SAC-PA Workshop – June 14-15, 2018

Privacy in the Age of the Internet of Things

Norman Sadeh

Carnegie Mellon University

www.normsadeh.org

usableprivacy.org   privacyassistant.org

explore.usableprivacy.org

Quick Show of Hands

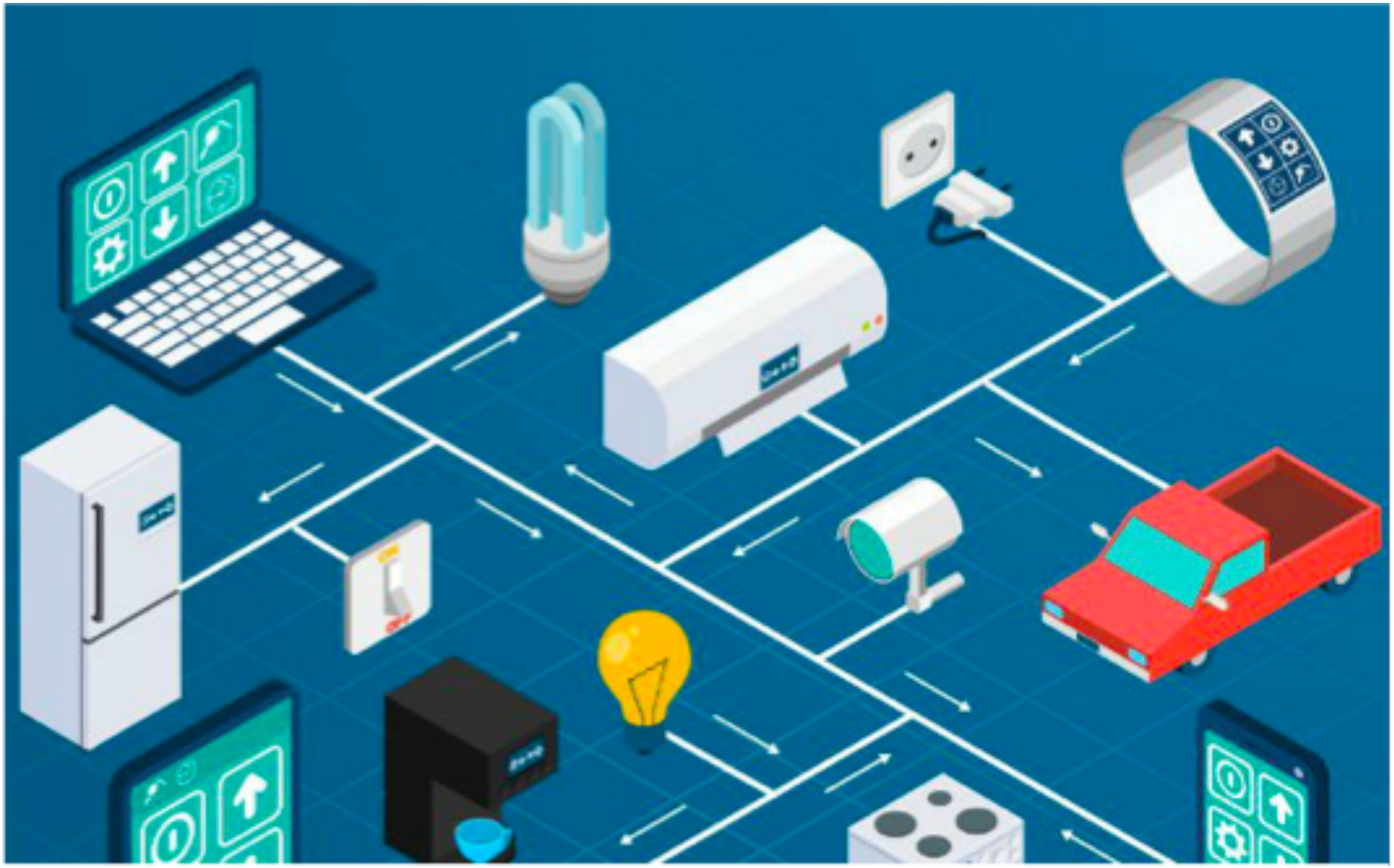
- Imagine that you are in the market to purchase a car insurance policy…

How Many of You Would Feel Comfortable Disclosing…

- How many miles you drive per year?
- How fast you drive…
  - Based on GPS…
- Where you go and when…
  - Based on GPS…
- Relevant health data…
  - Such as how many hours you sleep at night…
  - Based on data sensed by your wristwatch…

Internet of Things & Big Data

Increasingly diverse, complex and opaque dataflows



http://www.iamwire.com/2017/01/iot-ai/148265

Information Privacy

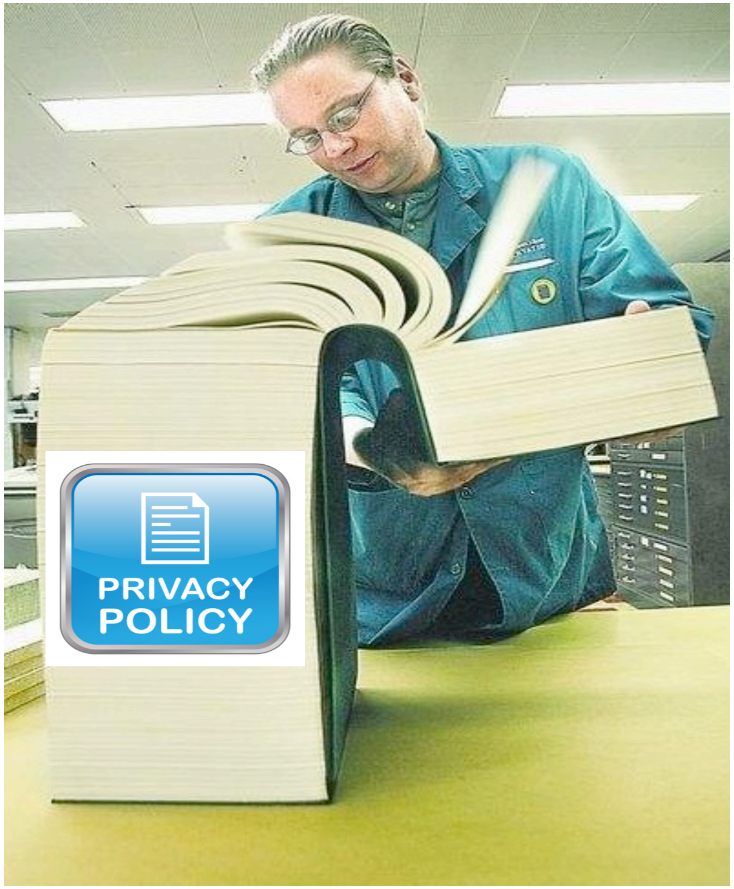
- The claim that ***certain information should not be collected by government or businesses*** – *or possibly* ***only under special circumstances*** and subject to various rules
  - **individuals have some control over the collection and use of information about them**

Legal Landscape

- **A number of privacy laws around the world:**
  - US: State, federal and local laws
    - Federal level: Patchwork of sectoral laws and laws that pertain to data collected by the government
  - EU: General Data Protection Regulation (GDPR)
- All these laws share some **commonalities: They set minimum requirements to:**
  - **Inform** users about data collection and use practices
  - Provide users with some type of **choice**

In practice…

- **Notice and choice** is **broken**
  - No time to read policies
  - Policies difficult to understand
  - No time or motivation to configure settings
- **91%** of people report feeling they **have lost control over their information**



Pew Survey 2014 http://www.pewinternet.org/2014/11/12/public-privacy-perceptions/

Mobile and IoT: A Number of Complicating Factors

- A typical mobile phone user with 50 mobile apps each requesting 3 permissions would have to **configure 150 settings**
- IoT: Technology is often **“invisible”**
- **Reading policies is even less practical**
- Explosion in the number of apps and devices
- Developers often **lack the necessary sophistication**

