PRIVACY CON
FEDERAL TRADE COMMISSION
DC // 1.12.17

Opening Remarks

Chairwoman Edith Ramirez

PRIVACYCON

Session 1: Internet of Things and Big Data

PRIVACYCON

A Smart Home is No Castle: Privacy Vulnerabilities of Encrypted IoT Traffic

Noah Apthorpe, Dillon Reisman, Nick Feamster

Center for Information Technology Policy, Princeton University

This research was funded by the Department of Defense through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program, a Google Faculty Research Award, and the National Science Foundation

PRIVACYCON

Acknowledgements

- Arvind Narayanan, Princeton CITP
- Nina Taft, Google Inc.
- Alex Yue, Princeton CITP
- Caio Burgardt, Princeton CITP & Federal University of Pernambuco

What is a smart home?

ISP

Internet

Google

amazon

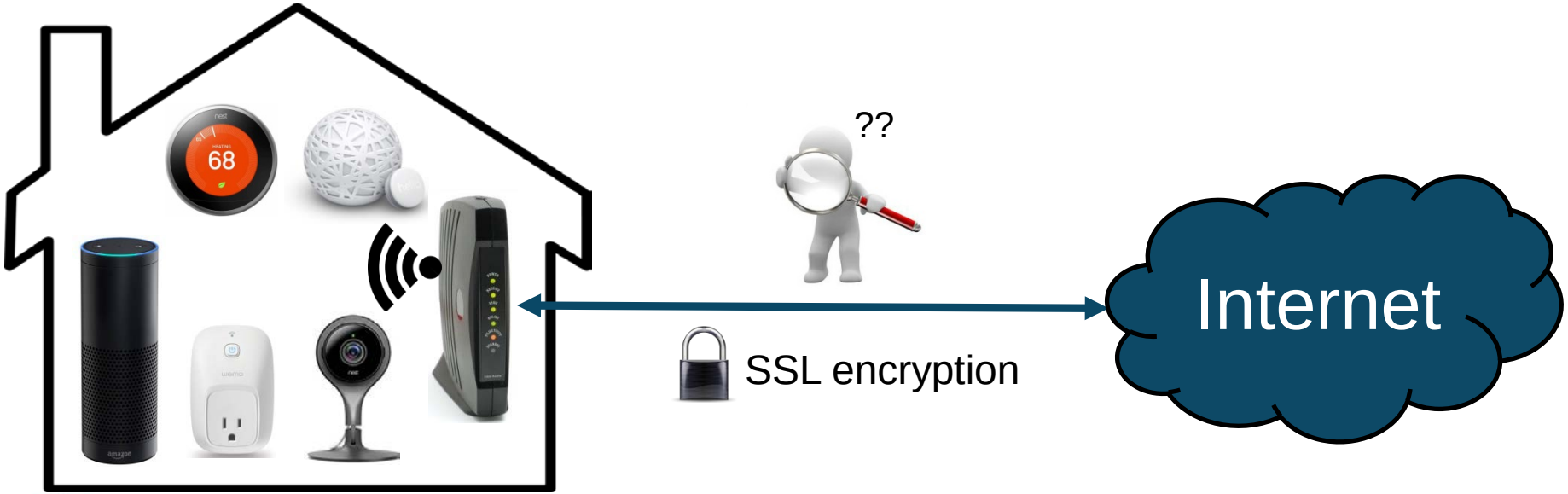
belkin

hello

Smart home privacy

- Smart home Internet traffic can contain sensitive consumer information
  - Physical presence in home
  - Sleeping habits
  - Media consumption
  - Appliance usage

