ICISSP 2022

8th International Conference on Information Security Systems and Privacy

Why Has Usability Become Privacy's Biggest Challenge and What Can We Do About It?

Norman Sadeh
Professor of Computer Science
Carnegie Mellon University

https://normsadeh.org

© Copyright 2022, Norman Sadeh, All rights reserved.

Privacy Threats Are Everywhere

Source: CSO online

Data-Hungry Economy

- AI/ML
- IoT sensors everywhere
- Myriads of APIs and dataflows

and also:

- unscrupulous/ignorant data collectors/processors

© Copyright 2022, Norman Sadeh, All rights reserved.

New Regulations

- Increasingly more specific data practice disclosure requirements
- Increasingly specific data subject rights
- Emerging, yet loosely specified, usability expectations
- New, significantly steeper financial penalties

© Copyright 2022, Norman Sadeh, All rights reserved.

"*All is for the Best in the Best of Possible Worlds*" (Voltaire)

Privacy Policy Before GDPR

COMPANY	WORD COUNT	READING TIME*	READING LEVEL
Google	2722	12:54	14
Facebook	5420	26:24	11
Reddit	5524	26:42	12
Amazon	2627	12:48	12
Wikipedia	2881	13:49	15
Yahoo	1611	7:44	14
Twitter	3764	18:37	13
eBay	5244	25:27	18
Instagram	2981	14:26	12
Netflix	3046	15:10	15

* Minutes:Seconds

Progress?

Privacy Policy After GDPR

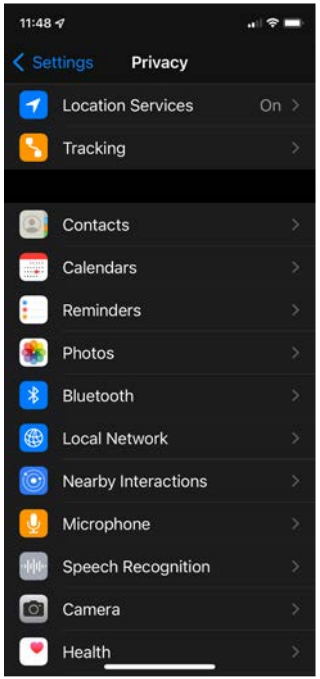
COMPANY	WORD COUNT	READING TIME*	READING LEVEL
Google	4036	19:11	14
Facebook	4233	20:41	13
Reddit	3414	16:39	12
Amazon	3837	18:24	13
Wikipedia	5617	27:06	14
Yahoo	2225	11:12	13
Twitter	4880	22:25	13
eBay	5666	27:32	20
Instagram	4221	20:38	13
Netflix	3417	16:39	16

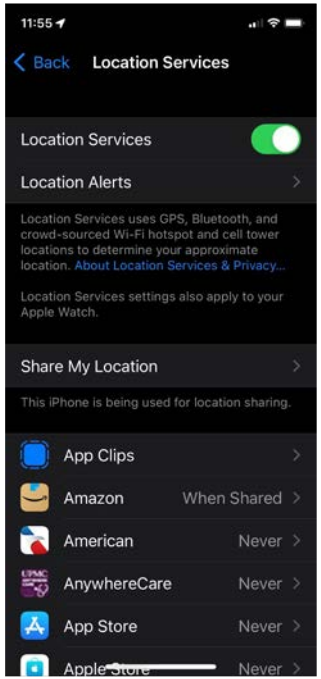
* Minutes:Seconds

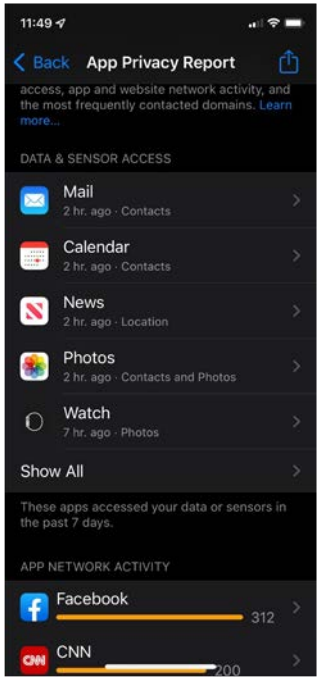
Source: https://www.varonis.com/blog/gdpr-privacy-policy

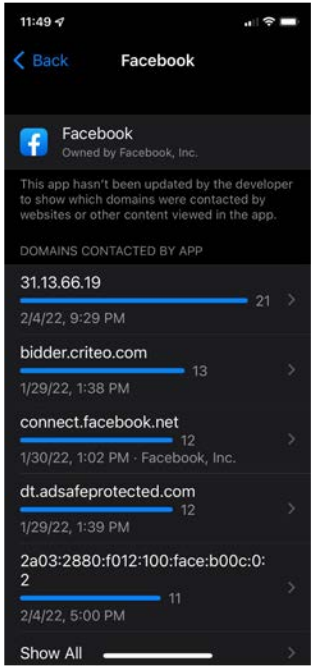
© Copyright 2022, Norman Sadeh, All rights reserved.

…And Who Has the Time to Review & Manage All these Privacy Settings?









© Copyright 2022, Norman Sadeh, All rights reserved.

The Human Bottleneck

Lack of:

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

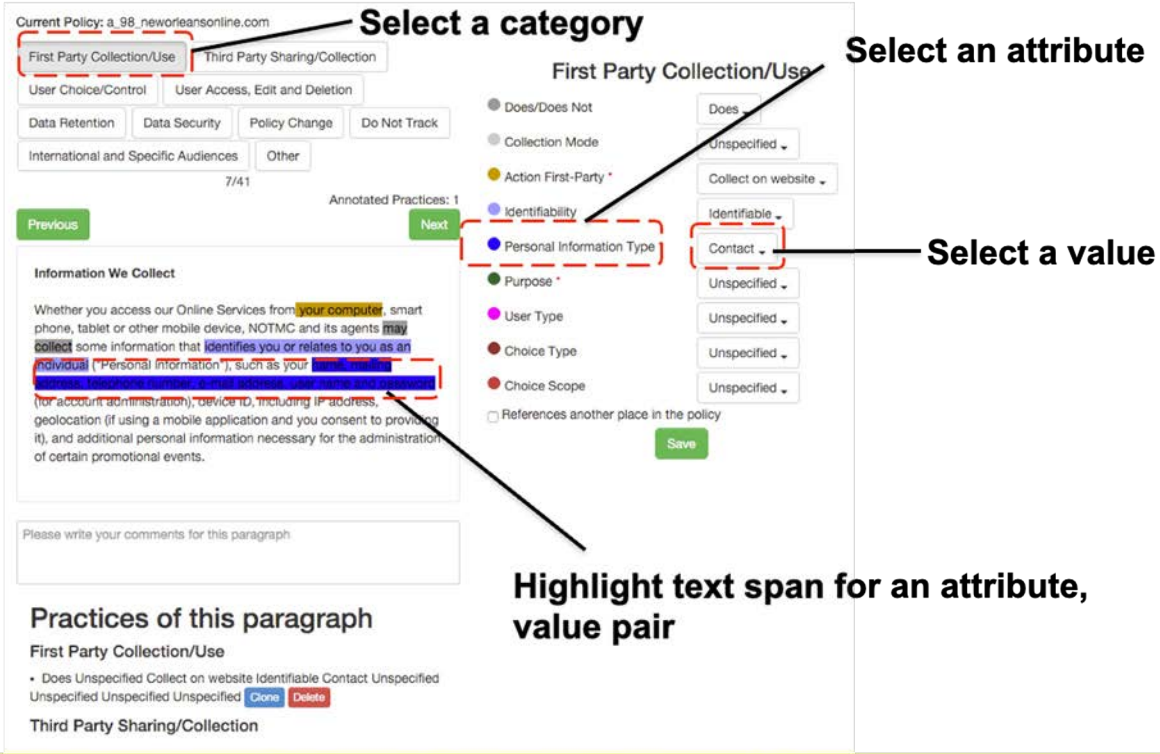
Source: https://www.datanami.com/2016/09/13/sas-goes-back-future-cognitive-computing-viya/

© Copyright 2022, Norman Sadeh, All rights reserved.

What If Computers Understood the Text of Privacy Policies?

© Copyright 2022, Norman Sadeh, All rights reserved.