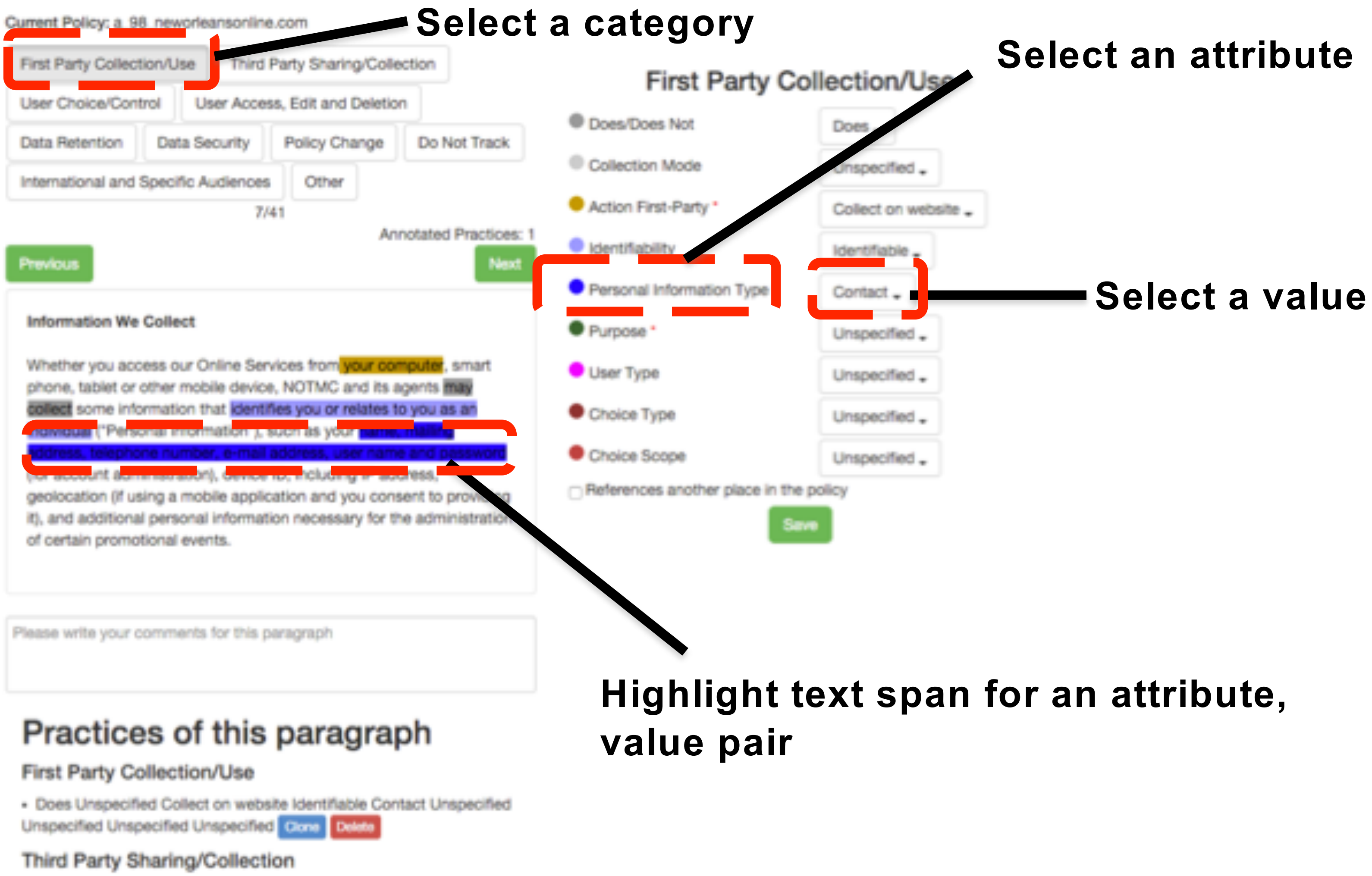
“Modeling Users’ Mobile App Privacy Preferences: Restoring Usablility in a Sea of Permission Settings”, J. Lin, B. Liu, N. Sadeh, J. Hong, Proc. of the USENIX Symposium on Usable Privacy and Security, SOUPS 2014, Jul. 2014

What If….

- **Computers understood privacy policies?**
  - Machine-readable policies have been proposed but have not gained traction
- **Computers understood** what we **care about** and what we **already know/expect**

Could We Teach Computers to Read Privacy Policies?

Annotation Tool



Current Policy: a_98_neworleansonline.com

First Party Collection/Use | Third Party Sharing/Collection | User Choice/Control | User Access, Edit and Deletion | Data Retention | Data Security | Policy Change | Do Not Track | International and Specific Audiences | Other

7/41

Annotated Practices: 1

Previous | Next

Information We Collect

Whether you access our Online Services from your computer, smart phone, tablet or other mobile device, NOTMC and its agents may collect some information that identifies you or relates to you as an individual ("Personal Information"), such as your name, mailing address, telephone number, e-mail address, user name and password (for account administration), device ID, including IP address, geolocation (if using a mobile application and you consent to providing it), and additional personal information necessary for the administration of certain promotional events.

Please write your comments for this paragraph

First Party Collection/Use

- ● Does/Does Not: Does
- ● Collection Mode: Unspecified
- ● Action First-Party *: Collect on website
- ● Identifiability: Identifiable
- ● Personal Information Type: Contact
- ● Purpose *: Unspecified
- ● User Type: Unspecified
- ● Choice Type: Unspecified
- ● Choice Scope: Unspecified

□ References another place in the policy

Save

Practices of this paragraph

First Party Collection/Use

- Does Unspecified Collect on website Identifiable Contact Unspecified Unspecified Unspecified Unspecified Clone Delete

Third Party Sharing/Collection

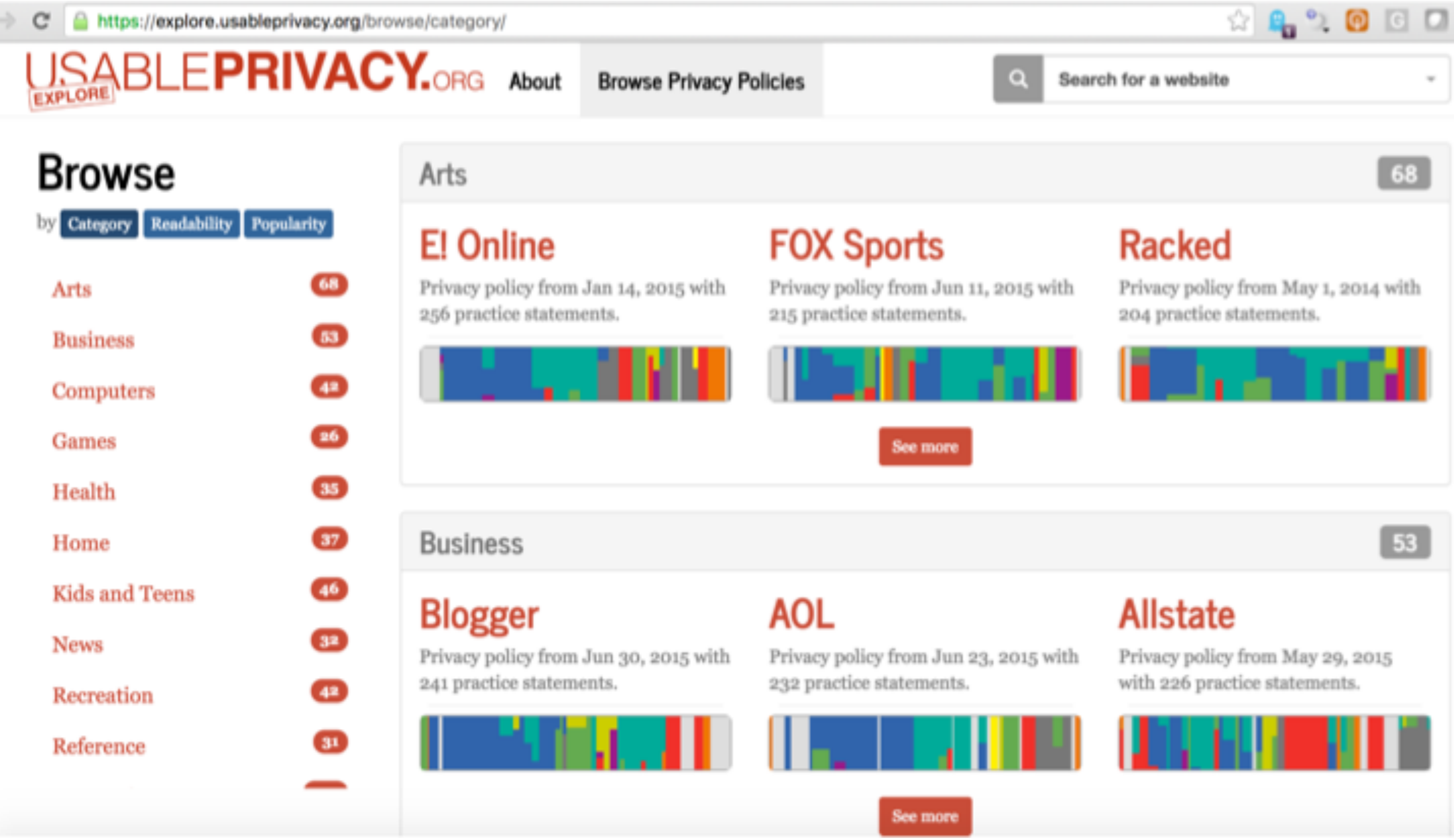
Select a category

Select an attribute

Select a value

Highlight text span for an attribute, value pair

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



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

USABLEPRIVACY.ORG EXPLORE

About

Browse Privacy Policies

Search for a website

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.