Smart home privacy

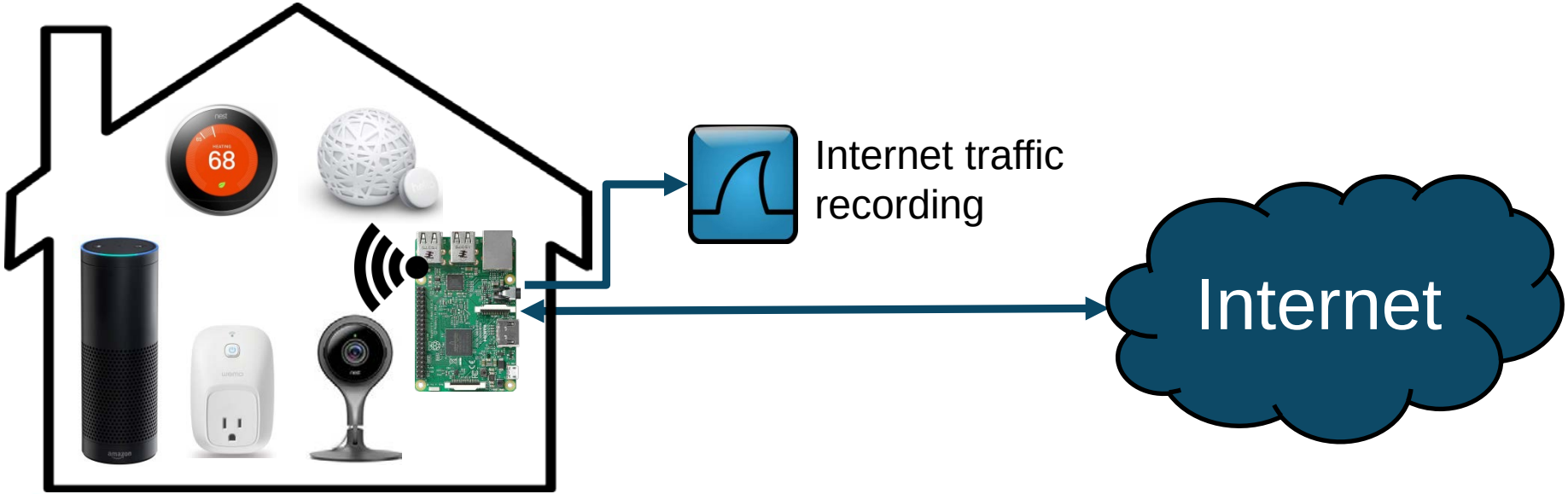


Smart home privacy

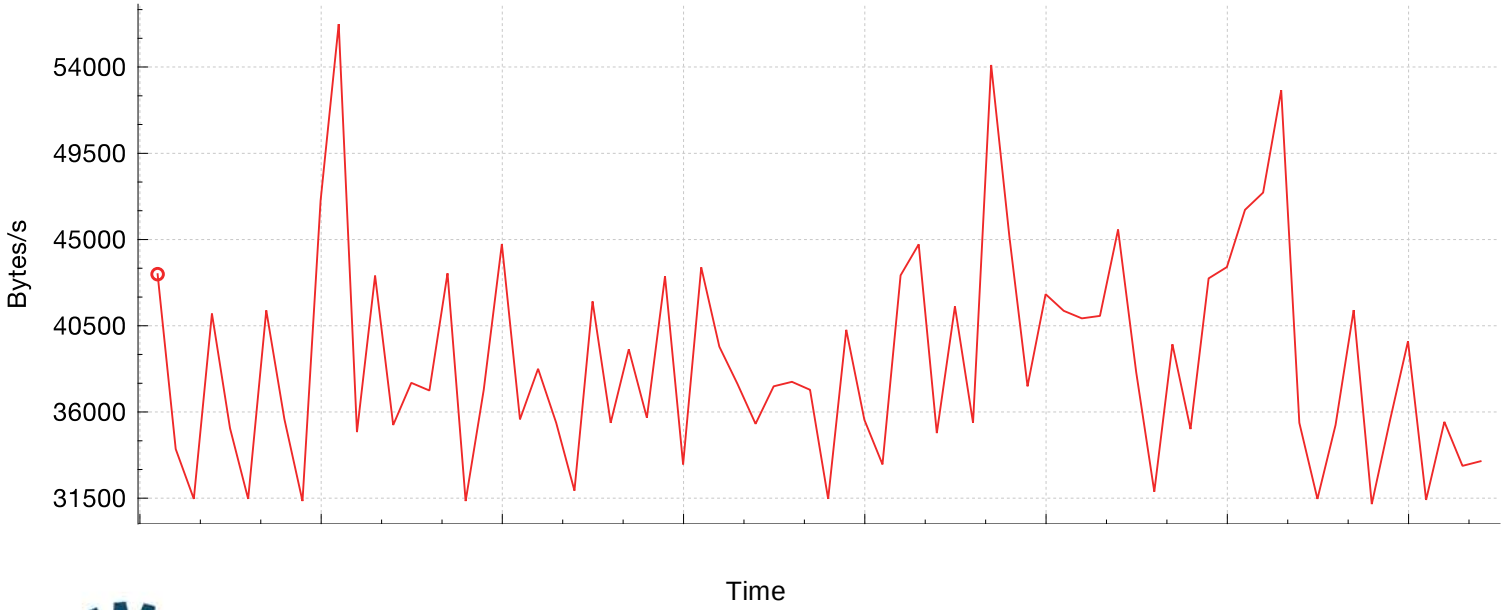
Smart home Internet traffic patterns reveal consumers' private in-home behaviors