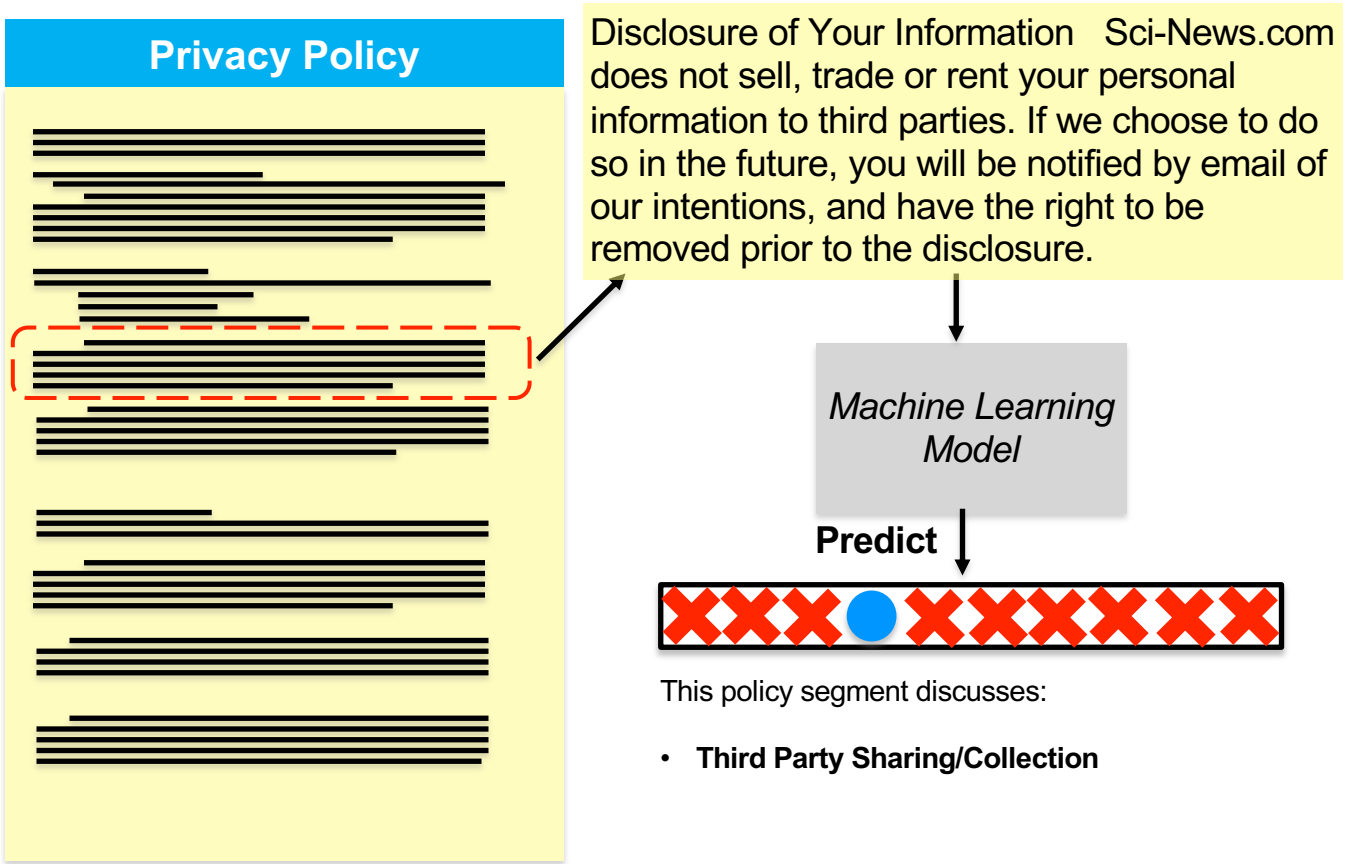
Annotation Tool



S. Wilson, F. Schaub, A. Dara, F. Liu, S. Cherivirala, P.G. Leon, M.S. Andersen, S. Zimmeck, K. Sathyendra, N.C. Russell, T.B. Norton, E. Hovy, J.R. Reidenberg, N. Sadeh, "The Creation and Analysis of a Website Privacy Policy Corpus", ACL '16: Annual Meeting of the Association for Computational Linguistics, Aug 2016

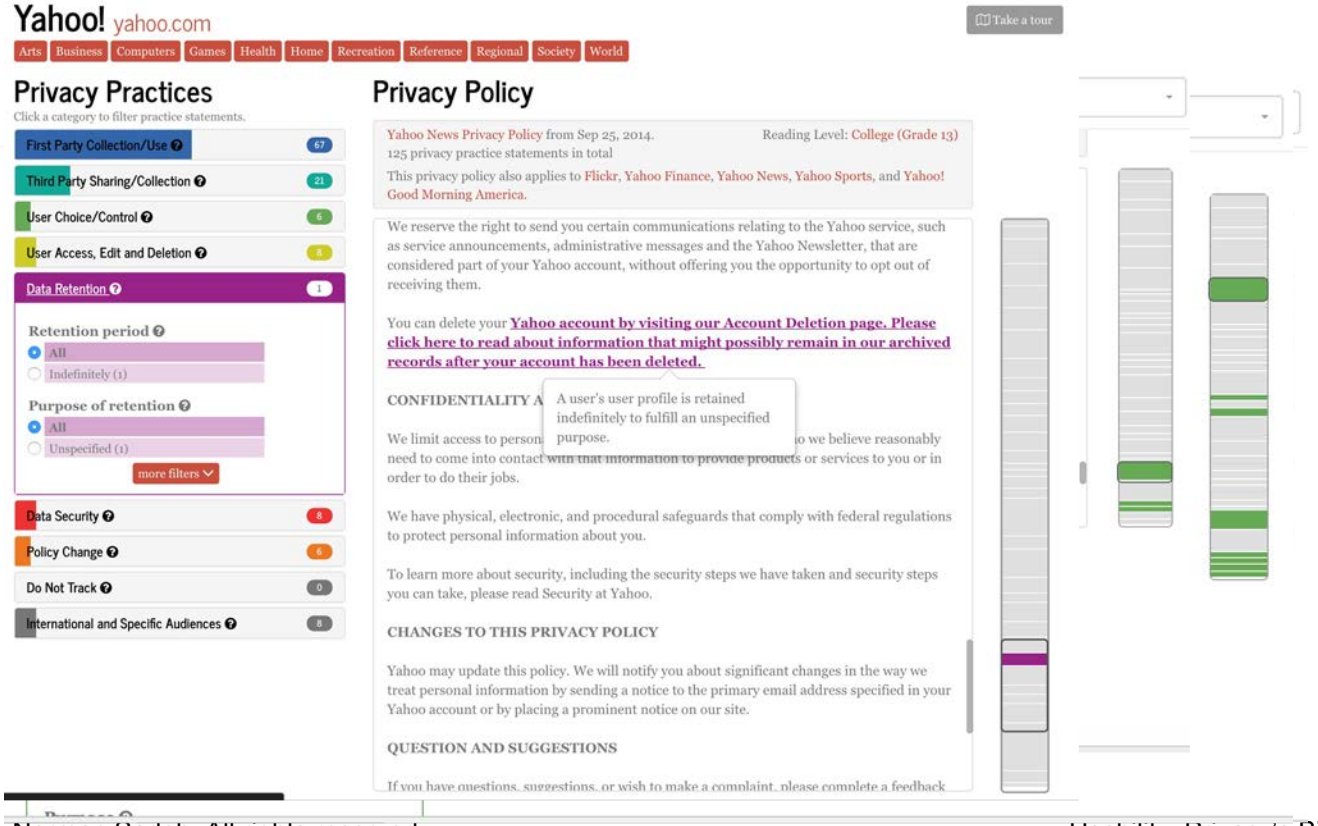
© Copyright 2022, Norman Sadeh, All rights reserved.

