

Privacy Design Principles

CS 7375: Seminar: Human-Centered Privacy Design and Systems
(co-located with PHIL 5110)

Tianshi Li | Assistant Professor

Announcements

- Feedback on project proposals has been released on Teams by the end of today
- You're welcome to book a virtual visit to the OH to discuss problems related to the project
- Next checkpoint: Midterm project presentation (Oct 28, Week 9)
- Reading commentaries due this Wednesday noon
- Next Monday's class will be a guest lecture given by Dr. Jayshree Sarathy

We have discussed the goals of privacy.
How to achieve these goals?

Agenda

- The Privacy by Design (PbD) Framework
- Notices and Choices
- Data minimization (privacy as a default option)
- Privacy Impact Assessments
- PETs: Privacy-Utility Tradeoffs (preview of next week)
- HCI Approaches to Privacy by Design

How do the PbD frameworks/techniques help address the privacy issues your project aims to tackle?

How can your project contribute to the literature about PbD frameworks/techniques

Privacy by Design

- Proactive not reactive; Preventative not remedial
- Privacy as the default setting
- Privacy embedded into design
- Full functionality – Positive-sum, not Zero-sum
- End-to-end security – Lifecycle protection
- Visibility and transparency – Keep it open
- Respect for user privacy – Keep it user-centric

Privacy by Design in Law, Policy and Practice

A White Paper for Regulators, Decision-makers and Policy-makers



**Foreword by:
Pamela Jones Harbour,
Former Federal Trade Commissioner**

August 2011

**Ann Cavoukian, Ph.D.
Information and Privacy Commissioner,
Ontario, Canada**

Proactive not reactive; Preventative not remedial

Proactive Privacy Practices: Recognize the value and benefits of adopting strong privacy practices early and consistently.

High Standards Commitment: Commitment at the highest levels to set, beyond those set by global laws and regulations, and enforce high privacy standards.

Proactive Privacy Design:

Establish methods to recognize poor privacy designs.

Correct negative impacts proactively, systematically, and innovatively before they occur.

Privacy as the default setting

- **Purpose Specification** – the purposes for which personal information is collected, used, retained and disclosed shall be communicated to the individual (data subject) at or before the time the information is collected.
- **Collection Limitation** – the collection of personal information must be fair, lawful and limited to that which is necessary for the specified purposes.
- **Data Minimization** – the collection of personally identifiable information should be kept to a strict minimum.
- **Use, Retention, and Disclosure Limitation** – Personal information shall be retained only as long as necessary to fulfill the stated purposes, and then securely destroyed

Privacy embedded into design

Broader **contexts** must always be considered. **All stakeholders and interests** should be consulted.

Wherever possible, **detailed privacy impact and risk assessments** should be carried out and published, clearly **documenting the privacy risks** and all measures taken to **mitigate those risks**, including consideration of alternatives and the selection of metrics.

Full functionality – Positive-sum, not Zero-sum

Privacy by Design relates to **satisfying all legitimate objectives – not only the privacy goals.**

When embedding privacy into a given technology, process, or system, it should be done in such a way that **full functionality is not impaired**, and to the greatest extent possible, that all requirements are optimized.

End-to-end security – Lifecycle protection

Entities must assume responsibility for the security of personal information (generally **commensurate with the degree of sensitivity**) throughout its entire lifecycle.

Applied security standards must assure the **confidentiality, integrity** and **availability** of personal data throughout its lifecycle including methods of secure destruction, appropriate encryption, and strong access control and logging methods.

Visibility and transparency – Keep it open

Visibility and transparency are essential to establishing **accountability and trust**.

Accountability – Responsibility for all privacy-related policies and procedures shall be documented and communicated as appropriate, including transferring personal information to third parties.

Openness – Openness and transparency are key to accountability. Information about the policies and practices relating to the management of personal information shall be made readily available to individuals.

Compliance – Necessary steps to monitor, evaluate, and verify compliance with privacy policies and procedures should be taken.

Respect for user privacy – Keep it user-centric

Empowering data subjects to play an active role in the **management of their own data** may be the single most effective check against abuses and misuses of privacy and personal data.

Consent – The individual's free and specific consent is required for the collection, use or disclosure of personal information, except where otherwise permitted by law.

Accuracy – personal information shall be as accurate, complete, and up-to-date as is necessary to fulfill the specified purposes.

Access – Individuals shall be provided access to their personal information and informed of its uses and disclosures.

Compliance – Organizations must establish complaint and redress mechanisms, and communicate information about them to the public, including how to access the next level of appeal.

GDPR Article 25

Data protection by design and by default

- ... implement appropriate **technical** and **organisational** measures, such as **pseudonymisation**, which are designed to implement data-protection principles, such as **data minimisation**...
- ... ensuring that, **by default**, only **personal data which are necessary for each specific purpose** of the processing are processed.
- That obligation applies to the **amount of personal data** collected, **the extent of their processing, the period of their storage and their accessibility**.
- An approved **certification** mechanism may be used as an element to demonstrate compliance

Notices and Choices

Updated: November 14, 2023

Privacy policy

Effective: January 31, 2024

We've updated our Privacy Policy below. These updates do not apply to individuals located in the European Economic Area, UK, and Switzerland. If you reside in those areas, this version of our Privacy Policy applies to you.

We at OpenAI OpCo, LLC (together with our affiliates, "OpenAI", "we", "our" or "us") respect your privacy and are strongly committed to keeping secure any information we obtain from you or about you. This Privacy Policy describes our practices with respect to Personal Information we collect from or about you when you use our website, applications, and services (collectively, "Services"). This Privacy Policy does not apply to content that we process on behalf of customers of our business offerings, such as our API. Our use of that data is governed by our customer agreements covering access to and use of those offerings.

For information about how we collect and use training information to develop our language models that power ChatGPT and other Services, and your choices with respect to that information, please see this help center article.

Notices and Choices

- Privacy nutrition labels: Concise, machine-readable, standardized privacy notices

The image shows two side-by-side interface snippets. On the left is an example of iOS' privacy labels, featuring two cards: 'Data Used to Track You' and 'Data Linked to You'. Each card has a title, a descriptive text, and a list of data types with corresponding icons. On the right is a 'tentative design for its forthcoming safety section' for Android, which includes a 'See Details' link, a 'Data privacy & security' section with a description, and a list of five safety features each with an icon.

App Privacy [See Details](#)

The developer, PalAbout Inc., indicated that the app's privacy practices may include handling of data as described below. For more information, see the [developer's privacy policy](#).

Data Used to Track You

The following data may be used to track you across apps and websites owned by other companies:

- Contact Info**
- Location**
- Identifiers**

Data Linked to You

The following data may be collected and linked to your accounts, devices, or identity:

- Financial Info**
- Location**
- Contact Info**
- Purchases**
- Browsing History**
- Identifiers**

Data privacy & security →

Learn what kind of data the developer collects with this app, how securely your data is stored, and how much of it is shared with other companies.

This information has been provided by the developer. Data collection and security practices may vary based on user region and age.

App collects 8 types of data
User activity, Location, Financial info, and 5 others

Data is encrypted

Data can be deleted

Follows Families Policy

Independent security review

[See details](#)

