by at least two participants sufficiently reduced the number of choices to a manageable amount for the next stage’s participants.

A limitation of this approach is that participants in the verify phase would only see a subset of proposed fixes. The best proposed fixes might not always be associated with the most popular entity of whom the collective wishes to make demands. However, without prioritization, the collective’s attention would be spread too thin. We elected to prioritize entity popularity so that the process could be, in theory, entirely self-contained, deterministic, and not requiring moderator input.

Content-wise, participants were better at providing detailed actions for Equifax and OPM to take, such as specifying exact amounts for compensation or step-by-step timelines for them to communicate with people affected in the breaches. On the other hand, they

struggled to generalize these demands into regulations that would prevent scenarios like the inciting PVEI from recurring in the future, perhaps because average users do not have extensive knowledge of what could go into S&P regulations. In the same vein, participants might be good at identifying themes of concern—take, for example, participants in the Instagram profiling scenario wanting Instagram to “modify their algorithm”—but may require expert stewardship to translate these desires into actionable demands.

3.3.5 Verify: Prioritizing Demands

Collective Output

For each scenario, we gathered the two concerns that received the highest share of votes, as well as the two top-ranked actions for those concerns. Complete results are in Table 3.2.

For the Equifax data breach scenario, participants were most concerned with a lack of existing regulation that could hold Equifax accountable and prevent similar future incidents at other institutions. To address this, participants wanted legislation in place that would set security standards, harsher punishment for Equifax to pay more in reparations to those affected, along with investigations into who was involved with and responsible for the breach. For the Instagram profiling scenario, participants most disliked the lack of transparency around Instagram’s recommendation algorithms. Participants wanted Instagram to update their algorithm to better reflect their expectations: specifically, basing recommendations on what a user posts rather than inferences about their identity. They also wanted a formal apology from Instagram. In response to the OPM data breach scenario, participants prioritized clear communication from OPM, and were concerned that the author of the account did not know what specifically had been “messed with.” To address these concerns, they wanted explicit disclosures about the nature and extent of the data breach, detailed plans for compensation, and general transparency from OPM.

Findings

In this phase, we did not solicit answers from participants about their emotional motivation, but rather aimed to converge on the most popular demands, i.e., answering **RQ1.2 (Scaffolding)**.

We started from 400 concerns across the three scenarios. Fix-phase participants assigned accountable entities and demands to these concerns. We first isolated only concerns from the most popular entities, resulting in 110 concerns and demands; we further advanced only concerns that had at least two votes from the fix-phase, for 23 concerns and 49 demands in the verify phase. Verify-phase participants then condensed these further into 6 concerns and 12 demands. The collective effectively synthesized a total of 400 concerns down to six. (Again, a diagram of this FFV process can be found in Fig. 3.1.) Since participants were faced with a limited set of choices, they could converge on a small set of popular concerns and demands. Both the most popular concerns and most popular demands from the end of this stage were specific and impassioned. While we did not solicit any free-text responses from participants in this phase, we can surmise that participants heeded our instructions and took into consideration the specificity, effectiveness, and desirability of the demands based on the demands they ultimately selected. Indeed, demands containing low-effort text such as “Investigations and policies to prevent” (Equifax data breach scenario), “Improve their software” (Instagram profiling scenario), and “Be upfront” (OPM data breach scenario) did not receive many votes.

Overall, the compensatory demands that participants selected suggested the desire for broad, systemic changes but were low on implementation details. Indeed, participants wanted regulatory reform or algorithmic changes, but, unsurprisingly, could not (or did not) specifically articulate what would go into these regulations or what changes to make. Note that this lack of specificity could be because of the lack of expert knowledge, but it could also be because we recruited participants to complete a short compensated

survey. In practice, self-motivated, self-organized collectives may also be more motivated to be detailed in their responses.

Still, to bridge this gap, we might consider enrolling S&P experts to act as stewards for the collective—i.e., to interpret or translate the collective’s voice into actionable steps for the real world. We explore this possibility with an expert panel.

Table 3.2: The final output of concerns and demands voted on by participants in the Verify phase.

Scenario	Concern	Rank 1 Solution	Rank 2 Solution
Equifax	“How can a company continue operation as normal after such a major security event? How does our legislation let something as serious as its citizens’ identities being leaked pass without reform? I feel the entire situation described was a high level case of lunacy, and I feel I’m probably not the only one who feels that way.”	“Regulations and legislation addressing the larger issues that allowed the breach and mistreatment of customers to happen in the first place. Coming down hard on Equifax and requiring them to increase security and paying reparations to those they harmed.”	“I would want to see an investigation into the persons who perpetrated the act as well as the security measures that should have been taken. I would like to know why the data wasn’t so secure.”
	“There was no clarity on what has changed to prevent the same thing happening in the future, so this will continue to happen and we will suffer for it.”	“I want the government to try to set better standards and regulations to prevent something like this happening again and to minimize the damage if it happens again”	“Laws and/or regulations to prevent this type of data breach from happening again.”
Instagram	“It directly opposes the idea of an algorithm that Instagram uses to recommend posts. Logically, if most of his posts are about baking and cooking, that would be what his recommendations fill up with. However, the recommended posts have nothing to do with that.”	“Update their supposed algorithms so that recommended posts are related to most of the content posted by the person.”	“Work on their algorithm and address why this is happening with a formal apology”
	“I would guess there is an algorithm to blame. But no one should have to deal with that reduction of their identity.”	“I would like them to make sure the algorithm is fixed to work better and see if they are responsible or the consumer. Being more transparent about how the algorithm works could also help.”	“I would want them to improve their algorithm so that this does not continue to happen to this man, or anyone else.”
OPM	“The fact that [OPM] didn’t explain what happened is terrible. I would want to know why my info got messed with, what exactly got messed with, what problems I can expect, what I can do to save myself.”	“I would want OPM to be thorough and transparent in their communications. It is their responsibility to inform the people affected with all pertinent information”	“The employer should be mandated by law to make specific disclosures about the nature and extent of the data breach and should be required to pay for an independent review of its IT security policy and procedures”
	“After a large data breach like this, I would expect the company to detail out to all users of their service how they have updated their security protocols, and how they plan on making things right and better for me and the pain that I was sent through because of their mistake.”	“OPM would be detailing out all of their mistakes and processes to rectify these mistakes. Detailing plans for compensation and timeline for this. Additional regulatory process in place that would catch future issues like this.”	“I would want OPM to revise their data security protocols and communicate any changes to all customers. They owe people an apology and a better compensation that the 5 year protection plan that people didn’t ask for.”

3.4 The Stewardship RQ: Expert Panel

Participants who used our FFV prototype, henceforth known as “participants” or “the collective”, converged on a few salient concerns and demands for redress. However, to effect real change, these demands must be presented clearly to the institutions accountable for compensatory action (be it a regulatory body or the offending institution itself). Our next research question (**RQ2**), thus, focuses on understanding the role of experts in representing and guiding the collective: *How do privacy experts view the privacy demands generated by a collective, and how might they be effective stewards for the collective in translating their demands into actionable recourse?* To answer this question, we asked a panel of security and privacy (S&P) experts to evaluate the collective’s top-ranked concerns and solutions.

3.4.1 Procedure and Recruitment

We recruited eight non-compensated experts in privacy and security whose expertise and experience spanned academia, government, law, and industry, by reaching out to them directly through Twitter, email, and LinkedIn. We reached out to five more experts, but, as might be expected, many experts were oversubscribed and unable to commit. Demographic details of those who participated are found in Table 3.3. All experts had completed at least a bachelor’s degree in computer science, and all were based in the United States.

Three experts participated in a 30-minute interview conducted over BlueJeans, and five filled out a survey of free-text responses hosted on Qualtrics. We offered all experts the opportunity to participate in the interview, but most preferred the asynchronous questionnaire. Experts were assigned the set of top-ranked demands from Table 3.2 for the scenario most closely aligned with their expertise and asked to translate those demands into tangible tasks for entities in the real world, as well as assess the impact of these demands. They were also asked if they agreed with the rankings that the verify-phase participants had decided on, and whether they thought there were more appropriate solutions.

Table 3.3: Demographics of experts. A = academia, I = industry, G = government, L = law.
 *E4 did not wish to be identified.

Expert	Gender	Age	S&P Experience
E1	Female	45-54	A, I, G, L
E2	Male	35-44	A
E3	Female	45-54	A, I, G
E4	*	*	I
E5	Male	25-34	A
E6	Female	25-34	A, I
E7	Female	25-34	A, I
E8	Male	35-44	I

3.4.2 Findings

Expert responses can be broken down into two themes: (i) alignment and (ii) misalignment between experts and the collective. First, both experts and the collective recognized the harmful effects of the violations described in the three scenarios: for the most part, experts empathized with the collective’s sense of helplessness, even if they were more informed about the technicalities behind data breaches and algorithmic personalization. Second, experts dismissed the punitive demands our participants wanted offending institutions to take. For example, experts felt that quantifying harm and attributing blame in these scenarios would be difficult and unrealistic, and that the collective-generated demands could result in unintended negative consequences.

Expert-Collective Alignment.

The panel agreed that several participant concerns and demands were pressing to address and appropriate to enact, respectively.

For the Equifax scenario, while most of the panel felt that generalized laws and regulations would be too slow to implement, several felt that FTC consent decrees, where an institution would be legally obligated to abide by certain terms and regulations or pay a heavy penalty, could help ensure good behavior from the institution for a set period of time. Some also agreed that setting federal-level legal standards for security best practices

and strict punishments for non-compliance would help prevent this from happening in the future. Other experts were less optimistic that these legal standards could be established in a manner that benefits consumers: E7 conceded, *“I think meaningful regulatory reform is unlikely to happen as long as the interest of policymakers and corporate are deeply intertwined.”*

The Instagram profiling scenario drew mixed reactions across the panel. While they largely agreed that the experience that the author of the account had was unfortunate, the panel also expressed that it was unlikely to be intentional on Instagram’s part and tried to guess at possible explanations that led to the PVEI described in the scenario. E4 surmised that while the author might not search for pictures of half-naked gay men himself, Instagram might infer that his identity is similar on average to other people who do want this behavior. E1 was curious if the author had searched for these pictures once but had simply forgotten.

The OPM data breach also drew mixed reactions. E1, who was also affected in the same breach, suggested that all of the solutions that people were demanding from OPM—transparent communications about who is affected, plans for compensation, strategies for rectifying the mistake—had likely already been implemented. They argued that it was the responsibility of the author of the inciting PVEI to keep up-to-date with the communications that OPM sent out in the aftermath. E2 rebutted: *“The information might exist, but people are not necessarily seeking it out, or it’s not directly presented to them in ways that are actually meaningful to them. I’m sure there’s like a 50-page report out there about what went wrong, but most people won’t see that or look at that.”* Expert stewards may be helpful, thus, in not just refining demands but also in pointing collectives towards sources of information that may help address their concerns.

In general, the panel sympathized with the powerlessness that the collective felt with regard to the two data breach scenarios. E2 mentioned that because Equifax did not require consumer consent to collect their data, and it was impossible for federal employees to avoid

interacting with OPM, average people could not speak out or demand more transparency from either institution. They added, “*OPM is basically your employer or entity that manages your employee data. So what are you going to do, quit because of this breach? It’s not like they have a heightened interest in being more transparent or forthright.*” E7 suggested that users could file complaints with the FTC, but admitted that this was only for the “*highly motivated... I know that’s a lot of burden on consumers when they already have limited time and emotional distress to deal with, but unfortunately, that’s the reality we need to work with.*”

Expert-Collective Misalignment: Unknown Harms and Unintended Consequences.

While experts were sympathetic to the collective’s anger and frustration, they also dismissed the collective-generated demands as infeasible or having the potential to cause unintended negative effects. For example, one peripheral demand for both the Equifax and OPM breaches was providing compensation for victims. Experts felt that calculating a specific amount to pay in damages would be difficult: from a legal standpoint, compensation must be commensurate with tangible, proven harm, but the extent of the harm caused by both breaches is both unknown and ongoing. Experts also felt it would be too difficult to attribute all harm that victims endured directly to Equifax or OPM because the breach happened due to negligence and not malicious intent. Similarly, in the Instagram scenario, experts noted that because there was no evidence Instagram directly intended to make the author of the account feel targeted or marginalized, there was little recourse to be had.

The panel also noted that collective-requested solutions may have undesirable and unintended consequences. For example, in the *Instagram profiling scenario*, FFV participants wanted a “formal apology” from Instagram to the author of the account. The panel felt that there were a few things that could inhibit Instagram from doing so: (1) There would have to be a specific line of communication between Instagram and the author; (2) Instagram would have to issue equivalent apologies every time something like this happened. E4 of-

ferred a worst case version of the formal apology solution: to keep up with all the feedback and adverse effects of algorithmic profiling, Instagram might even resort to algorithmically predicting ahead of time if users will negatively react to a targeted post.

The collective also wanted Instagram to modify their algorithm to recommend posts only based on what users post, rather than based on Instagram’s inferences of user identities. E4 felt that this solution could have negative effects: *“What people like to read and what people like to post are very, very different. Doing this [solution] means I am limited to seeing the things I can talk about. I think this can be a pretty dangerous thing. But on a more innocuous level, I like to read recipes, but I don’t like to post them because nobody wants to eat my food.”* E8 added that even if such changes could work, their effects would not be permanent: *“Realistically, those algorithms are tuned for engagement and if proposed changes would result in a drop in that metric I’d expect that, over time, the same issue would resurface again.”* E5 said bluntly, *“I am not quite sure how willing a platform designer will be to so drastically modify their system.”*

Summary

Our findings for **RQ2**—i.e., understanding the role of experts in guiding the collective—were multi-faceted. The expert panel empathized with the collective’s frustration and desire for broad change, but only dismissed their demands as unrealistic rather than conceive of alternative approaches that might still capture the spirit of participants’ demands. We hypothesize that this dismissal was partially due to the panel’s depth of knowledge about the existing legal and technical structures that could be utilized to effect change: the panel discussed how realizing collective demands would be difficult to implement (e.g., quantifying harm for reparations) or could have negative consequences (e.g., facilitating filter bubbles). Instead, the panel tended to favor incremental reform or punitive measures compatible with existing legal structures or even inaction, in spite of the emotionally resonant frustration underlying collective concerns with the inciting PVEIs. Indeed, we observed

the panel defending the offending institution and/or absolving it of responsibility owing to a presumption of good intentions. This misalignment between people and experts could contribute to people’s impression that they are unable to effect change, despite their privacy concerns [6].

3.5 Discussion

Today, there is a wide power chasm between individuals and the data harvesting institutions who collect, process and monetize their personal data. With our FFV sensitizing concept, we envisioned a future in which people can come together with one unifying voice to demand change when these institutions commit egregious privacy violations. Such a future is not necessarily far-fetched: as we found in our evaluation of our concept, a collective can empathize and advocate for strangers and rapidly converge on a small set of concerns and compensatory action; however, they require expert stewardship to translate their desires for broad, systemic change into actionable steps. On the other hand, S&P experts, while sympathetic to the sentiment of the collective, were generally pessimistic of the collective’s desire for systemic change. They preferred, instead, working within the system for incremental reform.

3.5.1 Towards Platforms to Facilitate Grassroots Privacy Collective Action

Prior work by Shaw et al. [12] has unpacked models of online collective action and proposed stages of computer-supported collective action (CSCA) that align with our findings. Our artifact focuses primarily on the second stage—*Generate and debate ideas*—of Shaw’s model, but also speaks to the first stage—*Identifying a problem*. However, there are still three other phases in the model: coordinating and preparing to take action, actually taking the action, and reflection after the action is taken. Das et al. scoped out a vision for future work spanning *all* of the CSCA phases in the context of end-user privacy [31], and Vincent et al. proposed a framework for some of the types of actions that users can take

as leverage against privacy-violating institutions [50]. Some tools for facilitating subversive privacy collective action already exist, but are not yet directly integrated into broader, coordinated efforts: e.g., AdNauseum [51] aims to protect users from tracking advertisers by silently clicking on blocked ads to send noisy data back to advertisers; similarly, TrackMeNot sends “ghost queries” to search engines to obfuscate users’ actual searches [52]. We thus envision a rich future design landscape that includes systems or processes that help people coordinate in a more sophisticated manner than a series of surveys; that allow people to effectively take leveraging action against PVEIs; and that clearly communicate the progress that they have made.

An open question for future work is how to move from our sensitizing FFV to a self-contained system through which collectives can effectively act in the real world. From our sensitizing concept, we saw the utility of canvassing collectives to take an inciting PVEI and giving them a platform to collectively compose demands for redress. Participants felt clearer about their own stances on PVEIs after being asked to extract concerns from them, more alert about the PVEIs in their own lives, and more incensed to take action of their own. We also witnessed a shared sense of social responsibility that people had for injustices faced by strangers.

Work by Abebe [53] has also examined broader trends of the role of computing in social change, which can be used to understand the potential for future online platforms for collective action. More specifically, our work fits several roles of computing in social change categorized by Abebe [53]. Using Abebe’s terms, as a *diagnostic*, a future platform that allows users to voice their concerns about institutional privacy violations would help us measure and understand what the public wants. As a *formalizer*, such a platform can concretize these concerns via collective-powered sensemaking, be it through voting, external governance, or sophisticated topic modelling. As a *rebuttal*, it highlights the growing gap between individual users and institutional priorities. And as *synecdoche*, it exposes glaring

tensions of reform versus revolution, with security experts who have a stake in upholding existing institutions on one side, and users on the other, respectively.

3.5.2 Reformist vs. Non-Reformist Reform

The philosopher Andre Gorz made a distinction between reformist and non-reformist reforms [54]. Reformist reform is the incremental updating of existing structures and is pursued with no intention of ultimately modifying the structure of society and institutions, and instead aims to keep a calm status quo. Many of the solutions raised by experts, which included FTC consent decrees, heavier fines, and federal best-practice standards, fall into the broad category of reformist reform. Such solutions bolster the punitive power of existing regulatory institutions (e.g., the FTC) or encourage institutions that are responsible for PVEIs to develop methods to circumvent fines (e.g., Equifax), rather than tip the power imbalance in favor of the people.

Non-reformist reform, in contrast, challenges entrenched power structures. It originates “not in terms of what is possible within the framework of a given system and administration, but in view of what should be made possible in terms of human needs and demands” [54]. As an example, some participants (and even one expert) were deeply unhappy with the lack of consent to financial surveillance by Equifax, and worried that lack of regulation meant that other corporations could also misbehave with little recourse. In response, they wished for reparations bequeathed directly to those who were affected, rather than fines collected by a government entity. In short, even if they personally held grander desires for systemic change in privacy, experts’ assessments of collective demands suggests that while the people want non-reformist reform, experts view only reformist reform as tenable.

3.5.3 The Role of Experts

Perhaps owing to this disconnect, some experts explicitly said they did not want to get involved with stewarding collectives or interpreting their desires owing to the perceived

lack of knowledge non-experts had about “how the world works”. Others had trouble seeing outside of the context of existing frameworks of heavy fines or legal settlements. Our findings illustrated a misalignment between non-expert end-users and S&P experts in their demands for how institutions can collect and manage personal data. Many S&P experts are themselves embedded in institutions that are responsible for PVEIs, even if their goal is to effect change from within. Others are experts precisely because of their in-depth knowledge of existing systems. Perhaps due to their immersion in these institutions, we found that the experts we interviewed tended to be more dismissive of general demands for sweeping action or large changes.

As Rahwan argues, however, it is impossible for any one person to be fully informed on all aspects of some policy question: public opinion, which shapes social norms and morals, should be used as a check on the “sovereign force” of experts, and influence the metrics by which expert performance is evaluated [55]. Extending calls from prior work on the privilege that researchers hold in designing for vulnerable populations [56, 57], we implore S&P experts who work on improving user privacy protections to consider the role they play in bolstering the power chasm between institutions and the individuals whose personal data they exploit.

Are we working towards reform or revolution—and how does that orientation align with the objectives of those for whom we advocate or design? We, as “S&P experts” ourselves, also recognize our own role in upholding this power chasm. One avenue future work might explore is developing a scaffolded process—much like the one we explored for collective demand generation here—that opens a line of communication between expert stewards and the collective to facilitate collaborative refinement of demands. This line needs to be a constant dialogue with devotion to repair and maintenance [58], rather than a one-off panel of experts. At the same time, in the same flavor as Irani and Silberman [59], being able to synthesize and funnel end-user privacy concerns into existing regulatory frameworks will not make us (and other experts) design saviors of user privacy; the experiences of end-users

themselves, and the work they contribute to collective movements, must remain the driving force of action.

Feminist ethicist Carol Gilligan also distinguishes an “Ethics of Justice” (EoJ) from an “Ethics of Care” (EoC) [60]. Dominant groups tend to prefer an EoJ, which focuses on generalizable standards, impartiality, and a respect for Western democratic ideals. In contrast, an EoC system emphasizes benevolence and the importance of a response to the individual. The panel—in their emphasis on fitting collective demands within existing legal and technical structures—demonstrated alignment with a dominant EoJ. But perhaps what people affected by PVEIs need is for S&P experts to adopt an EoC: people know that what they want is unrealistic, yet they want to be heard and they want change.