A First Task: Segment Annotation

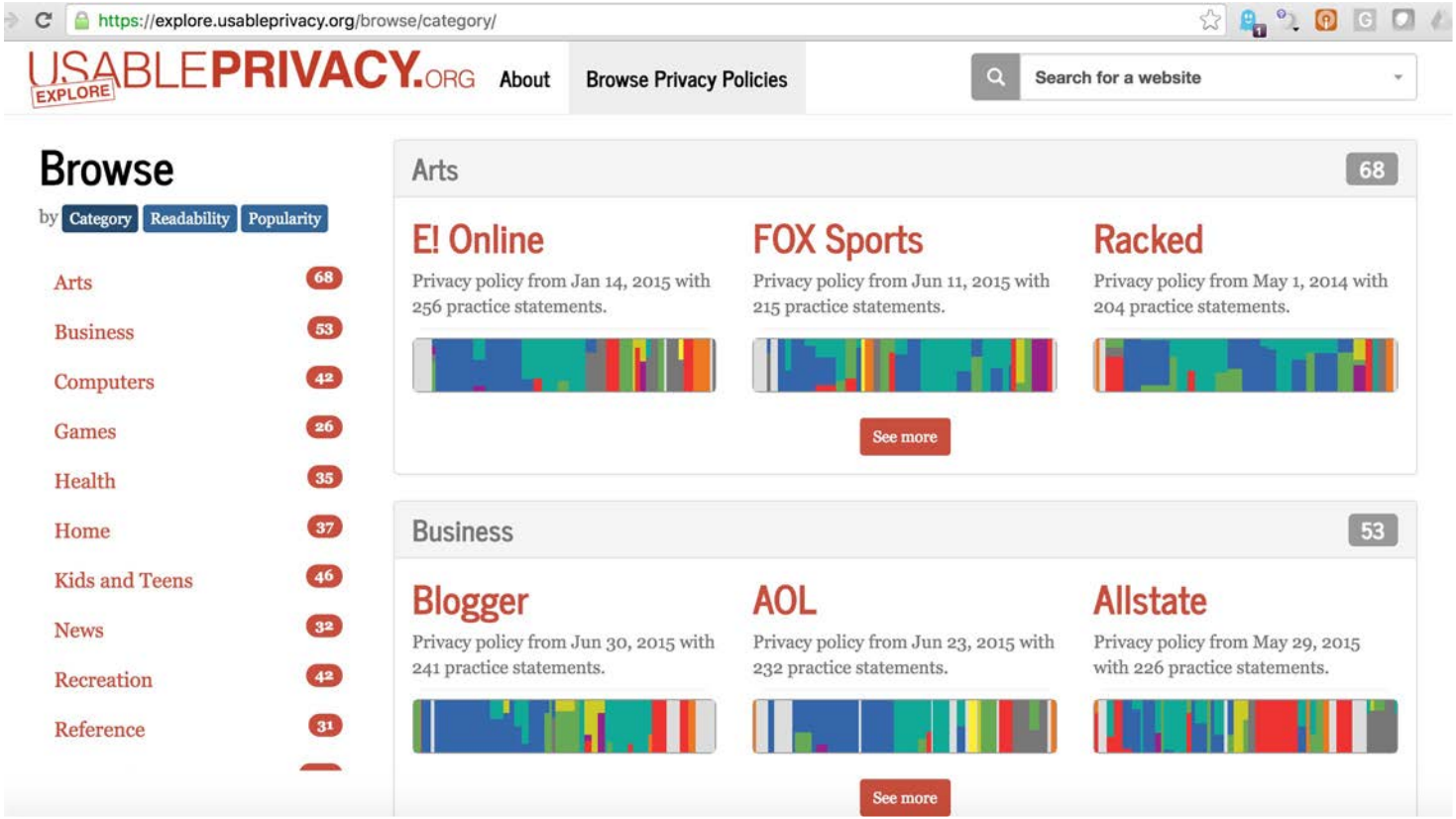


© Copyright 2022, Norman Sadeh, All rights reserved.

Automatic Identification of Data Practice Disclosures



© Copyright 2022, Norman Sadeh, All rights reserved.



© Copyright 2022, Norman Sadeh, All rights reserved.

Press Coverage – Notice the Irony

FastCompany informing their readers about their new policy

FastCompany's article about our research

© Copyright 2022, Norman Sadeh, All rights reserved.

Privacy Question Answering

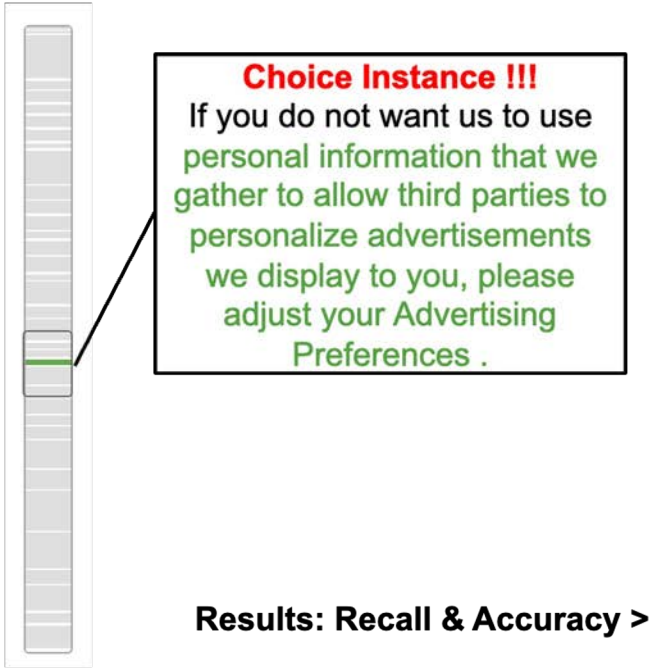
- **One-size-fits-all summaries of privacy policies only go so far**
- Different people have different questions at different points in time
- Could we develop privacy question answering functionality?
- A number of challenges
  - Can people accurately articulate their questions. If not, how can we help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

© Copyright 2022, Norman Sadeh, All rights reserved.

User Choice Instance Extraction



- User choices often buried deep in the text of long policies
- Is it possible to **automatically extract information** about such “choice instances” from privacy policies?
- Use Natural Language Toolkit tokenizer to subdivide segments into sentences & build classifiers

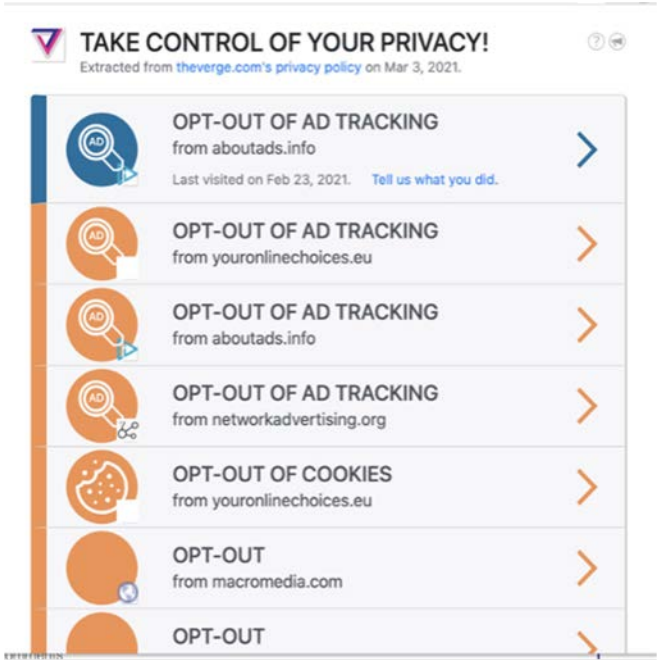
Results: Recall & Accuracy > 90%

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]

© Copyright 2022, Norman Sadeh, All rights reserved.

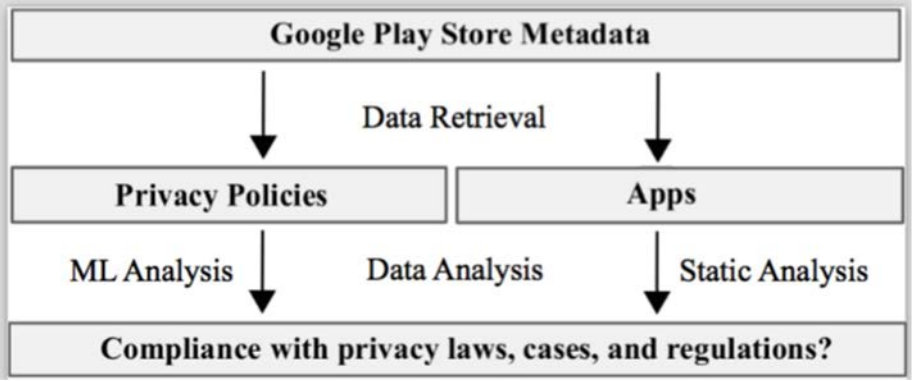
Opt-Out Easy Browser Extension

- Automatically identify and categorize opt-out choices in the text of privacy policies
- And present them in an easy-to-use interface to users as they browse the web
- Available in Google Chrome store and Firefox store - Watch our video



© Copyright 2022, Norman Sadeh, All rights reserved.

Can We Automatically Check for Potential Compliance Issues?

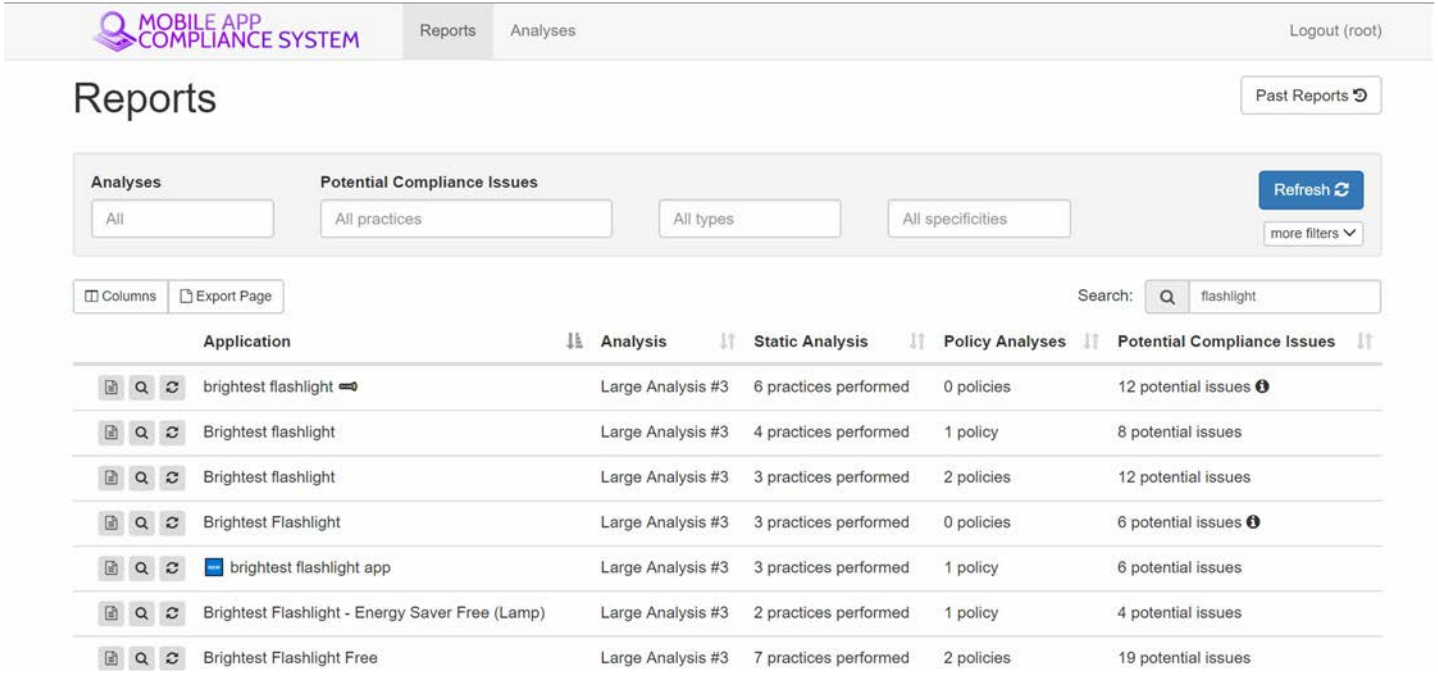


- Training **machine learning classifiers** to extract relevant policy statements
- Compare these statements against:
  - **Regulatory requirements**
  - What the software actually does
    - **Static and dynamic code analysis**

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

© Copyright 2022, Norman Sadeh, All rights reserved.