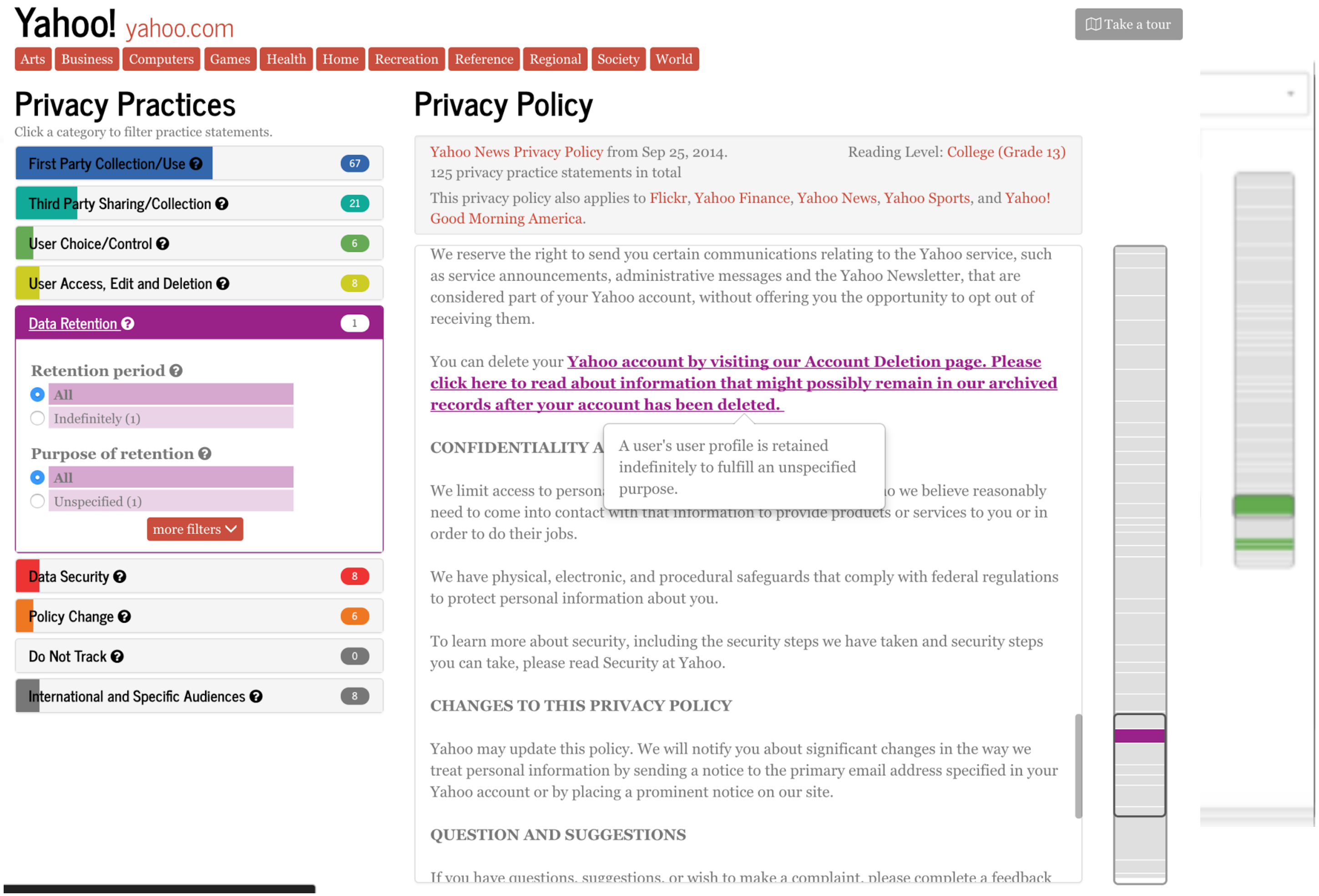
AOL

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

Allstate

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

See more



Yahoo! yahoo.com

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	8
Data Retention	1

Retention period

● All

○ Indefinitely (1)

Purpose of retention

● All

○ Unspecified (1)

Data Security	8
Policy Change	6
Do Not Track	0
International and Specific Audiences	8

Privacy Policy

Yahoo News Privacy Policy from Sep 25, 2014. Reading Level: College (Grade 13)

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.**

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

CONFIDENTIALITY A

We limit access to person o 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.

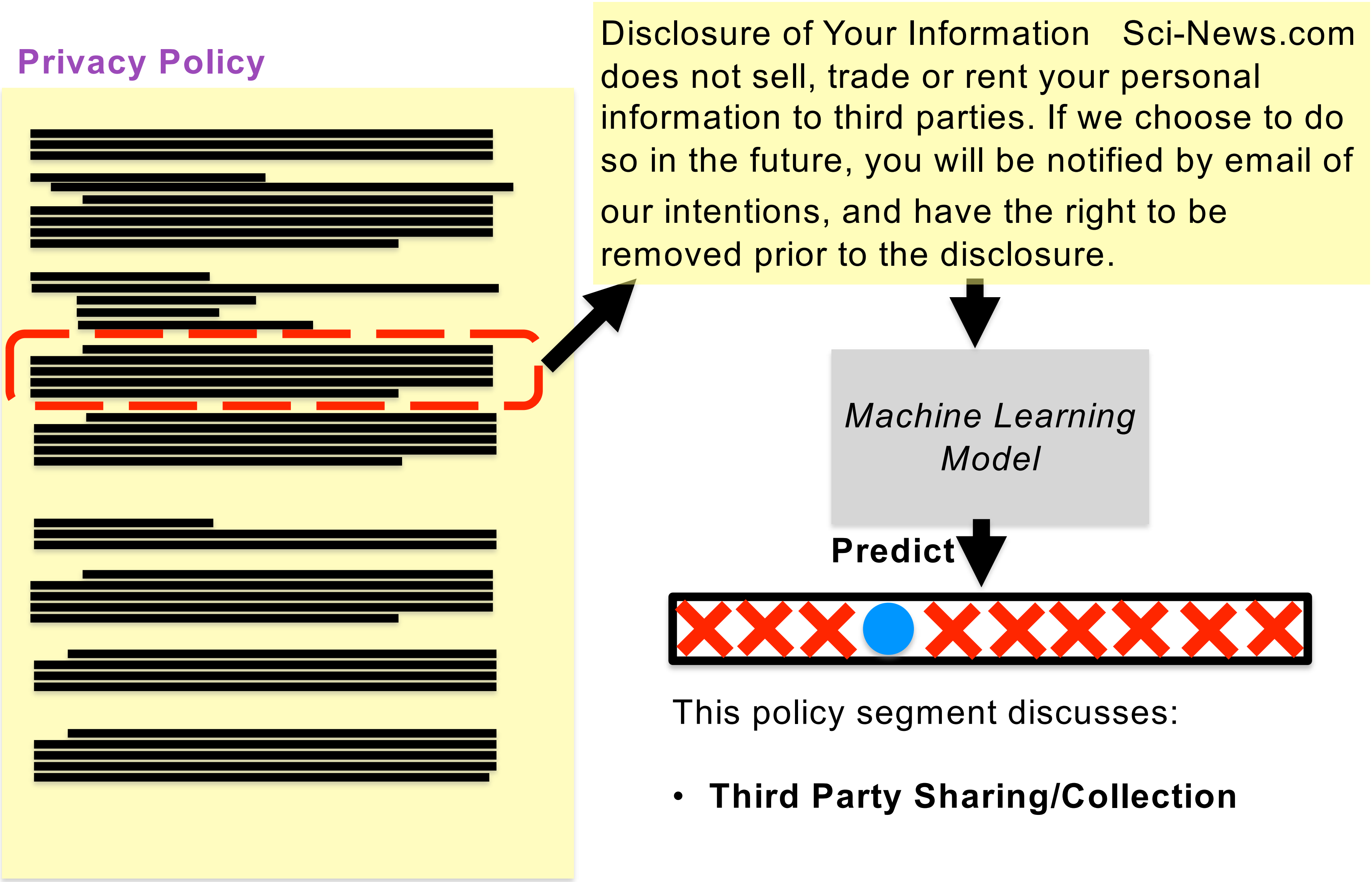
CHANGES TO THIS PRIVACY POLICY

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

A First Task: Segment Annotation



Privacy Policy

Disclosure of Your Information Sci-News.com does not sell, trade or rent your personal information to third parties. If we choose to do so in the future, you will be notified by email of our intentions, and have the right to be removed prior to the disclosure.

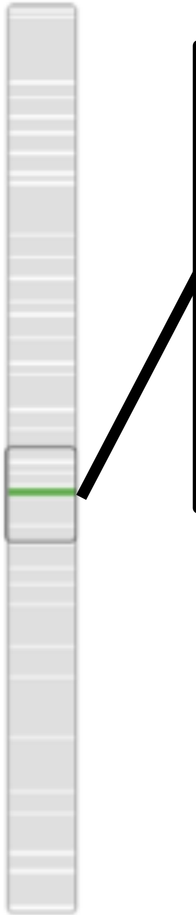
Machine Learning Model

Predict

This policy segment discusses:

- **Third Party Sharing/Collection**

Another Task: User Choice Instance Extraction



Choice Instance !!!

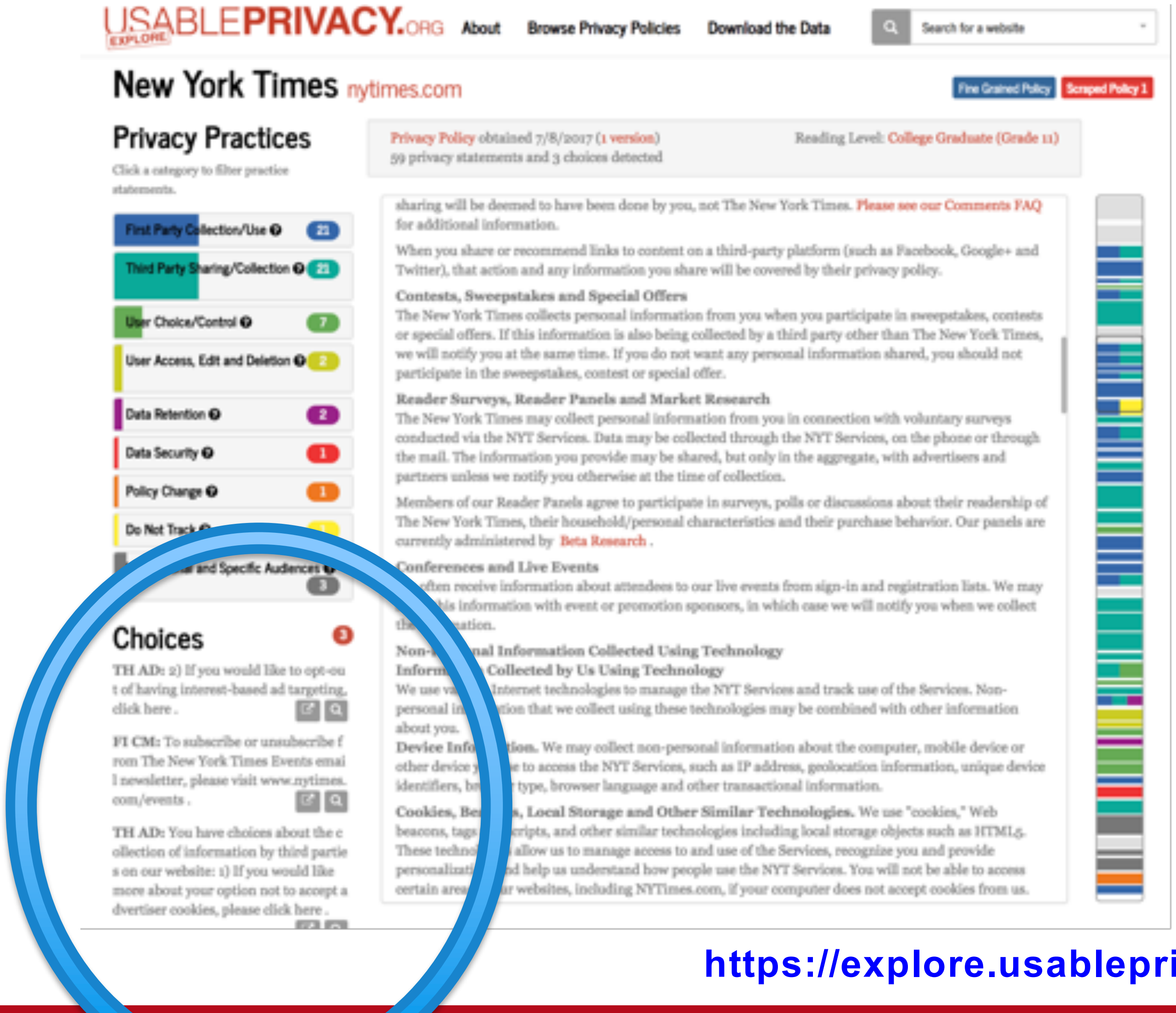
If you do not want us to use personal information that we gather to allow third parties to personalize advertisements we display to you, please adjust your Advertising Preferences .