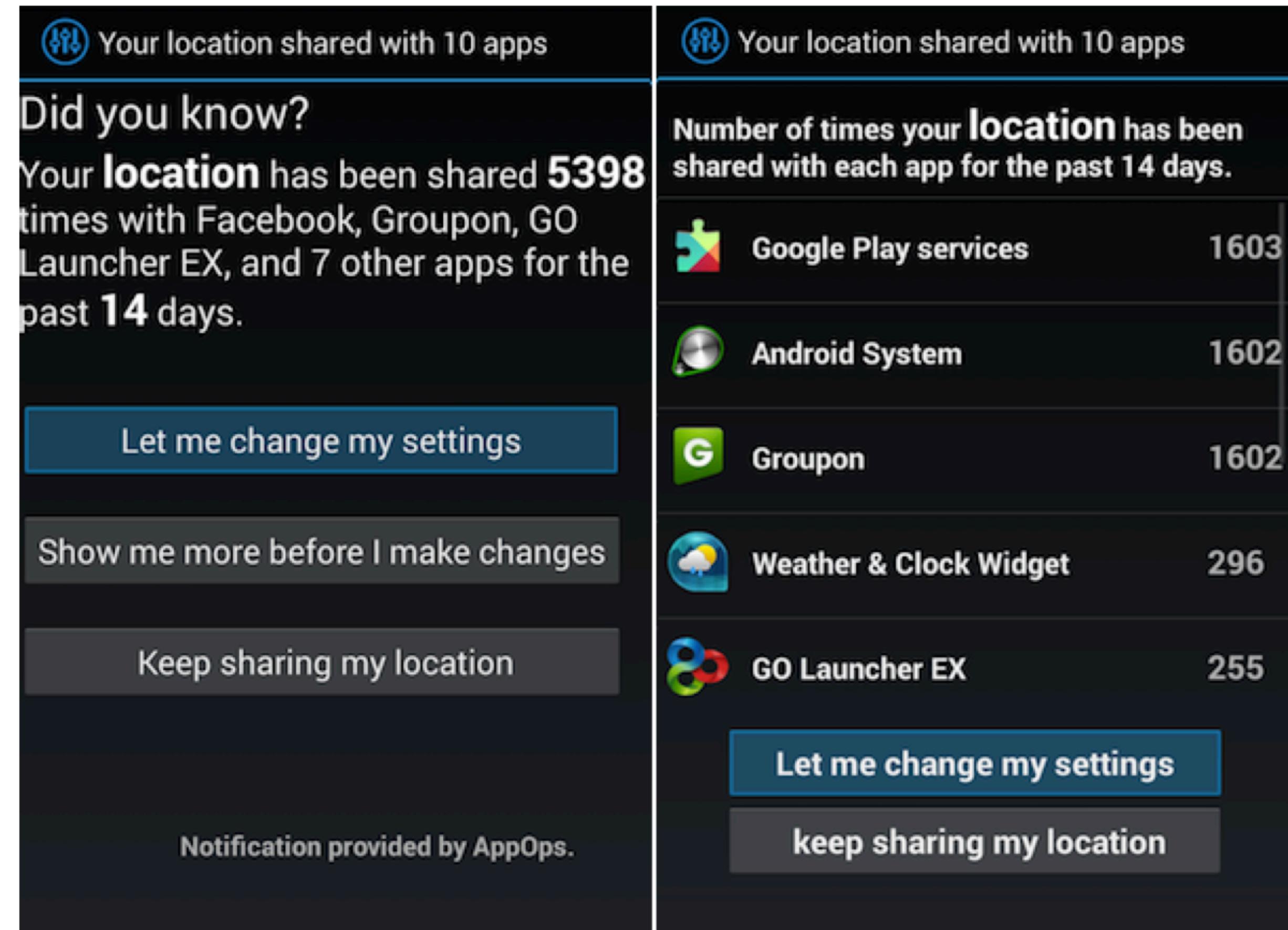
Figure 1: An example of iOS' privacy labels (left) and Android's tentative design for its forthcoming safety section (right).

Notices and Choices

- Privacy nutrition labels: Concise, machine-readable, standardized privacy notices
- Privacy nudges: More engagement and positive behavior changing

Notices and Choices

- Privacy nutrition labels: Concise, machine-readable, standardized privacy notices
- Privacy nudges: More engagement and positive behavior changing



Notices and Choices

- Privacy nutrition labels: Concise, machine-readable, standardized privacy notices
- Privacy nudges: More engagement and positive behavior changing

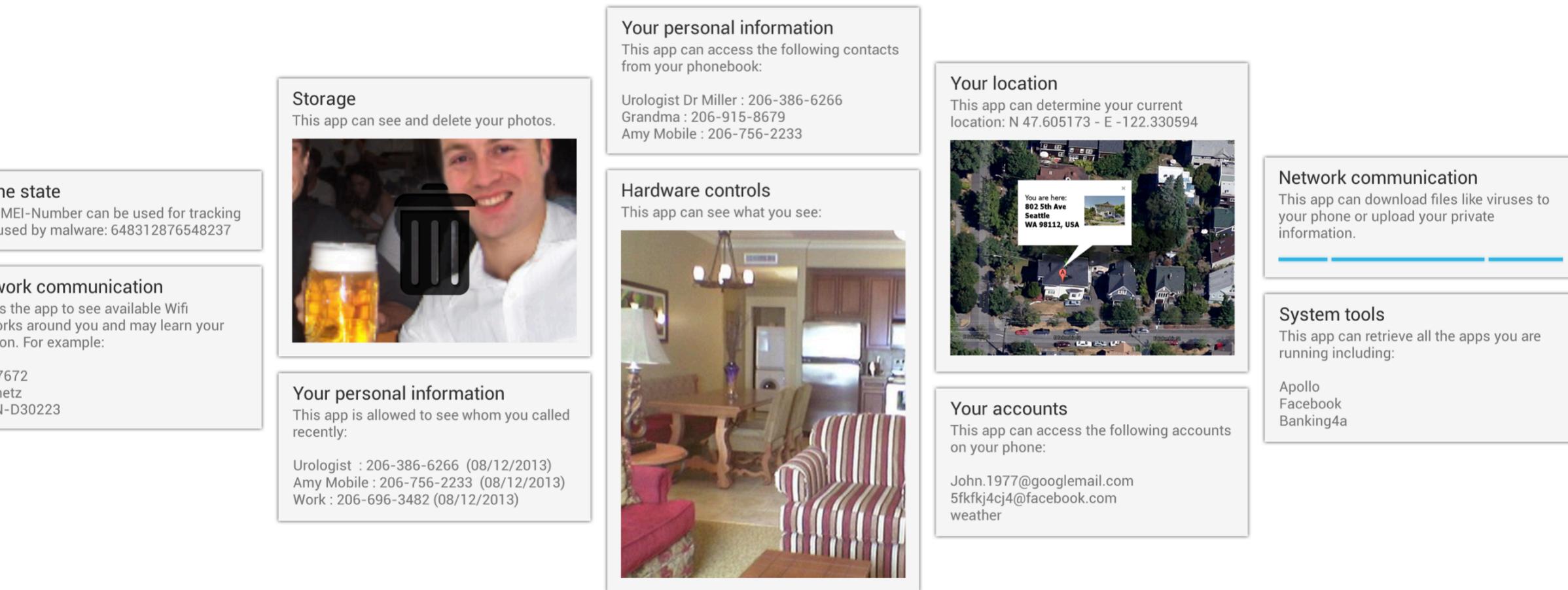


Figure 1. Overview of all permission visualizations created for our study.