Reports View with Search and Filters



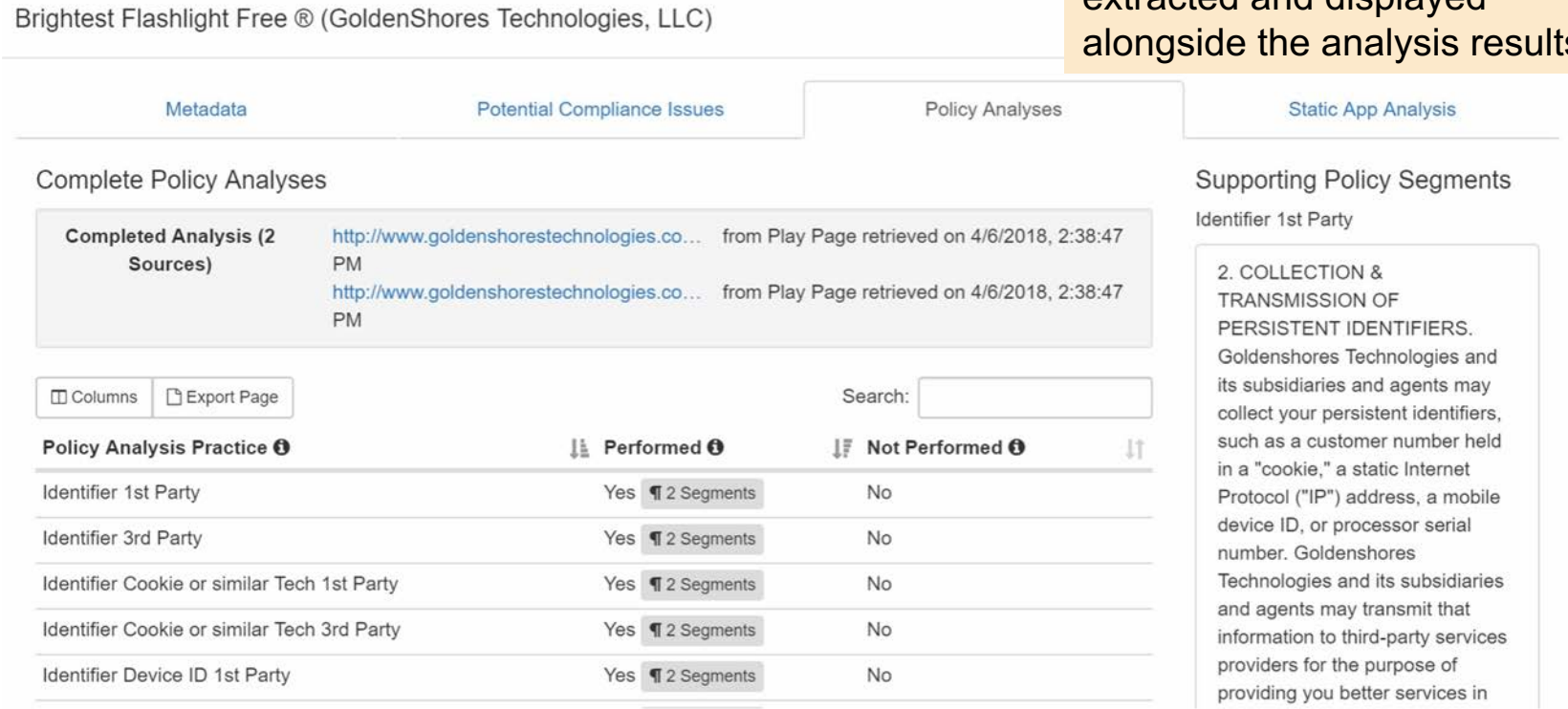
Search and Filter according to

- **Practices** (e.g., GPS location collection)
- **Type** (Is GPS location not mentioned in the policy or does the policy explicitly state it is not occurring?
- **Specificity** (Perhaps "location" suffices in some jurisdictions not specifically requiring to mention "GPS.")
- **Many more …**

© Copyright 2022, Norman Sadeh, All rights reserved.

Flashlight App: Policy Results

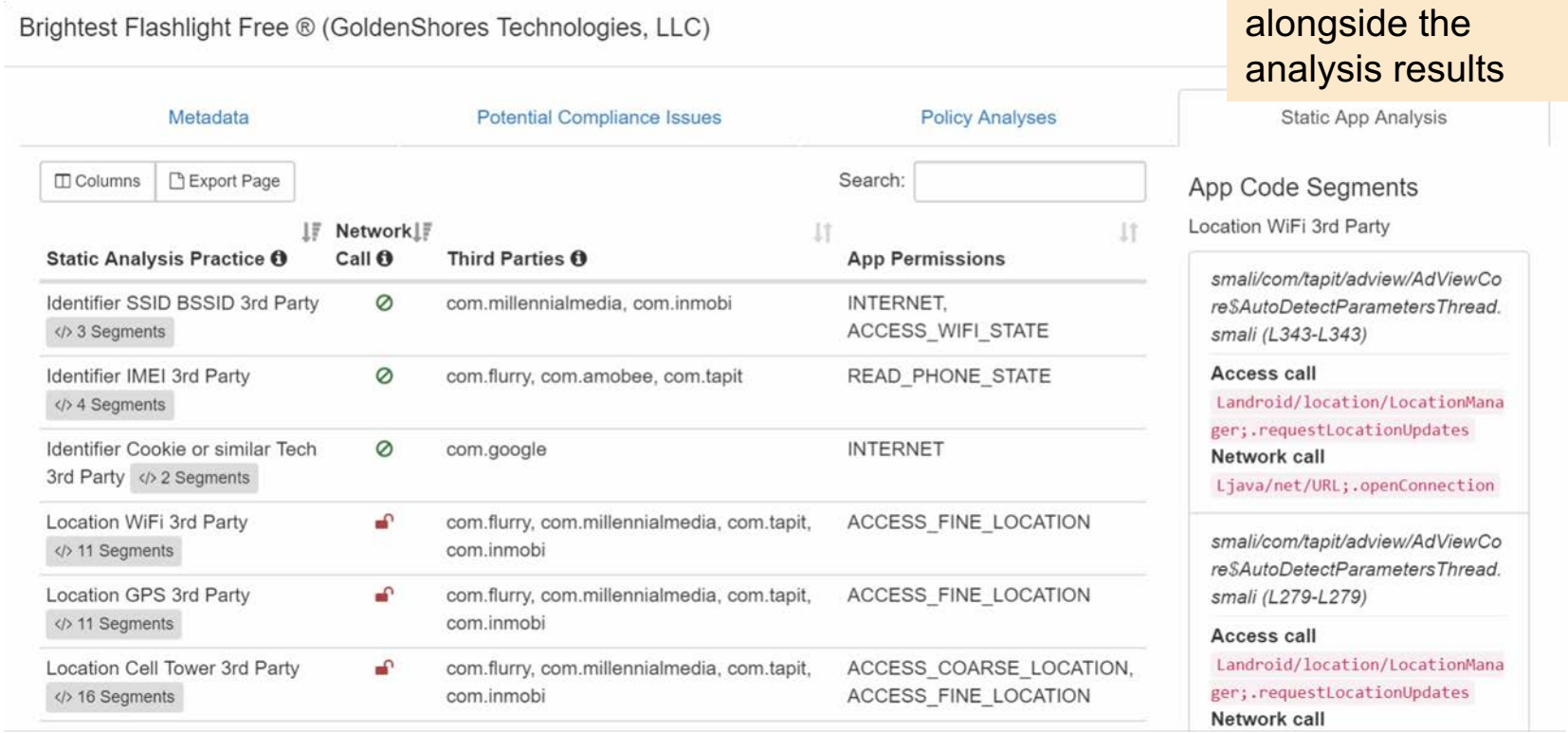
Relevant parts of the policies are extracted and displayed alongside the analysis results



© Copyright 2022, Norman Sadeh, All rights reserved.