3.5.4 Limitations

Our expert panel, though varied in background and industry, was comprised of only eight experts, all of whom were based in the United States. We juxtapose the diversity of their experiences against their near unilateral preference for working within existing institutional structures and against more sweeping action. However, we acknowledge that cultural norms around privacy regulations differ around the world; future work could explore differences in how experts respond to populist calls for systemic changes based on the regulatory contexts in which they operate.

Secondly, we ran only three scenarios through our FFV artifact, and these scenarios do not represent all the ways participants could have identified concerns or proposed demands. However, the emotionally-charged responses we did get from these scenarios support the argument that users lack representation as collectives working against institutions.

Finally, the use of Prolific itself could present limitations. For one, since our participants were paid, they do not necessarily reflect how a grassroots collective would act in the real world. And, while there have not been studies specifically comparing Prolific users’ security attitudes with “average” users, prior work on the representativeness of users on

Amazon Mechanical Turk (MTurk), a similar crowd work platform, has shown mixed results. For example, Kang et al. [61] found that MTurk users have higher privacy concerns and were better-educated about S&P than the larger U.S. public. In contrast, Redmiles et al. [62] found that MTurk users were fairly representative of the U.S. population in S&P experiences and education.

3.6 Conclusion

In this work, we explored how to design a system that facilitates privacy collective action by helping collectives affected by PVEIs generate a unified set of demands for redress. Specifically, we employed a three-stage set of online questionnaires, inspired by the Find-Fix-Verify crowd programming pattern [34], as a sensitizing concept to explore how non-expert collectives can generate concerns and compensatory demands in response to a triggering PVEI. We then presented the results of this artifact to a panel of S&P experts, whose responses helped us not only assess the collective's demands, but also uncover insights into how experts might better steward collectives towards effecting enduring change in privacy practices. Finally, we discussed how our results fit into existing paradigms of computing for social change, and how even well-meaning experts might serve as hurdles to further institutional privacy change. People are frustrated with how their personal data is collected, processed and monetized, but may not have the knowledge to effect meaningful change. Experts have that knowledge, but can be dismissive of those for whom they should advocate. A synergy between expert stewards and the crowds of non-experts fed-up with existing privacy protections could be the foundation for broader change that helps shift power over personal data to the people.

CHAPTER 4

INTERPRETING THE UNIFIED VOICE THROUGH A LENS OF HARM

Experts who sympathize with the collective should be points of resource for the collective to achieve its goals and obtain recognition of its harms. But how can we convince them of this role if they fundamentally don't believe that such harms are concrete enough or worthy of effort for redress? Are there alternative framings to allow both parties to better coordinate their efforts for action (i.e., CSCA stage 3)? To operationalize the collective's demands of apologies for wrongdoing, reparations, and recognition of harm, we can look at "harm" as a formal legal concept, as defining privacy harms is of increasing interest in legal scholarship.

In this chapter, I go over my **harms taxonomy** project, published at FAccT in 2023. This work answers **RQ2**: "How can identifying and taxonomizing harms from specific privacy violations help collectives and experts better coordinate their efforts?" It also explores stages 1 (identifying a problem) and 3 (coordinating and preparing to take action) of CSCA.

This work provides some initial empirical evidence for formally recognizing privacy harms in legal contexts. I delve into one kind of PVEI: online behavioral advertising (OBA), which in a screener survey, participants deemed most personally-violating. In the following sections, I describe how I collected hundreds of user-reported accounts of violating experiences with OBA and qualitatively analyzed the different resultant harms that people reported experiencing in their day-to-day lives.

4.1 Introduction

Surveillance capitalism unilaterally claims human experience as free raw material... It is obscene to suppose that this harm can be reduced to the obvious fact that users receive no fee for the raw material they supply...the essence of

the exploitation here is the rendering of our lives as behavioral data for the sake of others' improved control of us.

— Shoshana Zuboff [63]

Surveillance capitalism—the profit-driven collection and commodification of personal data by private corporations—has resulted in the gradual erosion of privacy, leaving people with “no exit, no voice, and no loyalty; only helplessness, resignation, and psychic numbing” [63]. Its key driver: online behavioral advertising (OBA), or “the practice of tracking an individual’s online activities in order to deliver advertising tailored to the individual’s interests.” [64]

While OBA has been touted as a way to efficiently match advertisers and users, people have myriad concerns about the practice. People dislike OBA for not only the specificity of its targeting, but also its abundance and ubiquity (not to mention finding it generally “creepy”) [21, 65, 66, 67]. All-in-all, there has been extensive documentation of the negative ways that people respond to OBA. However, we know comparatively less about how OBA materially harms people, especially through its entanglement with modern daily life. As Zuboff describes, “there are consequences to this diminishment of rights that we can neither see nor foretell” [63].

The concept of *harm* is both a colloquial and legal one: according to Black’s Law Dictionary [68], it is defined as “injury, loss, damage; material or tangible detriment”. Defining privacy harms is of increasing interest in legal scholarship [69, 70]. Whereas financial losses and physical injury can be clearly identifiable as harms in a court of law [70], privacy harms like those entailed by OBA are less well-understood and often unrecognized. For example, a user might be alarmed to see embarrassing personal shopping history pop up in targeted ads on a work device. Or, they might feel spied on when they see a targeted ad for a product they thought they only discussed out loud with a friend. Insidiously, however, these small, seemingly mundane events can accumulate into a loss of control over interper-

sonal context and being constantly surveilled without consent [63]—in other words, a lived experience of fear and powerlessness [71].

Other experiences with OBA can be more evidently harmful. Advertisers, implicitly or explicitly, can infer and target specific sensitivities and vulnerabilities to increase clicks and sales: e.g., mental and physical health conditions [72]; demographic characteristics like age, gender identity, sexual orientation, race, etc.; bereavement; and unhealthy body stigma [73]. Ads based on these personal and psychological vulnerabilities can entail harmful consequences: users might, e.g., call their self image into question or reveal details about their personal identity without their consent.

In this work, we ask, “**How does online behavioral advertising harm people?**” Through a survey of 420 participants online, we investigated and categorized people’s lived experiences of harm from OBA. Specifically, we asked participants to share with us a recent privacy-violating experience with OBA that they felt was personally impactful. In analyzing these accounts, we identified four main types of harms arising from OBA:

1. *Psychological distress*. Broad negative mental or cognitive effects related to OBA.
2. *Loss of autonomy*. Denial or limiting of opportunity to make own choices.
3. *Constriction of user behavior*. Alteration of user interactions with technical systems in response to other OBA harms.
4. *Algorithmic marginalization and traumatization*. Harms specific to personal characteristics (i.e., demographics) or vulnerabilities (e.g., sensitive medical information).

We then contextualize users’ tendency to normalize these harms within both the concept of “slow violence” [18, 73], and a legal landscape that struggles to recognize privacy harms as concrete injuries [70, 69]. We argue that FAcct, HCI, and privacy researchers have an imperative to consider these contexts in future work, so as to help legitimize these experiences as harms to be mitigated and worth redress. In our analysis, we also suggest

two potential first steps for future work: empirical measurement of the four types of OBA harms we identified, and documentation of protective actions that people have taken to evade these harms. In so doing, we can facilitate the formal recognition of OBA harms and institute processes to mitigate and redress these harms.

To summarize, this work contributes the following:

- A typology of privacy harms from OBA based on a large-scale empirical study.
- A discussion of how formal recognition of OBA’s privacy harms can be a first step to alleviate them.

4.2 Related Work

In this work, we examined how OBA can harm people. We build on prior work of not only user perceptions of online behavioral advertising—including attempts to change these perceptions, such as through user education and increased transparency—but also harm in socio-technical systems.

4.2.1 Online Behavioral Advertising

Online targeted ads are highly effective at engaging users to click. Broadly speaking, however, people have various reasons to dislike online targeted ads, finding them creepy, privacy-invasive, and disruptive [21, 74, 75, 76, 77, 65, 66]. We build on this by specifically examining negative *effects* of OBA on people’s lived experiences, beyond descriptive perceptions of or affective responses to OBA. We view these effects through a lens of harm, which we ground in literature discussed in more detail in the following subsection. Past work on the harms of ad targeting has primarily focused on targeting based on political interests, which can limit user exposure to diverse viewpoints [78, 79] (a phenomenon reinforced and exacerbated by the ad delivery mechanisms themselves [80]). However, as Gak et al. [73] point out, what an ad algorithm deems as “interests” can easily be some-

one’s vulnerability, e.g., sensitive health topics like weight-loss ads. In our work, however, we examine not just the harms of sensitive “interest”-based targeting, but also broad emotional and psychological harms to autonomy and the way people go about their day-to-day lives.

Mental models and folk theories also influence the way that people approach and respond to OBA. Yao et al. [77] found that user understanding of OBA can vary along three dimensions: who tracks the user’s information, where the information is stored, and how the ads are delivered. To address these perceptions, prior work has tried to increase user awareness and agency about OBA, primarily via greater algorithmic transparency [81, 82, 67, 83] and providing more user controls to hold advertisers accountable; two thirds of the FAccT field name itself—Accountability and Transparency—mirror this tendency. But these approaches can also harm people. For one, online privacy notices can be confusing or too long for users to read [84, 85]; users who *do* read them can become alarmed and carry greater psychological burdens with the knowledge that their privacy is being violated [67, 86, 87]. More user control also does not mean more privacy [88], but can rather burden users further [89]. Finally, as we will show in the sections to come, when people believe they have exhausted all possible avenues to evade targeted ads, they feel frustrated and trapped.

Further, even with transparency and awareness measures in place, people might not take advantage of such measures or trust in advertisers and corporations to fully protect their privacy. As Lee et al. [67] found, users reject viewing explanations for targeted ads due to a sense of helplessness and resignation: since they felt powerless to change anything about the targeting, they did not want to know more about it. As another example, news media has frequently debunked the myth that Facebook secretly listens to real-life conversations via users’ mobile phones and targets ads from those conversations; however, people persistently believe this rumor due to mistrust in Facebook and Meta. Das et al. help explain why through a recent review of barriers to end-user privacy and security be-

haviors [90]: awareness is only one barrier that must be overcome—people also have low motivation because they feel helpless, and have little ability to verify and control what data harvesters collect. In our work, we also explore the ways that this mistrust and helplessness burdens and harms people.

4.2.2 Harms and Socio-Technical Systems

We draw upon a rich history of literature in the fields of computer-supported cooperative work (CSCW), human-computer interaction (HCI), and FAccT that examines the relationship between computing and social justice, and how users can be oppressed and marginalized by such systems [91, 92, 72, 93, 94]. For example, Seberger et al. [95] distinguish between the “power to” do something technical that a particular app grants a user, and the “power over” the user that the app and its institutional back-end has over the user; this tension forms a user ambivalence to privacy that opens the door to more invasive data practices. Related work [96] found that this “affective discomfort” has become normalized in the user experience. In our work, we explore how OBA can inflict this persistent feeling and how it harms users.

OBA also leads to specific harms of its own. As discussed previously, people exhibit negative affective responses to OBA, disliking their repetitive nature [66] and the specificity of their targeting [21]. Milano et al. [97] also proposed a taxonomy of *potential* harms primarily based on the content and context of OBA: (1) bad content (e.g., using sexist stereotypes to promote a shaving product to men), (2) omission of essential content (e.g., hard-to-reach communities not seeing public health ads for a vaccine), (3) exploitative context (e.g., exploiting users’ personal vulnerabilities, similar to [73]), and (4) deprived context (e.g., a job-seeker not seeing job ads in their area). And, more recently, Gak et al. [73] previously extensively examined the specific relationships between targeted weight loss ads and users with histories of disordered eating, and the consequent harms of that relationship.

Prior work has not, to our knowledge, analyzed *user-reported* harms of OBA *generally*. We hope to show in our work that OBA, coupled with its inextricability from modern daily life, causes evident harms in people’s day-to-day lives. We situate our contributions in the broader landscape of *privacy harms*, theorized by both Calo [69] and Citron and Solove [70], who assert that harms from privacy violations are currently inconsistently recognized by courts, and that certain non-financial and non-physical harms from privacy violations should be as cognizable as financial and physical ones.

4.3 Methodology

While the literature is clear that people find OBA creepy, unsettling, and threatening, how OBA materially and negatively impacts lived experience remains unclear. Building on prior work systematizing the harms of socio-technical systems, we aimed to systematize the many concrete ways OBA can harm. We conducted an online survey on Prolific, a crowd-work platform, with 420 participants who had indicated in a screener questionnaire that they had previously experienced a privacy violation related to online targeted or behavioral advertising.

4.3.1 Recruitment, Ethics, and Compensation

We first screened 1275 potential participants by asking them if they had recently experienced feeling violated by OBA. These potential participants were adults located in the United States, fluent in English, and active users of Internet-based services like social media, a smartphone, or a smart home device. This screener took on average less than a minute to complete; participants were compensated 0.25 USD on the Prolific platform. Those who answered “yes” to the screener (N=420) were recruited to participate in the main study, a short survey hosted on Qualtrics, which took on average 5 minutes, for which participants were compensated 1.50 USD. Our study was approved by the Georgia Tech IRB.

Table 4.1: Demographics of participants, broken down by whether they chose to share an account of their experiences with online behavioral advertising (OBA).

Demographic	Group	Shared experience	Did not share
Age	18-24	86	31
	25-34	111	26
	35-44	47	19
	45-54	28	10
	55-64	27	13
	65+	11	6
Gender Identity	Female	174	57
	Male	119	42
	Genderqueer/Non-conforming	7	3
	Trans Male/Trans Man	8	1
	Different Identity	3	0
Ethnicity	White	245	85
	Black	17	6
	Asian	22	4
	Mixed	18	9
	Other	13	1
Education	Less than high school	12	1
	High school diploma	69	25
	Technical/community college	50	11
	Undergraduate degree	127	44
	Graduate degree	46	17
	Doctorate degree	7	5
Total	420	315	105

4.3.2 Survey

There were three main components to the survey. First, after reminding participants that they had previously told us that they had a recent privacy-violating experience involving OBA, we asked if they wanted to tell us about the experience in more detail. Because such experiences can be sensitive in nature and difficult to talk about, we gave participants the option not to tell us about the experience at all. Second, if a participant agreed to share their experience, to encourage richer qualitative contributions beyond simply describing it as “creepy”, we suggested details to include in their account of the experience: the parties and information involved, any actions they took in response to the incident, emotional reactions, changes in how they used the Internet, or why the incident was personally im-

pactful. If a participant chose not to share an account, to ensure participants were being equally compensated for the same amount of work, we asked them if there were other privacy harms or violations unrelated to OBA they would be willing to share instead. Finally, we asked the participants why they did or did not, respectively, choose to contribute to our study.

4.3.3 Analysis

To understand the many ways OBA can harm or burden people, we applied an inductive approach to qualitative data analysis. One member of the research team read through each of the accounts provided by participants and performed open coding, iteratively updating the codebook as necessary. The researcher then performed an initial round of axial coding to consolidate codes into different types of reactions to online targeted advertising, as well as any descriptions of the content of the ads or where the experience took place. A second researcher independently coded the data according to the codebook. Through multiple discussions, all members of the research team consolidated and synthesized the concepts into broader categories. The codebook, grouped by preliminary categories, can be viewed in Chapter A.

The qualitative accounts, by their privacy-violating nature, described how online targeted ads negatively affected participants. We thus grouped codes and concepts based on the nature of the negative effect the experience had on the participants; more specifically, we examined them through a lens of harm. As aforementioned, through our coding process and multiple iterative discussions, and taking inspiration from prior work [70, 73, 66], we developed four broad categories of harms that we summarize in Table 4.2 and discuss in detail in Section 4.4.

Table 4.2: Descriptions of the types of harms and their distinguishing characteristics.

Harm	General description	Distinguishing characteristics
Psychological Distress	Broad negative mental or cognitive effects related to OBA in general.	General emotional distress, i.e., painful or unpleasant feelings, or disruptions to peace of mind.
Loss of Autonomy	Denial or limiting of opportunity to make own choices.	Lack of control or consent over not only targeting, but also secondary contexts like interpersonal relationships and purchasing behaviors.
Constriction of User Behavior	Alteration of user interactions with technical systems in response to other OBA harms.	Losses in usability and utility of devices and services due to adopting additional privacy and security behaviors, as well as the time and resources associated with such evasive actions.
Algorithmic Marginalization and Traumatization	Harms specific to personal characteristics (i.e., demographics) or vulnerabilities (e.g., sensitive medical information).	Feelings of diminishment associated with highlighting user-specific characteristics, rather than broader senses of unease or overwhelming.

4.4 Findings

We first provide a demographic breakdown of participants based on whether they chose to contribute an account of their experiences. We then report on a broad overview of the online platforms where these accounts took place. Finally, in the bulk of the section, we discuss four broad categories of harms that can arise from online behavioral advertising: psychological distress, loss of autonomy, behavior constriction, and algorithmic marginalization and traumatization. We note that these harms are not mutually exclusive, but have distinguishing characteristics as described in Table 4.2.

4.4.1 Quantitative Breakdown

Participant Demographics

In total, 315 participants chose to share an account; 105 chose not to. A summary of the demographics of our participants can be found in Table 4.1. They are grouped by whether or not they chose to contribute an account of their experiences.

Where Accounts Took Place

Over half of participants mentioned specific companies or websites as the source of their experiences with OBA. Meta products were the most-mentioned site of violating experiences, with 84 participants mentioning an experience involving Facebook, and 55 mentioning Instagram. 49 participants mentioned Google, and 18 mentioned YouTube specifically. 20 participants mentioned Amazon or Alexa devices. Only 5 participants mentioned TikTok. 140 participants did not name any specific company or site; however, 34 of these participants mentioned seeing ads on “social media” generally, and 10 participants said the advertising was “everywhere”.

4.4.2 Psychological Distress

Psychological harms involve a wide range of negative mental responses, but typically fall into two primary buckets: emotional distress, i.e., painful or unpleasant feelings; and disturbance, i.e., disruption to peace of mind. In the following subsections, we discuss different examples of both types of psychological harms.

General Emotional Distress

As Citron and Solove argue [70], one of the most common types of harm caused by privacy violations is emotional distress. Our participants’ experiences provide empirical support for this claim: a fifth of accounts mentioned feeling unsettled by or uncomfortable with the

specificity of targeted ads. For example, P326 expressed discomfort with the uncertainty and lack of transparency on what data is collected and the inferences that could be made thereof: “[Google] knew I was interested [in a new phone] because I had said so, out loud, to my girlfriend on a private call. If they hear that, who knows what else they hear? And what could be done with that information?”

Disruption of Browsing Experience

Other participants expressed that the targeted ads disrupted their normal browsing experience. For example, P284 was angry and frustrated with seeing ads after they finished comparing an item’s price at Walmart: “I was done needing to see Walmart, and now it’s all over the searches and websites afterward.” Similarly, other participants felt that they wanted to search for things independently, rather than have that search be re-incorporated into an ad targeting experience: “You can’t just search anything anymore without being bombarded by ads” (P171). Some participants saw so many targeted ads that they had trouble distinguishing what was an ad: “I lose track of the number of real posts I see [on Facebook] vs. ads these days” (P160).

Information Redundancy

Sometimes, the sheer amount of targeted advertising from *specific* advertisers resulted in participants being oversaturated with redundant information. For example, P188 was frustrated with repeatedly seeing the exact same ad from Halara, a clothing retailer: “After the third time I felt very frustrated and bored. This ad was haunting me and honestly I don’t think I would buy from this company now. It annoyed me so much. I would mute my computer while it played and try to skip it as soon as possible.” Another participant who was targeted with ads from a guitar retailer said they were “inundated with advertising...for other guitars on unrelated sites repeatedly. This presumes upon my attention and cognitive/emotional space and angers me” (P243). This inundation could also translate to

a loss of time and effort in real life. P154 noted that even if they had blocked ads *online*, they still received corresponding physical marketing in the mail: *“I’ve cleared my cache and cookies since but I’ll be receiving snail mail for generations. I get angry about this marketing because it consumes my time to throw everything away...I have to shred every offer that comes in the mail. Waste of paper, waste of resources, waste of time and energy.”*

Questioning Own Browsing Behavior

Beyond feeling annoyed or overwhelmed by targeted ads, however, we also found that participants frequently tried to guess at where the ads came from. This echoes prior work [67] that found that users wanted explanations for ad targeting to confirm their own pre-conceptions of how their data was collected or the motives of advertisers. For example, multiple participants mentioned that ads “must have” come from their browsing and search history or their online chats with friends and family. P358 surmised that ads related to their personal shopping showed up on their work computer due to sometimes logging into their personal accounts at work: *“We are a Microsoft-based system [at work] but at times I have clients send me Google Drives, etc. which requires me to log into my Gmail. I suspect that is how these ads came to be on my work computer.”* Some participants approached ads like a mystery to be solved with breadcrumbs of all the places where they had encountered certain ads:

When I turned on another computer that I watch streaming television on, the same topic ads were on there, as well. I concluded that Google targets ads by your router’s IP...and then dispenses ads to all machines connected that IP, through your router. If you look up porn, be aware that porn ads could show up on their computer, because of the general use of ads pointed at your domain.