- 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

K.M. Sathyendra, F. Schaub, S. Wilson, N. Sadeh. *Automatic Extraction of Opt-Out Choices from Privacy Policies.* AAAI Fall Symposium on Privacy and Language Technologies. 2016.

K.M. Sathyendra, S. Wilson, F. Schaub, S. Zimmeck, N. Sadeh. *Identifying the Provision of Choices in Privacy Policies, EMNLP Conference, 2017 (accepted for publication)*

Annotated 7,000+ policies



New York Times nytimes.com

Privacy Practices

Click a category to filter practice statements.

First Party Collection/Use	21
Third Party Sharing/Collection	21
User Choice/Control	7
User Access, Edit and Deletion	2
Data Retention	2
Data Security	1
Policy Change	1
Do Not Track	
and Specific Audiences	3

Choices 3

TH AD: 2) If you would like to opt-ou t of having interest-based ad targeting, click here .

FI CM: To subscribe or unsubscribe f rom The New York Times Events emai l newsletter, please visit www.nytimes. com/events .

TH AD: You have choices about the c ollection of information by third partie s on our website: 1) If you would like more about your option not to accept a dvertiser cookies, please click here .

Privacy Policy obtained 7/8/2017 (1 version)
59 privacy statements and 3 choices detected

Reading Level: College Graduate (Grade 11)

sharing will be deemed to have been done by you, not The New York Times. Please see our Comments FAQ for additional information.

When you share or recommend links to content on a third-party platform (such as Facebook, Google+ and Twitter), that action and any information you share will be covered by their privacy policy.

Contests, Sweepstakes and Special Offers
The New York Times collects personal information from you when you participate in sweepstakes, contests or special offers. If this information is also being collected by a third party other than The New York Times, we will notify you at the same time. If you do not want any personal information shared, you should not participate in the sweepstakes, contest or special offer.

Reader Surveys, Reader Panels and Market Research
The New York Times may collect personal information from you in connection with voluntary surveys conducted via the NYT Services. Data may be collected through the NYT Services, on the phone or through the mail. The information you provide may be shared, but only in the aggregate, with advertisers and partners unless we notify you otherwise at the time of collection.

Members of our Reader Panels agree to participate in surveys, polls or discussions about their readership of The New York Times, their household/personal characteristics and their purchase behavior. Our panels are currently administered by Beta Research .

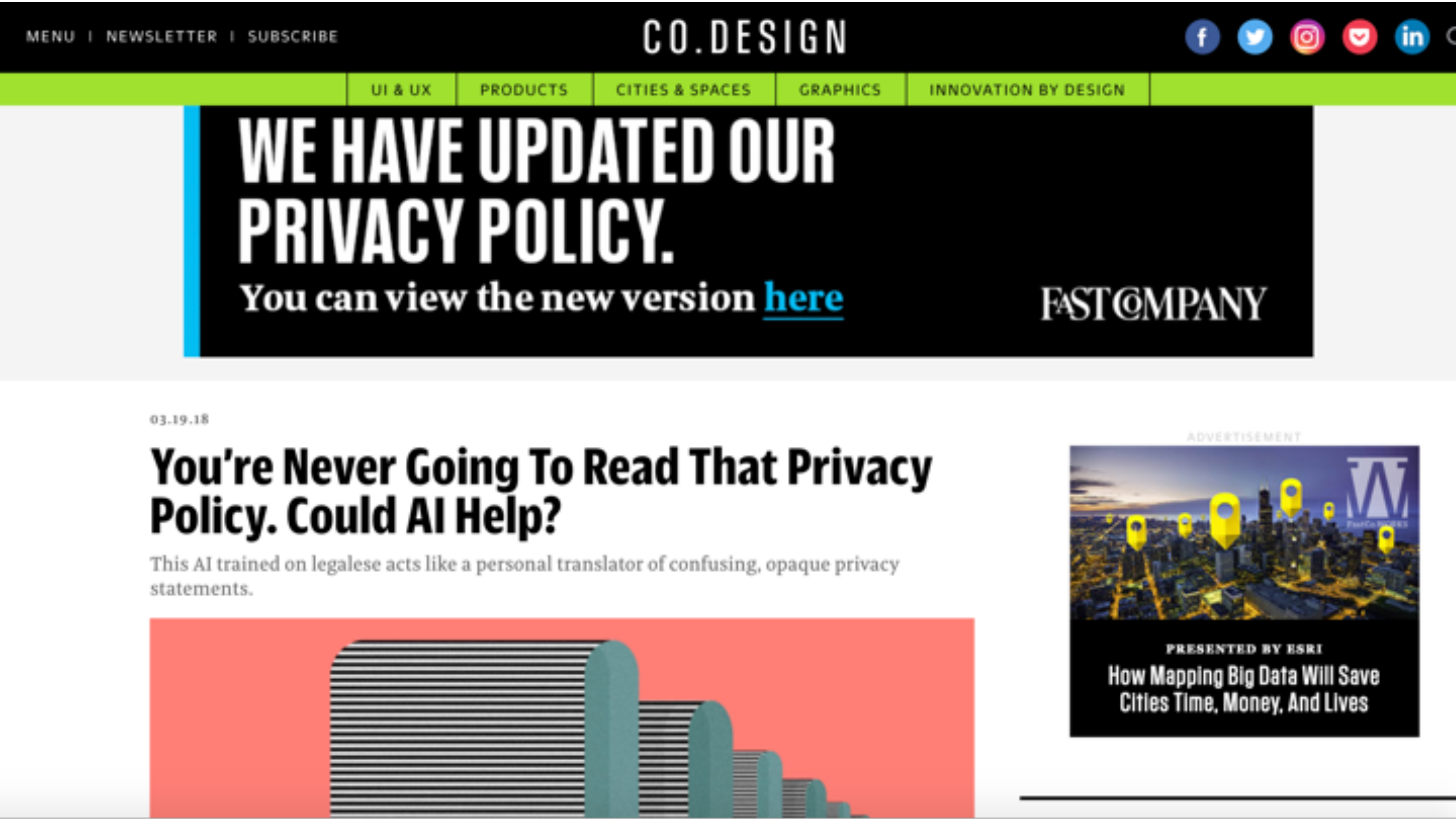
Conferences and Live Events
often receive information about attendees to our live events from sign-in and registration lists. We may this information with event or promotion sponsors, in which case we will notify you when we collect the ation.

Non- nal Information Collected Using Technology
Inform Collected by Us Using Technology
We use va Internet technologies to manage the NYT Services and track use of the Services. Non-personal in tion that we collect using these technologies may be combined with other information about you.
Device Info tion. We may collect non-personal information about the computer, mobile device or other device y e to access the NYT Services, such as IP address, geolocation information, unique device identifiers, br type, browser language and other transactional information.

Cookies, Be s, Local Storage and Other Similar Technologies. We use "cookies," Web beacons, tags cripts, and other similar technologies including local storage objects such as HTML5. These techno allow us to manage access to and use of the Services, recognize you and provide personalizati d help us understand how people use the NYT Services. You will not be able to access certain area r websites, including NYTimes.com, if your computer does not accept cookies from us.

https://explore.usableprivacy.org/

Press Coverage – Notice the Irony



WE HAVE UPDATED OUR PRIVACY POLICY.

You can view the new version here

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.

PRESENTED BY ESRI

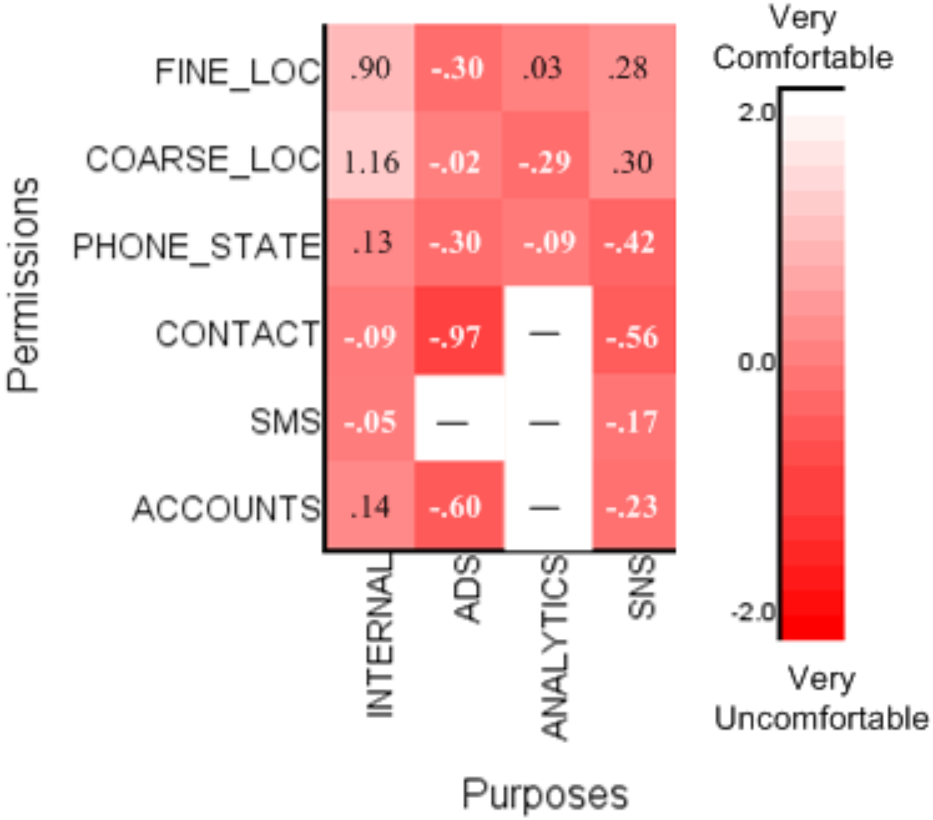
How Mapping Big Data Will Save Cities Time, Money, And Lives