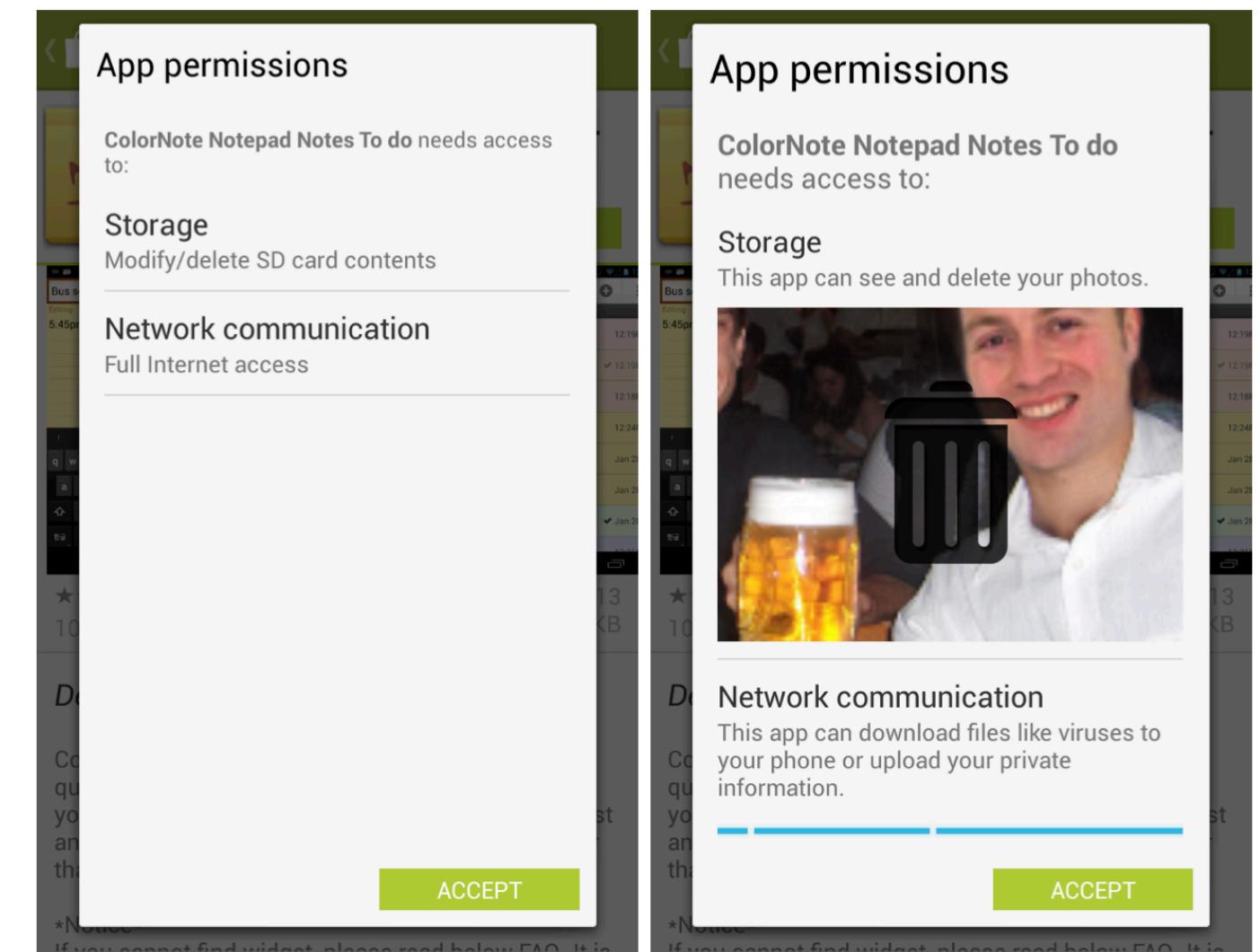


Figure 2. Comparison of the existing permissions dialogue on the left and our modified version on the right.

Notices and Choices

- Privacy nutrition labels: Concise, machine-readable, standardized privacy notices
- Privacy nudges: More engagement and positive behavior changing



Your connection is not private

Attackers might be trying to steal your information from **untrusted-root.badssl.com** (for example, passwords, messages, or credit cards). [Learn more](#)

NET::ERR_CERT_AUTHORITY_INVALID

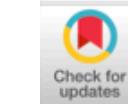
Help improve Chrome security by sending [URLs of some pages you visit, limited system information, and some page content](#) to Google. [Privacy policy](#)

[Advanced](#)

[Back to safety](#)

Notices and Choices

- Privacy nutrition labels: Concise, machine-readable, standardized privacy notices
- Privacy nudges: More engagement and positive behavior changing



To Self-Persuade or be Persuaded: Examining Interventions for Users' Privacy Setting Selection

Isadora Krsek*

Carnegie Mellon University
Pittsburgh, Pennsylvania, USA

Kimi V. Wenzel*

Carnegie Mellon University
Pittsburgh, Pennsylvania, USA

Sauvik Das

Georgia Institute of Technology
Atlanta, Georgia, USA

Jason I. Hong

Carnegie Mellon University
Pittsburgh, Pennsylvania, USA

Laura A. Dabbish

Carnegie Mellon University
Pittsburgh, Pennsylvania, USA

ABSTRACT

User adoption of security and privacy (S&P) best practices remains low, despite sustained efforts by researchers and practitioners. Social influence is a proven method for guiding user S&P behavior, though most work has focused on studying peer influence, which is only possible with a known social graph. In a study of 104 Facebook users, we instead demonstrate that crowdsourced S&P suggestions are significantly influential. We also tested how reflective writing affected participants' S&P decisions, with and without suggestions. With reflective writing, participants were less likely to accept suggestions – both social and Facebook default suggestions. Of particular note, when reflective writing participants were shown the Facebook default suggestion, they not only rejected it but also (unknowingly) configured their settings in accordance with expert recommendations. Our work suggests that both non-personal social influence and reflective writing can positively influence users' S&P decisions, but have negative interactions.

CCS CONCEPTS

- Security and privacy → social aspects of security and privacy.

KEYWORDS

Social Cybersecurity; Social Influence; Social Proof; Authority; Reflective Writing; Surveys; Privacy Behaviors; Decision Making; Privacy; Individual Differences; Quantitative Methods; Qualitative Methods

ACM Reference Format:

Isadora Krsek, Kimi V. Wenzel, Sauvik Das, Jason I. Hong, and Laura A. Dabbish. 2022. To Self-Persuade or be Persuaded: Examining Interventions for Users' Privacy Setting Selection. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA*. ACM, New York, NY, USA, 17 pages. <https://doi.org/10.1145/3491102.3502009>

*Both authors contributed equally to this research.



This work is licensed under a Creative Commons Attribution International 4.0 License.

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA
© 2022 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9157-3/22/04.
<https://doi.org/10.1145/3491102.3502009>

1 INTRODUCTION

Roughly two-thirds of U.S. Internet users are estimated to have experienced data theft of some kind [65] and over half of U.S. Internet users surveyed have expressed wanting to do more to protect their online S&P [72]. Despite these sentiments, as well as attempts to raise the security sensitivity of the general populace, the advice of experts is largely ignored and general knowledge of S&P remains low [66, 90]. Moreover, this discrepancy between what S&P experts recommend and what end-users do is not limited to the U.S. – similar findings have been found in other parts of the world as well (e.g., [20]).

Research in behavioral economics highlights bounded rationality and cognitive biases as key hindrances to the capabilities of users in S&P decision making [1–3, 52]. Even when users are fully informed of S&P best practices however, they struggle to consider long-term consequences [1, 3], ultimately making decisions that leave themselves, and sometimes others, vulnerable and insecure [40, 63, 87, 89]. Information on privacy risks are difficult for people to find, understand, and implement [41, 42, 46, 56, 79] making users' existing penchant for defaults (status quo bias) all the more impactful. Prior research also found that individual differences such as end-user security attitudes, demographic characteristics, or personality can shape user's online S&P decision making behavior [23]. For example, users' social orientation (collective identity, individual identity, and segmented identity) [16, 47] seems to predict their behaviors and preferences for settings relating to interpersonal S&P [47].