(P378)

Others expressed disgust at the tracking after after detailing their browsing behavior step by step. P248, for example, shared how they looked up a clothing brand in Chrome on

a work device, and then saw related ads on their personal device while using Firefox: *“I clicked on those items in a different browser in a different physical location. I felt absolutely stalked by this ad.”* Similarly, after retracing their digital breadcrumbs, some participants then expressed regret about mistakes they had made along the way when researching and discussing the related ad topics. P154 shared that they forgot to block cookies on a credit card website one time, and have regretted it ever since: *“I forgot to deselect the marketing cookies once on a credit card website...but now every single website has credit card offers.”*

Paranoia From Suspicion of Eavesdropping

When their investigations into the origins of targeted ads reached dead ends, participants often felt that the only possible explanation was that their devices were eavesdropping on them. Several participants felt that merely mentioning a product in a real life conversation with a friend or family member would result in seeing ads for that product, whether it be spices (P42), cat food (P49), electric toothbrushes (P56), hula hoops (P90), or press-on nails (P131). While the concept of smart and mobile devices—in particular, Facebook and Meta—eavesdropping on users via microphones has been frequently debunked in popular news media and prior work [98], the myth persists. The lack of transparency and trustworthiness surrounding these ad targeting practices[71] necessitates misguided guesswork on the part of the users, which results in concrete harms: constant suspicion, fear, and paranoia.

The immediacy and consistency with which targeted ads appear made participants suspicious of their microphone-enabled devices: *“An Alexa device was in the same room, but was off, or so we thought. On more than one occasion the items we discussed showed up almost immediately on our devices (email, internet ads, social media, etc.)”* (P20). Similarly, P60 shares:

I don’t have an Alexa or anything like that, but somehow my phone is apparently listening anyway? I don’t know what to think. These are private con-

versations! And I know that I haven't just entered any of that into the search window on my phone or computer. It has happened far too many times for it to be coincidence and all I can assume is nothing is private anymore. It's sickening.

Some more tech-savvy participants admitted that even though they were educated otherwise about microphone eavesdropping, they still felt concerned. For example, as P1 described, *“While I know rationally that these programs aren't listening, it is very unsettling that they are reading my data to target ads to me.”* Similarly, P36 said they closed the Instagram app on their phone as soon as they were done using it, even though *“I know it supposedly doesn't listen in but rather tracks you in other ways...but sometimes the ads are a little too targeted for comfort”*. P55 adds that the targeting is simply too accurate and immediate to ignore: *“While it is possible that it's a coincidence and social media shouldn't have access to my microphone to capture data and tailor advertisements to me, I can't help but feel creeped out and paranoid that I'm being recorded at all times.”* These persistent fears are direct vectors to psychological harm.

4.4.3 Loss of Autonomy

Autonomy harms involve accounts where participants were prevented from making their own choices, either via being directly denied these choices, being tricked into thinking their choices were freely made when they were not, or being limited in the choices they could make. These harms recall Zuboff [99], who wrote that in surveillance capitalism, *“the surest way [for advertisers] to predict behavior is to intervene at its source and shape it”*. In the following subsections, we discuss different types of autonomy harms.

Lack of Consent or Control over Targeting

One of the most common concerns that participants voiced was that they did not consent to being targeted for online advertising. Some participants, for example, felt violated when

they saw online advertisements based on purchases they'd made in a physical store. P77 shared how they had seen ads related to groceries they bought in a physical store, but had never expressly consented to connecting these purchases with any shopping website or app: *"I feel it should be against the law to...invade my privacy without express permission each time they want to do something like this."*

Others felt like they had no control over the nature of the targeting, even when the targeting was incorrect. For example, one participant was researching flooring materials on behalf of their mother, but still received endless ads related to it. They felt frustrated at being misunderstood:

The flooring isn't really for me, but just in my searching, I've had countless ads and emails sent to me about all kinds of flooring. I've even had other companies sending me info about flooring. I feel like I'm being attacked by salesman at a used car dealership and I really have no control over it. (P351)

Other participants noted that even though they understood their data was being collected online, it still felt like a breach of consent. Participants found the pervasiveness and specificity of the ads overwhelming, describing a mission or scope creep of sorts: *"It's not even what I'm doing anymore, it's everything I am thinking"* (P49).

Lack of Control over Self-Presentation

Several participants mentioned seeing ads on devices they used at work that were related to interests in their personal life, or vice versa. This exemplifies context collapse, or "how people, information, and norms from one context seep into the bounds of another" [48]. Participants who experienced context collapse felt they had little control over the consequences, usually in the form of the targeted ads unwittingly revealing private information about themselves.

One immediate harm of this phenomenon was social embarrassment, as P358 writes:

About a year ago, I had been shopping for lingerie for my honeymoon. This was only on my personal laptop. I had a coworker in my office looking up some information with me. I believe it was thesaurus.com or something like that, but I can't fully remember because what was ON the page was so mortifying. There, in front of my coworker, on my WORK computer, were specialized ads for lingerie. I was so embarrassed. I tried to ignore the ads that seemed to be disproportionately large on the screen. My coworker thankfully did not mention them, but now probably thinks I shop for lingerie while at work.

Other participants noted that even friends who talk to them about sensitive problems could influence the ads they saw and result in negative outcomes. For example, P124 mentioned speaking to a friend about the friend's pending divorce, and subsequently got ads related to divorce lawyers. This made the participant concerned that their own spouse would accidentally see these ads and misinterpret them: *"This could potentially lead to misunderstandings with my spouse. What if I was showing them something on my phone and a divorce attorney ad came up?"*

Relatedly, participants also experienced context collapse with family members directly. P363, who had shared their Facebook credentials with their mother, described how their mother saw ads from their feed for explicit content and surmised that those ads were based on P363's own browsing behavior. The ads thus revealed private information about P363 to their mother without their consent, making P363's relationship with their mother *"uncomfortable for a long time"*. Another participant shared that a surprise birthday gift for their husband was ruined when the husband was targeted with ads related to the participant's search history. As a result of this ruined surprise (a harm in itself), they began altering the way they used the Internet out of persistent concern that future surprises could be ruined too:

I have started only using my desktop at work to search for presents for people in my family. I'm paranoid even to buy gifts for my parents and in-laws on our

family computer, though they don't live with us and the chances they will use our family desktop to search the web is very small. But just in case one day they need to use the computer, I don't want them to see the gift I am searching for them! (P192)

Similarly, P265, who researched medical treatments on behalf of a friend, felt constantly reminded of their friend's heart problems. They also now felt burdened with protecting not only their own privacy, but that of their friend, too: *"I was not looking up the information for myself, but for a friend. I cannot go to most sites without seeing ads for TAVR (transcatheter aortic valve replacement) and now other heart problems. I will be careful about searching for sensitive information for both myself and my friends".* P239 offered a similarly sensitive account of searching for resources for their sister-in-law who was dealing with marital rape: *"I started seeing ads related to mental health to help rape victims. I do not know who actually made the ads. It was a constant reminder of the abuse that she went through."*

Encouraging Negative Purchasing Habits

A few participants felt compelled to buy things they did not need because they kept seeing ads for them. For example, P35 described themselves as *"impulsive with my money"*, and said that a stream of targeted ads *"makes it hard to use social media when I see ads for clothing that I want but cannot afford."*

Limiting Consumer Choice

On the other hand, for some participants, companies that used OBA were so off-putting that it compelled them to limit their choice set of things to buy and shop elsewhere, so as not to reward bad behavior. Several participants stated that the fact the advertiser was directly pandering to them made them not want to purchase anything from that advertiser. For example, a few participants felt that if an advertiser was spending so many resources

on marketing targeted toward them, it must be a signal of a deficiency in the products being advertised. P18 wondered, *“If they [an e-bike company] have the budget to spend so much on targeted advertising, what is wrong with their e-bikes? I wonder if they are charging too much or that the product is of much lower quality than their competitors.”* More bluntly, P408 said, *“I become so disinterested and put off by these practices I look at other brands, and I would NEVER click on such an ad regardless of my level of seriousness to purchase such a product.”*

4.4.4 Constriction of User Behavior

As we argued previously, users can face usability burdens when dealing with online targeted ads. Multiple participants mentioned taking privacy-protecting measures, such as disabling advertising-related tracking, deleting accounts on retail websites, and erasing browsing history. But even though they went through so much effort, participants felt they could not escape the ads. Despite having all *“privacy flags available set to the maximum”*, P167 said, *“the tracking persists. I feel powerless to prevent this from occurring.”* Not only are the effort and time taken to implement these seemingly futile measures an unwelcome burden for the average user to bear, but users are also penalized with a loss in usability and utility when they are forced to use services or devices in less-than-optimal ways.

Perceptions of microphone eavesdropping elicited specific actions from participants. One participant mentioned disabling Siri on their iPhone *“so that it was not able to listen at all hours”* (P19), limiting themselves from accessing the full functionality or convenience of their personal phone. We acknowledge that disabling Siri was an active choice on the part of the participant, and the immediate inconvenience of not being able to use Siri might seem like an innocuous harm. However, it’s easy to imagine a scenario where opting out of one tool can lead to more extreme consequences, or the cost of opting out is greater than simply switching off a button.

As one example of the former, P119 shared that after making therapy appointments online, they saw ads related to depression and PTSD. As a result, they stopped booking their appointments online and instead could only do so over the phone. While this might only appear to be a minor inconvenience on the surface, people with social anxiety or social phobia could find the idea of making a phone call paralyzing, and may rely on online booking services. (In Section 4.4.5, we discuss in detail harms that specifically come from ads with sensitive content or offensive profiling). And, as an example of the latter, P340 said that they had “*removed all [Amazon] Alexa devices from their home and shut down all web camera (sic)*”, entailing a not-insignificant amount of time unplugging and covering up all their devices, and not to mention the money lost on purchasing the devices in the first place.

Several other participants mentioned avoiding having conversations about potential purchases or changing the way they talk to their friends to steer clear of getting related ads. For example, as P259 put it, “*I take the approach ‘the walls have ears’ and typically act as if someone (like my boss, for example) were listening [in] on my conversation, because it’s clear that what I say in private might not actually be private anymore.*” P55 said, out of fear of eavesdropping, they started physically separating themselves from their phone: “*I don’t carry my phone with me into other rooms if I am hanging out with someone and if I need to search something up, I have to go and grab my phone from wherever I left it. I feel the need to keep distant in order to maintain some sort of privacy and to ease my paranoia.*”

4.4.5 Algorithmic Marginalization and Traumatization

OBA based on user interests can over-simplify those interests and hone in on user vulnerabilities. Echoing prior work [73], we found that when users identified as part of a sensitive “interest” group, they were particularly vocal about being violated. We distinguish these

harms from general psychological distress (Section 4.4.2) due to how these ads diminish people by highlighting their specific personal characteristics or vulnerabilities.

Violation of Boundaries

Some participants felt that certain information should simply be off-limits as a basis for targeting. For example, one participant dealing with the death of a loved one felt that targeted ads related to funeral services tried to exploit their private grief: *“It was inappropriate to intrude on our grieving with an attempt to get us to spend money on elaborate funeral services or gouging us for insurance”* (P130).

These limits also applied to medical histories. For one, even if participants felt that the medical treatments promoted in certain targeted ads were valid, they were still disturbed that data brokers knew about their medical history and targeted them for it. For example, P244 felt that they were shown ads related to substance abuse treatment programs because of their history of opiate abuse, and were upset with how Facebook concluded this about them: *“It reminded me of a very dark time that I would like to forget. I am not against the company or the treatment program, just how Facebook selected me for targeted advertising.”* P273 felt that seeing ads related to a medical condition on Twitter, Instagram, and Facebook, *“was akin to a HIPAA violation.”*

Amplification of Self-Consciousness

Multiple participants reported feeling discriminated against when they saw certain targeted ads. Some older female participants reported seeing ads related to menopause, and said they felt uncomfortable that so much attention was being drawn to their age. As P249 writes, *“Few women like to have this stage of life rubbed in their faces by a social network, for goodness’ sake...I didn’t feel ashamed; I felt really depressed. I don’t like to dwell on things that basically say to me, ‘Say, I hear you’re growing OLD, girl!’”* P366, who self-identified as a woman, said they started seeing ads for Botox and plastic surgeons *“EVERYWHERE”*

when they entered their age into Instagram; they felt this promoted ageism. To avoid these ads, P366 created a new account with a fake age (evoking Section 4.4.4).

Participants also felt misrepresented and hurt by ads referencing neurodiversity. P2 shared an experience with ads related to autism, which unsettled them not only due to the sensitive topic, but also misleading portrayals of autism:

I was recently diagnosed with autism. In the following week, I was getting tons of ads on Facebook about “holistic medications” and “lifestyle changes” that help people to be “less autistic.” Additionally, I was getting sponsored content from Autism Speaks and moms in the community. The ads were hurtful. Autism is a neurotype, not something that can be “cured”, especially by unregulated supplements or diets. I think targeted misinformation like this is extremely sinister.

Similarly, specific representations of neurodiversity in targeted ads made participants self-conscious about how they were perceived by others in real life, and caused them to alter their browsing behavior. One participant felt hurt due to an ad that portrayed people with ADHD as “closed off and goofy”, and said that it “showed [ADHD] to be more lighthearted than it actually was. Now I’m a bit more self-conscious thinking others have seen the advertisement and are judging me. Now I know to be more careful with what I search” (P126).

Traumatic Triggers

Ads related to eating disorders and body image can serve as constant painful triggers for participants, echoing prior work [73]. For example, P78, who actively participates on an anonymous eating disorder forum, started getting ads for both erectile dysfunction and weight loss programs. They felt that seeing these ads outside of the forum was “damaging to my mental and ultimately physical health, as they are constant reminders of my restrictive eating problem.” Similarly, P206, who discussed body image and weight issues with

friends and family members, noticed “*an increase in various weight loss ads across my social media platforms. Instagram has given me ad after ad for diet plans and paid exercise programs. I felt like every ad I saw was pointing out my insecurities and confirming my embarrassment in my appearance. Since this started, I have felt very unsafe on the internet.*”

In a similar vein, participants dealing with bereavement also felt that seeing ads related to funeral services prevented them from moving on with their grief. P341, completely exasperated, expressed regret at looking up headstones for their grandmother on Google: “*Well that was a mistake, because to this day I get reminded multiple times a day that I had to lay my grandmother to rest. I see multiple ads a day on Google, Facebook. I’m over it. I just want to be able to mourn and let it go. At this point I don’t want to ever Google anything again.*”

Fear of Social Exposure

Beyond the immediately obvious harms of reductive ads based on demographics, participants also worried about secondary social ramifications of such ads. For example, participants who were members of the LGBTQ community felt alarmed that data brokers had algorithmically profiled their gender identity and sexual orientation. For one, they expressed anxiety about how these characteristics were inferred in the first place, especially if they themselves had not yet come out. For example, P173 wrote:

I am closeted on Facebook (not out as trans, neutral name, lots of family members added, had not yet changed my pronouns on the site) and it started advertising products aimed specifically at trans men to me—[chest] binders, online medical services for HRT [hormone replacement therapy], etc. I had not shopped for any relevant products. I do not know why Facebook decided I was/am a trans man. I am anonymous on other Facebook products like In-

stagram, no mention of gender or pronouns there at all, let alone my correct pronouns.

As another, more immediate danger, P164 shared how they were afraid of exposure of their gender identity at work. This fear also constricted how they used their phone at work:

My coworkers do not know that I am transgender. Before a meeting we were all...scrolling through our phones. I got a targeted ad for a trans pride flag that was large and very visible. I immediately got scared that people behind or beside me would see so I quickly closed my phone and put it away. Now I don't open my phone when I'm in the same room with coworkers. I'm not trying to get outed in an unsafe way.

Participants who were already out were still concerned. P214, for example, described a conversation with their boyfriend about underwear unrelated to their sexual orientation, but seemingly led to targeted ads about underwear for gay men: *“It felt wrong, especially when being gay is criminalized in some countries and even demonized among communities in the U.S. It's potentially dangerous if someone else sees an ad that's targeted towards something private about you.”*

4.5 Legal Recognition and Formal Measurement of Privacy Harms

Our participants shared personal accounts of how online behavioral advertising harmed them by disrupting their peace of mind, eroding their autonomy, impeding on their day-to-day lives, and preying on their personal vulnerabilities. Yet, many still prefaced their responses with qualifiers like, “it may not seem like a hugely violating deal” (P12). As Seberger et al. [96] write, these violations do not become “less catastrophic or problematic upon regular repetition” but rather normalized; the authors reference the salience of the word “creepy” in the consumer vocabulary as a proxy for the undefined, constant institutional rebalancing of user convenience against violating data practices.

Manifestations of OBA Harms

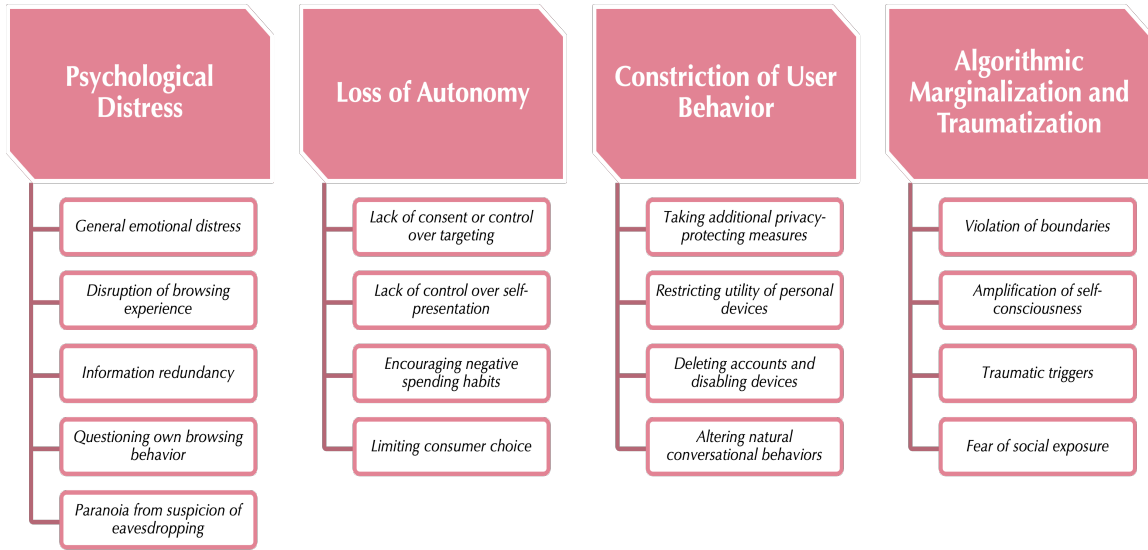


Figure 4.1: A summary of the ways we found OBA harms manifesting in people’s lives. Aggregated, collective evidence of such experiences may help establish such harms as concrete injuries with legal standing.

The mundanity of targeted ads makes it difficult for users to devote specific attention to fighting it; as Gak et al. [73] argue, the embeddedness of harmful targeted ads in a typical user’s digital experience typecasts them as “non-events” to which users become habituated and numb. Literary scholar Rob Nixon [18] coined the term “slow violence” to describe things like the normalization of seemingly small harms. Slow violence, as Nixon defines it, consists of “calamities that are slow and long lasting, calamities that patiently dispense their devastation while remaining outside our flickering attention spans”. Originally conceptualized by Nixon in reference to environmental degradation and climate change, slow violence was adapted by Gak et al. [73] to OBA.

Giving legal recognition to this slow violence can be a first step to mitigating privacy harms on a systemic level; designing formal ways to gather evidence of these harms can better support users in achieving this recognition.

4.5.1 Legally Recognizing Privacy Harms

As a start, as early as 1980, the Federal Trade Commission (FTC) has recognized that small harms, in aggregate, can be sufficiently substantial if suffered by a large number of people¹. However, as regulators can only address a small fraction of these privacy harms at any given time, more attention is given to flashier violations that are well-understood to cause harm to users, e.g., large-scale data breaches like the Equifax breach in 2017, which resulted in hefty FTC fines. But these one-off, ad-hoc solutions don't necessarily mesh with the slow violence of OBA. For example, Wu et al. [100] found that when a group of users collectively preferred an apology from Instagram as a solution to offensive algorithmic profiling—in other words, a recognition of harm—security and privacy experts dismissed the users' preferences as naive and fraught with implementation challenges. In this way, as Nixon [18] argues, the extended temporality of slow violence hides a more sinister foundation of social inequality: the people who experience the small harms of OBA and surveillance capitalism have unequal education, access, and power relative to the security and privacy experts who dismiss them.

One remedial measure is recognizing privacy harms, legally, in the same light as other sorts of harm. Currently, courts do not recognize privacy harms that don't involve tangible financial or physical injury. However, as Citron and Solove [70] write, "Individuals whose privacy has been violated need to hear the message that law is concerned with the harms they have suffered. Law's recognition of privacy harms tells individuals that their suffering is real and that their suffering is not just a fact of life that should be endured, but harm that should not be tolerated. Individuals can see themselves as harmed."