Flashlight App: Static Analysis Results

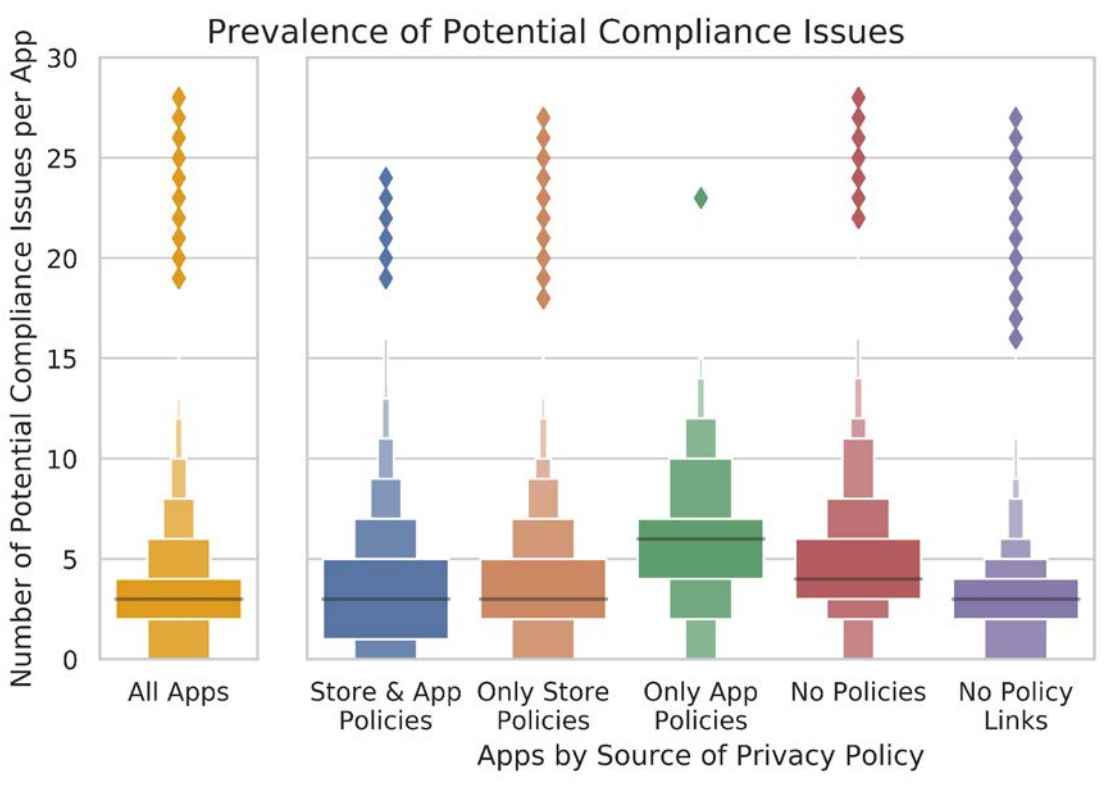
Relevant parts of the code are extracted and displayed alongside the analysis results



© Copyright 2022, Norman Sadeh, All rights reserved.

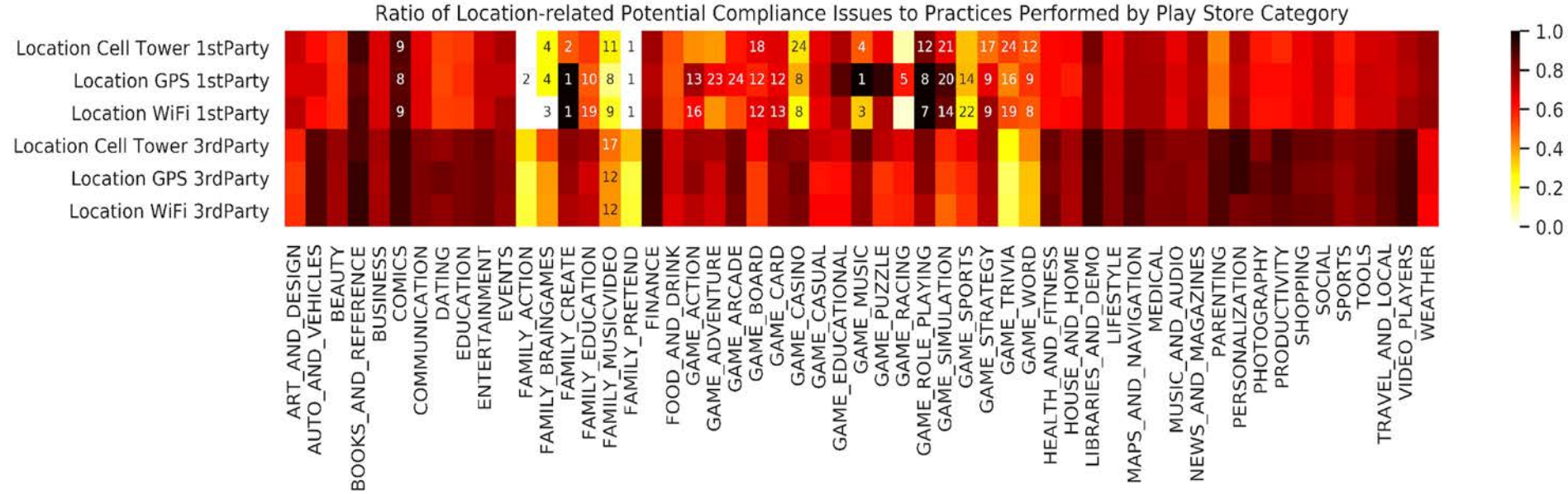
Analysis of over 1 million Android Apps in Google Play Store

- 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



© Copyright 2022, Norman Sadeh, All rights reserved.

Developers Struggle with 3rd Party APIs



- Lighter colors indicate greater transparency of practices. Darker colors indicate that practices are being performed but not disclosed.
- Cells with fewer than 25 apps performing the practice are annotated with the respective number of apps.

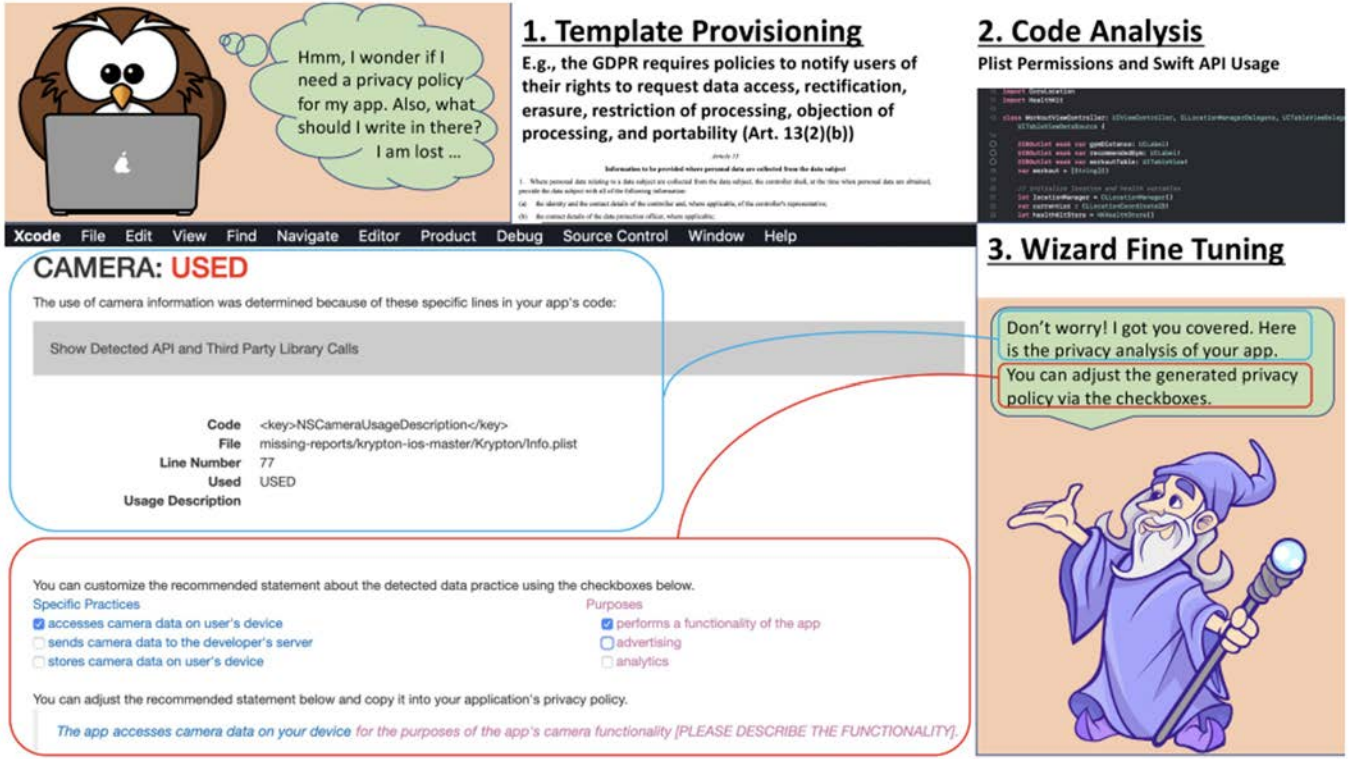
© Copyright 2022, Norman Sadeh, All rights reserved.

