



Privacy as a New Tech Sector

Norman Sadeh, PhD, CIPT

Professor of Computer Science

Co-Director, Privacy Engineering Program

Carnegie Mellon University

<https://normsadeh.org>

Imagine...



- ...that you are in the market to purchase a car insurance policy...
- **...the insurance companies you request quotes from want to know more about you...**

How Comfortable Are You Disclosing...

...how fast you drive?



credit: <https://driveknight.com/blog/safety/truck-driving-too-fast-for-conditions/>

- Based on a GPS unit you are asked to install in your car?
- *What if this information was obtained from a data broker collecting data from apps on your phone?*

How Comfortable Are You Disclosing...

...where you go and when?

- Based on GPS information?

What if this information was obtained from a data broker collecting data from apps on your phone?



Credit: mobileappdaily.com

How Comfortable Are You Disclosing...

...how many hours you sleep at night?

- *Based on info collected by your smartphone, smartwatch, activity bracelet*
- *...or your bed*



Credit: entrepreneur.com

How Comfortable Are You Disclosing...

...your health history?

- including history of possible substance abuse

Yes	No	Condition
		Diabetes
		Hypertension (high blood pressure)
		Adult or congenital (acquired) heart disease, surgery or procedure
		Family history of heart disease (related death of a family member)
		Stroke/TIA
		Asthma
		Lung/respiratory disease
		COPD
		Ear/eye/nose/throat
		Musculoskeletal

Credit: todaysrdh.com

What Have We Learned?

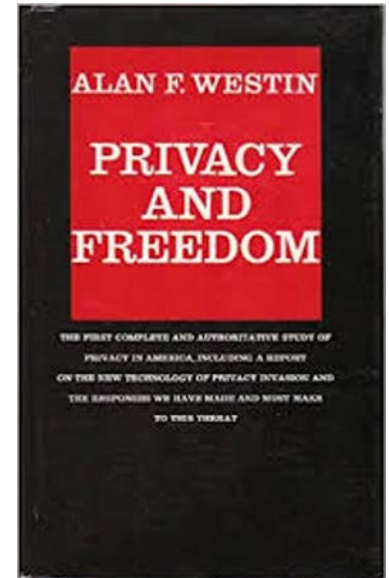
- **Not everyone feels the same way**
- **It's not just what is collected but also how it is collected (e.g. consent), for what purpose, and more**
- **Most of us have reservations about at least a subset of these scenarios...everyone cares about privacy**
- **All this data can readily be collected by a number of different actors using mobile & IoT technologies**

Information/Data Privacy

“...the desire of people to freely choose the circumstances and degree to which individuals will expose their attitudes and behavior to others”

Alan Westin, "Privacy and Freedom," 1967

Individuals should have some **control** over the **collection and use of information/data** about them → the so-called “**Notice and Choice**” framework



But How Does This Relate to Business?

- **Data-centric economy**
 - AI, machine learning/data mining, smartphones, Internet of Things
- New business practices centered around the collection and sharing/selling of people's data have prompted the **emergence of new significantly more stringent regulations** – in the US and abroad
 - These regulations come with significantly steeper fines
 - These regulations are also fueling the **emergence of a new Privacy Tech Sector**



Different Facets of Privacy Today



FTC Imposes \$5 Billion Penalty and Sweeping New Privacy Restrictions on Facebook

July 24, 2019

Following a yearlong investigation by the FTC, the Department of Justice will file a [complaint on behalf of the Commission alleging that Facebook repeatedly used deceptive disclosures and settings](#) to undermine users' privacy preferences in violation of its 2012 FTC order. These tactics allowed the company to share users' personal information with [third-party apps](#) that were downloaded by the user's Facebook "friends." The FTC alleges that many users were unaware that Facebook was sharing such information, and therefore did not take the steps needed to [opt-out of sharing](#).



Apple is turning privacy into a business advantage, not just a marketing slogan

PUBLISHED MON, JUN 7 2021 6:52 PM EDT | UPDATED TUE, JUN 8 2021 12:30 AM EDT



OneTrust named Inc 500 fastest growing company in 2020

Source: Inc 500

Understanding the Scope of New Privacy Regulations – Some Examples*

- **Data minimization** (GDPR and now CPRA)
- **Privacy by design and by default** (GDPR)
- **Opting out of the "sale"** of one's information (CCPA/CPRA – opt-in for people under 16 – parental consent under 13)
- Rights to **be informed**, to **access** one's data, request **rectification/erasure/copy** of one's data (GDPR/CCPA/CPRA)
- **Interpretability** (GDPR)
- Right to **object to automated decision making** (GDPR)