As one example, the targeted ads that could ostensibly result from microphone eavesdropping are, to users, indistinguishable from those based on cross-site tracking and algorithmic profiling. Both conceptualizations of OBA, regardless of actual practice, can lead to

¹FTC Policy Statement on Unfairness. December 17, 1980. <https://www.ftc.gov/legal-library/browse/ftc-policy-statement-unfairness>

the same harmful outcomes of fear, distress, and unease in users, causing them to alter the way they handle their devices and hold conversations in real life. Thus, would more user education and transparency about the latter method being closer to the truth than the former do anything to mitigate these feelings? Instead, a formal recognition that these harms carry a real burden can be used to actually redress those harms by, e.g., establishing legal precedent and allowing for the allocation of remedial resources (e.g., funding, headcount) to mitigate those harms.

4.5.2 Formally Measuring Harms

On these bases, we envision two intertwined areas of future work: concrete measurement of privacy harms, and documentation of the actions that users have taken to avoid harm.

As we've reiterated, courts have expressed doubt that there is enough evidence to demonstrate a concrete harm from privacy violations (e.g., *TransUnion LLC v. Ramirez* [101, 102], and *Spokeo Inc. v. Robins* [103]). Yet, in our work, we have collected hundreds of accounts of the concrete ways that people are harmed by OBA in their day to day lives. Future work could explore how to document these harms in a more systematic manner, perhaps through quantitative measurements. While not all harms can be easily quantified, it can be helpful to measure and make public harms that *are* quantifiable.

For example, one disruptive characteristic about OBA that participants brought up was the sheer amount of ads they saw. One first step toward concretizing this as a harm is to record the number of ads shown and how much time and data they take to load. Popular ad and cookie blockers (e.g., PrivacyBadger², uBlock Origin³) already show counts of trackers and ads found and blocked during a user's browsing experience. The Brave browser, which also blocks ads, goes one step further, calculating the amount of time and bandwidth it would have potentially taken to load them if they were not blocked; these metrics are displayed to users as bandwidth and hours "saved", purportedly analogous to 23 USD

²<https://privacybadger.org/>

³<https://ublockorigin.com/>

per month per user⁴. Beyond automatically measurable variables like ads blocked, we can envision third-party watchdogs (e.g., EFF, Consumer Reports) developing tools that help users collectively report on ads or ad practices that distress, constrict, and marginalize. Empirically and explicitly measuring all the different OBA harms we uncovered in this work, aggregated over a “sufficiently substantial” number of users, could present compelling evidence to government officials of a privacy harm.

Ashley Gorski, a staff attorney at the American Civil Liberties Union (ACLU), has argued that it is often easier to “produce evidence of protective measures than evidence of secret surveillance itself” [104]. When people stop using devices or software, or switch to sub-optimal services to avoid surveillance, they are taking actions that cost them; thus, as Gorski contends, they are being injured by that amount of time and money. In other words, actions that users have taken *against* OBA could qualify as an injury, evoking our “constriction of user behavior” harm type. Different evasive actions might also help signal the other OBA harm types we uncovered: for example, users who provide false information about themselves might wish to avoid algorithmic marginalization, whereas those who disable their smart home device microphones may wish to avoid the psychological distress of being constantly surveilled. Thus—concurrent with documenting harms generally—measuring the amount of time, money, effort, and other resources that users take to *avoid* OBA is another form of evidence for courts and legislative figures.

4.5.3 Limitations

Our work has a few limitations. For one, our participant pool was entirely located in the United States of America, limiting the cultural scope and policy relevance of our findings. Another limitation is the use of Prolific for recruitment: while some past work has found that users on Amazon Mechanical Turk (MTurk), a similar crowd-work platform, were fairly representative of the U.S. population in S&P experiences and education [62], other

⁴<https://brave.com/tips-and-tricks-for-brave-on-your-phone/>

work has found that they have higher privacy concerns and were better-educated about S&P than the larger U.S. public [61]. Future work could explore the extent to which users from different cultural and regulatory contexts—e.g., more collectivist countries [105], or the European Union, where the GDPR strictly enforces user consent—as well as different educational backgrounds, might hold different attitudes regarding what is harmful.

4.6 Conclusion

Online behavioral advertising is a key driver of surveillance capitalism [63], which has resulted people feeling concerned about the state of privacy but helpless to effect change [71]. The first step in empowering people is to provide a model and vocabulary for the many ways online behavioral advertising harms people. In this work, we provide this model of lived OBA harms. Through a survey of 420 participants online, we uncovered four key types of harms that users endure from OBA: psychological distress, loss of autonomy, constriction of user behavior, and algorithmic marginalization and traumatization. We then discussed how users can become inured to these harms over time, and how formal legal recognition of these harms can be a first step to mitigating them. Finally, we recommend that FAccT, HCI, and privacy researchers reconceptualize OBA as part of a bigger picture of privacy-encroaching, actively harmful societal practices, and provide initial guidance for how we might combat and mitigate these practices.

CHAPTER 5

IMAGINING FORMAL WAYS TO MEASURE AND RESPOND TO PRIVACY HARMS

Given a dataset of lived privacy harms from online behavioral advertising that people have reported, we can taxonomize them and give experts and non-experts alike the power to name their harms. Expressing that specific harms have been committed is the first step to formally having those harms remedied or redressed. However, an empirical *measurement* of those harms is also necessary as consistent and reliable evidence in a court of law. What's needed, then, is a way to chronicle evidence of privacy harms across a population over time.

Today, there exists no system or mechanism for the population-at-large to report on everyday privacy harms. Consequently, we also haven't fully imagined the downstream impacts of such a harm-reporting model, or what the success of collective reporting even looks like. In this chapter, I describe my **harm-reporting design fiction** project. I explore the design requirements and cultural ideals of a government-run system to empower people to collectively report on and make sense of experiences of privacy harm from OBA. With fictional inquiry, story completion, and comicboarding methods, via a 50-participant online survey, I found that participants had detailed conceptions of the user experience of such a tool, but wanted assurance that their labor and personal data would not be exploited further by the government if they engaged with the tool. I extrapolate these design insights to government harm-reporting tools in other domains, finding common design gaps that might disincentivize people to report experiences of harm, be they privacy-related or otherwise.

This work answers **RQ3**: "How can we envision a world where evidence of privacy harms is systematically collected and used successfully?" It also explores stages 2 (generating, debating, and selecting solutions), 4 (taking action), and 5 (following up, documenting, and assessing action taken) of CSCA.

5.1 Introduction

In recent years, there has been a rise in civic technology platforms from government agencies to support citizen complaint-reporting—e.g., in city 311 portals [106]—and gathering of collective public sentiment in other domains. As a broader example, the Consumer Financial Protection Bureau (CFPB) operates a complaints website where consumers can report problems with financial products and services and receive a response from the CFPB. Such platforms also exist in other federal-level institutions within the United States, such as the Federal Trade Commission (FTC), the Federal Communications Commission (FCC), and the Occupational Safety and Health Administration (OSHA). However, while these platforms receive and process millions of complaints, people continue to mistrust the government to adequately represent their interests in bringing privacy-violating corporations to justice [107] and remedying their harms more broadly [108].

One reason people might feel this way is that these existing systems primarily operate as intake forms, where the sole action people can take is submitting a complaint; this limited scope also fails to capture people’s nuanced expectations for seeking help from the government. Are there other opportunities for people to participate in a harm-reporting ecosystem? How might these possibilities change how we think about existing systems? We propose design fiction as a way of better understanding the collective needs and desires of people when using such systems. Through a mixture of fictional inquiry, story completion, and comicboarding methods deployed in an online survey, we explore the potentials of a fictional government-run reporting tool that empowers users to report on and make sense of privacy harms. To contextualize our design fictions, we focus on arguably the most pervasive end-user touchpoint for digital privacy harms: online behavioral advertising (OBA). We ask the following research questions:

RQ1 What are design requirements for everyday end-users to participate in collective civic reporting of privacy harms from online behavioral advertising?

RQ2 What cultural ideals do people have around redress for reported privacy harms from online behavioral advertising, and the parties that carry out this redress?

We find that participants had detailed conceptions of what the user interface and user experience of such a tool could look like, such as the ability to support multiple types of evidence of harm, and ticket numbers to follow-up on claims. They also wanted guarantees against privacy risks for contributing reports containing sensitive information, and had vocal expectations of speed, transparency, and accountability for such a tool. However, participants felt ambivalent about relying on volunteer labor to make sense of the data contributed through the tool: while some expressed the potential for pride in volunteering and a duty to help others, others worried that such a tool would simply be another way for their data and labor to be exploited by the government.

While our design fictions focused specifically on *privacy* harms from online behavioral advertising, participants also provided rich insights into their ideals for government reporting tools for *consumer* harms at large. We thus synthesized, from our findings, seven key design principles for supporting people's trust in Government Tools for Civic Harm-reporting (GoTCHas): (1) visible, upfront benefits; (2) timely, useful feedback; (3) contestability; (4) error prevention measures; (5) integration into everyday life; (6) consideration of social influence; and (7) diversity and flexibility of commitment. We use these principles to evaluate a sample of five existing GoTCHas in other domains, such as the CFPB's complaint reporting site, finding that these existing GoTCHas fail to offer both transparency into how they resolve people's complaints, as well as concrete benefits for people to contribute. We also find that existing GoTCHas lack consideration of social influence in people's motivations to contribute evidence and complaints, offer very limited ways for people to contribute or participate, and, as standalone websites, are poorly integrated into everyday life. These limitations can disincentivize people to contribute to these systems, further contributing to people's feelings of helplessness with respect to what the

government can and will do about digital harms—especially as they relate to privacy, but also more broadly. This work makes the following contributions:

- An adaptation of comicboarding and story completion methods to generate fictional futures where collective civic reporting of privacy harms is a reality;
- A rich understanding of people’s cultural ideals surrounding OBA, and the government’s capacity to meet those ideals;
- A set of seven design principles for building effective collective civic harm-reporting systems, generated from those user stories; and,
- A preliminary application of these principles to evaluating existing systems in other domains.

5.2 Related Work

As I found in Chapter 3, the demands that users generated and voted for—reparations for institutional privacy violations, modifications to the algorithms powering targeted ads, and formal apologies from offending institutions—were summarily dismissed by security and privacy experts across industry, academic, and policy environments. Specifically, experts noted that it could be difficult to directly attribute harm and malintent to violating institutions. This attribution problem is a core issue in the broader landscape of privacy harms, [69, 70], wherein harms from privacy violations are currently inconsistently recognized by courts. For example, flashier privacy violations, such as the Equifax data breach in 2017, have resulted in heavy fines from the FTC, in part because evidence of resultant harms on users is fairly easy to understand and define: financial losses for both institutions and consumers. However, this evidence is collected in an ad-hoc manner only once a data breach has come to light rather than as a *sustained* effort to support cases against *ongoing* violations and harms.

We argue that designing formal, sustained systems to gather evidence of these privacy harms on a *collective* level, echoing [109], is key to lending further legitimacy to them in both legal settings and beyond. Relatedly, Rakova et al. [110] have proposed a social imaginary framework, Terms-we-serve-with (TwSw), for addressing broader *algorithmic* harms, specifically through a lens of algorithmic *reparation*. The authors specifically highlight “complaint and algorithmic harms reporting” as one of five key dimensions of this framework, which can not only help us address existing current issues of harm, but also *anticipate* them and prepare for them in the future. We build upon this dimension of TwSw through our design fictions that we discuss in detail in later sections.

The mundanity of **online behavioral advertising (OBA)**, as I described in Chapter 4, makes it a unique case study for exploring how to report privacy harms as they manifest in-situ. Despite the challenge of OBA being a form of slow violence, people are still motivated to take small actions against OBA in the form of preventative measures—they use ad-blockers, VPNs, and private browsing sessions, and even avoid having real-life conversations near their personal mobile devices to avoid being “listened to”. Gorski [104] and later Korff [111] both have argued that in cases where it’s difficult for plaintiffs to provide evidence of surveillance due to the surveillance measures themselves being secret—such as in cases of national intelligence—it can be easier to demonstrate a “diversion of time or resources” instead. Both the harms themselves and the actions to evade them create a rich well of evidence to be gathered. In this work, we focus on gathering and reporting the former.

5.3 Design Context and Approach

While OBA harms might be amenable to a structured system for civic reporting, the design space for such a system is still broad and warrants additional exploratory work. In this section, we describe a rich landscape of collective civic harm-reporting in other domains, and why such existing systems might still be inadequate. We then propose a mixture of

design fiction methodologies to address the gaps of this context. Finally, we detail how we developed our design fictions.

5.3.1 Design Context: GoTCHas

We situate our work in a growing landscape of government-based civic reporting and evidence-gathering platforms in other domains within the United States of America. We term this class of systems as ***Government Tools for Civic Harm-reporting (GoTCHas)***.

At the broadest level, government agencies in the United States often solicit comments from the public on matters of legislation or to understand how to prioritize their resources, through the cross-agency federal website Regulations.gov. However, to the average person, this website is virtually unknown, and even if an individual has heard of it, it can be overwhelming to navigate: the home page features a catch-all search bar with few affordances for how to find a relevant topic to comment on, and consequently, thousands of search results. And, if a person has somehow found a relevant topic, there is no guidance for how to write a public comment, what information to include, or the format of the submission. As such, these public comments—which might not even be relevant to reporting harm at all—are primarily authored by policy professionals and experts, either as representatives of large corporations or from the federal government itself.

Two existing platforms for *non-expert* people to report Internet harms and evidence of those harms come from the Federal Trade Commission (FTC), the primary regulator of privacy and data security within the United States. The FTC operates ReportFraud.ftc.gov and IdentityTheft.gov, for people to report “fraud, scams, and bad business practices” and “report identity theft and get a recovery plan”, respectively. On both websites, consumers are guided through a questionnaire about the type of complaint they are submitting, and then asked to include specific documents to support their complaint. On ReportFraud.ftc.gov, the FTC claims it will share fraud reports with law enforcement agencies and provide helpful tips for consumers to protect themselves from fraud in the future. On IdentityTheft.gov,

there are tutorials for consumers to create plans and checklists for the recovery process, including in cases of data breaches. The FTC also publishes aggregated infographics of trends in fraud reporting, as well as whether consumers got money back from the resolution of those fraud cases.

Another well-established example of a civic reporting tool comes from the Consumer Financial Protection Bureau (CFPB), which operates a website for consumers to submit complaints about financial services and products, such as bank accounts, credit cards, mortgages, and other types of loans. Similar to the FTC, consumers on the CFPB complaints website fill out a guided form and can attach specific types of supporting evidence for their complaints. The CFPB claims that it will submit the complaint to the offending company on behalf of the consumer, or to another appropriate federal agency for a response usually within 15 days, but up to 60 days. On the opposite end, the CFPB publishes a publicly-accessible, anonymized database of these complaints that consumers submit, with an interactive visualization dashboard and API to download the data. The CFPB database goes one step further than the FTC's reports, however, and allows consumers to read and download individual complaints that others have submitted, including their anonymized personal accounts and associated evidence.

Examples from other domains include the Federal Communications Commission (FCC) Consumer Inquiries and Complaint Center, which, in a similar manner, accepts consumer complaints about telecommunications services like internet service and television broadcasts. The FCC's website also provides a form for privacy complaints, but these are only limited to those related to privacy concerns about internet or telephone service providers. The Occupational Safety and Health Administration (OSHA) also operates a similar complaint website for people to report workplace health and safety issues, with a publicly accessible database. At the local level, municipal governments often operate 311 data portals where residents can submit complaints about non-emergency incidents, such as noise com-

plaints, road blockages, streetlight outages, etc.; these portals often have also have publicly accessible, downloadable databases of these complaints similar to the CFPB.

While these existing GoTCHas can all claim receipt of (and, to an extent, addressing) millions of complaints, the percentage of Americans who felt they could trust the government to “do the right thing” has not surpassed 30% since 2007 [107]. The context of privacy offers a similarly bleak outlook: only about 3 out of 10 Americans feel that the government will hold the CEOs of social media companies accountable if they misuse or compromise users’ data, and 6 out of 10 are skeptical that the actions they take to protect their privacy will make any difference [112]. These issues of mistrust can persist because existing GoTCHas take what Corbett and Le Dantec [108] argue is a “traditional” HCI approach of defining trust as a cognitive-based decision based on transparency and information—e.g., offering detailed timelines and accessible open data to instill trust—when a more nuanced, relational approach to trust is necessary with civic technology [113].

5.3.2 Design Approach

To support a more nuanced, relational approach to trust in GoTCHas, specifically with regards to privacy and OBA, we hope to answer the following research questions:

RQ1 What are design requirements for people to participate in collective civic reporting of privacy harms from OBA?

RQ2 What cultural ideals do people have around redress for reported privacy harms from OBA, and the parties that carry out this redress?

Because the senses of resignation, powerlessness, and pessimism that people can have about GoTCHas and OBA can feel nearly intractable, we employ a mixture of design fiction methods to answer the above RQs, in hopes of encouraging people to feel more creative about the possibilities of GoTCHa design. As Bleecker [114] first introduced, fiction can “be a purposeful, deliberate, direct participant in the practices of science fact” that allows

us to understand, explore, and question alternate futures. Within the context of user security and privacy, especially in challenging the hegemony of large tech institutions, there has already been some design fiction work. For example, Wong et al. [115] employed design workbooks—collections of conceptual designs—to first explore questions about user privacy stemming from a science fiction novel. More recently, Møller et al. [116] also explored the lack of agency of unemployed individuals over consent to data sharing through a fictitious job placement app that collects highly invasive personal data.

In the realm of design fiction, we employed a combination of three participatory design approaches in our study: **fictional inquiry, story completion, and comicboarding**. As an immediate descendent of design fiction, *fictional inquiry* is the practice of using partially fictional settings, artifacts, and circumstances to create a space for conducting collaborative design activities [117]. In this space, people are urged to imagine desirable futures and consider their everyday impacts. *Story completion* draws parallels with design fiction, asking participants to write or complete a fictional story given a hypothetical seed scenario [118]. With the story completion method (SCM), people can share their thoughts about an idea without the burden of imagining or inserting themselves into the situation, allowing them to be more imaginative with the possibilities of the design, and be encouraged to think outside of immediate impacts on their own lives to how different scenarios affect other people. Within HCI, SCM has been employed to identify potential thematic futures around human-VR pornography [119], human-robot [120], and human-AI voice assistant [121] interactions. To support these two approaches, we adapted *comicboarding*, a co-design method that provides a structure of comic strips and partially-completed content as a scaffold for users to come up with novel ideas. Comicboarding can be especially helpful in cases where people are not accustomed to brainstorming [122], e.g., when users feel powerless about their privacy.

In this work, we created two comicboards illustrating a fictional government tool that supports collective civic reporting of privacy harms from OBA, with an empty panel in each

board, where participants could write fictional stories about the tool. We hoped this format could sufficiently introduce complex technical concepts, but also be open-ended enough to elicit rich design requirements and cultural ideals for both this fictional tool and its impacts.

5.3.3 Comicboard Development

We began our comicboard development by brainstorming textual descriptions of the ways users could interact with such a tool, eventually consolidating these descriptions into two roles that lay-users could play in the life-cycle of a collective civic harm-reporting system. We loosely based these roles on a composite of the first few stages of the data management life cycle [123], such as data generation and collection, as well as cleaning and processing. We chose to focus on these roles as they were more likely to be accessible to regular users or easily teachable, as opposed to roles like data analysis, interpretation, storage, and sharing, which might require additional levels of technical expertise and access.

We viewed the first role, data generation and collection, as a reflection of existing GoTCHas that we noted in the previous subsection. The *fictional* component of this role is that the complaints filed are related to *privacy harms, specifically in the context of OBA*, rather than other existing supported violations in other domains. We propose a second role of data cleaning and processing as a natural augmentation of GoTCHas. Platforms that support crowd-worker-based annotation tasks—e.g., Amazon Mechanical Turk and Prolific—are already prevalent in the research community, but have not been widely explored as a way to increase citizen engagement with GoTCHas. Prior work [100] has suggested that giving people a small stake in a collective process of voicing demands—e.g., through responding to others’ concerns and quality-checking others’ contributions—in response to privacy harms can instill a sense of empathy in them for strangers, and make them feel more impassioned about the subject. The broad fictional design that we will explore, then, is the linkage of these two roles within one new system.

We initially described these roles through short text summaries, following Jin et al.'s [124] recommendations for low-cost privacy assessment methods; however, through initial pilot testing, we found that pilot participants were quickly bored or overwhelmed due to having to keep track of a number of stakeholders and technical concepts. Through a series of discussions as a research team, we pivoted to create rough drawings of how fictional characters might inhabit these roles, as well as short captions for these sketches. With additional rounds of pilot-testing, we further refined the content of these drawings and captions so that they would lead to responses relevant to our research questions, while remaining both easily legible, and open-ended enough for rich responses.

A key tension in our design process, as aforementioned, was balancing the dense technical nature of concepts like algorithmic ad delivery systems, privacy harms, and responsible regulatory bodies, with our goal of eliciting creative, imaginative stories from our participants. For example, we eliminated references to official government agencies like the FTC, because they required too much exposition and could confuse and overwhelm participants. Instead, we referenced “the government” broadly, to evoke an institutional authority that could lend the tool legitimacy.