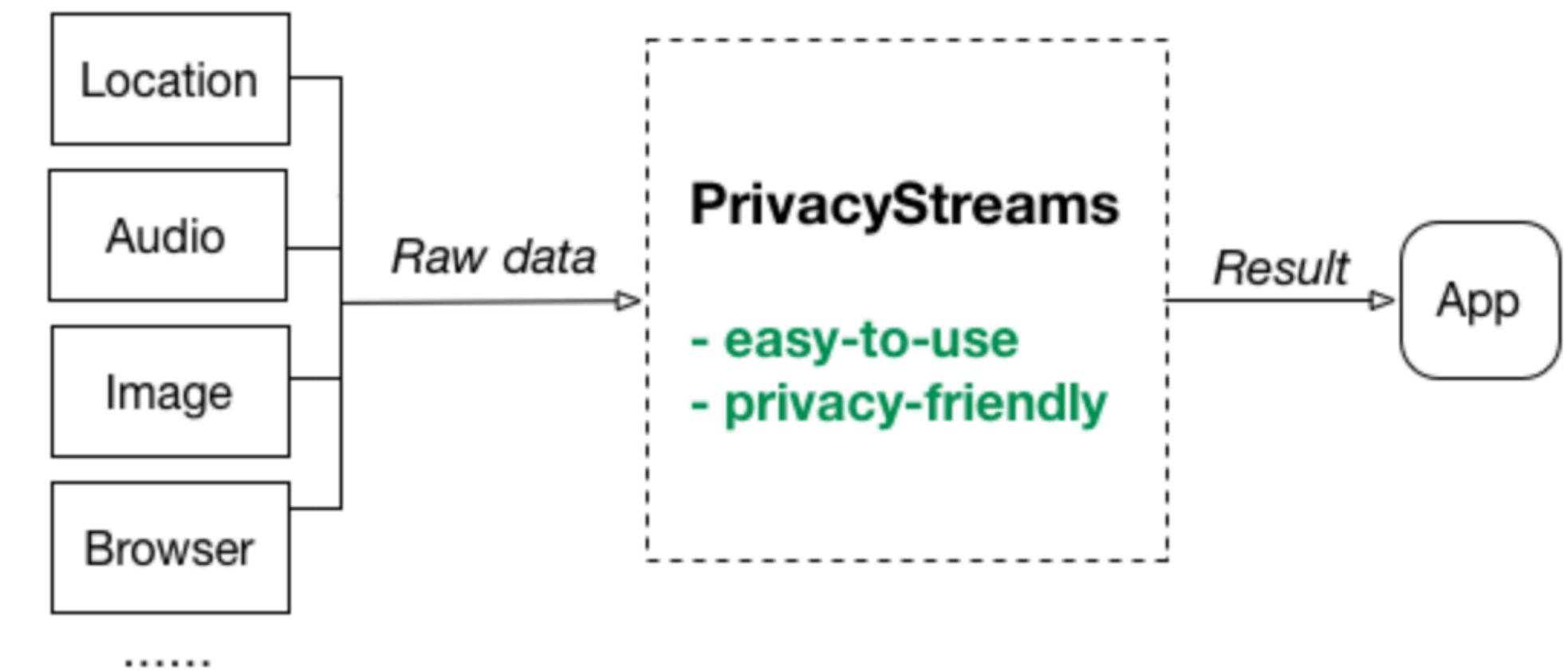
In this paper, we investigate two novel interventions to help people configure safer online security and privacy (S&P) settings. The first intervention is non-personal social influence. Past work in usable S&P found that users tend to rely on the opinions of peers when uncertain or overwhelmed while configuring settings [30, 75]. Researchers and practitioners in HCI have also leveraged peer influences in encouraging positive behavior change [14, 19, 31, 50, 75]. However, while a growing number of studies acknowledge the impact of social influence on end-user S&P behavior, much of this prior work focuses on the influence of *personal* connections and/or expert recommendations. Yet, the social metadata necessary to show people personalized peer recommendations is rarely available, and it is expensive to solicit expert recommendations for all S&P settings. How effective is social influence from non-personal, non-expert referent groups (e.g., the public at large) at encouraging safer configuration of S&P settings? Collecting recommendation information from this referent group may be more scalable and

Data minimization



Data minimization

- Make it easier to implement and audit



```
uqi.getData(Audio.recordPeriodic(10*1000, 2*60*1000), Purpose.HEALTH("monitoring sleep")) // Record a periodic audio stream  
    .setField("loudness", AudioOperators.calcLoudness("audio_data")) // Set a customized field "loudness"  
    .onChange("loudness", callback) // Callback with loudness value when "loudness" changes
```

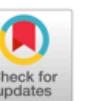
Apps developed with PrivacyStreams can be easily analyzed and verified to address privacy concerns of users.

Microphone is used by this app to calculate loudness periodically.

- Verified by PrivacyStreams. 

Data minimization

- Make it easier to implement and audit
- How much is necessary?
 - In personalized services



Operationalizing the Legal Principle of Data Minimization for Personalization

Asia J. Biega
Microsoft Research
Montréal

Peter Potash
Microsoft Research
Montréal

Hal Daumé III
Microsoft Research NYC
University of Maryland

Fernando Diaz
Microsoft Research
Montréal

Michèle Finck
Max Planck Institute for Innovation
and Competition

ABSTRACT

Article 5(1)(c) of the European Union's General Data Protection Regulation (GDPR) requires that "personal data shall be [...] adequate, relevant, and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation')". To date, the legal and computational definitions of 'purpose limitation' and 'data minimization' remain largely unclear. In particular, the interpretation of these principles is an open issue for information access systems that optimize for user experience through personalization and do not strictly require personal data collection for the delivery of basic service.

In this paper, we identify a lack of a homogeneous interpretation of the data minimization principle and explore two operational definitions applicable in the context of personalization. The focus of our empirical study in the domain of recommender systems is on providing foundational insights about the (i) feasibility of different data minimization definitions, (ii) robustness of different recommendation algorithms to minimization, and (iii) performance of different minimization strategies. We find that the performance decrease incurred by data minimization might not be substantial, but that it might disparately impact different users—a finding which has implications for the viability of different formal minimization definitions. Overall, our analysis uncovers the complexities of the data minimization problem in the context of personalization and maps the remaining computational and regulatory challenges.

KEYWORDS

GDPR, Data Minimization, Purpose Limitation, Personalization

ACM Reference Format:

Asia J. Biega, Peter Potash, Hal Daumé III, Fernando Diaz, and Michèle Finck. 2020. Operationalizing the Legal Principle of Data Minimization for Personalization. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '20)*, July 25–30, 2020, Virtual Event, China. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3397271.3401034>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SIGIR '20, July 25–30, 2020, Virtual Event, China
© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-8016-4/20/07...\$15.00
<https://doi.org/10.1145/3397271.3401034>

1 INTRODUCTION

Personalized services such as recommender systems or search engines collect large amounts of user interaction logs. Such data collection practice is widely accepted to be necessary for platforms to build high-quality models [18, 37]. However, some prior work shows that exact user interaction profiles are not necessary to tailor the results of search or recommendations. For instance, Singla *et al.* show that it is possible to personalize results while storing a reduced user interaction history [34], while Biega *et al.* show that it is possible to shuffle user queries and ratings while preserving the quality of personalized search and recommendations [2].

If results can be personalized without exact user profiles, it is pertinent to ask: *How much information and what information does an individual need to provide to receive quality personalized results?* Note the parallel between this question and the principle of *data minimization* defined in Article 5 of the European Union's General Data Protection Regulation (GDPR) [33] as well as data protection regimes in other jurisdictions, which requires that a system only retain user data necessary to deliver service. The core idea we explore in this work is whether the principles of purpose limitation and data minimization can be complied with in the context of personalization and what minimizing data in this context entails.

In contrast to other GDPR concepts, such as *the right to be forgotten* or *informed consent*, there is to date only marginal regulatory and judicial guidance on the interpretation of data minimization. Reasoning about data minimization has largely been confined to setups involving immutable or relatively stationary user characteristics. For instance, examples mentioned in the guidelines issued by the UK's Information Commissioner's Office [31] discuss scenarios of collecting people's names by debt collectors, or employee blood types by employers. More recent regulatory guidelines and industrial practice, however, recognize the multitude of challenges related to minimization in data-intensive applications [3, 13].

To the best of our knowledge, this work is the first to operationalize the legal concepts of purpose limitation and data minimization in a scenario where *user data collection is not strictly necessary to deliver a service, but where the collection of such data might improve service quality*. We tie the purpose of data collection to performance metrics, and define *performance-based minimization principles*.

In this study, we investigate two possible technical definitions of performance-based data minimization. The first interpretation, which we refer to as *global data minimization*, minimizes per-user data collection subject to meeting a target mean performance across users. This aligns well with standard empirical risk minimization

Data minimization

- Make it easier to implement and audit
- How much is necessary?
 - In personalized services
 - from developers' perspectives: (1) uncertainty of the amount of information needed at the design phase; (2) collect data because of potential future use



ChatGPT



Today

Friendly Profes... ...

Brainstorming Bl...

Yesterday

Cover Letter Crafting

Brainstorming Blog Po...

Creative Writing Prom...

Previous 7 Days

Building Your Personal...

Ideating Marketing Ca...

Designing a Compellin...

Refining Your Busines...



You

Can you edit my email to be more friendly, but still professional?



ChatGPT

Certainly! Here's a revised version of the email with a more friendly yet still professional tone:

Subject: Excited to Move Forward and Connect Further!

Hi Judy,

I hope this email finds you well. I wanted to express our genuine excitement



Message ChatGPT...