**Many other regulations besides GDPR and CCPA/CPRA, though these are 2 are particularly influential*

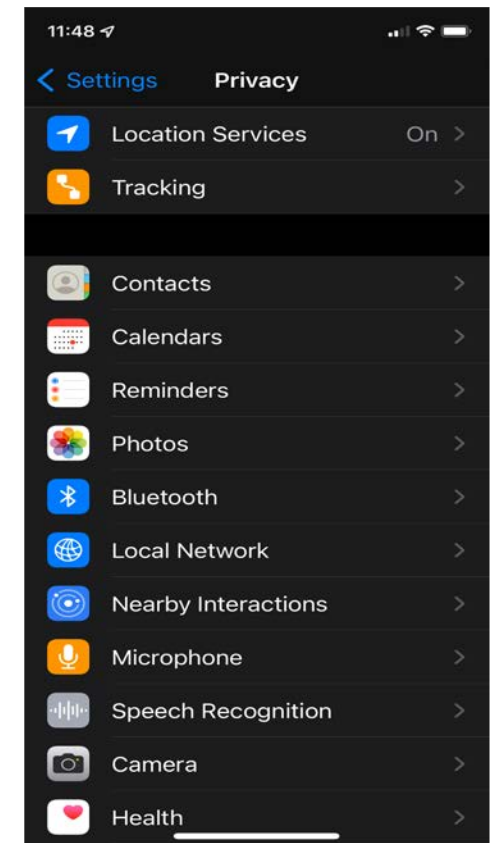
Impact of New Regulations

- Everyone has been **scrambling** to get up to speed
- Significant increase in demand for **privacy professionals**
- Many new regulatory requirements remain **subject to interpretation...**
 - Will need time to settle + **best practices are bound to evolve**
- Emergence of **new tech sector**
 - From software engineering tools to differential privacy tools to consumer-facing tools (e.g., browsers, search engines, messaging tools, VPNs and more)
- **Some requirements remain aspirational...technical gaps & room for innovation**

Our Research at CMU focuses on the “Human Bottleneck” in Privacy

Usability as the Biggest Privacy Challenge

- Regulations like GDPR or CCPA/CPRA represent significant progress towards protecting people's privacy
- Yet in some ways they have also made privacy even more challenging
 - **Longer privacy policies**
 - **More privacy choices to manage**
- The above is compounded by the exploitation of **cognitive and behavioral biases** (e.g. “dark patterns”)



The Human Bottleneck

Lack of:

- **Expertise/Understanding**
- **Time**
- **Attention**
- **Motivation**
- **etc.**



Source: <https://www.datanami.com/2016/09/13/sas-goes-back-future-cognitive-computing-viya/>

What If Computers Understood the Text of Privacy Policies?



National Science Foundation Frontier project started in 2013 – Collaboration between CMU, Fordham School of Law, Stanford Center for Internet and Society, University of Michigan, Columbia University, Penn State University, University of Cincinnati

Automatic Identification of Data Practice Disclosures

Yahoo! yahoo.com

Arts Business Computers Games Health Home Recreation Reference Regional Society World

Privacy Practices

Click a category to filter practice statements.

- First Party Collection/Use 67
- Third Party Sharing/Collection 21
- User Choice/Control 6
- User Access, Edit and Deletion 6
- Data Retention 1**
- Data Security 8
- Policy Change 6
- Do Not Track 0
- International and Specific Audiences 8

Data Retention ?

Retention period ?

☒ All
☐ Indefinitely (1)

Purpose of retention ?

☒ All
☐ Unspecified (1)

more filters v

Privacy Policy

Yahoo News Privacy Policy from Sep 25, 2014.
125 privacy practice statements in total

This privacy policy also applies to [Flickr](#), [Yahoo Finance](#), [Yahoo News](#), [Yahoo Sports](#), and [Yahoo! Good Morning America](#).

We reserve the right to send you certain communications relating to the Yahoo service, such as service announcements, administrative messages and the Yahoo Newsletter, that are considered part of your Yahoo account, without offering you the opportunity to opt out of receiving them.

You can delete your **Yahoo account by visiting our Account Deletion page. Please click here to read about information that might possibly remain in our archived records after your account has been deleted.**

CONFIDENTIALITY AND SECURITY

A user's user profile is retained indefinitely to fulfill an unspecified purpose.

We limit access to personal information to those employees who we believe reasonably need to come into contact with that information to provide products or services to you or in order to do their jobs.

We have physical, electronic, and procedural safeguards that comply with federal regulations to protect personal information about you.

To learn more about security, including the security steps we have taken and security steps you can take, please read [Security at Yahoo](#).

CHANGES TO THIS PRIVACY POLICY

Yahoo may update this policy. We will notify you about significant changes in the way we treat personal information by sending a notice to the primary email address specified in your Yahoo account or by placing a prominent notice on our site.