5.3.4 Comicboard Content

We developed two sets of comicboards, each with four panels, that describe the two potential roles that lay-users could play in the life cycle of a GoTCHa for user privacy.

The first set of panels (herein referred to as the “contribution” board) addresses the role of data generation and collection. Specifically, it follows the story of a character named Alex, who is concerned about their privacy in relation to OBA, and is told about a tool from the government where they can report their violating experiences with it. Subsequently, Alex decides they want to contribute a report of their experiences to this tool. The final fourth panel does not contain an image, and asks participants the question, “What happens

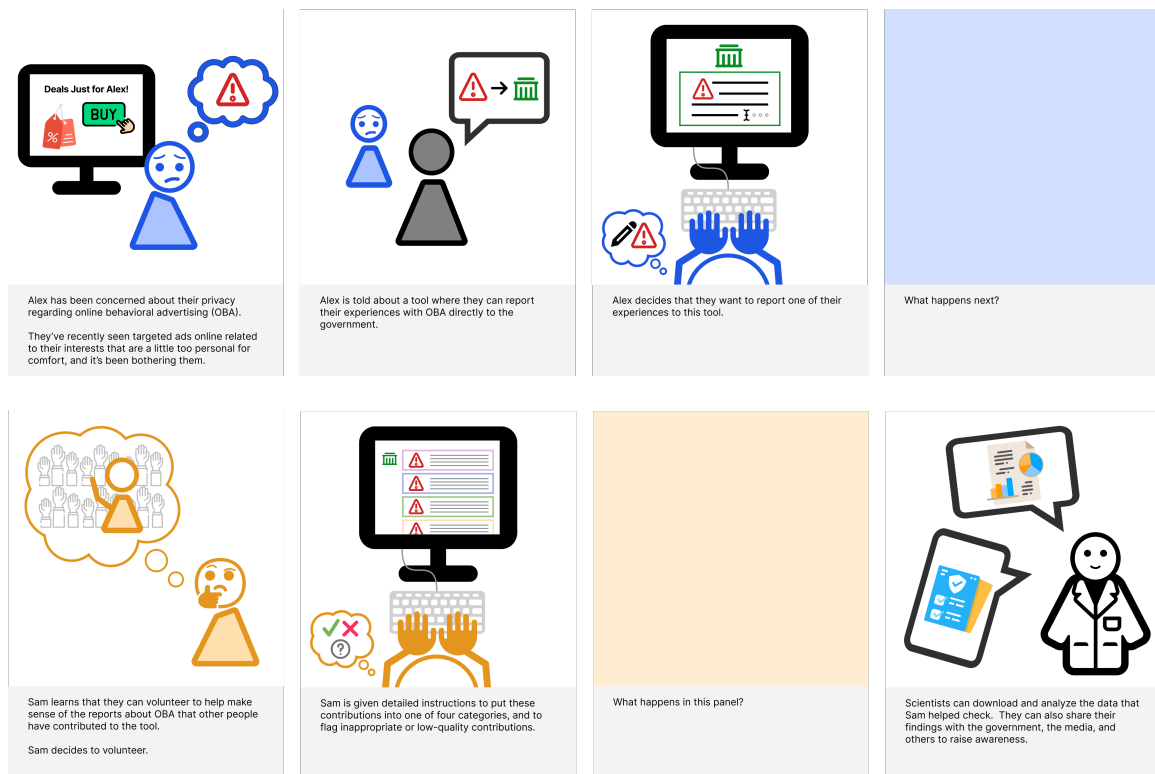


Figure 5.1: The final two comicboards we developed. We refer to the first comicboard as the "contribution" board, and the second as the "annotation" board.

next?" We left the *final* panel empty because we wanted participants to consider all the unknown possibilities for Alex's contribution.

The second set of panels (herein referred to as the "annotation" board) covers the role of data cleaning and processing. Set in the same fictional world as the contribution board, the annotation board follows the story of Sam, who learns they can volunteer to make sense of the reports that other people have contributed to the tool. Sam decides to become a volunteer. The third panel does not contain an image, and asks participants the question, "What happens in this panel?" The final fourth panel explains that scientists can access the data the Sam helps annotate and clean, and can use it to create analyses and reports that will be shared with others to raise awareness. In this orange board, we purposefully left a panel empty in the *middle* of the story because we wanted participants to consider how Sam's volunteering could be directly tied to the work of privacy experts and policy professionals, and how they fit into a process with many stakeholders.

Both contribution and annotation boards can be viewed in Fig. 5.1.

5.4 Study Design

We presented our comicboards in an online survey instrument deployed on Qualtrics and shared on Prolific, a crowd-work platform. We recruited a total of 50 participants for the main study to share design requirements, cultural ideals, and fictional stories about collective civic reporting of OBA harms.

5.4.1 Recruitment, Ethics, and Compensation

We first screened 200 participants by asking them if they were aware of online behavioral advertising (OBA) or online targeted advertising. These potential participants were also asked to confirm they were adults located in the United States and fluent in English. This screener took on average less than a minute to complete, and participants were compensated 0.25 USD on the Prolific platform. Out of the participants who indicated they were aware of OBA, we recruited 50 to participate in our main study. (Participant demographics are reported in Table 5.1). The main survey took on average 19 minutes and 33 seconds to complete, for which participants were compensated 9 USD. Our study was approved by an institutional review board.

5.4.2 Survey Instrument

There were four main components to the survey instrument: a comprehension check, the first contribution comicboard, the second annotation comicboard, and a summary of all comicboards.

First, we provided participants with detailed instructions on the context of the study, explaining that they were about to read and see images related to a fictional world, and that they would be asked to write stories about this fictional world. We also explained that participants could be as creative as they wanted, and there were no right or wrong ways

to write these stories. We then asked participants to confirm their understanding of these instructions.

Then, we showed participants the contribution comicboard (Fig. 5.1), involving a fictional character, Alex, and asked them to share any features or information they felt was necessary for Alex to make a successful contribution to the government tool. We then asked participants about any hopes, concerns, or other thoughts Alex had about the contribution process or the tool itself. Finally, we asked participants to write a short story completing the empty panel in the comicboard. We placed the story completion task after the first two questions to prime participants to include richer details in their final stories.

Third, we showed participants the annotation comicboard, involving a second fictional character, Sam (again, see Fig. 5.1). Participants were asked about any obstacles Sam might encounter in annotating others' contributions, as well as what Sam thought about the other people who've interacted with the tool or will do so in the future. Then, once again, we asked participants to write a short story completing the empty panel in the comicboard.

Finally, we showed participants all of the comicboards they had seen previously in the study, as well as all of the responses and stories they had written. Participants were asked about the extent to which they would like to live in the fictional world they wrote about, and to explain why or why not, similar to past work [119, 121].

5.4.3 Analysis

Similar to prior comicboarding and storyboarding studies in HCI [125, 126], we intentionally analyzed participants' responses *across* storyboards and questions, rather than analyzing specific responses to each question or per story written. As with AI systems [126], systems supporting user privacy, especially in the context of privacy harms, always have multiple, interconnected stakeholders and involve countless intertwined decisions. Our comicboards themselves were self-referential and inherently related, and we specifically asked participants to reflect across all comicboards.

Table 5.1: Demographics of Prolific participants in the main study.

Demographic	Group	n
Age	18-24	10
	25-34	21
	35-44	14
	45-54	2
	55-64	2
Gender Identity	Female	24
	Male	26
Ethnicity	White	41
	Black	5
	Asian	3
	Mixed	1
Employment Status	Employed Full-Time	17
	Employed Part-Time	4
	Student	8
	Unemployed	4
	Homemaker, retired, or disabled	3
	Not reported	19
Total		50

We adopted a reflexive approach to thematic analysis. Two members of the research team first conducted open coding on a sample of 20 participants’ responses and generated a total of 75 initial codes. Then, the two researchers discussed disagreements and consolidated codes, agreeing on a codebook with 19 codes and applying this codebook to the full dataset of responses. (Cohen’s κ values for these codes ranged from 0.420 to 0.916, with a mean of 0.700 and a mean standard error of 0.049.) Finally, we generated broader themes from these codes with the rest of the research team. We present these themes in the following section.

5.5 Study Findings

We group contents of participants’ responses into five key themes. The first theme, the **UI or UX elements** of the tool, refers to stories that referenced the visual elements of the tool, specific types of evidence that should be supported or ingested by the tool, information about the tool or the submission generally, and any other features participants felt

the tool should have. The second theme, **post-contribution expectations**, deals with stories that mentioned Alex's ideals for their experience after contributing a report, including speed, transparency, and accountability, as well as specific design elements related to those factors. The third theme, **costs of contributing**, covers stories that mentioned downsides of the tool entailed by both contributing and annotating; the fourth theme, **benefits of contributing**, deals with the opposite. Finally, the fifth theme, **outlook on the future**, explores participants' optimism or pessimism about the tool and its place in society.

5.5.1 UI and UX Elements of Tool

A vast majority (N=40) of participants mentioned specific evidence of harm that should be supported or ingested by the tool in their stories and responses, which included both details about the specific harmful targeted ads, as well as *why* those ads were harmful. For example, as a start, as P43 summarized, "*Alex would need to be able to provide images of the advertising he is getting as well as images or just information about his...personal life and how this advertising is infringing on it.*" Participants also brought up direct links to the ads, as well as information about both the advertiser and the platform where the ad was served. Finally, some participants also noted that Alex would enter personal information about themselves, such as a Social Security Number (e.g., P36), browsing and search history (e.g., P23), and contact information so they could receive updates about how their contribution was being processed (more on this in the following subsection).

To support all these types of information, some participants (N=10) mentioned specific visual elements of the tool, including the method for delivering their contribution and associated evidence. For example, while P47 mentions that Alex would submit their contribution by email, other participants suggested that Alex would type into an online form with a large longform text box. Participants also mentioned other modalities for submissions, such as a mix of both multiple choice and free responses to ensure uniform data, as well as places to upload images, along with "*basic features found in any word processing*

application” (P7). For the actual process of writing up a contribution, participants hoped Alex would be educated about the proper terminology to use in their contribution (P21), guidelines on word length and level of detail (P33), as well as tips for what Alex should do in the meantime while their contribution was being processed (P44).

However, a few participants also noted that the tool, as built by the government, would likely have poor documentation and guidance for appropriate submissions: *“There wasn’t a lot of detail on the website to guide [Alex] through the process”* (P14). As a similar dig toward government-based technology, P21 wrote,

“Alex went to the website and was surprised to find that the front page of it looked fairly modern. However, when they went any deeper than the front page, the website looked like it was designed in 2002. They were able to figure out how to lodge a complaint, but they are not confident that it will accomplish the goal set forth. There was a lot of legalese on the website. It made it confusing at times, and they are not sure that they got their point across entirely.” —P21

Finally, a few participants also expected that the tool would provide additional features, such as periodic security and privacy tips (P10, P18, P35), background monitoring of the ads Alex saw, ad-blocking capabilities integrated directly into the tool (P20, P37), as well information about local community groups Alex can join for more involvement (P25).

5.5.2 Post-Contribution Expectations

Several participants (N=33) directly wrote in their stories about Alex’s expectations for the tool and the government after they had made their contributions, primarily focused on three qualities: speed, transparency, and accountability.

Eleven participants referenced some sort of **speed** element (or lack thereof) related to Alex’s expectations of feedback from the tool, ranging from immediate feedback; to waiting a few weeks, months, or years; to never hearing back at all. Those that wrote about

Alex experiencing longer wait times or never hearing back also mentioned that while Alex expected this from the government, they were still disappointed and resigned. For example, as P41 wrote, *“Nothing seemed to come of his report. Two months later Alex noticed really no difference and just gave up on the whole thing, realizing it doesn’t matter and everybody already has all his information anyways.”* And, as P19 succinctly wrote, *“Alex couldn’t believe he waited 6 months for nothing.”*

Relatedly, participants also expressed concern about a lack of **transparency** about what would happen to their contribution after they finished submitting it. Specifically, 8 participants brought up concerns over these contributions being collected for nefarious purposes, coupled with a lack of reassurance from the government. For example, P5 noted that Alex might not feel confident submitting so much personal information to a government website, and had *“feelings of insecurity over whether or not the tool itself is going to invade their privacy in some way”*; similarly, P45 wrote that Alex might fear retaliation for reporting, and did not have enough information to feel assured.

In a similar vein, participants wrote about a lack of **accountability** from the government after Alex submitted their contribution, particularly via their perceptions of the government as impersonal. For one, a few participants felt concerned that Alex’s contribution would be dismissed since the harms Alex’s reported from OBA were not severe enough to warrant further investigation. As a step further, both P16 and P4 mentioned that the government might use opaque automated tools to filter out contributions, so Alex’s contribution might not even be read by a human. P16 also noted that there was little recourse for this automated decision, and Alex couldn’t appeal the government’s decision to dismiss their contribution.

Participants’ expectations of a lack of speed, transparency, and accountability from the tool and the government also presented as an overall skepticism of the government’s abilities to demonstrate concrete help. As P44 wrote, *“The government is very large and is known for moving slowly; though they might be working as fast as they can, it might not feel*

that way to Alex.” As an escalation, participants also wrote about how a negative experience could influence Alex to permanently distrust the government: “[The tool suggests] to run some safety programs to help protect [Alex’s] data, but no word is given on taking action against the advertising companies. Frustrated, Alex follows the suggestion but vows to never trust the government to help his problems again” (P14).

As a solution to these concerns, several participants highlighted the importance of “a place on the tool to leave their information so they can be gotten back to by whoever runs the tool” (P41), i.e., personal contact information for further follow up. Participants mentioned that the tool should communicate to Alex an expected timeline for receiving a response from the government, and include features such as ticket numbers and progress bars to provide transparency. To support accountability, a few participants also mentioned that being provided with the contact information of an actual human government representative would make them feel more at ease about the legitimacy of the tool.

5.5.3 Costs of Contributing and Volunteering

A large majority of participants (N=43) also mentioned costs or downsides associated with contributing and volunteering. These costs were largely associated with the annotation role, possibly because our participants were crowd-workers themselves and were especially attuned to the struggles that online research participants might face. From the perspective of Sam, the volunteer annotator character in the orange board, participants (N=42) primarily wrote about two costs: volunteer fatigue and uncertainty over correct annotation.

Fatigue from crowd work and the psychological harms of online content moderation have been well-documented in existing literature [127, 128], as has the power imbalance between the researchers who solicit annotations and the annotators themselves [129]. Our participants gave responses that mirror these prior findings; P1 paints a particularly grim picture of Sam’s life:

“Sam doesn’t know why he volunteered for this, this is more difficult than he thought—he wished this was a paid position. The scientists keep the volunteers locked in a office until they complete so many reviews. Sam is trying to find a way out so he can escape this weird and terrible place. Sam vows to never volunteer for anything ever again, he will only be paid for his work.” —P1

Participants also frequently brought up the banality of the task, and how low quality or incomprehensible contributions could make it difficult for Sam to concentrate on the task. P25 also outlines two challenges that Sam might face as annotator:

“The amount of time that it takes: it ends up becoming a major drain on him due to the amount of time that he invests into it. The toll that it takes flagging inappropriate content: the things that he sees are scarring and end up causing problems for him.” —P25

The nuanced nature of these contributions also meant that some participants (N=21) felt Sam would feel pressured to annotate and check everything perfectly, and would be worried about making a mistake, especially if Sam is not given adequate training and instructions. For example, P28 writes: *“picking out the good contributions is no problem, but the ones on the fence of helpful or not could be a tough choice....[Sam] could go home concerned with some of his choices.”*

The potentially sensitive content of these contributions also drew mentions of costs from the perspectives of both Alex and Sam. As mentioned in the previous subsection, participants felt that Alex might be concerned about contributing private information that might be misused for nefarious purposes (especially if exposed in a data breach and the consequent risks, as P47 notes). And, on behalf of Sam, participants noted there might be pressure to maintain the privacy of the contributions they read:

“Sam also needs to balance the need for user privacy with the requirements of a proper investigation, which can be a delicate and complex task. Lastly,

addressing concerns and offering follow-ups to submitters who wish to remain anonymous presents its own set of challenges, as contact options may be limited.” —P45

Finally, on a more selfish level, some participants felt that Sam simply might not want to know that much information about other people’s lives. For example, P33 writes that there is so much detailed personal information in the contributions that Sam encounters that they feel like they are invading others’ privacy just by reading, let alone annotating.

5.5.4 Benefits of Contributing and Volunteering

While most participants (N=42) mentioned some sort of perceived benefit for either Alex or Sam in contributing and volunteering, these benefits were largely abstract. Participants noted both collective abstract benefits (N=30) and individual ones (N=21) for Alex and Sam. Collective benefits that participants mentioned included a sense of community with others, broad data privacy laws that support user control, and a general wish for “collective good”, rooted in an ideal that the government will listen to the people and have their interests in mind. In particular, being able to read others’ contributions, as Sam does, injects a sense of empathy in Sam:

“[Sam] comes across Alex’s submission and really can tell that Alex is disturbed by how personal his ad targeting is. Sam sorts the data as instructed, hoping that Alex (and everyone else who submitted) will get some kind of closure with the issue, because the ad targeting has just gotten out of hand.” — P49

Individual benefits comprised of a general sense of “doing the right thing” and feeling good about oneself for helping others. A few participants also envisioned a future where Alex becomes a privacy rights activist, seeded by Alex’s initial contribution to the tool. For example, as P1 writes, *“He doesn’t know it yet, but reporting this one website sets him on*

a path to get involved in politics and laws in the future. He also gets involved with disputes of AI in the online world.”

Other participants wrote about the sense of satisfaction that Sam might derive from putting in hours of work volunteering: *“After a few more hours of work, the pride she felt when she sorted her last submission was incomparable. She knew she did something that mattered, and that made her happy”* (P42). While this satisfaction and pride was largely associated with simply volunteering at all—e.g., *“Sam feels he has a duty to help the people”* (P49)—a few participants characterized Sam as finding new purpose in their role as a volunteer, and being extremely dedicated to optimizing and improving the annotation process. For example, P14 described Sam’s elation after devoting hundreds of hours to building an automated annotation tool: *“These long hours seemed to have finally paid off, with the most recent pass-through generating a nearly 99% success rate!”*

Only 15 participants mentioned possible concrete or immediately tangible benefits from the tool. These primarily consisted of seeing fewer or zero ads or monetary compensation for exposure to harmful ads. While a few participants constructed a fictional world where Sam was compensated for annotation, either through money (P27) or small trinkets (P36), several participants included some variation of *“Sam wishes they were being paid for annotation”* in their stories, evoking aforementioned costs associated with volunteer fatigue.

5.5.5 Outlook on Future of Tool and OBA

Beyond desired or anticipated costs and benefits of the tool, participants made several meta-level observations about the future of the tool, including Alex and Sam’s views on other contributors, and general outlooks on society.

Fewer than half of participants (N=22) felt optimistic about the future of the tool and its place in society; only 18 participants responded that they would like to live in the fictional world they had written about. These participants shared that they would appreciate a world in which the government listens to their privacy concerns and takes actions in response.

Echoing the previous subsection, participants also expressed gratitude toward other people who took the time and effort to care for each other, and noted that this gratitude could be a motivation for continuous contribution to the tool. As P32 remarks, “*the people in the stories are much more considerate than people are in real life.*” P13 hypothesized that in the future, there would be fewer contributors to the tool because there would be fewer problems to report, due to the success of the tool.

Overall, however, participants tended to express pessimism about the tool’s abilities to change anything in response to OBA harms, especially from Sam’s perspective after reading through others’ contributions. In particular, many participants felt that reading others’ contributions made Sam feel even *worse* about the future of OBA harms, since there was no evidence of positive changes. For example, as P16 exemplifies:

“Sam begins with a lot of gusto and determination but soon finds that the complaints are all quite similar. Some are superficial and of not much concern but others are extremely personal and Sam feels embarrassed for these people he is learning about. He really wants to help them but he realizes that the same type of complaints that he submitted are submitted by the thousands and nothing seems to be changing. He realizes that this is just another way for those in power to get more information from everyday people to use against them.”

—P16

Relatedly, participants also expressed negative judgments about people who had contributed to the tool. For example, some participants felt that people who contributed reports were delusional and wasting their time: “*Sam thinks people are generally using the tool like its going to cure a virus. That’s not the case at all and Sam feels bad for those people*” (P41). These negative perception extended to participants’ views of the annotators as well, who might purposefully submit low-quality annotations because they felt contributors were unintelligent. For example, P27, who wrote that Sam thought contributors were “*id-*

lots for complaining”, also described Sam as uncaring and motivated solely by monetary compensation to complete their annotations:

“Sam decides to randomly categorize the results, using an ad hoc algorithm based on the first letter of the person’s complaint. Since he is paid per complaint categorized, he is able to process the complaints quickly, and make much more money than if he completed the task as intended.” —P27

These negative outlooks were compounded by a handful of participants (N=7) who felt that the fictional world represented in the comicboards was boring and incomplete, as *“the only thing this world seems to focus on is internet privacy and OBA. I do believe these topics are important, but I want to live in a world where there is more than that”* (P46). As P6 wrote bleakly, *“There’s nothing but computers and scientists.”*

5.5.6 Overall Findings

We summarize our overall findings in references to our research questions below. Table 5.2 also presents how our RQs relate to the five qualitative themes we detailed above.