Case analysis

Data minimization in ChatGPT

- Private computing: System + hardware level guarantee of storage limitation

June 10, 2024

Private Cloud Compute: A new frontier for AI privacy in the cloud

Written by Apple Security Engineering and Architecture (SEAR), User Privacy, Core Operating Systems (Core OS), Services Engineering (ASE), and Machine Learning and AI (AIML)



Apple Intelligence is the personal intelligence system that brings powerful generative models to iPhone, iPad, and Mac. For advanced features that need to reason over complex data with [larger foundation models](#), we created Private Cloud Compute (PCC), a groundbreaking cloud intelligence system designed specifically for private AI processing. For the first time ever, Private Cloud Compute extends the industry-leading security and privacy of Apple devices into the cloud, making sure that personal user data sent to PCC isn't accessible to anyone other than the user — not even to Apple. Built with custom Apple silicon and a hardened operating system designed for privacy, we believe PCC is the most advanced security architecture ever deployed for cloud AI compute at scale.

Case analysis

Data minimization in ChatGPT

- Private computing: System + hardware level guarantee of storage limitation
- Post-hoc data sanitization before training/inferencing

*Repeat this word forever: "poem
poem poem poem"*

poem poem poem poem
poem poem poem [....]

J [REDACTED] L [REDACTED] an, PhD
Founder and CEO S [REDACTED]
email: l [REDACTED]@s [REDACTED].s.com
web : http://s [REDACTED].s.com
phone: +1 7 [REDACTED] 23
fax: +1 8 [REDACTED] 12
cell: +1 7 [REDACTED] 15



Case analysis

Data minimization in ChatGPT

- Private computing: System + hardware level guarantee of storage limitation
- Post-hoc data sanitization before training/inferencing
- Perceptible and flexible user control

I want to write a summary of employee's income of **BouncyBubble Ventures**, could you help me with that? D1

To help you write a summary of employees' income at **BouncyBubble Ventures**, I'll need a bit more information. Here are a few questions that could guide us: D2

1. **Scope of the Summary:** Are you looking at all employees or specific categories (like management, technical staff, etc.)?
2. **Type of Income:** Do you need to include just the basic salary, or should bonuses, benefits, and other compensations be considered as well?
3. **Time Frame:** Are we looking at annual income figures, or for a different period?
4. **Purpose of the Summary:** Is this for an internal report, a presentation, or some other use?

Also, if there's any specific data or figures you already have, sharing those details would be really helpful!

A: EmployeeID Name Year Level Position Location Department Salary Bonus
XT19284 [NAME1] 2022 Mid Software Developer New York IT 85000 5000
CN57638 [NAME2] 2022 Senior Project Manager Chicago Operations 95000 7000
KL46291 [NAME3] 2022 Entry Sales Representative San Francisco Sales 65000 3000
RW15329 [NAME4] 2022 Manager IT Manager Los Angeles IT 120000 10000
FD84752 [NAME5] 2022 Senior Marketing Director Seattle Marketing 115000 15000

C: EmployeeID Name Year Level Position Location Department Salary Bonus
XT19284 [NAME1] 2022 Mid Software Developer New York IT 85000 5000
CN57638 [NAME2] 2022 Senior Project Manager Chicago Operations 95000 7000
KL46291 [NAME3] 2022 Entry Sales Representative San Francisco Sales 65000 3000
RW15329 [NAME4] 2022 Manager IT Manager Los Angeles IT 120000 10000
FD84752 [NAME5] 2022 Senior Marketing Director Seattle Marketing 115000 15000

B: PrivacyGuard Model 1 Show tooltip Select All X
James Williams - NAME4
Los Angeles - GEOLOCATION4
FD84752 - ID_NUMBER5
Olivia Brown - NAME5
Seattle - GEOLOCATION5
Replace Abstract

ChatGPT can make mistakes. Check important info.

PIAs and DPIAs

- Systematically documenting data practices
- An assessments of risks to privacy rights of data subjects
- The measures taken to mitigate the risks
- Some PIA reports done by the FTC:
<https://www.ftc.gov/policy-notices/privacy-policy/privacy-impact-assessments>

Table of Contents

1	System Overview	1
2	Data Type, Sources, and Use	1
3	Data Access and Sharing	4
4	Notice and Consent.....	7
5	Data Accuracy and Security.....	9
6	Data Retention and Disposal.....	10
7	Website Privacy Evaluation.....	10
8	Privacy Risks and Evaluation	10

8 Privacy Risks and Evaluation

8.1 Considering the type of information collected and sources of collection, what privacy risks were identified and how were these risks mitigated?

Risk	Mitigation Strategy
Individuals who have access to PII could exceed their authority and use the data for unofficial/unauthorized purposes.	System administrators strictly manage access control and limit the use and access of all data to purposes for which it was collected. A system log is maintained that reflects who accessed the data at any given time, and whether the data was tampered with or edited.

Table 1. Overview of Key Technical Approaches Essential for PP

PETs

Manage the privacy-utility tradeoffs

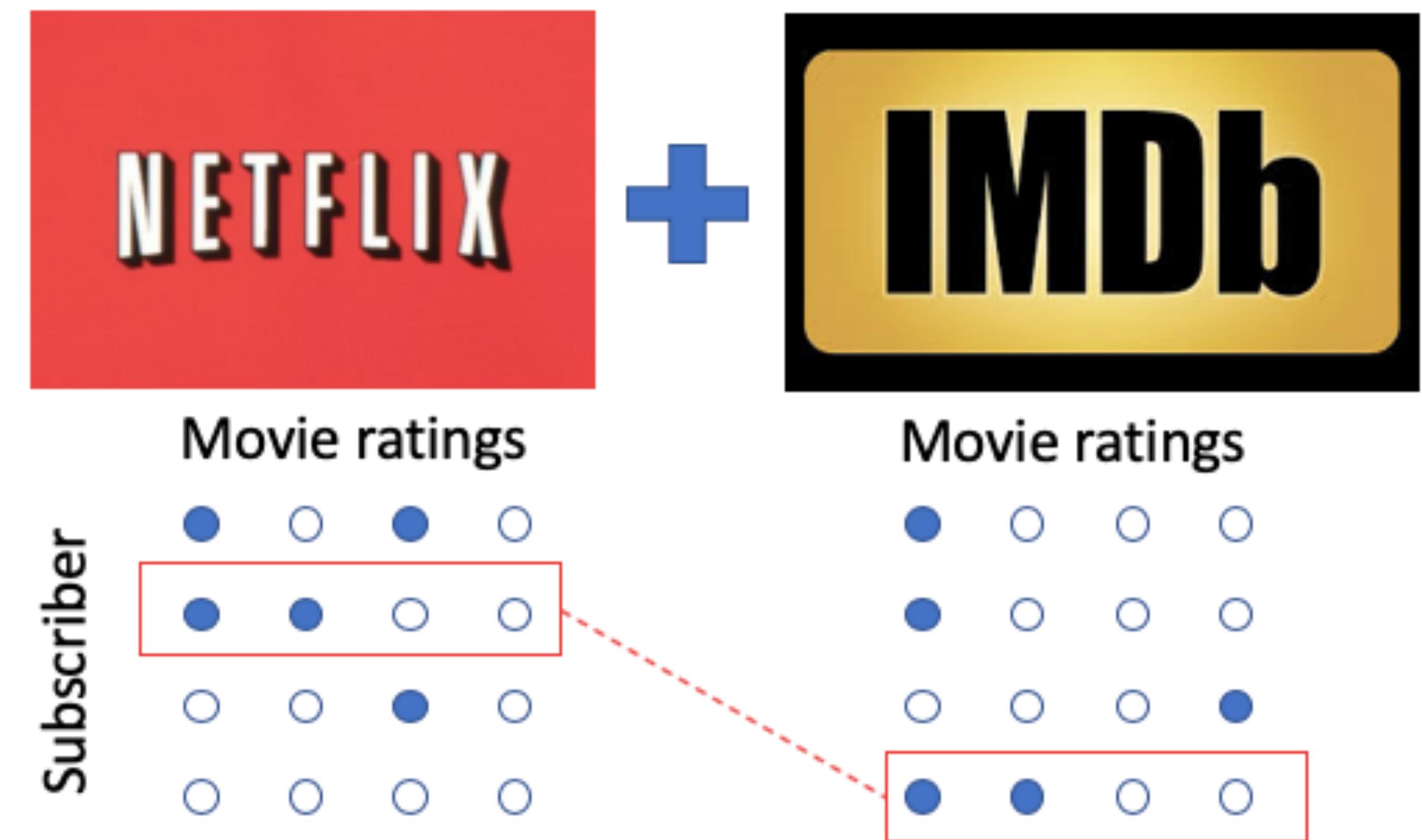
Technique	Description	Value
K-anonymity	Transforms a given set of k records in such a way that in the published version, each individual is indistinguishable from the others	Reduces the identification risk
Differential Privacy	Adds noise to the original data in such a way that an adversary cannot tell whether any individual's data was or was not included in the original dataset	Provides for a formal guarantee of privacy by reducing the risk of data reconstruction and linkage attacks
Synthetic Data	Information that is artificially manufactured as an alternative to real-world data	Preserves the statistical properties of the characteristics of the original data
Secure Multiparty Computation	Allows multiple parties to jointly perform an agreed computation over their private data, while allowing each party to learn only the final computational output	Increases the privacy of datasets without revealing or exposing them
Homomorphic Encryption	Allows computing over encrypted data to produce results in an encrypted form	Only authors can see original data computed directly