QUESTION AND SUGGESTIONS

If you have questions, suggestions, or wish to make a complaint, please complete a feedback form.

https://explore.usableprivacy.org/browse/category/

USABLEPRIVACY.ORG EXPLORE [About](#) [Browse Privacy Policies](#)

Browse

by [Category](#) [Readability](#) [Popularity](#)

- Arts **68**
- Business **53**
- Computers **42**
- Games **26**
- Health **35**
- Home **37**
- Kids and Teens **46**
- News **32**
- Recreation **42**
- Reference **31**

Arts **68**

E! Online
Privacy policy from Jan 14, 2015 with 256 practice statements.

FOX Sports
Privacy policy from Jun 11, 2015 with 215 practice statements.

Racked
Privacy policy from May 1, 2014 with 204 practice statements.

[See more](#)

Business **53**

Blogger
Privacy policy from Jun 30, 2015 with 241 practice statements.

AOL
Privacy policy from Jun 23, 2015 with 232 practice statements.

Allstate
Privacy policy from May 29, 2015 with 226 practice statements.

[See more](#)

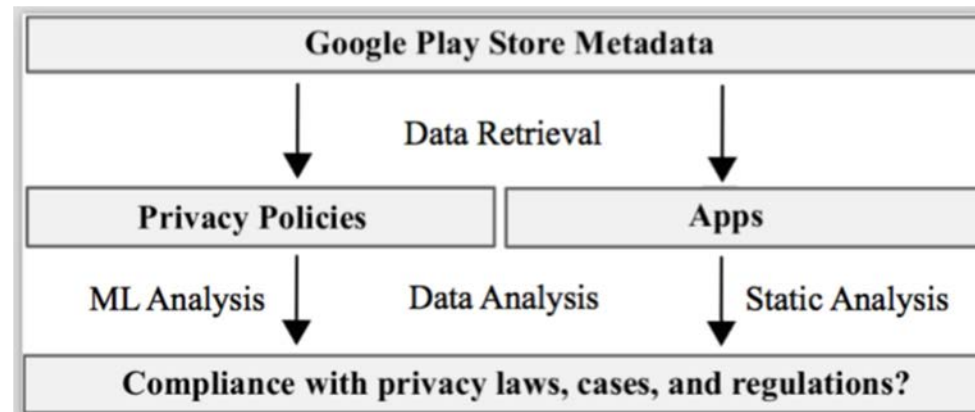
Opt-Out Easy Browser Extension

Browser extension tools to automatically **identify and categorize opt-out choices** buried deep in the text of privacy policies – available in **Google chrome and Firefox stores**



Vinayshekhar Bannihatti Kumar, Roger Iyengar, Namita Nisal, Yuanyuan Feng, Hana Habib, Peter Story, Sushain Cherivirala, Margaret Hagan, Lorrie Faith Cranor, Shomir Wilson, Florian Schaub, Norman Sadeh, "**Finding a Choice in a Haystack: Automatic Extraction of Opt-Out Statements from Privacy Policy Text**", WWW '20, Apr 2020 [\[pdf\]](#)

Can We Automatically Check for Potential Compliance Issues?



- Training **machine learning classifiers** to extract relevant policy statements
- Compare these statements against:
 - **Regulatory requirements**
 - What the software actually does
 - **Static and dynamic code analysis**

Zimmeck, S., Story, P., Smullen, D., Ravichander, A., Wang, Z., Reidenberg, J., Russell, N.C., Sadeh, N., "MAPS: Scaling Privacy Compliance Analysis to a Million Apps," in Proceedings on Privacy Enhancing Technologies, Vol. 3, pp. 66-86, 2019.
<https://doi.org/10.2478/popets-2019-0037>