Screenshot taken on May 31, 2018

Question

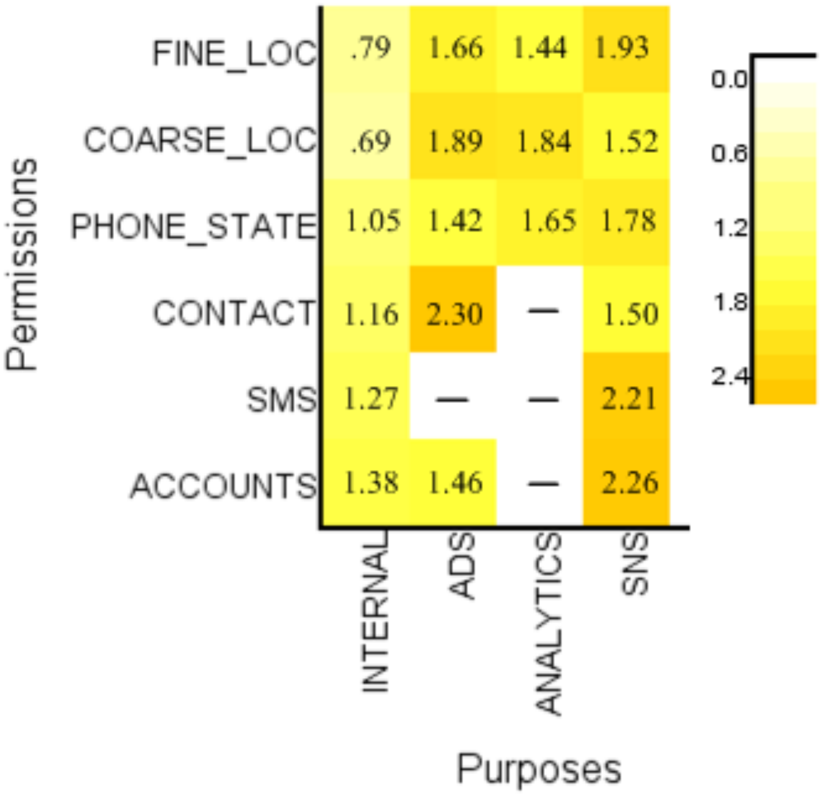
- How about helping end users?
- **Could we learn people's privacy expectations and preferences?**
  - To **selectively notify them about relevant privacy practices**
  - To **help them configure privacy settings**

One Size-Fits-All Defaults Don’t Work



Users’ Average Preferences

White → comfortable
Red → uncomfortable

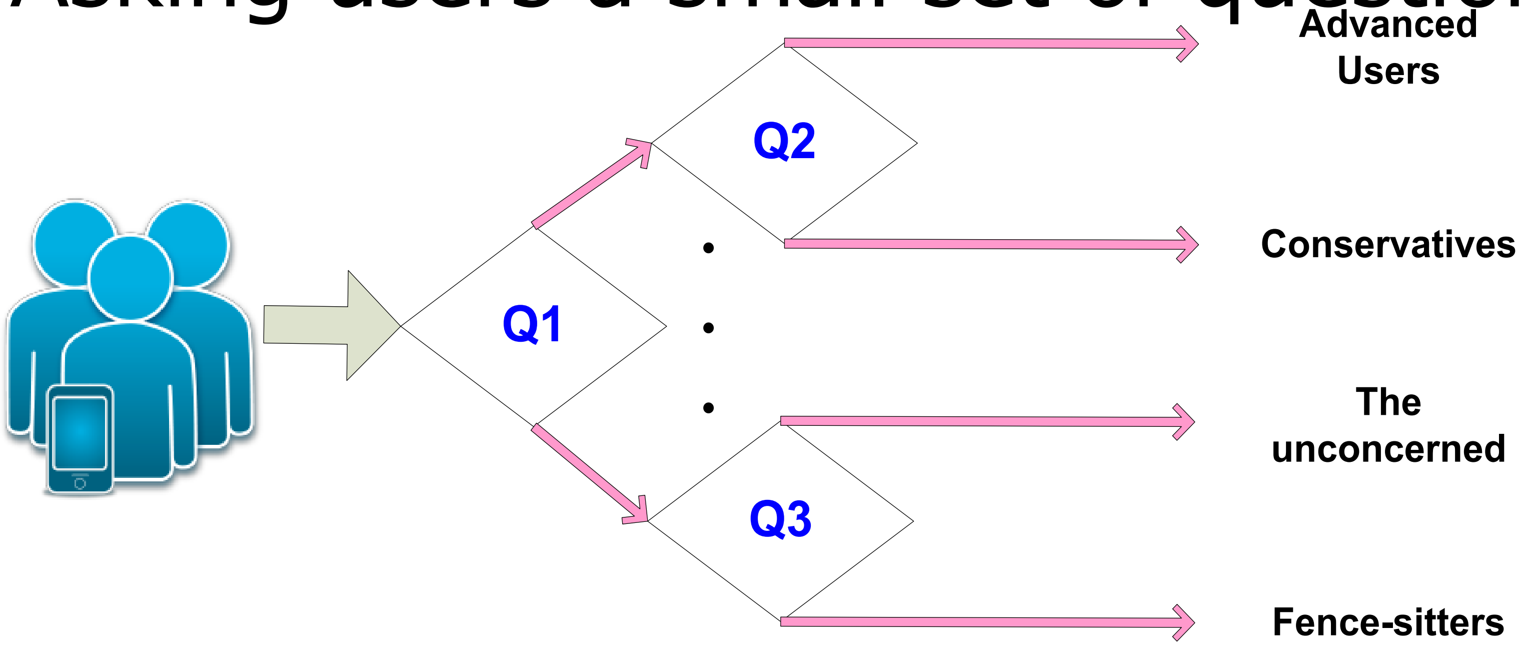


Variances among Users

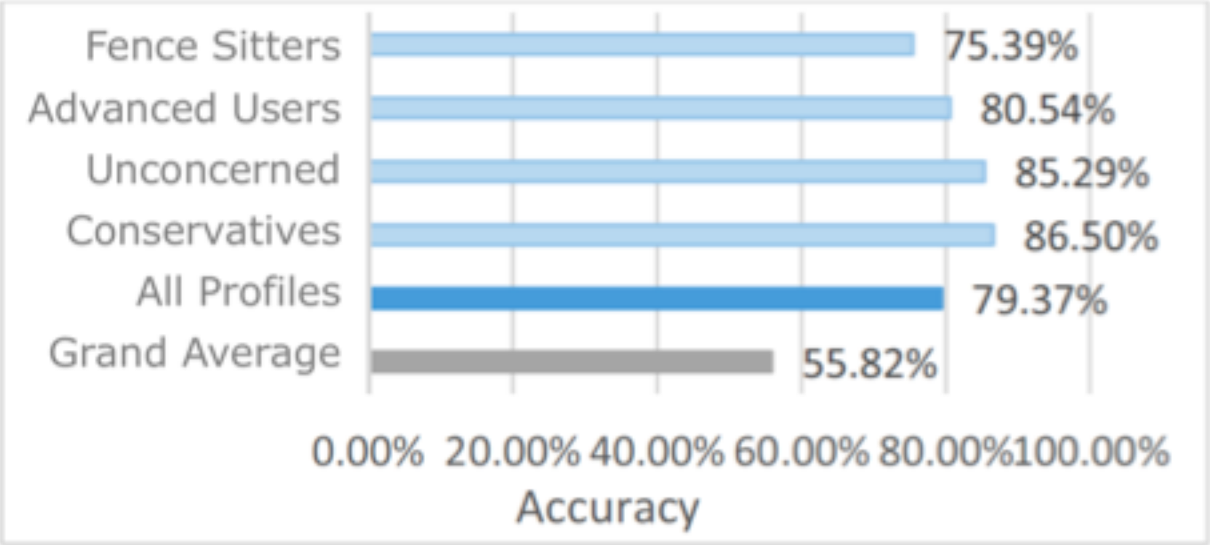
Darker yellow → larger variance

Identifying a User's Privacy Profile

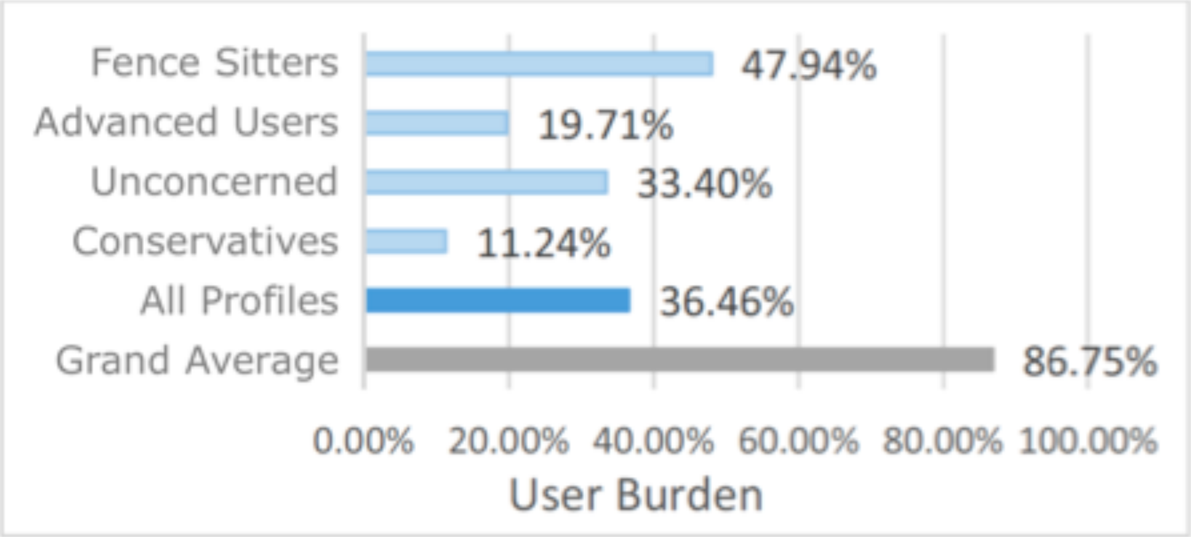
- Using Clustering techniques
- Asking users a small set of questions