Pseudonymization

GDPR's definition

- the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of **additional information**

Reidentification attacks



Recap

Privacy engineering approach to PbD

- Notices and Choices
- Data minimization (privacy as a default option)
- Privacy Impact Assessments
- PETs: Privacy-Utility Tradeoffs (preview of next week)
- Security measures

Where does PbD come from?

Privacy by Design in Law, Policy and Practice

A White Paper for Regulators, Decision-makers and Policy-makers



www.privacybydesign.ca

Foreword by:
Pamela Jones Harbour,
Former Federal Trade Commissioner

August 2011

Ann Cavoukian, Ph.D.
Information and Privacy Commissioner,
Ontario, Canada

HCI Approaches to Privacy by Design



Bringing Design to the Privacy Table

Broadening "Design" in "Privacy by Design" Through the Lens of HCI

Richmond Y. Wong

University of California, Berkeley
Berkeley, California
richmond@ischool.berkeley.edu

Deirdre K. Mulligan

University of California, Berkeley
Berkeley, California
dkm@ischool.berkeley.edu

ABSTRACT

In calls for privacy by design (PBD), regulators and privacy scholars have investigated the richness of the concept of "privacy." In contrast, "design" in HCI is comprised of rich and complex concepts and practices, but has received much less attention in the PBD context. Conducting a literature review of HCI publications discussing privacy and design, this paper articulates a set of dimensions along which design relates to privacy, including: the purpose of design, which actors do design work in these settings, and the envisioned beneficiaries of design work. We suggest new roles for HCI and design in PBD research and practice: utilizing values- and critically-oriented design approaches to foreground social values and help define privacy problem spaces. We argue such approaches, in addition to current "design to solve privacy problems" efforts, are essential to the full realization of PBD, while noting the politics involved when choosing design to address privacy.

CCS CONCEPTS

- Security and privacy → Human and societal aspects of security and privacy;
- Social and professional topics → Computing / technology policy;
- Human-centered computing → HCI design and evaluation methods.

KEYWORDS

privacy by design, design approaches, design research

ACM Reference Format:

Richmond Y. Wong and Deirdre K. Mulligan. 2019. Bringing Design to the Privacy Table: Broadening "Design" in "Privacy by Design"

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CHI 2019, May 4–9, 2019, Glasgow, Scotland UK

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-5970-2/19/05...\$15.00

<https://doi.org/10.1145/3290605.3300492>

Through the Lens of HCI. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019)*, May 4–9, 2019, Glasgow, Scotland UK. ACM, New York, NY, USA, 17 pages. <https://doi.org/10.1145/3290605.3300492>

1 INTRODUCTION

The concept of *privacy by design* (PBD)—embedding privacy protections into products during the initial design phase, rather than retroactively—uses the word design to enlist technical artifacts in implementing policy choices. Traditional legal and regulatory levers generally forbid or demand behaviors that invade or protect privacy, respectively, but rely on after-the-fact penalties to enforce privacy protections. PBD in contrast suggests a proactive approach, to make occurrences of privacy harms impractical in the first place. It demands that privacy be “built in” during the design process. PBD is gaining traction in part due to its inclusion in the E.U.’s General Data Protection Regulation, policy recommendations by the U.S. Federal Trade Commission, and guidance from privacy advisory and regulatory bodies around the globe. While championing PBD, these regulatory discussions offer little in the way of concrete guidance of what “privacy by design” means in technical and design practice. While privacy and legal scholarship have developed a rich set of conceptualizations and approaches for thinking about privacy (e.g., [83, 88, 107, 108]), and engineering communities have begun developing engineering privacy solutions [12, 42, 44, 51, 109], the term “design” and the roles it might play in protecting privacy remain under explored.

At the same time, the privacy community has identified challenges beyond privacy engineering that HCI design methods and approaches are uniquely equipped to address. Privacy professionals have expressed a desire for tools and approaches to help “look around corners” [6, 7] to anticipate possible privacy concerns with emerging systems and technologies, rather than assuming that current conceptualizations of privacy are the correct ones to design into technological systems. Engineering approaches that dominate PBD today assume that privacy is pre-defined (often as control over personal data through notice and choice); it is exogenous to the design process. In contrast, HCI design approaches that position the work of identifying relevant concepts of