RQ1: What are design requirements for people to participate in collective civic reporting of privacy harms from OBA?

Participants had clear conceptions of what a fictional tool for privacy harms from OBA would look like, heavily guided by their prior experiences interacting with government websites. They envisioned support for multiple kinds of ad-specific data, detailed guidelines for submission, and clear communication about feedback and claim statuses; along the way, they wanted assurance for protecting sensitive personal information. Participants also highlighted a need for greater protections and benefits for volunteers (or perhaps doing away with the volunteer nature of the tool altogether): they frequently mentioned immediate compensation or tangible benefits directly tied to contributing or annotating, as well as

Table 5.2: Summary of qualitative findings as related to our research questions.

Theme	RQ1: Design Requirements	RQ2: Cultural Ideals
UI and UX elements of tool	<ul style="list-style-type: none"> • Ingesting details related to specific harmful ads, rather than OBA as a generally harmful phenomenon • Detailed submission guidelines 	<ul style="list-style-type: none"> • Interim assistance while contributions are being processed (e.g., security and privacy tips)
Post-contribution expectations	<ul style="list-style-type: none"> • Evidence of progress or success • Form for leaving contact information • Timelines, ticket numbers, progress bars 	<ul style="list-style-type: none"> • Speed • Transparency • Accountability • A human reading and responding to contributions
Costs of contributing	<ul style="list-style-type: none"> • Protections for volunteers (e.g., fatigue, sensitive contributions) • Protections for interpersonal privacy risks (e.g., reading others' contributions) 	<ul style="list-style-type: none"> • Trade-offs between requiring highly personal information and convincing people to contribute • Grassroots versus government oversight of tool
Benefits of contributing	<ul style="list-style-type: none"> • Immediate compensation (e.g., money) or tangible benefits (e.g., a coffee mug) 	<ul style="list-style-type: none"> • Duty to help others • Pride in volunteering • Sense of community
Outlook on the future	<ul style="list-style-type: none"> • Insurance against bad actors within collective of contributors 	<ul style="list-style-type: none"> • Privacy harms from OBA are only secondary concerns

support for volunteer fatigue. Finally, participants wanted internal safeguards against bad actors who might disrupt or slow down the reporting process.

RQ2: What cultural ideals do people have around redress for reported privacy harms from OBA, and the parties that carry out this redress?

While participants noted that privacy harms from OBA were not necessarily at the forefront of their concerns, they also felt that the tool could inspire a sense of community and duty to help others around them, especially via the opportunity to read through others' reports. Consequently, people might find pride in volunteering their time and efforts to the tool and feel encouraged to participate in privacy activism more broadly. However, these abstract positive feelings alone might not fully offset potential costs of contributing: in particular, participants worried that the fictional tool could just be another way for their personal data and labor to be exploited by the government. Relatedly, participants expressed clear expectations for how they hoped the government would treat their concerns: they wanted speedy but personal responses that did not leave them feeling like just another number in a pile of reports.

5.6 Design Principles for Instilling Trust in GoTCHAs

While our design fictions were specifically tied to online behavioral advertising and privacy harms, participants frequently spoke about their ideals for the government's operational capacity to help people more broadly when referring to the fictional tool. Their responses give us valuable insights into what people think about reporting and processing not just *privacy* harms, but also *consumer* harms more broadly. In this section, we present seven design principles for instilling trust in GoTCHAs, derived from the five themes that we found in our analysis: **(1) visible, upfront benefits; (2) timely, useful feedback; (3) contestability; (4) error prevention measures; (5) integration into everyday life; (6) consideration of social influence; and (7) commitment diversity and flexibility.** Several of these prin-

Table 5.3: Initial evaluation of existing GoTCHas using our seven design principles.

✓: Satisfies this design principle.

✗: Does not address this principle at all.

†: Partially satisfies this principle.

*: Design principle is based on a Nielsen heuristic.

Design Principle	Regulations.gov	FTC (Fraud)	FTC (Identity Theft)	CFPB	FCC
<i>Visible, upfront benefits</i>	✗	✓†	✓†	✓†	✓†
<i>Timely, useful feedback*</i>	✗	✓†	✓†	✓†	✗
<i>Contestability</i>	✗	✓†	✓†	✓†	✓†
<i>Error prevention measures*</i>	✗	✓	✓	✓	✗
<i>Integration into everyday life</i>	✗	✗	✓†	✗	✗
<i>Consideration of social influence</i>	✗	✗	✗	✗	✗
<i>Commitment diversity & flexibility</i>	✗	✗	✗	✗	✗

principles draw upon Nielsen’s 10 usability heuristics for user interface design [130], widely regarded in the field of human-computer interaction as a canonical reference for low-cost evaluation of design choices. In the following subsections, we define each of our principles. We then apply these principles in a preliminary evaluation of a limited sample of existing GoTCHAs—Regulations.gov, the FTC fraud reporting site, the FTC identity theft reporting site, and the respective complaint sites for the CFPB and FCC—summarized in Table 5.3.

5.6.1 Visible, Upfront Benefits

Several participants noted that they would be unmotivated to report harms in real life, especially since it was difficult to see how their contributions would directly lead to a positive outcome or concrete benefit. The GoTCHAs we evaluated do little to strengthen this tenuous link. The FTC’s identity theft site, for example, clearly tells people the immediate next steps of reporting cases of identity theft: after submission, the system can create a case for the reporter, and the reporter will be given a checklist of actions they can take on their own while their case is being processed. However, it’s not clear what outcomes a reporter might expect from the process; while the checklist is helpful, it does not directly address any damages or policy changes people might expect from reporting. The FTC’s fraud site,

on the other hand, informs people from the outset that they will not resolve their individual case, but hints that people might be able to recoup financial losses from falling prey to scams later on. These relationships—between reporting action and expected outcome—are further complicated by the possibility of cases being delegated to third-parties who may or may not fully address people’s concerns (e.g., law enforcement, other government agencies, external companies) . To motivate contributions, then, GoTCHas should present and communicate **visible, upfront benefits** to people: examples might include descriptions of financial remuneration in prior resolved cases, discussions of how these contributions are being concretely used to put together policy briefs, or access to online tools and resources while contributors wait for a response.

5.6.2 Timely, Useful Feedback

Our participants highlighted the importance of hearing back about their contributions in a reasonable time frame as a bare-minimum requirement for their trust in the fictional tool’s ability to help people. A guaranteed quick response might get Alex to consider contributing, but they certainly wouldn’t get their hopes up for one, much less expect the response to be actually helpful. But the response needs to be not only quick, but also personal and informative about what exactly was happening with Alex’s contribution—not from an automated no-reply address. We liken this requirement to Nielsen’s [130] heuristic of “visibility of system status”—which argues that when users can understand a system’s state, they feel in control and trust in the system’s ability to do what they ask it—and propose **timely, useful feedback** as a necessary characteristic of GoTCHas. Currently, none of the GoTCHas we evaluated fully satisfy this principle. While a few make commitments to timeliness (e.g., within 60 days) and inform people of what they might *expect* in a response, there are no guarantees of this timeline, and the aforementioned risk of delegation to third-parties, who might not be beholden to these commitments, exacerbates this problem.

5.6.3 Contestability

Multiple participants expressed concerns about Alex receiving impersonal decisions from the government about their complaint that they could not appeal, e.g., “*an automatically generated email...saying that according to the standards of the government, the advertisements [Alex] was being shown were perfectly legal and there is nothing more they can do.*” Responses like these could further contribute to people feeling powerless to effect change, and may discourage them from contributing reports in the future. Giving people a space to respond to these potential risks and decisions can help them feel more confident in their choice to make contributions in the first place. Leaving room for **contestability** in GoTCHAs—inspired by Rakova et al. [110]—sets up an ongoing dialogue between people and the government and can signal a commitment on the GoTCHa’s part to pursue systemic changes, rather than ad hoc ones.

Relatedly, participants also worried about Alex’s contributions being used against them—i.e., if the detailed personal information in their contributions was used to target them in other ways. Participants forecasted a sense of betrayal and powerlessness if this were to happen. While it is nigh impossible to guarantee the security and privacy of a system [131], GoTCHAs can instill more trust in people by clarifying how they can be held accountable as data stewards. While all of the GoTCHAs in the sample link to their respective agencies’ privacy policies as a way to assure people that their privacy will be protected and their contributions will not be misused, there is little discussion of recourse if this data is violated in, e.g., a data breach. There is also little interactability with the privacy policy on the public’s part. Thus, none of the GoTCHAs we evaluated fully satisfy this design principle.

5.6.4 Error Prevention Measures

When participants wrote about Sam, the character in the annotator role, they frequently expressed how Sam could be frustrated by incomprehensible, irrelevant, or low-quality contributions. One immediate alleviation for this problem is **error prevention measures**—

drawn from Nielsen [130], again—that clearly define what kind of contributions a GoTCHa can ingest, with formatting built into the tool itself. Both FTC websites and the CFPB satisfy this design principle by including step-by-step, clearly-defined questionnaires to standardize the types of complaints people can submit. This results, in the case of the CFPB, in a tidy database of complaints that is easily searchable and accessible to the public; such a structured and comprehensible data output can also aid contestability. The FCC, on the other hand, fails in this principle: submission formats vary widely based on the type of complaint (i.e., phone, Internet, television, etc.), and the submission sites are more akin to open-ended email composition boxes rather than professionally-organized intake forms, with neither guidance on what details people should include in their submissions, nor limits on word count or file types to attach. Regulations.gov is equally open-ended, with only a small text box, file upload link, and contact information field; there are no affordances for what people should contribute, or to stop them from commenting on the wrong topic.

5.6.5 Integration into Everyday Life

Currently, all the GoTCHas we evaluated operate as standalone websites that require people exit their daily lives to visit if they want to report a harm. In other words, it's unlikely that that people will report harms unless they are individually aware of these platforms and especially motivated to do so, e.g., only in severe cases. Simultaneously, the FTC itself can primarily only respond to these “flashier” events, rather than setting up sustained efforts to the more-ubiquitous but smaller harms, due to being perpetually underresourced [132]. However, people's unrelenting exposure to these “small” harms can mean that they become habituated or numb to its effects over time, accepting them as a fact of life [18]. One way we might address this is by making *reporting* as accessible and mundane as the harms themselves have become, through **integration into everyday life**. According to a Statista poll¹, over a quarter of users in the United States use an ad-blocking tool when

¹<https://www.statista.com/topics/3201/ad-blocking>

browsing, which typically takes the form of an unobtrusive browser extension that runs in the background. We can imagine GoTCHas being combined with such relevant existing tools, as our participants suggested. Or, the GoTCHas themselves might run in the background in a similar manner and deliver periodic short questionnaires to people, perhaps through an Experience Sampling Method approach, taking a continuous “pulse” of harm measurements.

5.6.6 Consideration of Social Influence

Social influence can strongly affect people’s motivation to act, especially when uncertain [133]. If it is clear to people that many others report OBA harms, then people may feel empowered to do the same. In contrast, if it is clear that others largely ignore these harms, people may feel disincentivized to file reports themselves. Participants mentioned both positive and negative opinions of the other people who might use the fictional tool, which in turn affected their outlook for the tool and desire to use it. Many participants in our survey expressed solidarity and empathy with other users of our hypothetical reporting tool, contributing to a sense of community, while others expressed pessimism on considering the harms of OBAs or the perceived futility of relying on our reporting tool to address them. These reactions all prompt a necessary **consideration of social influence** when designing GoTCHas, especially if the goal is to encourage and sustain civic participation. Currently, however, none of the systems in our sample do much to address this design principle beyond simply *having* a publicly accessible database of consumer complaints that people can read. These systems could instead encourage community building outside of the GoTCHa, perhaps through offering contact points of local offices and advocacy organizations, connecting people with support networks, or supporting new and existing online communities dedicated to addressing these issues.

5.6.7 Commitment Diversity and Flexibility

Currently, while GoTCHas allow self-motivated individuals to access and download contribution data for personal use, the sole primary public-facing role that people can take on when interacting with these systems is that of a reporter or complainant. In other words, currently, if people want to interface with GoTCHas, they can really only do so through filling out the respective web form and waiting for a response. However, our participants described both Alex and Sam’s roles in the fictional world with varying levels of dedication and responsibility. Alex might contribute a report and forget about it, or they might feel emboldened to become a privacy legislation activist and recruit others to become contributors. Sam might annotate contributions just out of boredom, or they might devote hundreds of hours toward programming a more efficient annotation system to improve the performance of the fictional tool. Or, Alex and Sam might vacillate between any of these spectra of activity, or even decide to seek employment with the government and become policy professionals themselves. Where appropriate, then, GoTCHas should *formally* provide their users with the opportunity for **commitment diversity and flexibility**; we envision a plethora of roles that people could play in these systems, such as not only recruiters and developers (as aforementioned), but also analysts, modelers, and contesters of contributions [134], as well as broadcasters of eventual findings. This allowance for different types of public participation may also allow motivated contributors to help under-resourced government agencies process mass volumes of reports more quickly.

5.7 Limitations and Future Work

In this section, we discuss four possible limitations of our work that could serve as areas for future work in the design of GoTCHas: generalizability of GoTCHas; counter-data action; our design fiction method and its impact on participant outlook; and our participant sample and study context.

5.7.1 Application to Non-Government Harm-Reporting Tools

While in this work, we exclusively explored the concept of a *government*-hosted tool for citizen harm-reporting, we argue many of these design principles can apply to tools created by other parties as well. For example, the website Top Class Actions² aggregates details of ongoing class action lawsuits, including those related to privacy harms, such as data breaches and misuse of consumer data, and provides links for people to file claims and join classes. Adapting such platforms to be more in-situ—for example, by providing people with a periodic digest of class actions they might be interested in joining—could be a promising area of future impact, particularly in the face of oft-slow-moving government changes.

5.7.2 Counter-Data Action

While we highlighted contestability as a key design principle of GoTCHas, there is still an inherent power imbalance between the people who report harms and the government that responds to them. To this end, Dalton and Thatcher [135], and later, Currie et al. [136], have proposed “counter-data action” as a paradigm of resistance, for critically understanding and re-interpreting “politically dominant datasets”. Relatedly, Meng and DiSalvo [137] situate counter-data action in a long legacy of Black activism dating back to W.E.B. DuBois’s empirical studies from the Atlanta Sociological Laboratory, which challenged dominant perceptions of urban Black Americans [138]. And, Palacios et al. [139] have similarly explored a framework for rural and tribal communities within the United States to collectively gather Internet broadband measurements to contest federal data from the FCC. Future work could thus further explore how GoTCHas could support and create room for grassroots ownership and contestability.

²<https://topclassactions.com/>

5.7.3 Design Fiction and Participant Outlook

Participants in our study were, on average, pessimistic about the future of online behavioral advertising, and unenthusiastic about the capacity of a government tool to address OBA's harms. Perhaps OBA is already so prevalent and bleak that no single solution can inspire hope for change, or perhaps the overall harms of OBA are simply not striking enough for people to care about in a vacuum (again, recalling [18]). While both of these hypotheses might prove true, the setup of our comicboards themselves might have also encouraged more negative responses. For example, our comicboards were deliberately open-ended and sparse, because we specifically wanted participants to be as creative as possible in their story writing; however, this might have led them to fill in the blank with their existing perceptions of user privacy and the government, which generally trend negative in the United States [71]. In future work, we could consider seeding participants with examples of potential positive outcomes from GoTCHas, e.g., reparations or specific policy changes, in order to encourage more idealistic stories to which we can aspire through design.

5.7.4 Participant Sample and Study Context

Our participants were recruited from Prolific, an online crowdwork platform. Even though past work has found that long-term exposure to repetitive tasks does not affect worker “quality” [127], trauma-like symptoms have been reported in online content moderators—a role similar and related to crowdwork—who are repeatedly exposed to harmful and offensive content [128, 140]. It is thus not out of the question that our participants may have been uniquely sensitive and attuned to the issues of volunteer fatigue and exposure to harmful content. Future work could explore the co-design of GoTCHas with crowd workers and content moderators, while heeding recommendations from Irani and Silberman's experiences with Turkopticon [59]. Additionally, because our scope was limited to GoTCHas within the United States of America, examining systems in other regulatory environments or cultural contexts could prove fruitful in the future.

5.8 Conclusion

In this work, we explored—through a blend of fictional inquiry, story completion, and comicboarding—fictional futures in which a government-supported tool could facilitate people reporting on the privacy harms they experience from online behavioral advertising. In doing so, through an online survey, we found that participants had detailed conceptions of the user experience of such a tool, but wanted safeguards to prevent them from being exploited further for their data by the government itself. Consequently, participants also expressed a broad and deep distrust in the government’s capacity to appropriately bring about mitigations for these harms. We extrapolated these design findings to existing government complaint-reporting tools in other domains, finding that they, too, lacked key qualities to instill trust; such systems are ripe for future design exploration using the design principles we propose as a starting point.

CHAPTER 6

DISCUSSION

Developing tools and interfaces to increase user control and better educate users on how their data is being used is necessary but not sufficient to combat the slow violence of ubiquitous privacy harms. As Seberger et al. [96] argue, these “solutions” may simplify privacy problems into bite-size, solvable pieces, but they fail to address the larger problem of the normalization of affective discomfort, which “perpetuates the associated conditions of exploitation and legitimizes invasive data practices that are detrimental to the dignity of *people*”. Echoing prior arguments by Herley [49, 131] on why users choose not to take certain security advice, I contend the same for privacy: privacy dashboards, private browsing, and cookie blocking will not fully assuage people’s privacy concerns. As I have repeatedly noted, regardless of what users do to enhance their privacy and security, the harms of privacy violations were already inescapable.

Instead of designing for discrete interactions between people and specific apps, the usable privacy and security (UPS) community should recognize these interactions and harms as being inextricable from larger societal concerns. Participants from work covered in Chapter 4 have already grappled with this challenge:

My concerns with targeted and behavioral advertising aren’t so much with any specific event but with the concept as a whole. The internet has become such a fundamental aspect of modern life that I can’t just remove myself from the equation without significant impact. —P167 [141]

To that end, I argue, specifically, that the UPS community should acknowledge the aggregate scale and collective impact of such harms and design accordingly. In this chapter, I detail several long-standing UPS problems that have been left unanswered due to *not tak-*

ing a collective action approach, and how they have contributed to a sense of powerlessness and resignation in users. In turn, I also summarize how my work has attempted to answer these problems. I then discuss how future collective action approaches to UPS can build upon the design insights I have uncovered.

6.1 Long-Standing Problematic Constants in Usable Privacy and Security

As I introduced in the beginning of this thesis, there are several problems in UPS that have often been assumed to be environmental constants that researchers and users must work around, rather than address head on. I briefly summarize and provide context for each of the following problems and demonstrate how my completed work has addressed them:

P1: Expert-User Misalignment A mismatch between what experts think users should do and feel is realistic in terms of change, and what users actually want to address their problems and concerns;

P2: The Privacy Paradox A frequently-observed and historically-cited phenomenon where individuals appear to act against their privacy preferences, i.e., they express concern over their privacy but do not take action to protect it; and

P3: Slow Violence The extended temporal nature of certain kinds of privacy harms that feel unimportant or difficult to prioritize individually, but that can have significant effects collectively.

6.1.1 P1: Expert-User Misalignment

One persistent problem in UPS has been the mismatch between what security and privacy experts believe to be best practices for protecting people's S&P, and what users themselves actually do. The UPS field, after all, originally stemmed from a curiosity from S&P researchers about why users make errors when using S&P tools, characterizing users in its nascent years as unmotivated and the "weakest link" [142]. As Ion et al. [143] found in

2015, in the United States, whereas S&P experts might recommend that users prioritize installing software updates, using two-factor authentication, and using a password manager to protect their S&P, regular users primarily tended to use antivirus software, visit only known websites, and change their passwords frequently. (Busse et al. [144] later replicated this study for a European population and found largely the same outcomes).

This difference can lead to an awkward tension in the relationships between experts and non-experts, S&P or otherwise. For example, as Poole et al. [145] found, tech experts who were initially enthusiastic about helping non-experts with their computer problems quickly became bored of doing so; simultaneously, they wanted to maintain an aura of expertise, and felt uncomfortable when they could not solve the non-experts' problems. And, as Adams and Sasse wrote in their now-canonical 1999 work, "Users Are Not the Enemy" [146], security experts within organizations tend to characterize non-expert users as, "at best...a security risk that needs to be controlled and managed, at worst...the enemy within." I observed similarly in my interviews with S&P experts in Chapter 3 [100]: one expert felt jaded about non-experts and said they "did not understand how the world works", and others had difficulty envisioning solutions outside of existing legal frameworks. However, these experts are deeply entrenched in the institutions that mass-inflict privacy harms upon people, and ultimately may be disincentivized to advocate for non-reformist changes [54] as a matter of protecting their expert status.