Other Collaborations

- Collaboration with California Attorney General's office
- COPPA report compiled for Federal Trade Commission
  - Focusing on location, apps with a large number of downloads, and companies based in the US
- CDT report on mobile apps for connected cars
- Work with large European electronics manufacturer – checking for GDPR compliance of mobile apps

© Copyright 2022, Norman Sadeh, All rights reserved.

Tools for Developers



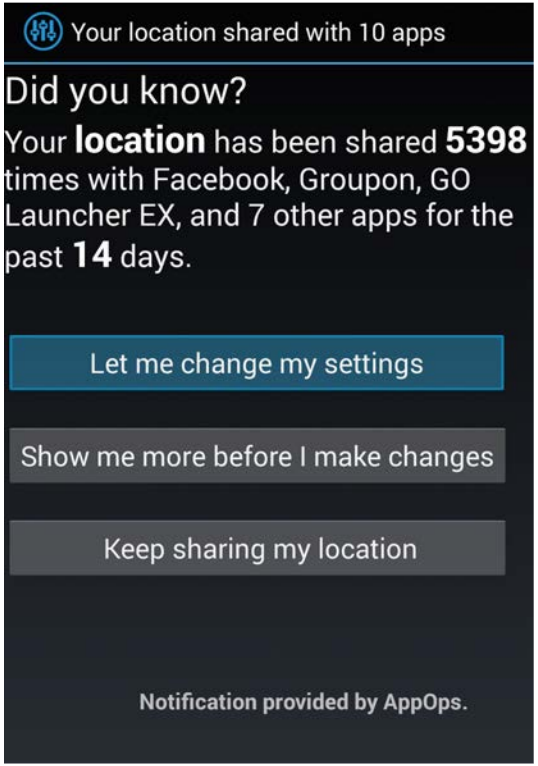
Compliance traceability: Privacy policies as software development artifacts, S Zimmeck, P Story, R Goldstein, D Baraka, S Li, Y Feng, N Sadeh, Privacy, Usability and Transparency Workshop (PUT 2019) at PoPETs 2019 conference

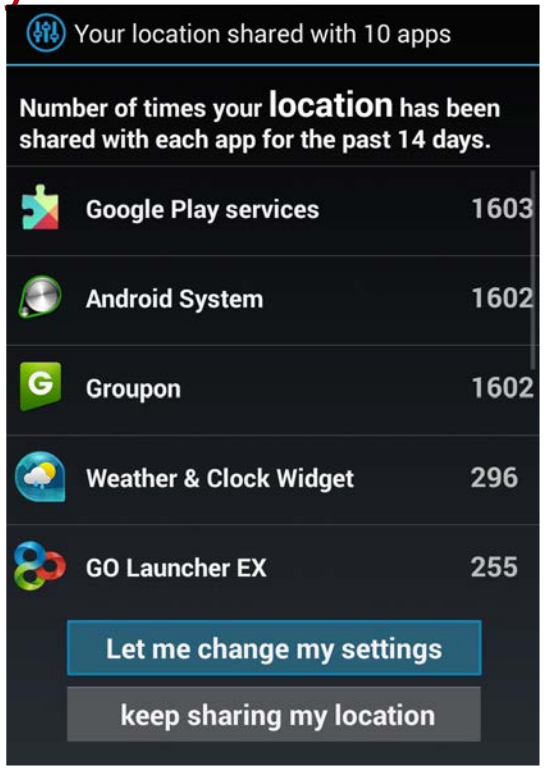
© Copyright 2022, Norman Sadeh, All rights reserved.

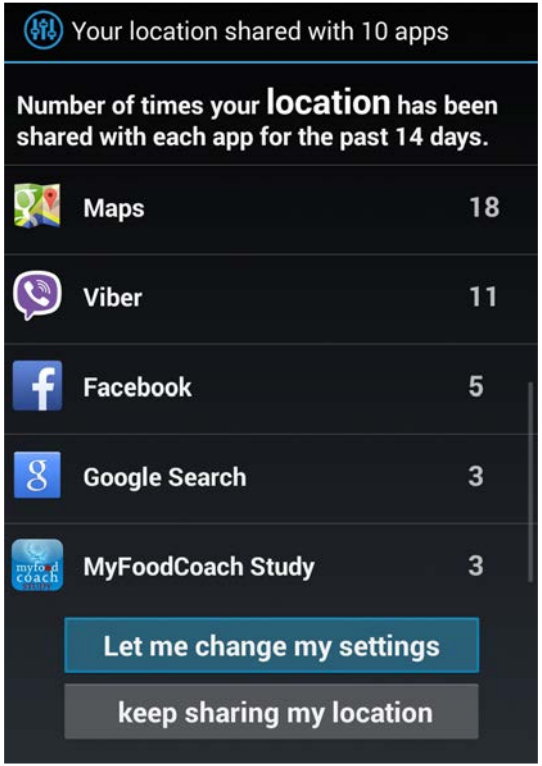
Could Computers Also Help Motivate People to Take Advantage of Privacy Settings?

© Copyright 2022, Norman Sadeh, All rights reserved.

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



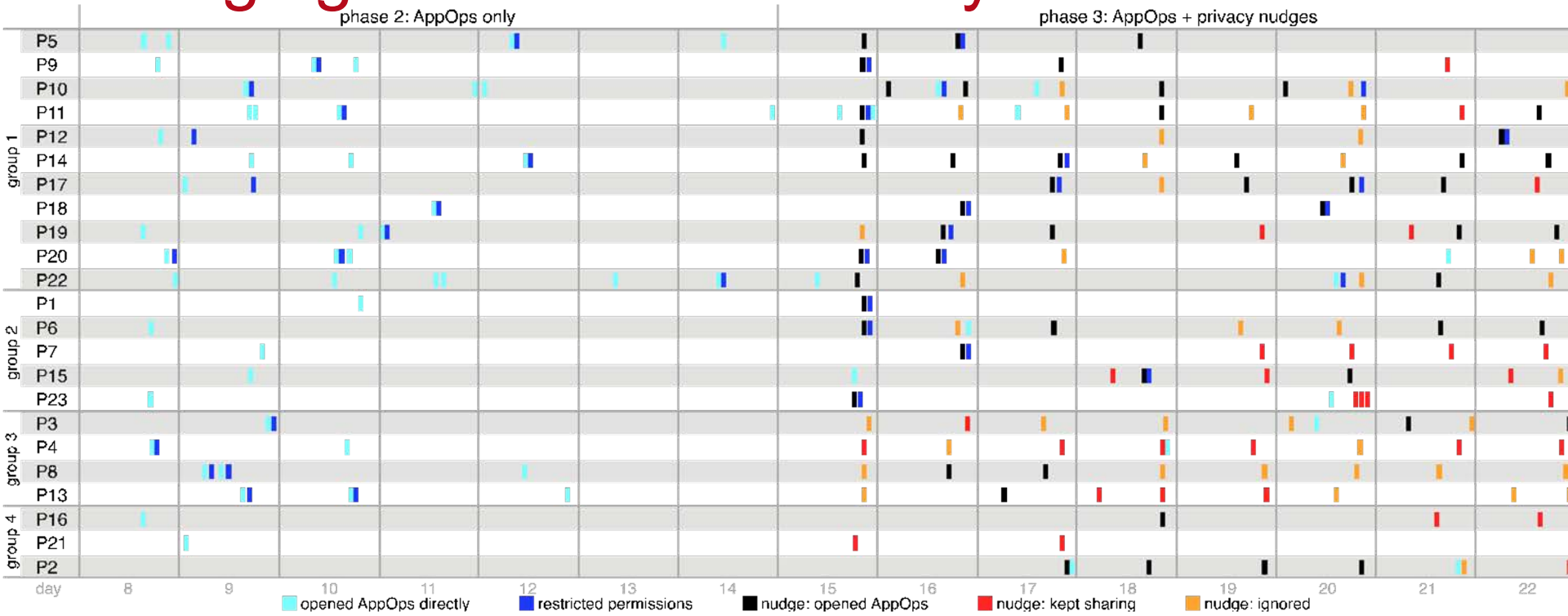




H. Almuhimedi,F. Schaub, N. Sadeh, I. Adjerid, A Acquisti, J. Gluck, L. Cranor, Y. Agarwal, "Your Location Has Been Shared 5,398 Times!: A Field Study On Mobile App Privacy Nudging," ACM CHI 2015.

© Copyright 2022, Norman Sadeh, All rights reserved.