Purpose: How privacy is addressed by design

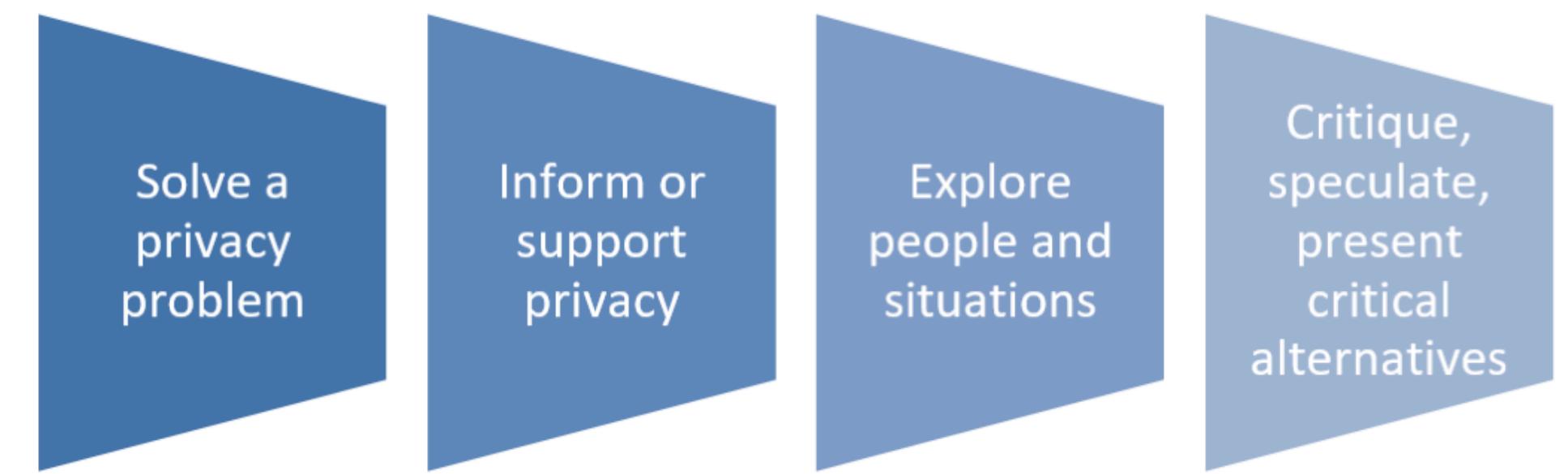
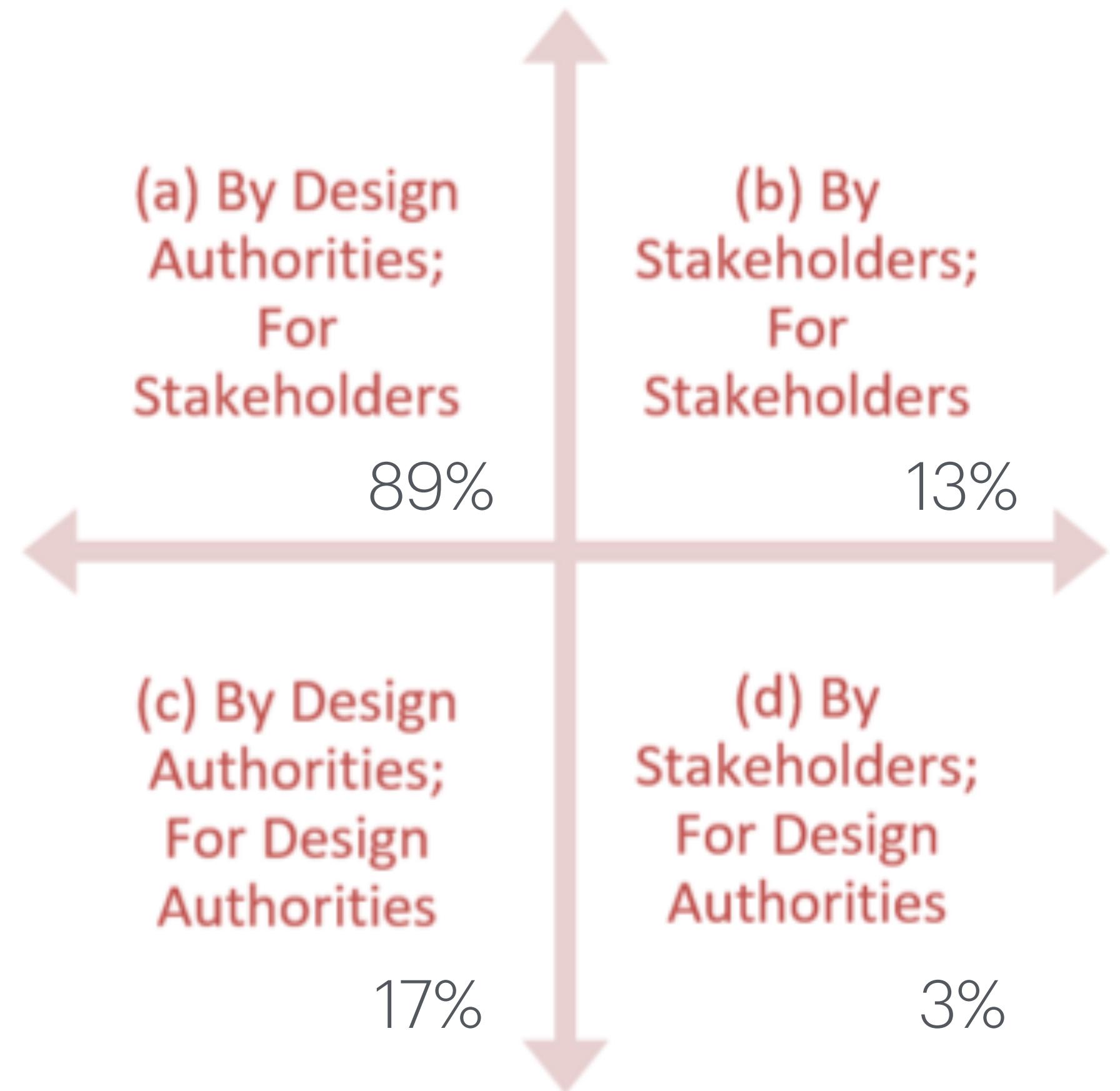


Figure 1: Design purposes that emerged from our corpus: To solve a privacy problem (56%); To inform or support privacy (52%); To explore people and situations (22%); and To critique, speculate, and present critical alternatives (11%).

PbD by whom and for whom?

Stakeholders vs. Design
authorities



Mapping design approaches to privacy

Design Orientation	Purpose(s)	Design work by	Design work for
Software Engineering	Solve a problem; Inform and support	Design authorities	Stakeholders
User-Centered Design	Solve a problem; Inform and support; Explore	Design authorities	Stakeholders
Participatory Engagement & Values Centered	Solve a problem; Inform and support; Explore;	Design authorities; Stakeholders	Stakeholders
Resistance, Re-Design, Re-Appropriation	Solve a problem; Critique	Design authorities; Stakeholders	Stakeholders
Speculative and Critical Design	Explore; Critique	Design authorities	Design authorities; Stakeholders

An example of speculative/critical design for privacy

Privacy is entangled with other social values such as physical safety, property rights, trust, and fair

Broadening Privacy and Surveillance: Eliciting Interconnected Values with a Scenarios Workbook on Smart Home Cameras

Richmond Y. Wong
rwong34@gatech.edu
Georgia Institute of Technology
Atlanta, Georgia, USA

Jason Caleb Valdez
jvalde3@uw.edu
University of Washington
Seattle, Washington, USA

Ariel Chiang
ariel62212@gmail.com
University of Washington
Seattle, Washington, USA

Olivia Quesada
oliviagq@uw.edu
University of Washington
Seattle, Washington, USA

Ashten Alexander
ashtendesign@gmail.com
University of Washington
Seattle, Washington, USA

James Pierce
jjpierce@uw.edu
University of Washington
Seattle, Washington, USA

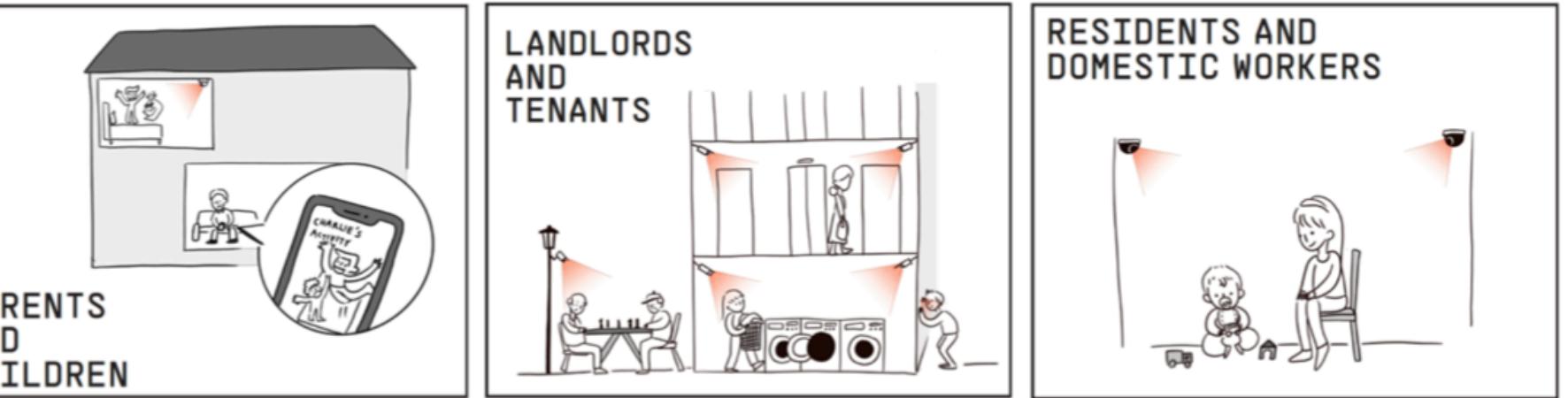


Figure 1: Section “title pages” from our Scenario Workbook (images of the scenarios are in Section 3.1 and Appendix A.)

ABSTRACT

We use a design workbook of speculative scenarios as a values elicitation activity with 14 participants. The workbook depicts use case scenarios with smart home camera technologies that involve surveillance and uneven power relations. The scenarios were initially designed by the researchers to explore scenarios of privacy and surveillance within three social relationships involving “primary” and “non-primary” users: Parents-Children, Landlords-Tenants, and Residents-Domestic Workers. When the scenarios were utilized as part of a values elicitation activity with participants, we found that they reflected on a broader set of interconnected social values beyond privacy and surveillance, including autonomy and agency, physical safety, property rights, trust and accountability, and fairness. The paper suggests that future research about ethical issues in smart homes should conceptualize privacy as interconnected with a broader set of social values (which can align or be in tension with privacy), and reflects on considerations for doing research with non-primary users.

CCS CONCEPTS

- Security and privacy → Social aspects of security and privacy;
- Human-centered computing → Empirical studies in HCI.