But it is infeasible to teach millions of non-expert users about complex legal processes or algorithmic delivery systems, so that a grassroots movement may collectively rise up and make more expert-friendly demands. Instead, it's up to us, as UPS experts, to do some of the heavy-lifting and reinterpret non-expert people's demands into a language that experts can work with. In Chapter 4, I identified the concept of "harm" as one such translational mechanism. While defining privacy harms is of rapid increasing interest in legal scholarship [69, 70], the concept is not yet well-defined or particularly dominant in users' perceptions of their relationship with privacy-violating institutions. In other words,

there is not yet evidence that users necessarily even understand that they have been harmed. Future work could explore how exposing users to the concept of “privacy harm”—e.g., presenting them with the taxonomy presented in Chapter 4—might influence how they articulate their privacy demands to an audience of experts and perceive their relationship with privacy-violating institutions.

6.1.2 P2: The Privacy Paradox

The adversarial relationship between experts and users is exacerbated by the so-called “privacy paradox”. As Barry Brown [14] first noted in 2001, while users express *concerns* over their privacy, they seem to take very little *action* to protect their privacy. Myriad studies [17] throughout the past two decades have uncovered cases where users have stated they were very interested in protecting their personal data, but were also not willing pay small amounts of money to do so (or, vice versa, were willing to give up their personal information for small discounts). Historically, UPS researchers have explained the paradox through one of two ways: either users truly do not value privacy as much as they say the do, or their privacy decisions are distorted by information asymmetry and thus do not correctly represent how much they value privacy. However, providing better user controls or promoting individual privacy self-management, which UPS work has historically striven for, is an illusory solution to broader issues of mass institutional surveillance. As Solove [147] argues, privacy itself has intrinsic value beyond whether or not people choose to trade their personal data:

“The result of increasing the amount of privacy self-management is akin to doling out yet more homework, heaping on more tasks that people lack the time or ability to do. The privacy paradox is a myth, born out of this vicious cycle when people express concerns about their privacy, are given a dose of privacy self-management in response, fail to succeed at the impossible project of privacy self-management, and then become disillusioned and resigned.”

One promising path to break out of this cycle, which I have explored in this dissertation, is to understand the *privacy-promoting* tasks people are willing to do in this environment, even if they are not directly aligned with being *privacy-protecting*. In the work presented in Chapter 3, I found that even when feeling powerless, people enthusiastically spoke out about their distaste for privacy-violating institutions without additional prompting; they were also empathetic when reading about others' PVEIs, and appreciated the opportunity to recall and share their own PVEIs. And in Chapter 5, I proposed commitment diversity and flexibility as a key design principle for privacy harm reporting tools, so that people can view themselves as contributing to a larger movement over which they have collective ownership, rather than being solely motivated through protecting individual privacy.

6.1.3 P3: Slow Violence

One final reason people may feel so powerless and resigned over their privacy is that although the number of violations they experience may be overwhelming, the individual effects of the violations are difficult to observe, so people cannot devote the time, effort, and resources to addressing them all. As I briefly defined in Chapters 4 and 5, the term “slow violence” [18] refers to incremental and accretive events that may be near invisible to people when they are initially experienced, but inflict significant harms when aggregated over long periods of time and across populations. Nixon [18] has suggested, as one antidote to slow violence, “to devise arresting stories, images, and symbols adequate to the pervasive but elusive violence of delayed effects”. In other words, gathering specific evidence of this violence, and publicly reporting on it in an emotionally-appealing way—recalling Bennett and Segerberg’s “personalized action frames” [44]—may be an effective vector for public resistance against slow violence.

While the concept of slow violence has long been applied in the context of industrial pollution, environmental degradation, and climate change, it has largely been historically ignored in the UPS community in favor of aforementioned areas of research. Only in

recent years has slow violence been referenced in UPS-adjacent contexts—specifically, in domains of state surveillance and policing [148], digital technology abuse [149], and sensitive categories of OBA [73]. I extend this line of research through first creating a taxonomy of OBA harms—summarized in Chapter 4 and published in [141]—so that they can, as Citron and Solove write, “be tackled and remedied in a meaningful way” by both legal communities and other parties [70].

Human geography scholar Thom Davies has simultaneously posited “slow observation” as a counterpoint to slow violence: those who live with the sustained brutality of slow violence can also, over time, slowly make their own observations and construct their own deeper understandings of the violence [150], in contrast with Nixon’s direct calls for formal symbolism and representation of such evidence. In between these two concepts, I recommended in Chapter 5 to integrate the collection and collation of people’s observations and narratives of privacy harm into their day-to-day lives; in this way, we can more formally scaffold how people already go about building their understandings of privacy harm. Similar to how smartwatches and fitness trackers enable users to be actively engaged in their understandings of their physical health, while granting them the ability to communicate with their doctors with high-level concrete data, we might imagine that taking regular measurements of privacy harm reports can empower people to speak more definitively about their experiences with professional privacy advocates and policy experts.

6.2 Remaining Challenges and Reflections

The chapters of this dissertation were originally inspired by the five stages of Shaw et al.’s [12] model of computer-supported collective action (CSCA): (1) Identifying a problem; (2) Generating, debating and selecting solutions; (3) Coordinating and preparing to take action; (4) Taking action; and, (5) Following up, documenting and assessing action taken. In **creating a unified voice of privacy concerns** in Chapter 3, I explored stages 1 and 2. I further explored stage 1 again, along with stage 3, in Chapter 4, where I **interpreted the**

unified voice in existing legal contexts. Finally, I explored stages 2, 4, and 5 in Chapter 5, where I **imagined formal ways to measure and respond to privacy harms.**

As evidenced by my mapping, the stages of CSCA are neither perfectly ordinal nor terminal. We might, for example, now know that it is straightforward to organize and distill a collective’s amorphous privacy demands through a loosely scaffolded survey and voting mechanism, but the demands are not specific or actionable enough yet to execute. Or, perhaps we might now have a deeper comprehension of what certain kinds of privacy harms look like, but not a system for collecting evidence of such harms. Or, we now better understand the specific roles that people might be willing to take in a collective action system for privacy, but we still don’t necessarily know the broader impacts of those actions. In this section, I outline three potential areas for future work to spin off of my initial cycle through CSCA.

6.2.1 What Is Collective “Action”?

When one hears the term “collective action”, visions of crowds taking to the streets in protests, holding up signs at the doors of capitol building steps, and withholding money and labor may come to mind. In this dissertation, I did not explore the design of tools that would directly enable any of these actions, and in the grand scheme of things, it’s unlikely that user privacy will ever become a primary concern for the majority of users for such events to happen. (Participants from the study in Chapter 5 expressed this all too clearly). However, collectively contributing evidence of harm is a kind of action in itself. As a parallel to Vincent et al.’s concept of “conscious data contribution” [151, 50]—where users exert leverage against a large tech company by contributing data to competitors of that company—users can also demonstrate their power through collectively sharing their experiences of harm (e.g., the #MeToo movement).

While I scoped out an initial design space for one kind of potential collective action—harm reporting—future work should also explore how to support other, already-feasible

types of action on a collective level. For example, Vincent et al. [152, 50] also outlined collective data poisoning and striking as promising types of action more in line with traditional views of collective action. And, after the overturning of *Roe v. Wade*, Song et al. [153] explored how users on social media participated in “collective privacy sensemaking” and strategized about various fertility- and period-tracking privacy issues. Relatedly, future work could also explore how users can work together to identify when companies have committed privacy *wrongs* [154], and perhaps even actively retaliate [155].

6.2.2 The Value of a Persistent Dataset of Privacy Harms

As I noted above, users may demonstrate collective power and leverage through collective sharing experiences of privacy harm. However, whereas contributing data to a competitor of a big tech company may directly impact the profits and growth of that company, the impacts of a collective dataset of privacy harms has not yet been explored. If the government is the party collecting and curating the dataset, then the benefits to to it are obvious: regulatory agencies such as the FTC, who have been historically under-resourced, can have a clearer idea of where to direct and how to manage its resources. And, if nothing else, users benefit by simply having a forum to vent their frustrations with online privacy, with a potential for receiving compensatory damages. However, tech companies might also find value in such a dataset; for example, a company might point to its *lack* of presence in the dataset as a signal that it respects its users’ privacy. We might imagine a persistent dataset of privacy harms as a way to certify privacy-promoting companies and alert users of misbehaving ones.

At the same time, in setting up such a dataset, we must be careful about who has access to and ownership or authorship over the dataset. For example, some participants from the **harm-reporting design fiction** work felt wary of associating too much personal information with their reports of harm, since they were concerned that the government could use their contributions to re-target them with even more invasive ads. Is there any party that users would trust enough to act as stewards over their contributions to the dataset? Are

there different levels of user-acceptability for granting access to different parties? Future work should explore the tradeoffs associated with such access control.

6.2.3 Co-Design, Field Deployment, and Real-World Evaluation

As I touched on in the Limitations sections of each of my completed works in Chapters 3 to 5, all of my study participants have been recruited from the online crowdwork platform Prolific. While recent work [156] has found that Prolific is representative for research on user perceptions and experiences, it is no substitution for a field-deployed tool with real life participants, especially in the context of privacy, where the impacts of both harm and progress are already difficult to observe and prove. Future work should explore co-designing privacy harm reporting and action tools with the various stakeholders mentioned in Section 6.2.2.

And, finally, as I heavily emphasized in my work in Chapter 5, there is already a rich landscape of civic harm reporting tools in other domains. As a start, future work should develop more refined criteria to systematically evaluate these GoTCHAs, beyond the seven design principles I proposed in that chapter. Future work could also explore co-designing and collaborating with the government agencies behind GoTCHAs to implement not only my seven design principles, but also the collective sensemaking process I explored in Chapter 3. One can easily imagine a Find-Fix-Verify process being applied to a Regulations.gov public comments page, with the most-pressing or in-demand comments being distilled to the top for the government to prioritize. In this way, we can also support more flexibility and diversity in the roles that people can play in civic engagement.

CHAPTER 7

CONCLUSION

Nearly a century before they had the right to vote, women working in textile mills in Lowell, Massachusetts, were presented with wage cuts that would threaten their then-rare economic independence. While they were disenfranchised on a federal level, the Lowell mill girls shared an understanding of their collective labor power: they sustained a strike for months and organized to financially support each other when they were not earning wages.

In 2024, however, collective power in the context of user privacy is still a foreign idea. People accept, as facts of life, being behaviorally-manipulated into passive consumption and feeling continuously creeped out by invasive, personalized content. They receive paltry remedies when their personal information is abused or leaked by the parties who are supposed to protect it, and have no way to contest the resultant harms. In other words, people have been disenfranchised by Big Tech. In my research, I've attempted to design interventions to break out of this learned helplessness, through planting the seed that users have collective power to influence privacy on a societal scale.

First, I showed in Chapter 3 that while it is possible for a crowd of people to distill their own views about institutional privacy violations into a representative set of demands for redress, these views didn't mesh well with the opinions of security and privacy experts, who found them unrealistic in existing regulatory environments. I also demonstrated that users wanted recognition of and apology for privacy harms by large institutions. Then, in Chapter 4, I described how analyzing and taxonomizing the privacy harms that people face in one specific context—online behavioral advertising—can help situate such harms more cognizably in a legal context. I then argued that a systematic, collective gathering of evidence of these harms could be a viable path for users to take against privacy-violating institutions. And in Chapter 5, I used design fiction methods to illustrate a potential world

where such evidence-gathering systems might be successful, and applied the design insights gained from this fictional privacy system to evaluate systems in other domains.

Finally, I reflected in Chapter 6 on how, in taking a collective action framing, I specifically eschewed focus on how users can take individual protective S&P actions for themselves—as has been traditional in the usable privacy and security community; I argued that my collective action framing can open up a novel path for promoting user privacy in the future. Upon this foundation, I am excited to continue to explore the societal impacts of collective action and civic participation in user privacy.

Appendices

APPENDIX A

OBA HARMS CODEBOOK

Table A.1: Codebook used in analysis of survey responses. The left column, “Initial Category”, refers to non-prescriptive categories we used as references in early discussions of the data. These groupings loosely formed a basis for the typology of harms.

Initial Category	Label
Emotional reactions	distrust in institutions feeling like I didn't consent frustrated, annoyed I'm already proactive about privacy invaded overwhelmed by number of ads paranoid surprised unsettled, uncomfortable
Emotional/psychological changes	changed perception of advertiser feeling effects on mental health feeling like my thoughts are being predicted influenced my purchasing behavior knowing mics are not on but still feels like it made me more proactive about privacy wanting to guess at how ads got their data work/personal life crossover
Physical actions	blocking future ads deleting app/service/account disabling voice assistant stop having conversations near device stop using app/service/account
Ad content origin guesses	my demographics my location data my physical/IRL purchases my searches/browsing my smart home devices talking out loud/mic usage talking on social media
Sensitive or offensive content	ad contains misinfo/scam ad refers to someone else's shopping generally sensitive content generally hurtful/offensive content incorrect targeting medical/health ad: general medical/health ad: mental health

REFERENCES

- [1] A. Rusbridger and E. MacAskill, “Edward Snowden interview: The edited transcript,” *The Guardian*, Jul. 2014.
- [2] “Restore the Fourth - opposing unconstitutional mass government surveillance,” *Restorethe4th.com*, 2020.
- [3] “Stop watching us,” *StopWatching.Us*,
- [4] “Equifax releases details on cybersecurity incident, announces personnel changes,” *Equifax*, Sep. 2017.
- [5] Y. Zou, A. H. Mhaidli, A. McCall, and F. Schaub, ““I’ve got nothing to lose”: Consumers’ risk perceptions and protective actions after the Equifax data breach,” in *Fourteenth Symposium on Usable Privacy and Security ({SOUPS} 2018)*, 2018, pp. 197–216.
- [6] B. Auxier, L. Rainie, M. Anderson, A. Perrin, M. Kumar, and E. Turner, “Americans and privacy: Concerned, confused and feeling lack of control over their personal information,” *Pew Research Center*, Nov. 2019.
- [7] G. Marwell and P. Oliver, in *The Critical Mass in Collective Action* (Studies in Rationality and Social Change), Studies in Rationality and Social Change. Cambridge University Press, 1993, pp. i–iv.
- [8] S. Webb and B. Webb, *The history of trade unionism*. Longmans, Green, 1920.
- [9] J. L. Greenslade, “Labor unions and the Sherman Act: Rethinking labor’s nonstatutory exemption,” *Loy. LAL Rev.*, vol. 22, p. 151, 1988.
- [10] B. Adler, “California passes strict internet privacy law with implications for the country,” *NPR*, Jun. 2018.
- [11] B. Kaiser, “Tell Facebook: Our data is our property #OwnYourData,” *Change.org*, 2017.
- [12] A. Shaw *et al.*, “Computer supported collective action,” *Interactions*, vol. 21, no. 2, pp. 74–77, 2014.
- [13] N. Salehi, L. C. Irani, M. S. Bernstein, A. Alkhatib, E. Ogbe, and K. Milland, “We are dynamo: Overcoming stalling and friction in collective action for crowd workers,” in *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, 2015, pp. 1621–1630.

- [14] B. Brown, “Studying the internet experience,” *HP laboratories technical report HPL*, vol. 49, 2001.
- [15] A. Acquisti, “Privacy in electronic commerce and the economics of immediate gratification,” in *Proceedings of the 5th ACM conference on Electronic commerce*, 2004, pp. 21–29.
- [16] S. B. Barnes, “A privacy paradox: Social networking in the united states,” *First Monday*, 2006.
- [17] S. Barth and M. D. De Jong, “The privacy paradox—investigating discrepancies between expressed privacy concerns and actual online behavior—a systematic literature review,” *Telematics and informatics*, vol. 34, no. 7, pp. 1038–1058, 2017.
- [18] R. Nixon, *Slow Violence and the Environmentalism of the Poor*. Harvard University Press, 2011.
- [19] M. Golla *et al.*, ““What was that site doing with my Facebook password?” Designing password-reuse notifications,” in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, 2018, pp. 1549–1566.
- [20] P. Mayer, Y. Zou, F. Schaub, and A. J. Aviv, ““ now i’m a bit angry:” individuals’ awareness, perception, and responses to data breaches that affected them,” in *30th {USENIX} Security Symposium ({USENIX} Security 21)*, 2021.
- [21] B. Ur, P. G. Leon, L. F. Cranor, R. Shay, and Y. Wang, “Smart, useful, scary, creepy: Perceptions of online behavioral advertising,” in *proceedings of the eighth symposium on usable privacy and security*, 2012, pp. 1–15.
- [22] S. Gaw and E. W. Felten, “Password management strategies for online accounts,” in *Proceedings of the second symposium on Usable privacy and security*, 2006, pp. 44–55.
- [23] S. Das, A. D. Kramer, L. A. Dabbish, and J. I. Hong, “The role of social influence in security feature adoption,” in *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, 2015, pp. 1416–1426.
- [24] E. M. Redmiles, A. R. Malone, and M. L. Mazurek, “I think they’re trying to tell me something: Advice sources and selection for digital security,” in *2016 IEEE Symposium on Security and Privacy (SP)*, IEEE, 2016, pp. 272–288.
- [25] S. Das, J. Lo, L. Dabbish, and J. I. Hong, “Breaking! a typology of security and privacy news and how it’s shared,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–12.

- [26] S. Das, L. A. Dabbish, and J. I. Hong, “A typology of perceived triggers for end-user security and privacy behaviors,” in *Fifteenth Symposium on Usable Privacy and Security* ({SOUPS} 2019), 2019.
- [27] S.-W. Huang, M. Suh, B. M. Hill, and G. Hsieh, “How activists are both born and made: An analysis of users on Change.org,” in *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, 2015, pp. 211–220.
- [28] “Don’t let EQUIFAX escape liability!” *Change.org*, 2017.
- [29] *R/privacy - if cops can watch us, we should watch them. i scraped court records to find dirty cops*. 2020.
- [30] *R/privacy - I think I accidentally started a movement - policing the police by scraping court data*, 2020.
- [31] S. Das, W. K. Edwards, D. Kennedy-Mayo, P. Swire, and Y. Wu, “Privacy for the people? exploring collective action as a mechanism to shift power to consumers in end-user privacy,” *IEEE Security & Privacy*, vol. 19, no. 5, pp. 66–70, 2021.
- [32] T. Kriplean, J. Morgan, D. Freelon, A. Borning, and L. Bennett, “Supporting reflective public thought with considerit,” in *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, 2012, pp. 265–274.
- [33] H. Zhang *et al.*, “Wedo: End-to-end computer supported collective action,” in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, 2014.
- [34] M. S. Bernstein *et al.*, “Soylent: A word processor with a crowd inside,” in *Proceedings of the 23rd annual ACM symposium on User interface software and technology*, 2010, pp. 313–322.
- [35] G. A. Bowen, “Grounded theory and sensitizing concepts,” *International journal of qualitative methods*, vol. 5, no. 3, pp. 12–23, 2006.
- [36] Q. Yang, N. Banovic, and J. Zimmerman, “Mapping machine learning advances from hci research to reveal starting places for design innovation,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–11.
- [37] D. Watson and L. A. Clark, “The PANAS-X: Manual for the positive and negative affect schedule, expanded form,” 1999.
- [38] V. Braun and V. Clarke, “Reflecting on reflexive thematic analysis,” *Qualitative Research in Sport, Exercise and Health*, vol. 11, no. 4, pp. 589–597, 2019.

- [39] N. Perlroth, “Yahoo says hackers stole data on 500 million users in 2014,” *The New York Times*, Sep. 2016.
- [40] E. Nakashima, “Hacks of opm databases compromised 22.1 million people, federal authorities say,” *The Washington Post*, Apr. 2019.
- [41] E. Stolterman and M. Wiberg, “Concept-driven interaction design research,” *Human-Computer Interaction*, vol. 25, no. 2, pp. 95–118, 2010.
- [42] S. Das, T. H.-J. Kim, L. A. Dabbish, and J. I. Hong, “The effect of social influence on security sensitivity,” in *10th Symposium On Usable Privacy and Security ({SOUPS} 2014)*, 2014, pp. 143–157.
- [43] E. Rader, R. Wash, and B. Brooks, “Stories as informal lessons about security,” in *Proceedings of the Eighth Symposium on Usable Privacy and Security*, 2012, pp. 1–17.
- [44] W. L. Bennett and A. Segerberg, “The logic of connective action: Digital media and the personalization of contentious politics,” *Information, communication & society*, vol. 15, no. 5, pp. 739–768, 2012.
- [45] “Cybersecurity Resource Center: Cybersecurity incidents,” *U.S. Office of Personnel Management*,
- [46] V. Clarke, V. Braun, and N. Hayfield, “Thematic analysis,” *Qualitative psychology: A practical guide to research methods*, pp. 222–248, 2015.
- [47] P. Ekman, “Basic emotions,” *Handbook of cognition and emotion*, vol. 98, no. 45-60, p. 16,
- [48] J. L. Davis and N. Jurgenson, “Context collapse: Theorizing context collusions and collisions,” *Information, communication & society*, vol. 17, no. 4, pp. 476–485, 2014.
- [49] C. Herley, “So long, and no thanks for the externalities: The rational rejection of security advice by users,” in *Proceedings of the 2009 workshop on New security paradigms workshop*, 2009, pp. 133–144.
- [50] N. Vincent, H. Li, N. Tilly, S. Chancellor, and B. Hecht, “Data leverage: A framework for empowering the public in its relationship with technology companies,” in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 2021, pp. 215–227.
- [51] D. C. Howe and H. Nissenbaum, “Engineering privacy and protest: A case study of adnauseam.,” in *IWPE@ SP*, 2017, pp. 57–64.