Results with Just 4 Clusters

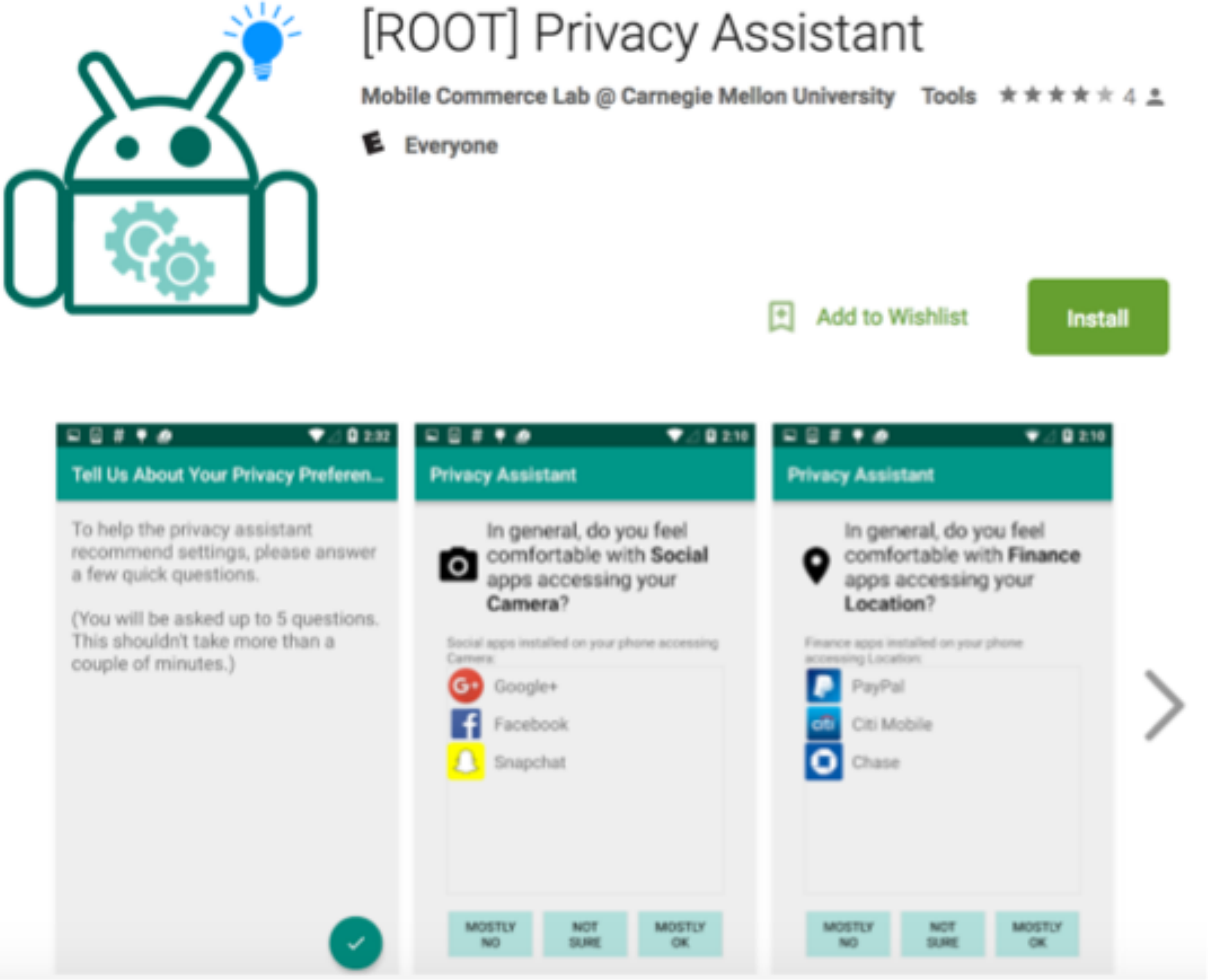


Accuracy:
One size fits all: 55.8%
4 Profiles: 79.4%



User Burden:
One size fits all: 86.8%
4 Profiles: 36.5%

Now Available on Google Play (rooted Android Phones 5 and up)



[ROOT] Privacy Assistant

Mobile Commerce Lab @ Carnegie Mellon University Tools ★★★★★ 4

Everyone

Tell Us About Your Privacy Preferen...

To help the privacy assistant recommend settings, please answer a few quick questions.

(You will be asked up to 5 questions. This shouldn't take more than a couple of minutes.)

Privacy Assistant

In general, do you feel comfortable with **Social** apps accessing your **Camera**?

Social apps installed on your phone accessing Camera:

- Google+
- Facebook
- Snapchat

MOSTLY NO | NOT SURE | MOSTLY OK

Privacy Assistant

In general, do you feel comfortable with **Finance** apps accessing your **Location**?

Finance apps installed on your phone accessing Location:

- PayPal
- Citi Mobile
- Chase

MOSTLY NO | NOT SURE | MOSTLY OK

What About IoT?

Overall Vision: **Personalized Privacy Assistants**

- Learn models of what users already expect & what they want to be informed about, how to communicate with them (when, how often, how), how to configure their settings
  - Or just allow users to manually configure settings
- **Selectively enter into dialogues** with users and **nudge them** towards safer practices
- **Extend privacy profiles across many environments:** from your smartphone, to your browser, to your smart home to your social networking account, etc.

Privacy Infrastructure for IoT*

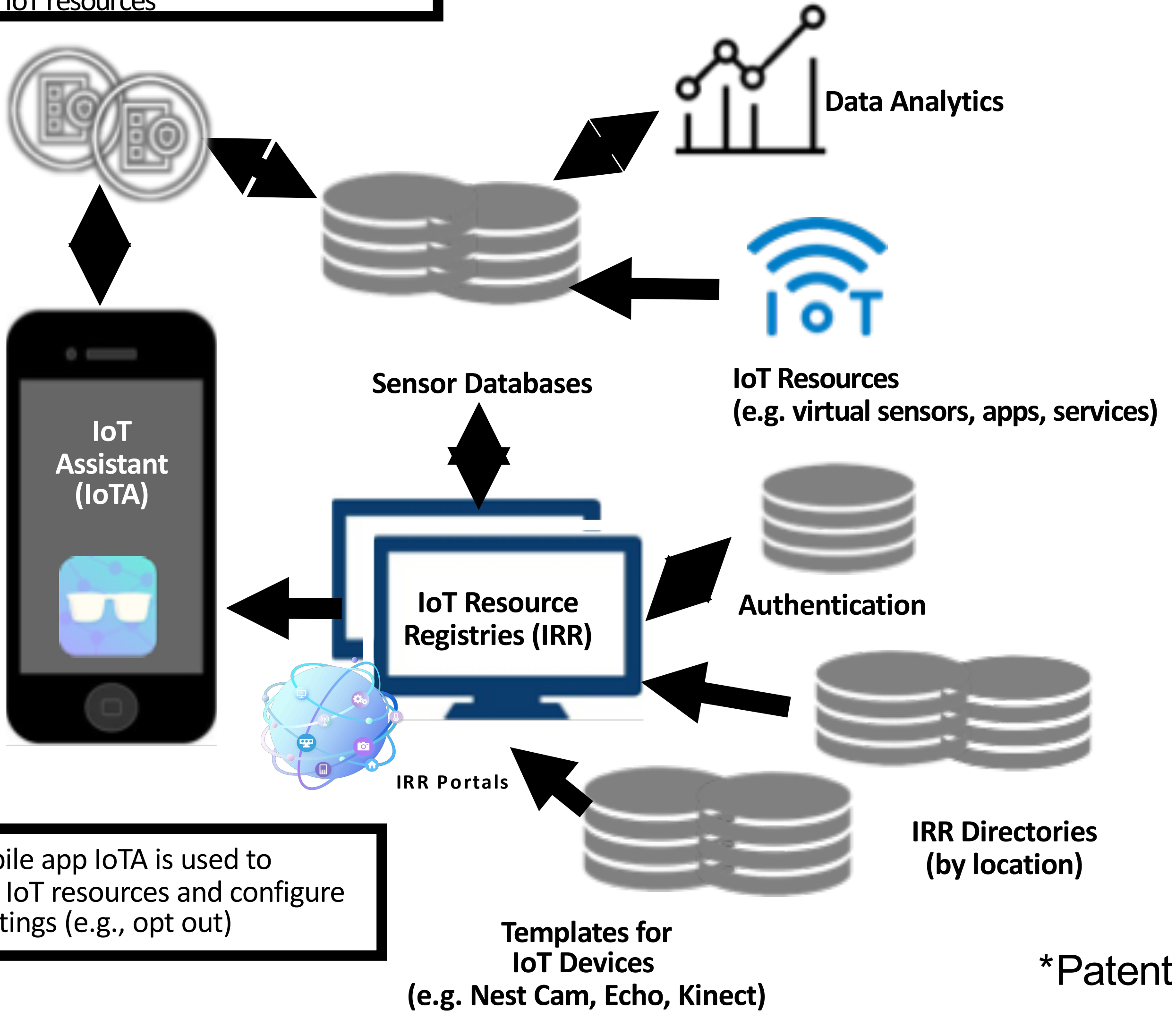
- **Registration** of IoT resources and their privacy policies – **IoT Resource Registry** (IRR) **& Portal**
  - Policies are in a machine readable format
  - Resources include: sensors (e.g., virtual sensors), applications, and services
  - Series of drop down menus, but also use of templates
- **Discovery of IoT resources and their policies**
- User notification via **IoT Assistant implemented as mobile app**
- Protocols to securely read and configure privacy settings