Privacy Question Answering

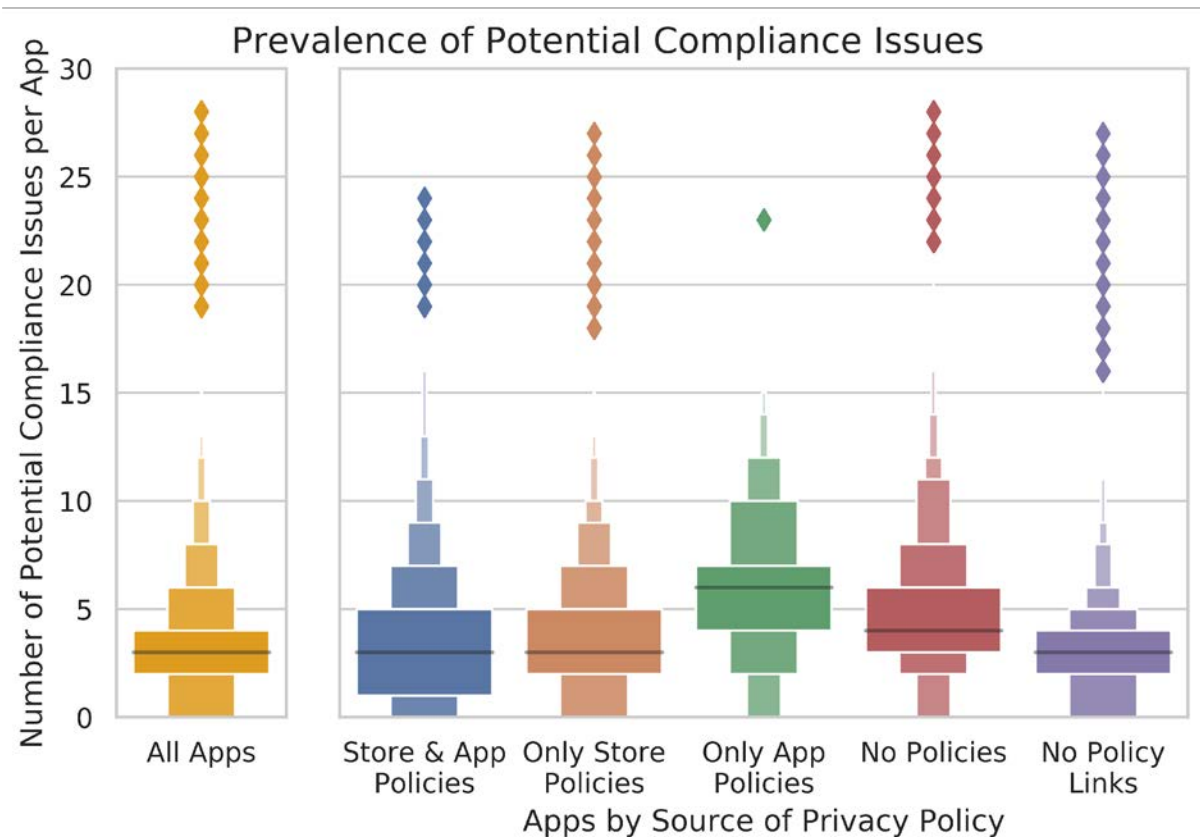
- **One-size-fits-all summaries of privacy policies only go so far**
- Different people have different questions at different points in time
- Could we develop privacy question answering functionality?
- A number of challenges
 - Can people accurately articulate their questions. If not, how can we help them?
 - How do we provide useful answers – vague policies, inaccurate classifiers
 - etc.

Question answering for privacy policies: Combining computational and legal perspectives. A Ravichander, AW Black, S Wilson, T Norton, N Sadeh, EMNL 2019 Conference, arXiv preprint arXiv:1911.00841

Breaking Down Walls of Text: How Can NLP Benefit Consumer Privacy?. A Ravichander, AW Black, T Norton, S Wilson, N Sadeh, ACL/IJCNLP 2021. <http://dx.doi.org/10.18653/v1/2021.acl-long.319>

Analysis of over 1 million Android Apps in Google Play Store

- Average number of potential compliance issues per app is 3.47 and the median is 3
- Requires manual vetting – both policy and app behavior to confirm potential compliance issue



Press Coverage – Notice the Irony

FastCompany
informing their readers
about their new policy



FastCompany's article
about our research



03.19.18

You're Never Going To Read That Privacy Policy. Could AI Help?


This AI trained on legalese acts like a personal translator of confusing, opaque privacy statements.



ADVERTISEMENT



Tools for Developers



Hmm, I wonder if I need a privacy policy for my app. Also, what should I write in there? I am lost ...

1. Template Provisioning

E.g., the GDPR requires policies to notify users of their rights to request data access, rectification, erasure, restriction of processing, objection of processing, and portability (Art. 13(2)(b))

Information to be provided where personal data are collected from the data subject

Article 13

1. Where personal data relating to a data subject are collected from the data subject, the controller shall, at the time when personal data are obtained, provide the data subject with all of the following information:

(a) the identity and the contact details of the controller and, where applicable, of the controller's representative;

(b) the contact details of the data protection officer, where applicable;

2. Code Analysis

Plist Permissions and Swift API Usage

```
import Foundation
import UIKit

class ViewController: UIViewController, UIImagePickerControllerDelegate, UINavigationControllerDelegate {


    @IBOutlet weak var galleryImage: UIImage!
    @IBOutlet weak var photoLibraryImage: UIImage!
    @IBOutlet weak var mediaLibraryImage: UIImage!
    var network = [String]()

    // Initialization (optional) and receive callback
    let UIImagePickerController = UIImagePickerController()
    var cameraImage = UIImagePickerController()
    let mediaLibraryImage = UIImagePickerController()
    let photoLibraryImage = UIImagePickerController()
}
```

3. Wizard Fine Tuning

Don't worry! I got you covered. Here is the privacy analysis of your app.