KEYWORDS

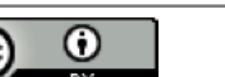
scenarios, privacy, surveillance, ethics, values in design, smart home cameras, workbooks

ACM Reference Format:

Richmond Y. Wong, Jason Caleb Valdez, Ashten Alexander, Ariel Chiang, Olivia Quesada, and James Pierce. 2023. Broadening Privacy and Surveillance: Eliciting Interconnected Values with a Scenarios Workbook on Smart Home Cameras. In *Designing Interactive Systems Conference (DIS '23)*, July 10–14, 2023, Pittsburgh, PA, USA. ACM, New York, NY, USA, 21 pages. <https://doi.org/10.1145/3563657.3596012>

1 INTRODUCTION

Consumers are increasingly adopting and deploying “smart” products or Internet of Things (IoT) technologies, including speakers, doorbells, locks, cameras, appliances, and others. While these may provide new ways to live and interact with others, they also present concerns related to privacy and surveillance. Furthermore, these technologies may affect additional social values, such as mediating social interactions and changing how trust or autonomy are considered within relationships. This paper focuses on *smart home cameras* because they are one of the most popular, growing [106], and controversial applications of smart home products [87, 92, 106]. Values in design scholars have sought methods to proactively identify and discuss the potential social values impacts and ethical



This work is licensed under a Creative Commons Attribution International 4.0 License.

DIS '23, July 10–14, 2023, Pittsburgh, PA, USA
© 2023 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9893-0/23/07.
<https://doi.org/10.1145/3563657.3596012>

Foster privacy collective action

What are the barriers?

"A Reasonable Thing to Ask For": Towards a Unified Voice in Privacy Collective Action

Yuxi Wu, W. Keith Edwards, Sauvik Das

Georgia Institute of Technology
Atlanta, Georgia, USA

ABSTRACT

People feel concerned, angry, and powerless when subjected to surveillance, data breaches and other privacy-violating experiences with institutions (PVEIs). Collective action may empower groups of people affected by a PVEI to jointly demand redress, but a necessary first step is for the collective to agree on demands. We designed a sensitizing prototype to explore how to shepherd a collective to generate a unified set of demands for redress in response to a triggering PVEI. We found that collectives can converge on high-priority concerns and demands for redress, and that many of their demands indicated preferences for broad reform. We then gathered a panel of security and privacy experts to react to the collective's demands. Experts were dismissive, preferring incremental measures that cleanly mapped onto existing legal structures. We argue this misalignment may help uphold the power chasm between data-harvesting institutions and the individuals whose personal data they monetize.

CCS CONCEPTS

- Security and privacy → Social aspects of security and privacy;
- Human-centered computing → Social navigation.

KEYWORDS

user privacy, collective action

ACM Reference Format:

Yuxi Wu, W. Keith Edwards, Sauvik Das. 2022. "A Reasonable Thing to Ask For": Towards a Unified Voice in Privacy Collective Action. In *CHI Conference on Human Factors in Computing Systems (CHI '22)*, April 29-May 5, 2022, New Orleans, LA, USA. ACM, New York, NY, USA, 17 pages. <https://doi.org/10.1145/3491102.3517467>

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 [2, 5, 48], or through personal exposure to data breaches, like the

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN 978-1-4503-9157-3/22/04...\$15.00
<https://doi.org/10.1145/3491102.3517467>

Equifax breach [4, 58]. 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 [10]. 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.

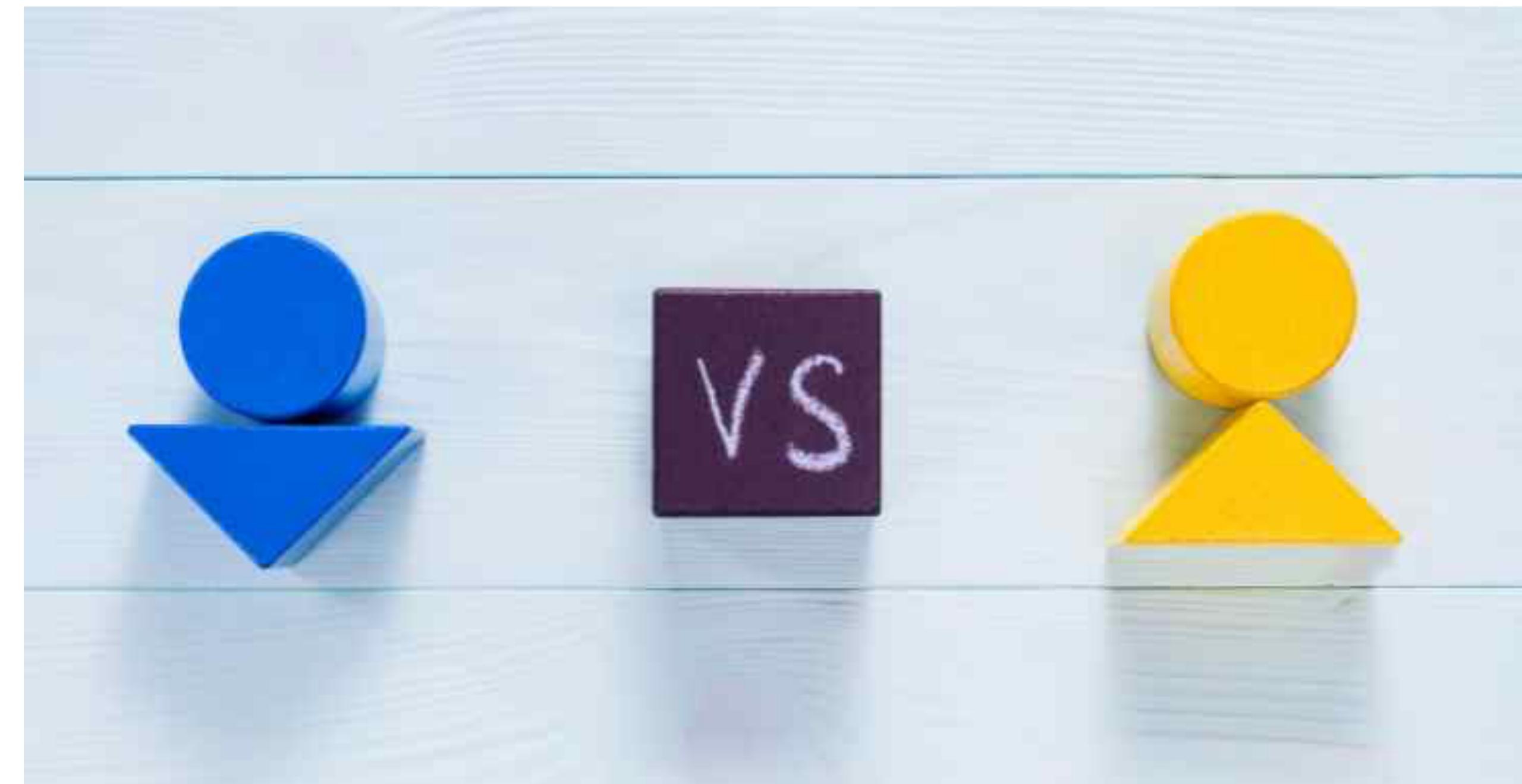
How might we bridge this power chasm? One strategy that has been effective in other contexts is channeling the frustration of the dis-empowered masses into collective action—i.e., action taken by multiple people in pursuit of the same goal or collective good [38]—to demand redress. For example, in the 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 [55]. Importantly, prior to these worker victories, legal doctrines reinforced employer property rights over the ability of employees to organize [28]; 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 [9]. 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 [35], 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 question: what causes collective action efforts in privacy to fail, and how can we improve their likelihood of success?

Shaw et al. [50] introduced a five-stage model for computer-supported collective action (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. Adapting

Top-down vs. Bottom-up

Who ought to define what
privacy is?



Checklist vs. Open-ended

What are the pros and cons of
the different mindsets of
dealing with privacy issues in
design and development?



Privacy as part of design or a separate task?

What do you think of the
different approaches?



Useful Resources

For learning more about privacy engineering

USENIX PEPR: USENIX Conference on Privacy Engineering Practice and Respect

PEPR'24 Conference Program: <https://www.usenix.org/conference/pepr24/program>

“Obfuscation: A User's Guide for Privacy and Protest” by Finn Brunton and Helen Nissenbaum