*Patent pending

Overall Architecture*

Policy Enforcement Points (PEP)

- Stores resource-specific and user-specific privacy policy settings
- Enforces settings for data collected by IoT resources



Data Analytics

Sensor Databases

IoT Resources
(e.g. virtual sensors, apps, services)

IoT Assistant (IoTA)

IoT Resource Registries (IRR)

Authentication

IRR Portals

IRR Directories
(by location)

Templates for
IoT Devices
(e.g. Nest Cam, Echo, Kinect)

The mobile app IoTA is used to discover IoT resources and configure their settings (e.g., opt out)

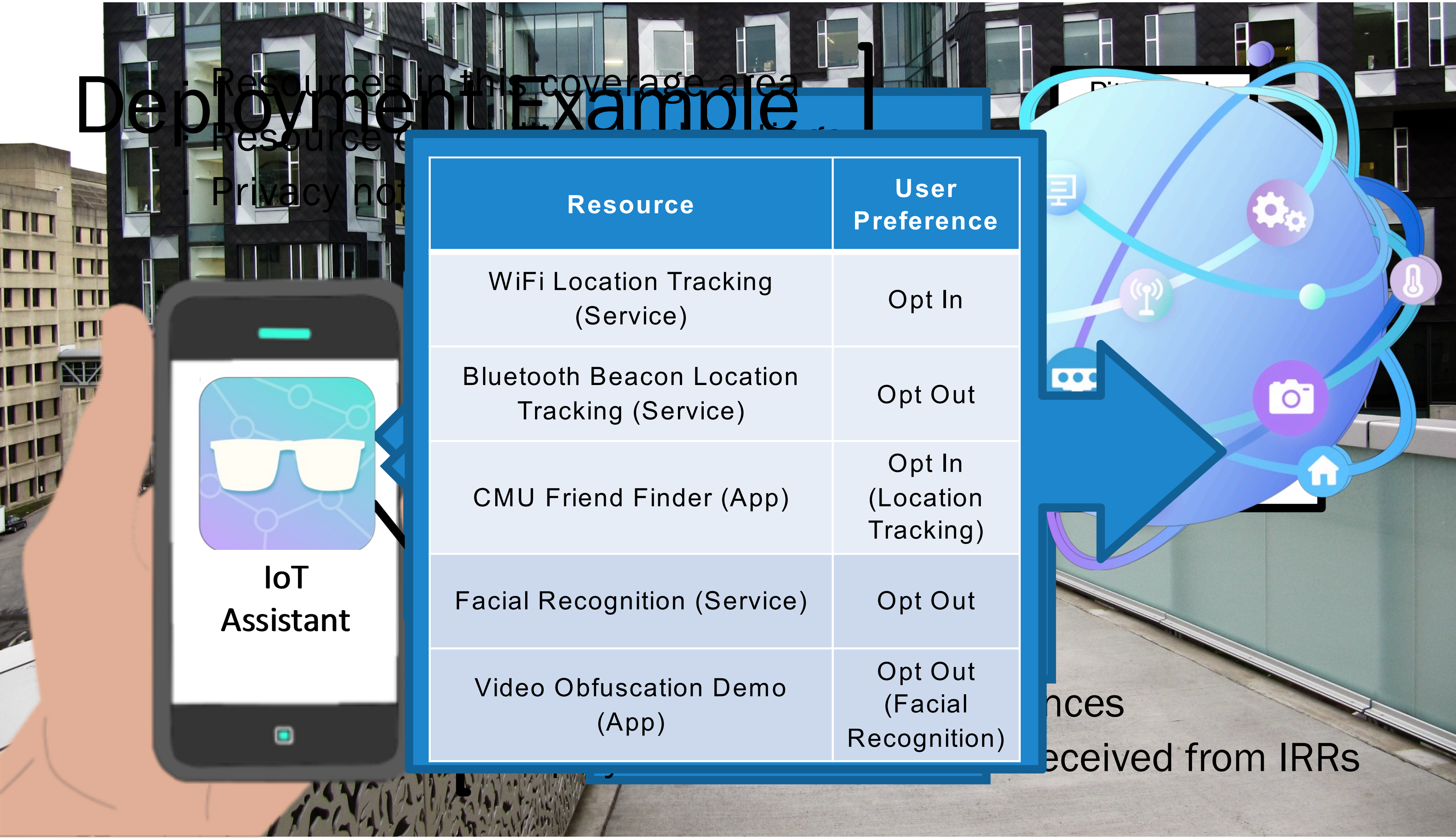
*Patent pending

Deployment Example

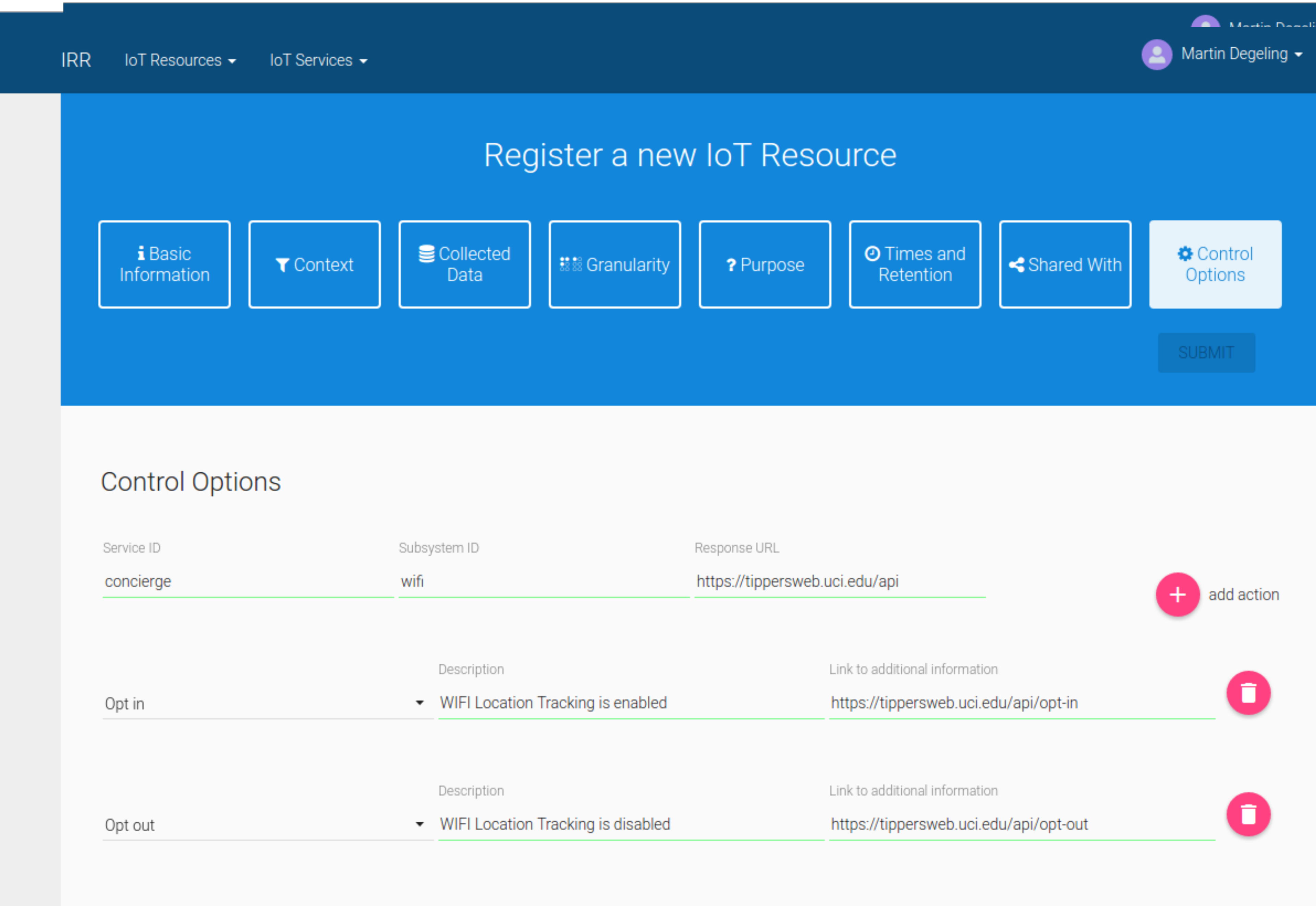
- Resources in this coverage area
- Resource c
- Privacy not

nces

eceived from IRRs



Resource	User Preference
WiFi Location Tracking (Service)	Opt In
Bluetooth Beacon Location Tracking (Service)	Opt Out
CMU Friend Finder (App)	Opt In (Location Tracking)
Facial Recognition (Service)	Opt Out
Video Obfuscation Demo (App)	Opt Out (Facial Recognition)