You can adjust the generated privacy policy via the checkboxes.



CAMERA: USED

The use of camera information was determined because of these specific lines in your app's code:

Show Detected API and Third Party Library Calls

Code	File	Line Number	Used	Usage Description
<key>NSCameraUsageDescription</key>	missing-reports/krypton-ios-master/Krypton/Info.plist	77	USED	

You can customize the recommended statement about the detected data practice using the checkboxes below.

Specific Practices

- ☒ accesses camera data on user's device
- ☐ sends camera data to the developer's server
- ☐ stores camera data on user's device

Purposes

- ☒ performs a functionality of the app
- ☐ advertising
- ☐ analytics

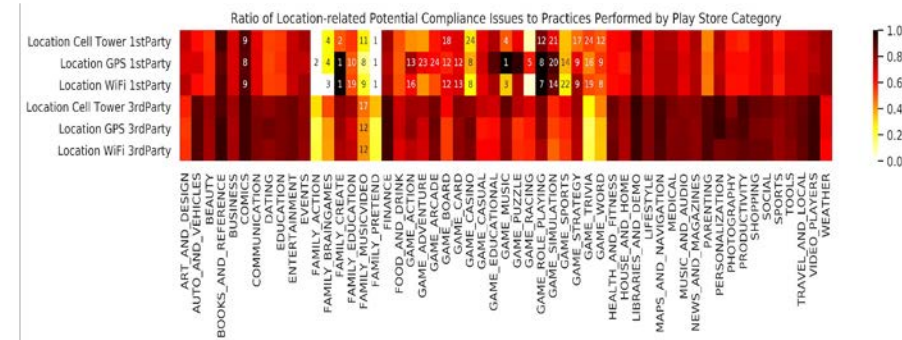
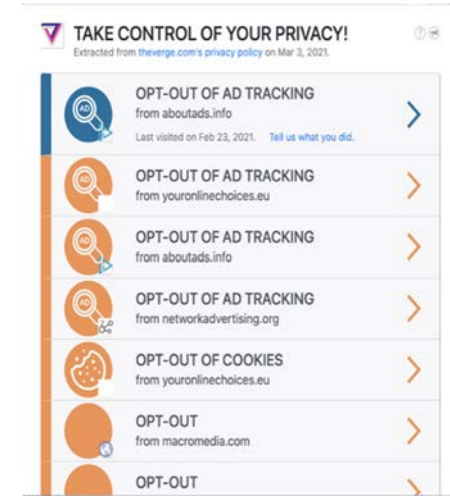
You can adjust the recommended statement below and copy it into your application's privacy policy.

The app accesses camera data on your device for the purposes of the app's camera functionality (PLEASE DESCRIBE THE FUNCTIONALITY).

Compliance traceability: Privacy policies as software development artifacts, S Zimmeck, P Story, R Goldstein, D Baraka, S Li, Y Feng, N Sadeh, Privacy, Usability and Transparency Workshop (PUT 2019) at PoPETs 2019 conference

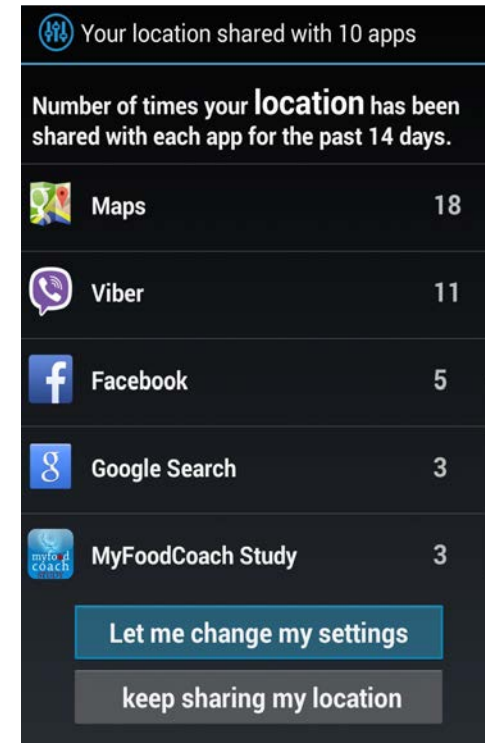
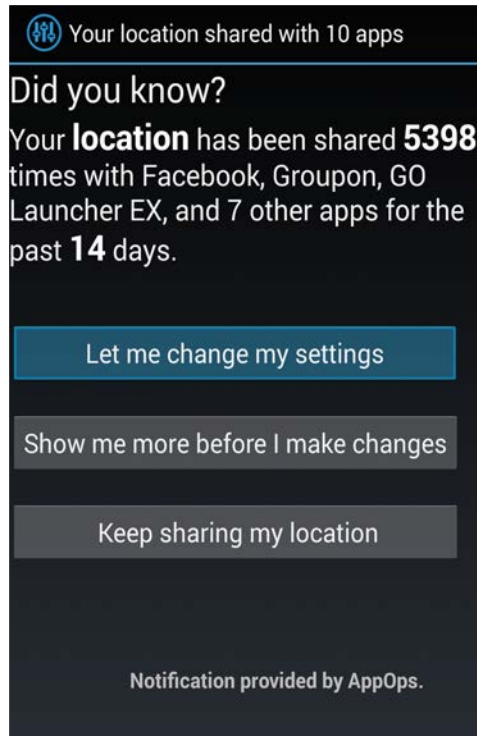
Examples of Techniques and Tools