- [52] H. Nissenbaum and H. Daniel, “Trackmenot: Resisting surveillance in web search,” 2009.
- [53] R. Abebe, S. Barocas, J. Kleinberg, K. Levy, M. Raghavan, and D. G. Robinson, “Roles for computing in social change,” in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020, pp. 252–260.
- [54] A. Gorz, M. Nicolaus, V. Ortiz, V. González, and B. P. (Boston)., *Strategy for Labor: A Radical Proposal* (Beacon paperback ; BP282). Beacon Press, 1967.
- [55] I. Rahwan, “Society-in-the-loop: Programming the algorithmic social contract,” *Ethics and Information Technology*, vol. 20, no. 1, pp. 5–14, 2018.
- [56] N. McDonald *et al.*, “Privacy and power: Acknowledging the importance of privacy research and design for vulnerable populations,” in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–8.
- [57] D. Freed *et al.*, ““ is my phone hacked?” analyzing clinical computer security interventions with survivors of intimate partner violence,” *Proceedings of the ACM on Human-Computer Interaction*, vol. 3, no. CSCW, pp. 1–24, 2019.
- [58] L. Irani and M. S. Silberman, “From critical design to critical infrastructure: Lessons from turkopticon,” *Interactions*, vol. 21, no. 4, pp. 32–35, 2014.
- [59] L. C. Irani and M. S. Silberman, “Stories we tell about labor: Turkopticon and the trouble with” design”,” in *Proceedings of the 2016 CHI conference on human factors in computing systems*, 2016, pp. 4573–4586.
- [60] C. Gilligan, *In a different voice: Psychological theory and women’s development*. Harvard University Press, 1993.
- [61] R. Kang, S. Brown, L. Dabbish, and S. Kiesler, “Privacy attitudes of mechanical turk workers and the us public,” in *10th Symposium On Usable Privacy and Security ({SOUPS} 2014)*, 2014, pp. 37–49.
- [62] E. M. Redmiles, S. Kross, and M. L. Mazurek, “How well do my results generalize? comparing security and privacy survey results from mturk, web, and telephone samples,” in *2019 IEEE Symposium on Security and Privacy (SP)*, IEEE, 2019, pp. 1326–1343.
- [63] S. Zuboff, *The age of surveillance capitalism: The fight for a human future at the new frontier of power: Barack Obama’s books of 2019*. Profile books, 2019.
- [64] Federal Trade Commission, “FTC Staff Report: Self-regulatory principles for on-line behavioral advertising,” Tech. Rep., 2009.

- [65] J. Hanson, M. Wei, S. Veys, M. Kugler, L. Strahilevitz, and B. Ur, “Taking data out of context to hyper-personalize ads: Crowdworkers’ privacy perceptions and decisions to disclose private information,” in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–13.
- [66] E. Zeng, T. Kohno, and F. Roesner, “What makes a “bad” ad? user perceptions of problematic online advertising,” in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–24.
- [67] H.-P. Lee *et al.*, “When and why do people want ad targeting explanations? evidence from a four-week, mixed-methods field study,” in *2023 IEEE Symposium on Security and Privacy (SP)*, IEEE Computer Society, 2022, pp. 923–940.
- [68] “Harm,” in *Black’s Law Dictionary*, B. A. Garner, Ed., Thomson Reuters, 2019.
- [69] R. Calo, “The boundaries of privacy harm,” *Ind. LJ*, vol. 86, p. 1131, 2011.
- [70] D. K. Citron and D. J. Solove, “Privacy harms,” *BUL Rev.*, vol. 102, p. 793, 2022.
- [71] B. Auxier, L. Rainie, M. Anderson, A. Perrin, M. Kumar, and E. Turner, “Americans and privacy: Concerned, confused and feeling lack of control over their personal information,” 2019.
- [72] S. Chancellor, M. L. Birnbaum, E. D. Caine, V. M. Silenzio, and M. De Choudhury, “A taxonomy of ethical tensions in inferring mental health states from social media,” in *Proceedings of the conference on fairness, accountability, and transparency*, 2019, pp. 79–88.
- [73] L. Gak, S. Olojo, and N. Salehi, “The distressing ads that persist: Uncovering the harms of targeted weight-loss ads among users with histories of disordered eating,” *Proceedings of the ACM on Human-Computer Interaction*, vol. 6, no. CSCW2, pp. 1–23, 2022.
- [74] T. H. Baek and M. Morimoto, “Stay away from me,” *Journal of advertising*, vol. 41, no. 1, pp. 59–76, 2012.
- [75] J. Turow, M. X. Delli Carpini, N. A. Draper, and R. Howard-Williams, “Americans roundly reject tailored political advertising,” 2012.
- [76] E. G. Smit, G. Van Noort, and H. A. Voorveld, “Understanding online behavioural advertising: User knowledge, privacy concerns and online coping behaviour in europe,” *Computers in human behavior*, vol. 32, pp. 15–22, 2014.

- [77] Y. Yao, D. Lo Re, and Y. Wang, “Folk models of online behavioral advertising,” in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 2017, pp. 1957–1969.
- [78] C. J. Bennett and D. Lyon, “Data-driven elections: Implications and challenges for democratic societies,” *Internet Policy Review*, vol. 8, no. 4, 2019.
- [79] C. J. Bennett and J. Gordon, “Understanding the “micro” in political micro-targeting: An analysis of facebook digital advertising in the 2019 federal canadian election,” *Canadian Journal of Communication*, vol. 46, no. 3, pp. 431–459, 2021.
- [80] M. Ali, P. Sapiezynski, M. Bogen, A. Korolova, A. Mislove, and A. Rieke, “Discrimination through optimization: How facebook’s ad delivery can lead to biased outcomes,” *Proceedings of the ACM on human-computer interaction*, vol. 3, no. CSCW, pp. 1–30, 2019.
- [81] B. Weinshel *et al.*, “Oh, the places you’ve been! user reactions to longitudinal transparency about third-party web tracking and inferencing,” in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, 2019, pp. 149–166.
- [82] M. Wei *et al.*, “What twitter knows: Characterizing ad targeting practices, user perceptions, and ad explanations through users’ own twitter data,” in *Proceedings of the 29th USENIX Conference on Security Symposium*, 2020, pp. 145–162.
- [83] C. Norval, K. Cornelius, J. Cobbe, and J. Singh, “Disclosure by design: Designing information disclosures to support meaningful transparency and accountability,” in *2022 ACM Conference on Fairness, Accountability, and Transparency*, 2022, pp. 679–690.
- [84] A. M. McDonald and L. F. Cranor, “The cost of reading privacy policies 2008 privacy year in review. i,” *S: A Journal of Law and Policy for the Information Society*, vol. 4, no. 3, p. 2008, 2009.
- [85] C. Jensen and C. Potts, “Privacy policies as decision-making tools: An evaluation of online privacy notices,” in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 2004, pp. 471–478.
- [86] S. Furnell and K.-L. Thomson, “Recognising and addressing ‘security fatigue’,” *Computer Fraud & Security*, vol. 2009, no. 11, pp. 7–11, 2009.
- [87] G. R. Milne, M. J. Culnan, and H. Greene, “A longitudinal assessment of online privacy notice readability,” *Journal of Public Policy & Marketing*, vol. 25, no. 2, pp. 238–249, 2006.

- [88] L. Brandimarte, A. Acquisti, and G. Loewenstein, “Misplaced confidences: Privacy and the control paradox,” *Social psychological and personality science*, vol. 4, no. 3, pp. 340–347, 2013.
- [89] D. J. Solove, “Introduction: Privacy self-management and the consent dilemma,” *Harv. L. Rev.*, vol. 126, p. 1880, 2012.
- [90] S. Das, C. Faklaris, J. I. Hong, L. A. Dabbish, *et al.*, “The security & privacy acceptance framework (spaf),” *Foundations and Trends® in Privacy and Security*, vol. 5, no. 1-2, pp. 1–143, 2022.
- [91] M. Altman, A. Wood, and E. Vayena, “A harm-reduction framework for algorithmic fairness,” *IEEE Security & Privacy*, vol. 16, no. 3, pp. 34–45, 2018.
- [92] S. Chancellor *et al.*, “The relationships between data, power, and justice in cscw research,” in *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*, 2019, pp. 102–105.
- [93] D. Baker, A. Hanna, and E. Denton, “Algorithmically encoded identities: Reframing human classification,” in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020, pp. 681–681.
- [94] J. Moore, “Towards a more representative politics in the ethics of computer science,” in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020, pp. 414–424.
- [95] J. S. Seberger, M. Llavore, N. N. Wyant, I. Shklovski, and S. Patil, “Empowering resignation: There’s an app for that,” in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–18.
- [96] J. S. Seberger, I. Shklovski, E. Swiatek, and S. Patil, “Still creepy after all these years: The normalization of affective discomfort in app use,” in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 2022, pp. 1–19.
- [97] S. Milano, B. Mittelstadt, S. Wachter, and C. Russell, “Epistemic fragmentation poses a threat to the governance of online targeting,” *Nature Machine Intelligence*, vol. 3, no. 6, pp. 466–472, 2021.
- [98] E. Pan, J. Ren, M. Lindorfer, C. Wilson, and D. R. Choffnes, “Panoptispy: Characterizing audio and video exfiltration from android applications.,” *Proc. Priv. Enhancing Technol.*, vol. 2018, no. 4, pp. 33–50, 2018.
- [99] S. Zuboff, “Surveillance capitalism and the challenge of collective action,” in *New labor forum*, SAGE Publications Sage CA: Los Angeles, CA, vol. 28, 2019, pp. 10–29.

- [100] Y. Wu, W. K. Edwards, and S. Das, ““a reasonable thing to ask for””: Towards a unified voice in privacy collective action,” in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 2022, pp. 1–17.
- [101] D. J. Solove and D. K. Citron, “Standing and privacy harms: A critique of *transunion v. ramirez*,” *BUL Rev. Online*, vol. 101, p. 62, 2021.
- [102] *TransUnion LLC v. Ramirez*, 141 S. Ct. 2190. 2021.
- [103] *Spokeo, Inc. v. Robins*, 136 S. Ct. 1540, 1549. 2016.
- [104] A. Gorski, *The Biden administration’s sigint executive order, part ii*, Nov. 2022.
- [105] E. M. Redmiles, ““ should i worry?” a cross-cultural examination of account security incident response,” in *2019 IEEE Symposium on Security and Privacy (SP)*, IEEE, 2019, pp. 920–934.
- [106] S. Hartmann, A. Mainka, and W. G. Stock, “Citizen relationship management in local governments: The potential of 311 for public service delivery,” *Beyond bureaucracy: Towards sustainable governance informatisation*, pp. 337–353, 2017.
- [107] P. R. Center, “Americans’ views of government: Decades of distrust, enduring support for its role,” *Pew Research Center*, 2022.
- [108] E. Corbett and C. Le Dantec, “Designing civic technology with trust,” in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–17.
- [109] A. Gordon-Tapiero, A. Wood, and K. Ligett, “The case for establishing a collective perspective to address the harms of platform personalization,” in *Proceedings of the 2022 Symposium on Computer Science and Law*, 2022, pp. 119–130.
- [110] B. Rakova, R. Shelby, and M. Ma, “Terms-we-serve-with: Five dimensions for anticipating and repairing algorithmic harm,” *Big Data & Society*, vol. 10, no. 2, p. 20539517231211553, 2023.
- [111] D. Korff, “The inadequacy of the October 2022 new US presidential executive order on enhancing safeguards for United States signals intelligence activities,” *Available at SSRN 4495169*, 2022.
- [112] C. McClain, M. Faverio, M. Anderson, and E. Park, “How Americans view data privacy,” *Pew Research Center*, 2023.

- [113] M. Harding, B. Knowles, N. Davies, and M. Rouncefield, “Hci, civic engagement & trust,” in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 2833–2842.
- [114] J. Bleecker, “Design fiction: A short essay on design, science, fact, and fiction,” *Machine Learning and the City: Applications in Architecture and Urban Design*, pp. 561–578, 2022.
- [115] R. Y. Wong, D. K. Mulligan, E. Van Wyk, J. Pierce, and J. Chuang, “Eliciting values reflections by engaging privacy futures using design workbooks,” *Proceedings of the ACM on Human-Computer Interaction*, vol. 1, no. CSCW, pp. 1–26, 2017.
- [116] N. Holten Møller, T. Rask Nielsen, and C. Le Dantec, “Work of the unemployed: An inquiry into individuals’ experience of data usage in public services and possibilities for their agency,” in *Designing Interactive Systems Conference 2021*, 2021, pp. 438–448.
- [117] C. Dindler and O. S. Iversen, “Fictional inquiry—design collaboration in a shared narrative space,” *CoDesign*, vol. 3, no. 4, pp. 213–234, 2007.
- [118] V. Clarke, V. Braun, H. Frith, and N. Moller, *Editorial introduction to the special issue: Using story completion methods in qualitative research*, 2019.
- [119] M. Wood, G. Wood, and M. Balaam, ““ they’re just tixel pits, man” disputing the’reality’ of virtual reality pornography through the story completion method,” in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 5439–5451.
- [120] E. Cheon and N. M. Su, “Futuristic autobiographies: Weaving participant narratives to elicit values around robots,” in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018, pp. 388–397.
- [121] J. Cambre, S. Reig, Q. Kravitz, and C. Kulkarni, ““ all rise for the ai director” eliciting possible futures of voice technology through story completion,” in *Proceedings of the 2020 ACM designing interactive systems conference*, 2020, pp. 2051–2064.
- [122] N. Moraveji, J. Li, J. Ding, P. O’Kelley, and S. Woolf, “Comicboarding: Using comics as proxies for participatory design with children,” in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2007, pp. 1371–1374.
- [123] A. Ball, “Review of data management lifecycle models,” 2012.
- [124] H. Jin, H. Shen, M. Jain, S. Kumar, and J. I. Hong, “Lean privacy review: Collecting users’ privacy concerns of data practices at a low cost,” *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 28, no. 5, pp. 1–55, 2021.

- [125] A. Hiniker, K. Sobel, and B. Lee, “Co-designing with preschoolers using fictional inquiry and comicboarding,” in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 5767–5772.
- [126] T.-S. Kuo *et al.*, “Understanding frontline workers’ and unhoused individuals’ perspectives on ai used in homeless services,” in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023, pp. 1–17.
- [127] K. Hata, R. Krishna, L. Fei-Fei, and M. S. Bernstein, “A glimpse far into the future: Understanding long-term crowd worker quality,” in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 2017, pp. 889–901.
- [128] A. Arsht and D. Etcovitch, “The human cost of online content moderation,” *Harvard Journal of Law and Technology*, vol. 2, 2018.
- [129] C. Le Ludec, M. Cornet, and A. A. Casilli, “The problem with annotation. human labour and outsourcing between france and madagascar,” *Big Data & Society*, vol. 10, no. 2, p. 20 539 517 231 188 723, 2023.
- [130] J. Nielsen, “Enhancing the explanatory power of usability heuristics,” in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 1994, pp. 152–158.
- [131] C. Herley, “Unfalsifiability of security claims,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 23, pp. 6415–6420, 2016.
- [132] A. Z. Hutnik and L. R. VanDruff, “*not outgunned, just outmanned*” (for now): *Senate hearing on privacy law addresses under-resourced ftc*, Oct. 2021.
- [133] R. B. Cialdini, *Influence: Science and practice*. Pearson education Boston, MA, 2009, vol. 4.
- [134] A. R. Schrock, “Civic hacking as data activism and advocacy: A history from publicity to open government data,” *New media & society*, vol. 18, no. 4, pp. 581–599, 2016.
- [135] C. Dalton and J. Thatcher, “What does a critical data studies look like, and why do we care? seven points for a critical approach to ‘big data’,” *Society and Space*, vol. 29, 2014.
- [136] M. Currie, B. S. Paris, I. Pasquetto, and J. Pierre, “The conundrum of police officer-involved homicides: Counter-data in los angeles county,” *Big Data & Society*, vol. 3, no. 2, p. 2 053 951 716 663 566, 2016.

- [137] A. Meng and C. DiSalvo, “Grassroots resource mobilization through counter-data action,” *Big Data & Society*, vol. 5, no. 2, p. 2 053 951 718 796 862, 2018.
- [138] E. Wright II, “Web du bois and the atlanta university studies on the negro, revisited,” in *WEB Du Bois*, Routledge, 2017, pp. 75–90.
- [139] B. Palacios Abad, E. Belding, M. Vigil-Hayes, and E. Zegura, “Note: Towards community-empowered network data action,” in *ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies (COMPASS)*, 2022, pp. 585–588.
- [140] M. Steiger, T. J. Bharucha, S. Venkatagiri, M. J. Riedl, and M. Lease, “The psychological well-being of content moderators: The emotional labor of commercial moderation and avenues for improving support,” in *Proceedings of the 2021 CHI conference on human factors in computing systems*, 2021, pp. 1–14.
- [141] Y. Wu, S. Bice, W. K. Edwards, and S. Das, “The slow violence of surveillance capitalism: How online behavioral advertising harms people,” in *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, 2023, pp. 1826–1837.
- [142] A. Whitten and J. D. Tygar, “Why johnny can’t encrypt: A usability evaluation of pgp 5.0.,” in *USENIX security symposium*, vol. 348, 1999, pp. 169–184.
- [143] I. Ion, R. Reeder, and S. Consolvo, “{“... no} one can hack my {mind}”: Comparing expert and {non-expert} security practices,” in *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, 2015, pp. 327–346.
- [144] K. Busse, J. Schäfer, and M. Smith, “Replication: No one can hack my mind revisiting a study on expert and {non-expert} security practices and advice,” in *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, 2019, pp. 117–136.
- [145] E. S. Poole, M. Chetty, T. Morgan, R. E. Grinter, and W. K. Edwards, “Computer help at home: Methods and motivations for informal technical support,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2009, pp. 739–748.
- [146] A. Adams and M. A. Sasse, “Users are not the enemy,” *Communications of the ACM*, vol. 42, no. 12, pp. 40–46, 1999.
- [147] D. J. Solove, “The myth of the privacy paradox,” *Geo. Wash. L. Rev.*, vol. 89, p. 1, 2021.
- [148] R. Kramer and B. Remster, “The slow violence of contemporary policing,” *Annual Review of Criminology*, vol. 5, pp. 43–66, 2022.

- [149] R. Brydolf-Horwitz, “Embodied and entangled: Slow violence and harm via digital technologies,” *Environment and Planning C: Politics and Space*, vol. 40, no. 2, pp. 391–408, 2022.
- [150] T. Davies, “Toxic space and time: Slow violence, necropolitics, and petrochemical pollution,” *Annals of the American Association of Geographers*, vol. 108, no. 6, pp. 1537–1553, 2018.
- [151] N. Vincent and B. Hecht, “Can “conscious data contribution” help users to exert “data leverage” against technology companies?” *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW1, pp. 1–23, 2021.
- [152] N. Vincent, B. Hecht, and S. Sen, ““data strikes”: Evaluating the effectiveness of a new form of collective action against technology companies,” in *The World Wide Web Conference*, 2019, pp. 1931–1943.
- [153] Q. Song, R. Ma, Y. Kou, and X. Gui, ““our users’ privacy is paramount to us”: A discourse analysis of how period and fertility tracking app companies address the roe v wade overturn.,” in *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 2024, pp. 1–17.
- [154] J. R. Reidenberg, “Privacy wrongs in search of remedies,” *Hastings LJ*, vol. 54, p. 877, 2002.
- [155] B. Chao, “Privacy losses as wrongful gains,” *Iowa L. Rev.*, vol. 106, p. 555, 2020.
- [156] J. Tang, E. Birrell, and A. Lerner, “Replication: How well do my results generalize now? the external validity of online privacy and security surveys,” in *Eighteenth symposium on usable privacy and security (SOUPS 2022)*, 2022, pp. 367–385.

VITA

Yuxi Wu was born on January 20, 1995, in Wuhan, China, to parents Xiaohua and Yan. She immigrated to the United States at the age of 5 and grew up in Hershey, Pennsylvania, and Carmel, Indiana with her brother Ethan. Currently, Yuxi resides in Atlanta, Georgia, with her husband Benson, as she completes her PhD at Georgia Tech.

Yuxi's graduate research has been supported by a Georgia Tech President's Fellowship and a J.P. Morgan Chase AI Research Fellowship. Previously, she was an analytics manager for the Mobilization team of the Democratic National Committee. She earned a Master of Science in Computational Analysis and Public Policy from the University of Chicago's Harris School of Public Policy in 2018, and a Bachelor of Arts in Economics and Business from University College London in 2015.

In her free time, Yuxi enjoys long-distance running, film photography, and hiking sections of the Appalachian Trail with Benson.