Register a new IoT Resource

Control Options

Service ID: concierge

Subsystem ID: wifi

Response URL: https://tippersweb.uci.edu/api

Opt in

Description: WIFI Location Tracking is enabled

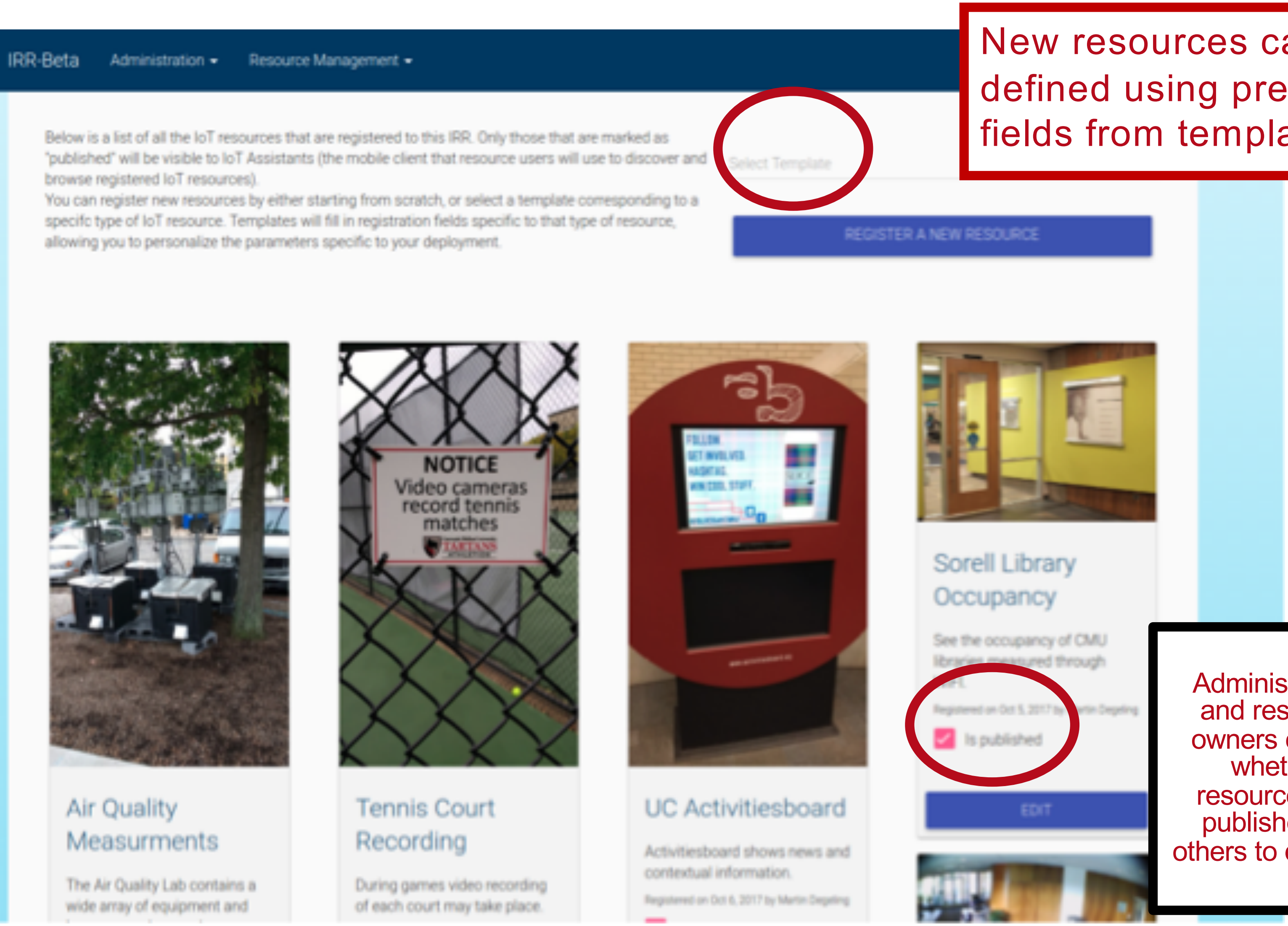
Link to additional information: https://tippersweb.uci.edu/api/opt-in

Opt out

Description: WIFI Location Tracking is disabled

Link to additional information: https://tippersweb.uci.edu/api/opt-out

Sample Entries & IoT Device Templates



Below is a list of all the IoT resources that are registered to this IRR. Only those that are marked as "published" will be visible to IoT Assistants (the mobile client that resource users will use to discover and browse registered IoT resources).
You can register new resources by either starting from scratch, or select a template corresponding to a specific type of IoT resource. Templates will fill in registration fields specific to that type of resource, allowing you to personalize the parameters specific to your deployment.

Air Quality Measurments

The Air Quality Lab contains a wide array of equipment and

Tennis Court Recording

During games video recording of each court may take place.

UC Activitiesboard

Activitiesboard shows news and contextual information.

Registered on Oct 6, 2017 by Martin Degeling

Sorell Library Occupancy

See the occupancy of CMU libraries measured through WiFi.

Registered on Oct 5, 2017 by Martin Degeling

☑ Is published

New resources can be defined using pre-filled fields from templates

Existing Templates:
- Echo
- Honeywell thermostat
- Kinect
- Google Home
- Cujo
- Nest Cam
- Wink Relay
- Zensors

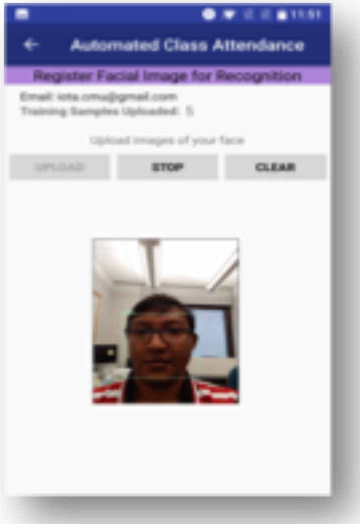
Administrators and resource owners control whether resources are published for others to discover

Where Do We Start?

- Smart cities (e.g. cameras)
- Malls (e.g., cameras, location tracking)
- University campuses – all sorts of IoT technologies
- Smart buildings (e.g. cameras, location, presence, HVAC)
- Smart homes (e.g. smart speakers)

Privacy-aware Video Streaming

Train Facial Features



Automated Class Attendance

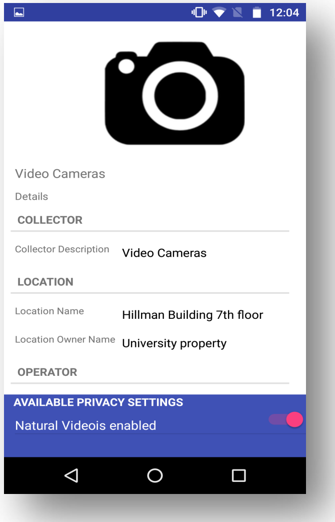
Register Facial Image for Recognition

Email: iota.cmu@gmail.com

Training Samples Uploaded: 5

Upload images of your face

Control Opt-in



Video Cameras

Details

COLLECTOR

Collector Description: Video Cameras

LOCATION

Location Name: Hillman Building 7th floor

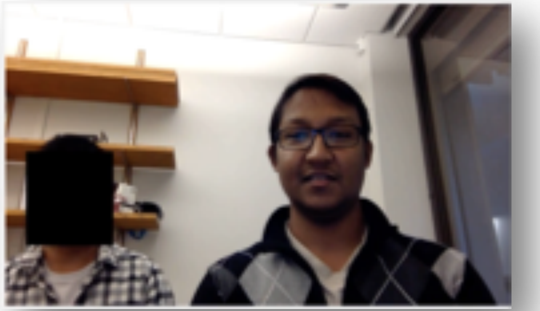
Location Owner Name: University property

OPERATOR

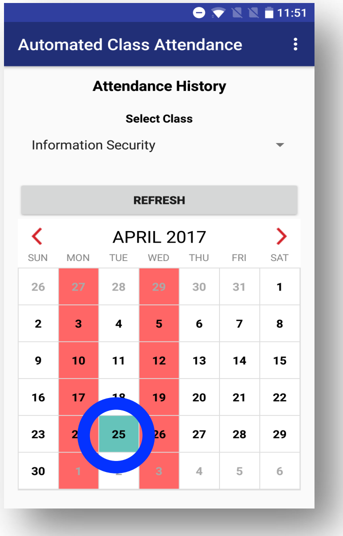
AVAILABLE PRIVACY SETTINGS

Natural Videois enabled

Live Video Stream



Monitor Class Attendance



Automated Class Attendance

Attendance History

Select Class

Information Security

APRIL 2017

SUN	MON	TUE	WED	THU	FRI	SAT
26	27	28	29	30	31	1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	1	2	3	4	5	6

Demo: https://goo.gl/gtpbpK

Current Status

- Deployed at UC Irvine
- Deployed at CMU
- First public release coming out this summer
- Includes tools to facilitate adoption
  - Tool to help manage IoT Resource Registries (e.g., administrator portal), tool to enter resources, templates for commercial off-the-shelf IoT resources
  - IoT resource registries hosted at CMU
  - Secure protocols for communicating with user-configurable privacy settings (e.g., opt-in, opt-out)

Concluding Remarks

- Privacy is a fundamental human right and people care about privacy
  - Regulations like COPPA, HIPAA but also GDPR
- Fundamental tension between privacy and usability
- Many IoT data collection processes are invisible/obscure and unexpected
- Notice and Choice in the IoT will require deployment of a **Privacy Infrastructure** that supports the discovery of IoT resources & their data practices
- First release this summer – **subscribe to our mailing list** for updates: https://www.privacyassistant.org/contact/

Acknowledgements: Work funded by the National Science Foundation, DARPA and Google

The **Usable Privacy Policy Project** and the **Personalized Privacy Assistant Project both** involve a collaborations with a number of individuals. See **usableprivacy.org** and **privacyassistant.org** for additional details incl. lists of collaborators and publications

Q&A