- Collection of NLP/ML **classifiers** capable of **automatically extracting and analyzing** a variety of privacy policy disclosure statements
- Browser extension tools to automatically **identify and categorize opt-out choices** buried deep in the text of privacy policies – available in **Google chrome and Firefox stores**
- **Mobile app privacy compliance tools** comparing privacy policy disclosures and actual practices revealed by the code of mobile apps – tools prototyped with **industry and regulators** – incl. **analysis of over 1 million Android apps**
- **Tool to help developers write privacy policies and create privacy labels** through static code analysis and wizard functionality
- **Privacy Question Answering assistant**



Could Computers Also Help Motivate People to Take Advantage of Privacy Settings?

Nudging Users: Surprise People with Something That Will Motivate Them to Pay Attention



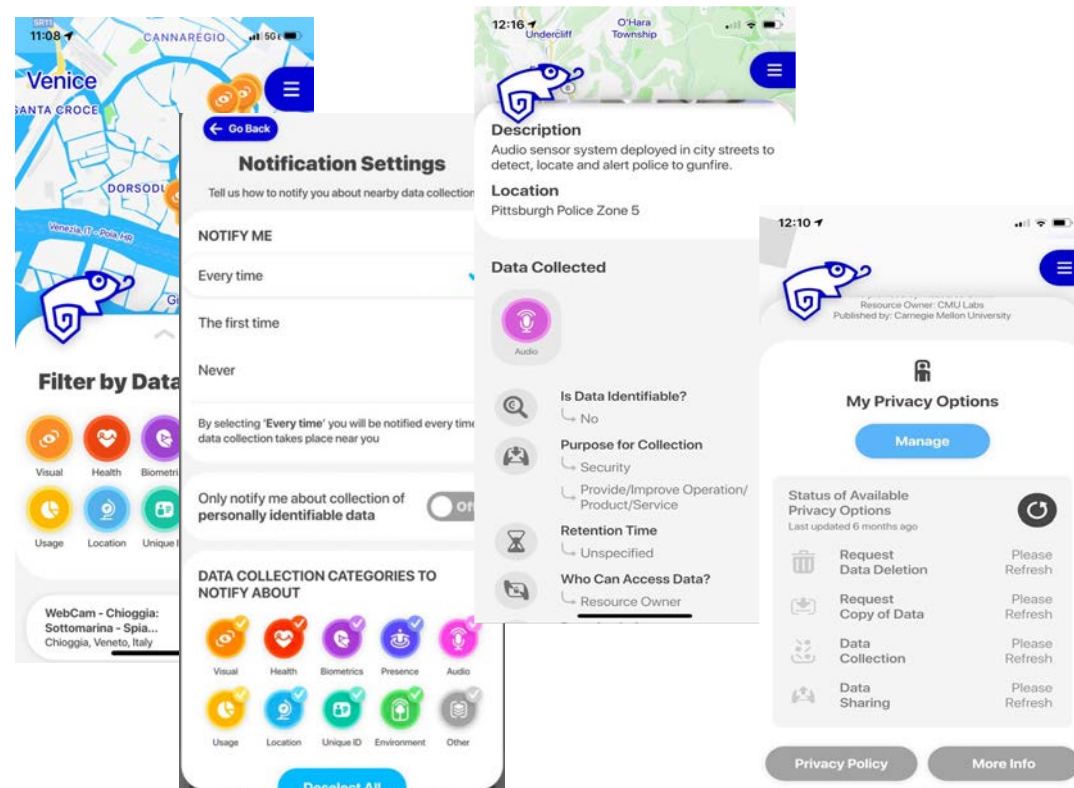
H. Almuhimedi, F. Schaub, N. Sadeh, I. Adjerid, A. Acquisti, J. Gluck, L. Cranor, Y. Agarwal, "Your Location Has Been Shared 5,398 Times!: A Field Study On Mobile App Privacy Nudging," ACM CHI 2015.

What If Computers Could Help Users Configure Privacy Settings?