Nudging Pilot – 3-week study



Week 2: Permission Manager Only

Week 3: Daily Nudges

- Permission Manager Helps (week 2)
- Nudges can make a big difference (week 3)

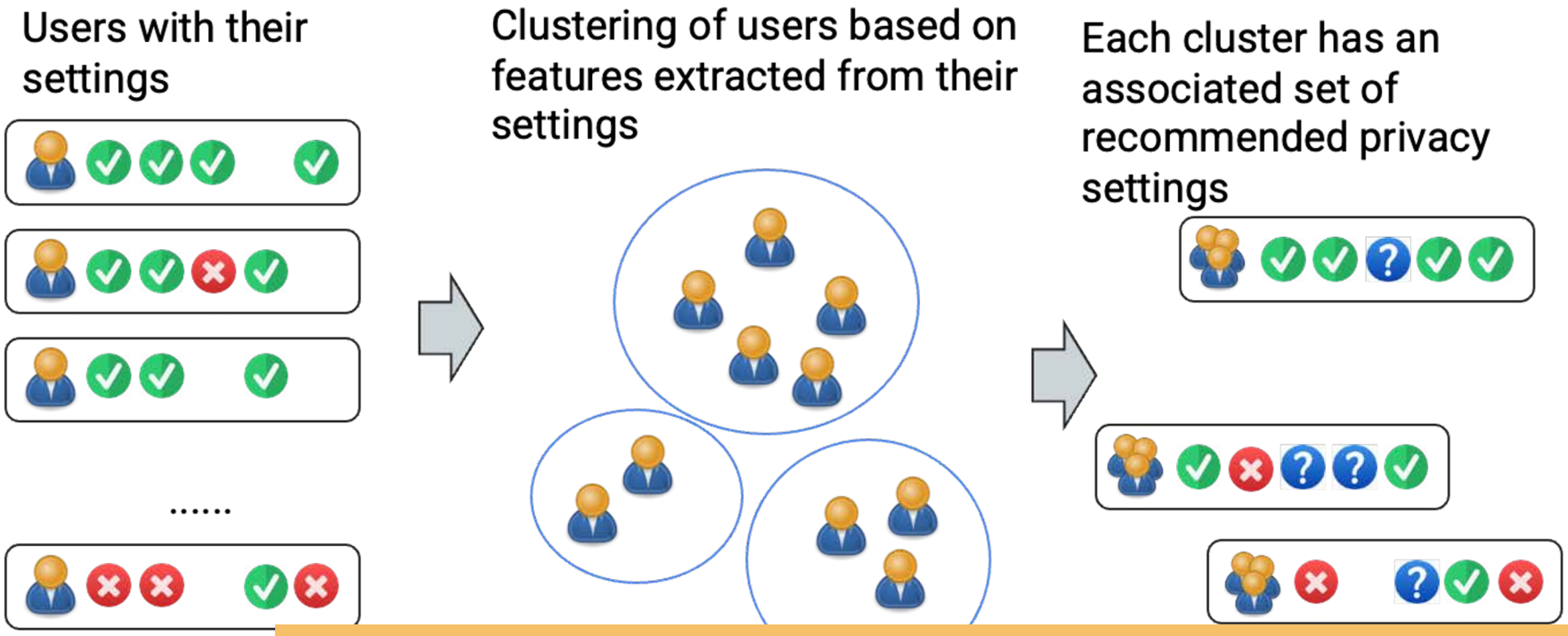
© Copyright 2022, Norman Sadeh, All rights reserved.

Could Computers Help People Manage their Privacy Controls?

Many privacy decisions are repetitive or similar… and tedious

© Copyright 2022, Norman Sadeh, All rights reserved.

Privacy Assistants - I



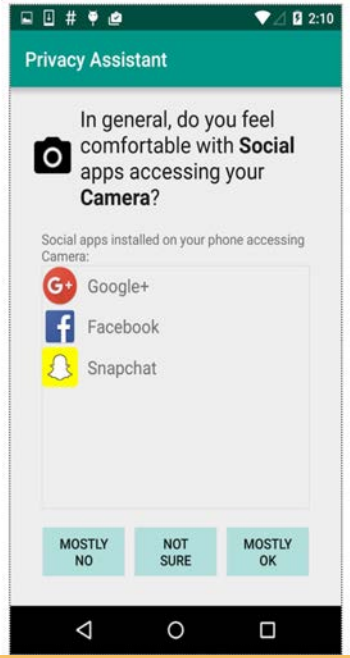
Even simple solution with small number of clusters achieves high levels of accuracy

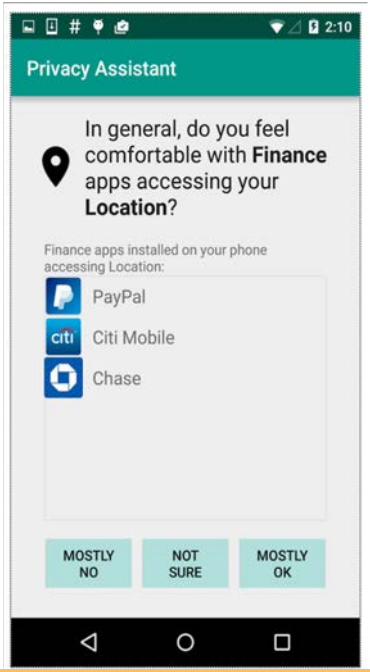
B. Liu, M.S. Andersen, F. Schaub, H. Almuhimedi, S. Zhang, N. Sadeh, A. Acquisti, and Y. Agarwal, **"Follow My Recommendations: A Personalized Assistant for Mobile App Permissions"**, SOUPS 2016- US patents 10,956,586

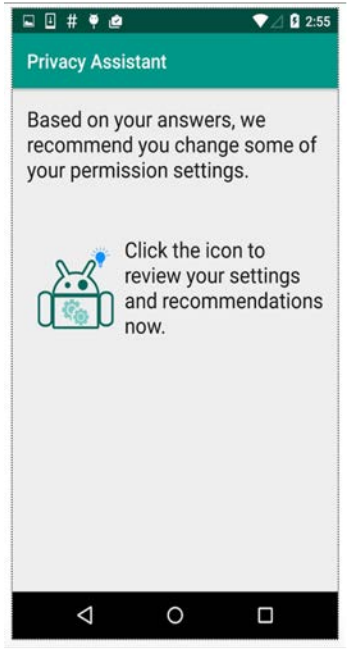
© Copyright 2022, Norman Sadeh, All rights reserved.

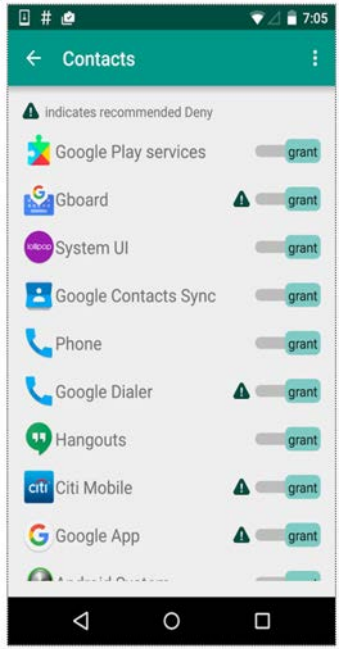
Privacy Assistants II

Generating recommendations rather than automating privacy decisions









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…

© Copyright 2022, Norman Sadeh, All rights reserved.

Why Recommendations?

Agency is a major part of privacy: users should remain in charge of their decisions…but **AI can help** them make these decisions and can help overcome fundamental **usability limitations**

- **Major requirement**: the recommendations have to be ***understandable*** and ***auditable***

© Copyright 2022, Norman Sadeh, All rights reserved.

Similar Results with Other Privacy Decisions

- Similar results with browser and IoT privacy decisions
- The challenge is that access to these privacy settings is generally not open
- **Would need regulation to make this possible…**
  - **…just think about the number of times you answer the same cookie questions…**

-S. Zhang, Y. Feng, A. Das, L. Bauer, L. Cranor, N. Sadeh, '**Understanding People's Privacy Attitudes Towards Video Analytics Technologies',** CMU Sch. of Comp. Sci. Tech Report, CMU-ISR-20-114.
-Daniel Smullen, Yuanyuan Feng, Shikun (Aerin) Zhang, Norman Sadeh, **"The Best of Both Worlds: Mitigating Trade-offs Between Accuracy and User Burden in Capturing Mobile App Privacy Preferences"**, Privacy Enhancing Technologies Symposium (PETS 2020), Sep 2019 [pdf]

© Copyright 2022, Norman Sadeh, All rights reserved.

How about the Internet of Things?