PRIVACYASSISTANT.ORG
the personalized privacy assistant project

Examples of Models, Techniques and Tools

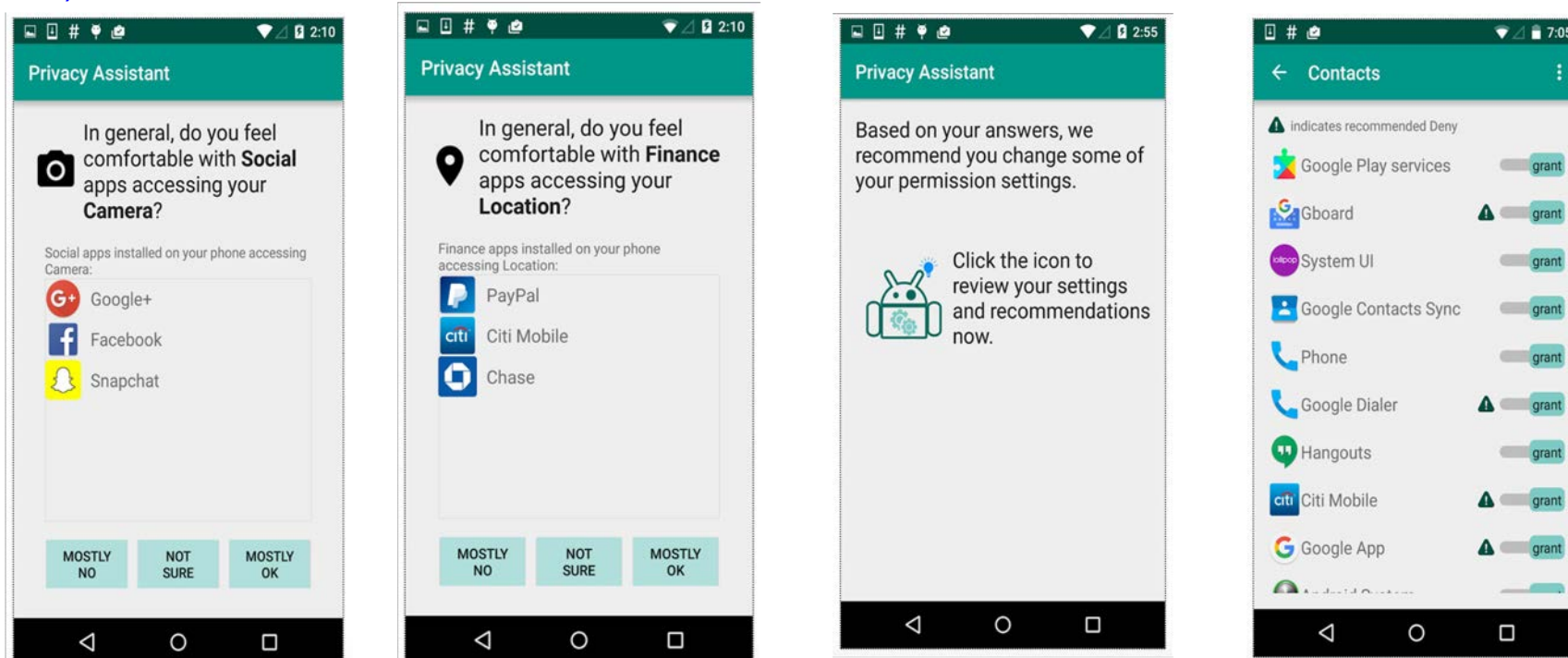
- **Mobile App Permission Manager, Mobile App Privacy Nudges, Privacy Preferences Modeling** – influenced development at Google and Apple (e.g. background privacy notices introduced in iOS13, expressiveness of permissions)
- **Mobile App Permission Assistant** – available in Google Play store for several years
- **Mobile App Privacy Labels** (2013CHI paper...2020 introduction in iOS14)
- Models of People's **Privacy Expectations and Notification Preferences** for video analytics
- **Privacy Infrastructure for the Internet of Things** – hosting over 100,000 discoverable IoT resource descriptions + APIs for privacy choices
- Design of **CCPA/CPRA Opt-Out Notice** adopted by California Attorney General



IoT Assistant app available in iOS and Google Play stores
See: <https://www.iotprivacy.io>

Privacy Assistants II

Generating recommendations rather than automating privacy decisions (Google Play store for several years)



**Vast majority of recommendations accepted by users and kept despite nudges to reconsider
Successfully deployed in Google Play store for several users – rooted Android phones only...**

IoT: Additional Usability Challenges



- How likely are you to notice this sign?
- Does this include facial recognition?
- What about facial expression or scene recognition?
- How long is the data retained?
- Do I get to opt in/opt out?
- Is this GDPR compliant?

Good Luck with this...

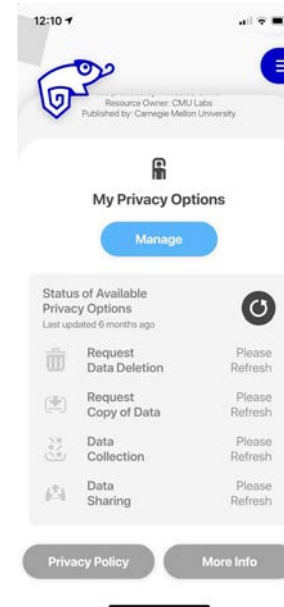
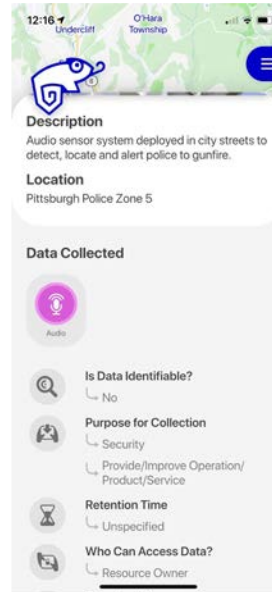
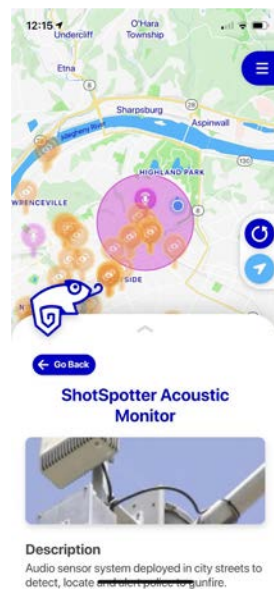


The notice required by the California Consumer Privacy Act, at knee height, at a B8ta outlet in San Francisco. PHOTOGRAPH: LAURYN HILL

QR code for CCPA compliance at knee height...

An IoT Privacy Infrastructure

- Portal to help people publicize the presence and data practices associated with IoT systems, incl. privacy controls (e.g. opt-in/out, deletion, etc.)
- IoT Assistant app (Android and iOS) for users to discover nearby IoT systems and their data practices, incl. accessing any available privacy controls



<https://www.iotprivacy.io>

US Patents 10,956,586 with additional USPTO and EPO patents pending

IoT Privacy Infrastructure

- The IoT Assistant App [video](#)
- IoT Privacy Infrastructure Overview [video](#)
- Hosting nearly 150,000 IoT system entries today

A. Das, M. Degeling, D. Smullen, and N. Sadeh, Personalized Privacy Assistants for the Internet of Things, 2018 IEEE Pervasive Computing: Special Issue - Securing the IoT, April 2018

Educating the Workforce - <https://privacy.cs.cmu.edu/>

First of its kind privacy engineering program in the world started in 2012

- Both full-time and part-time master's programs
- Professional certificate degree – 4 cohorts per year
- **Informed by our research**
- Currently training approx. 140 people per year

Carnegie Mellon University

Privacy Engineering Program

Masters Program

Certificate Program

Informing Public Policy

- We are regularly called to inform public policy discussions based on our research (mainly US but also EU)
- Current focus includes emphasizing unrealistic user burden associated with privacy choices (GDPR and CCPA/CPRA) and advocating, based on our research, for open APIs to allow privacy assistants to help users take advantage of their privacy rights (e.g., opt-in, opt-out, deletion, etc.)

Think about your **browser communicating your cookie preferences** to websites you visit, your **mobile app assistant configuring your permission preferences** for individual apps, your **IoT Assistant** selectively **notifying** you and communicating your **opt-in/opt-out preferences** rather than **requiring you to tediously enter the same choices over and over again** (e.g., videoanalytics, WiFi tracking, etc).



FTC Privacy Conference



International Conference of Data Protection and Privacy Commissioners (EU Parliament)

Privacy as a New Tech Sector

Concluding Remarks – Pri(vacy)Burgh?

- Privacy is a fundamental human right
- Data privacy is about controlling who collects our data and what they can do with it
- Regulations are becoming increasingly stringent
- Businesses have to rethink their data practices and ensure their workforce has the necessary training
- Privacy is challenging and new regulations are also **fueling a new privacy tech sector to fill the technical gaps**
- **Pittsburgh is a leading force in privacy tech**

USABLE PRIVACY.ORG
the usable privacy policy project

PRIVACYASSISTANT.ORG
the personalized privacy assistant project

More details at:

[*https://usableprivacy.org*](https://usableprivacy.org)

[*https://privacyassistant.org*](https://privacyassistant.org)

[*https://explore.usableprivacy.org*](https://explore.usableprivacy.org)

[*https://www.iotprivacy.io*](https://www.iotprivacy.io)

[*https://privacy.cs.cmu.edu/*](https://privacy.cs.cmu.edu/)

[*https://www.normsadeh.org/*](https://www.normsadeh.org/)