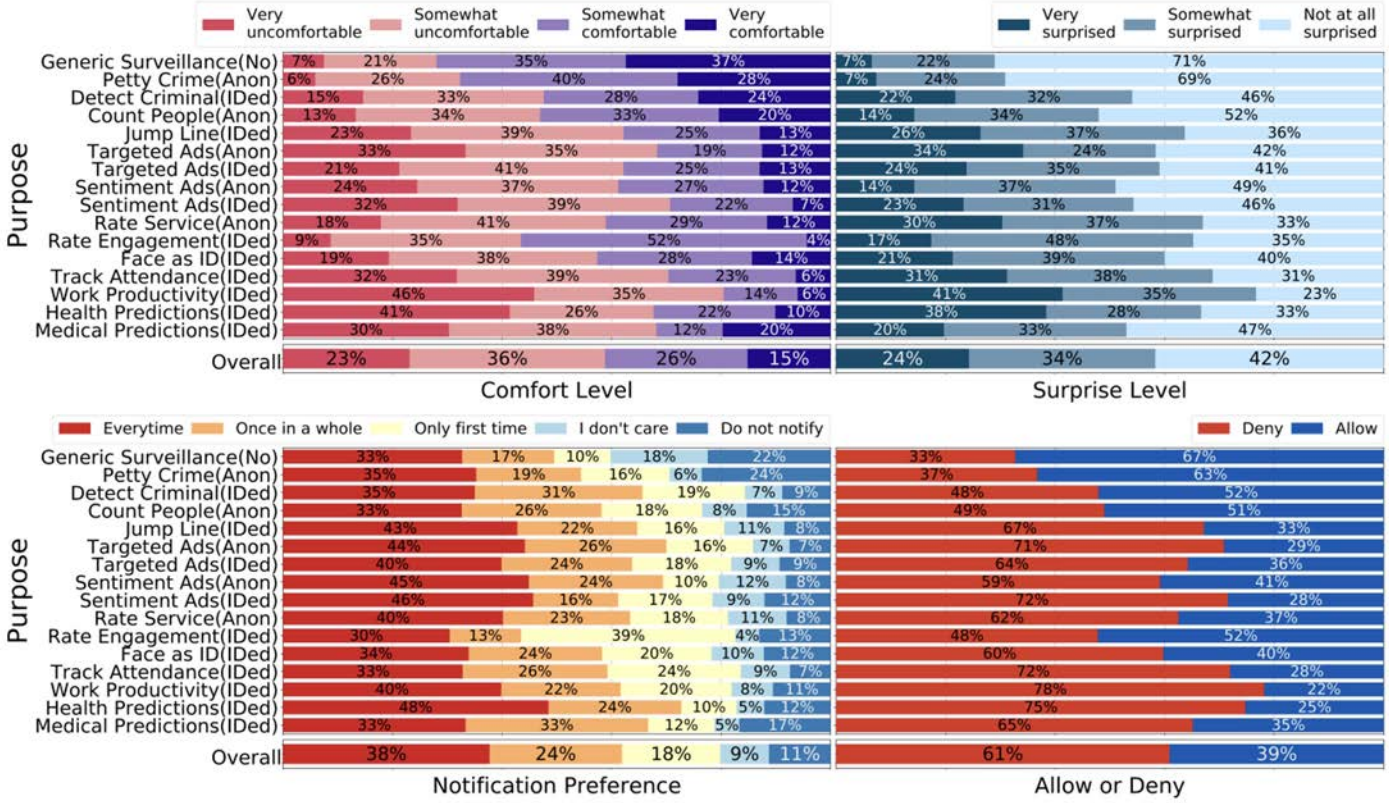
© Copyright 2022, Norman Sadeh, All rights reserved.

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?

© Copyright 2022, Norman Sadeh, All rights reserved.

In situ study of 123 people over 10 days in the context of their regular day-to-day activities

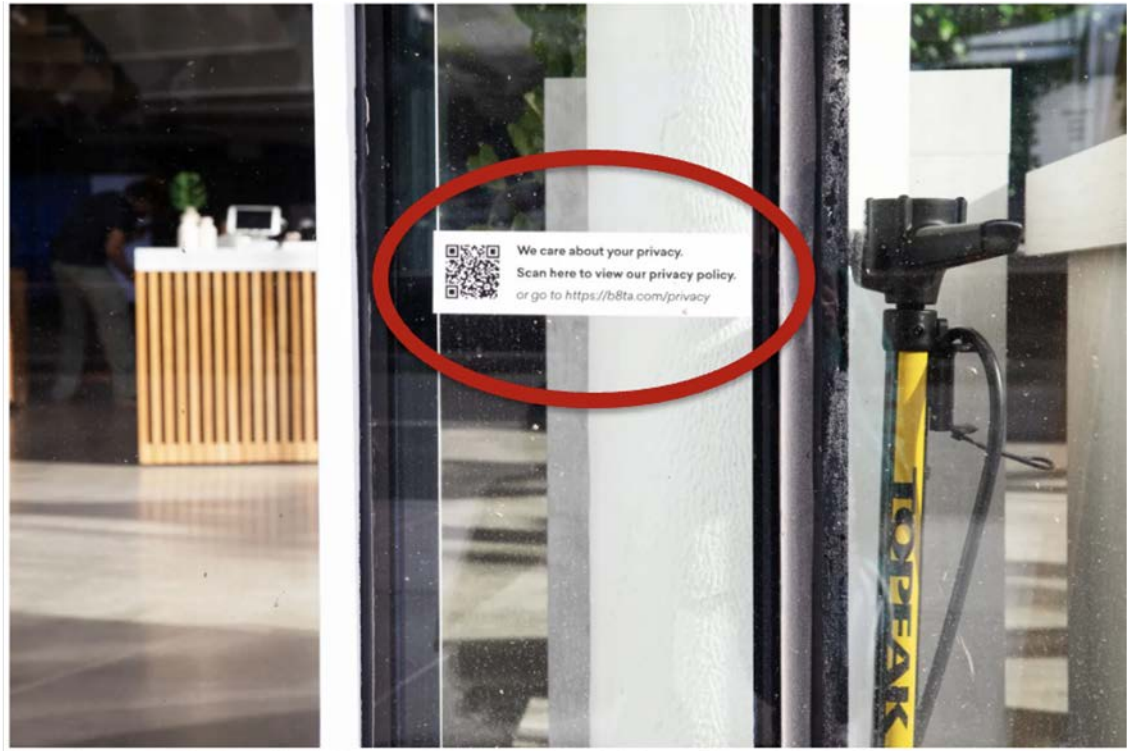


Deployment of a wide range of videoanalytics technologies:
Very clear need for effective notification and control functionality

Zhang, Y Feng, L Bauer, LF Cranor, A Das, and N Sadeh, **"“Did you know this camera tracks your mood?”: Understanding Privacy Expectations and Preferences in the Age of Video Analytics"**, Proceedings on Privacy Enhancing Technologies, 2, 1, Apr 2021 [pdf]

© Copyright 2022, Norman Sadeh, All rights reserved.

Good Luck with this…



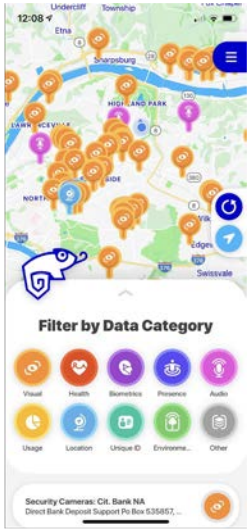
The notice required by the California Consumer Privacy Act, at knee height, at a B8ta outlet in San Francisco. PHOTOGRAPH: LAURYN HILL

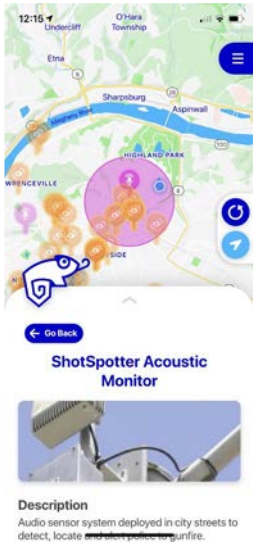
QR code for CCPA compliance at knee height…

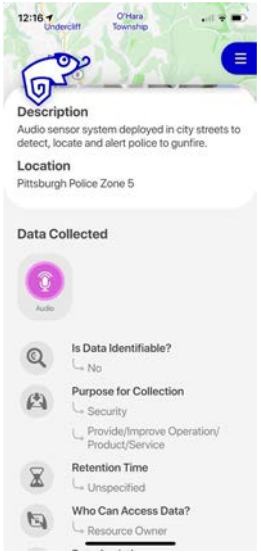
© Copyright 2022, Norman Sadeh, All rights reserved.

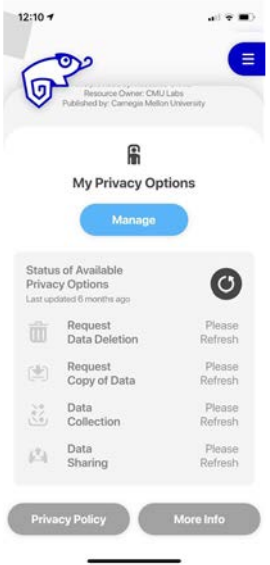
An IoT Privacy Inafrastructure

- 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

© Copyright 2022, Norman Sadeh, All rights reserved.

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

© Copyright 2022, Norman Sadeh, All rights reserved.

Concluding Remarks - I

- Privacy is becoming increasingly complex
  - Everyone is collecting our data, increasingly complex data flows
  - Smartphones, IoT, AI/ML
- New regulations have been introduced that are in great part motivated by these developments (e.g., GDPR, CCPA/CPRA)
- These regulations are an important step in the right direction
- **Yet, in the absence of better technologies, they make usability even more challenging**

© Copyright 2022, Norman Sadeh, All rights reserved.

Concluding Remarks - II

- My collaborators and I have been working on the development of technologies that aim to **mitigate these usability challenges**
- Our vision: **Privacy Assistants** that leverage
  - Techniques designed to empower people to take **advantage of more detailed privacy policy disclosures**
  - Techniques designed to motivate people and assist them with the **management** of an increasingly unmanageable number **of privacy decisions**
- Using techniques such as AI/ML or privacy nudges gives rise to complex **ethical issues** and requires **careful evaluation**
- Some of our efforts to help people manage their privacy settings call for **regulations that mandate privacy APIs** - **without these APIs privacy will remain unmanageable.**

© Copyright 2022, Norman Sadeh, All rights reserved.

Q&A

© Copyright 2022, Norman Sadeh, All rights reserved.

The **Usable Privacy Policy Project** and the **Personalized Privacy Assistant Project** involve collaborations with a number of individuals

More details at:

https://usableprivacy.org

https://privacyassistant.org

https://explore.usableprivacy.org

https://www.iotprivacy.io

© Copyright 2022, Norman Sadeh, All rights reserved.