even when traffic is encrypted

Smart home laboratory

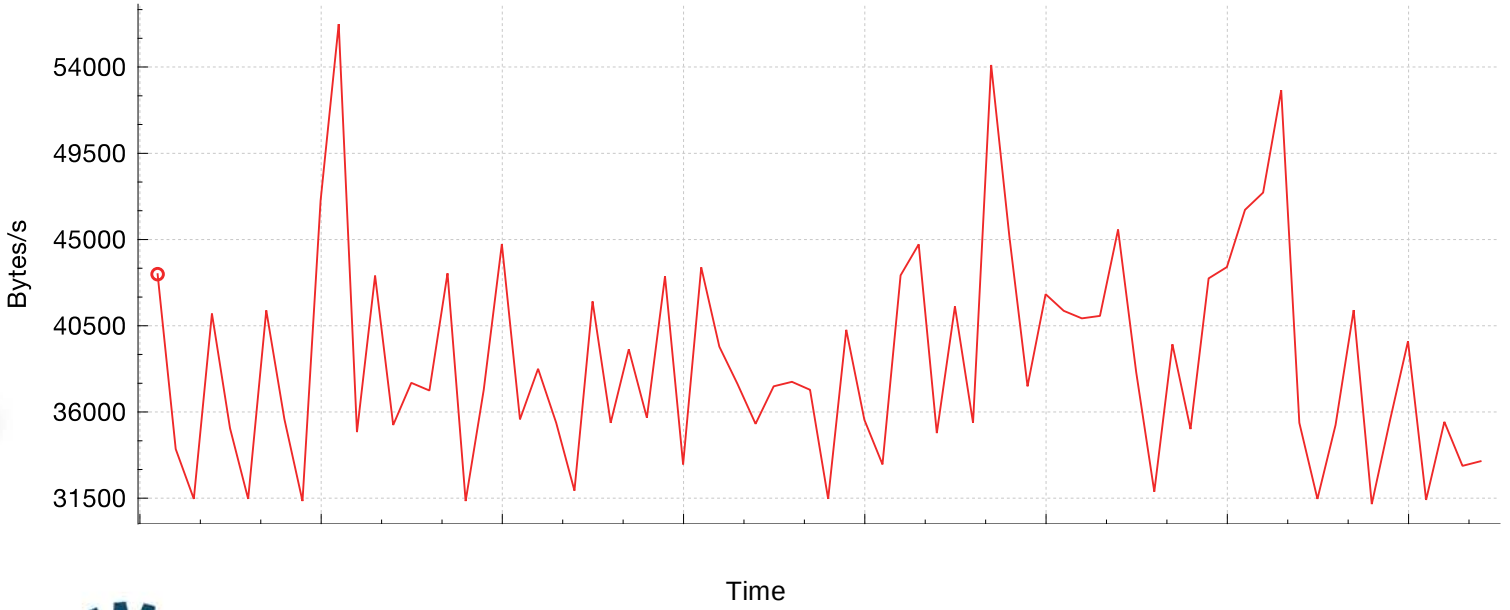


1. Obtain Internet traffic from device



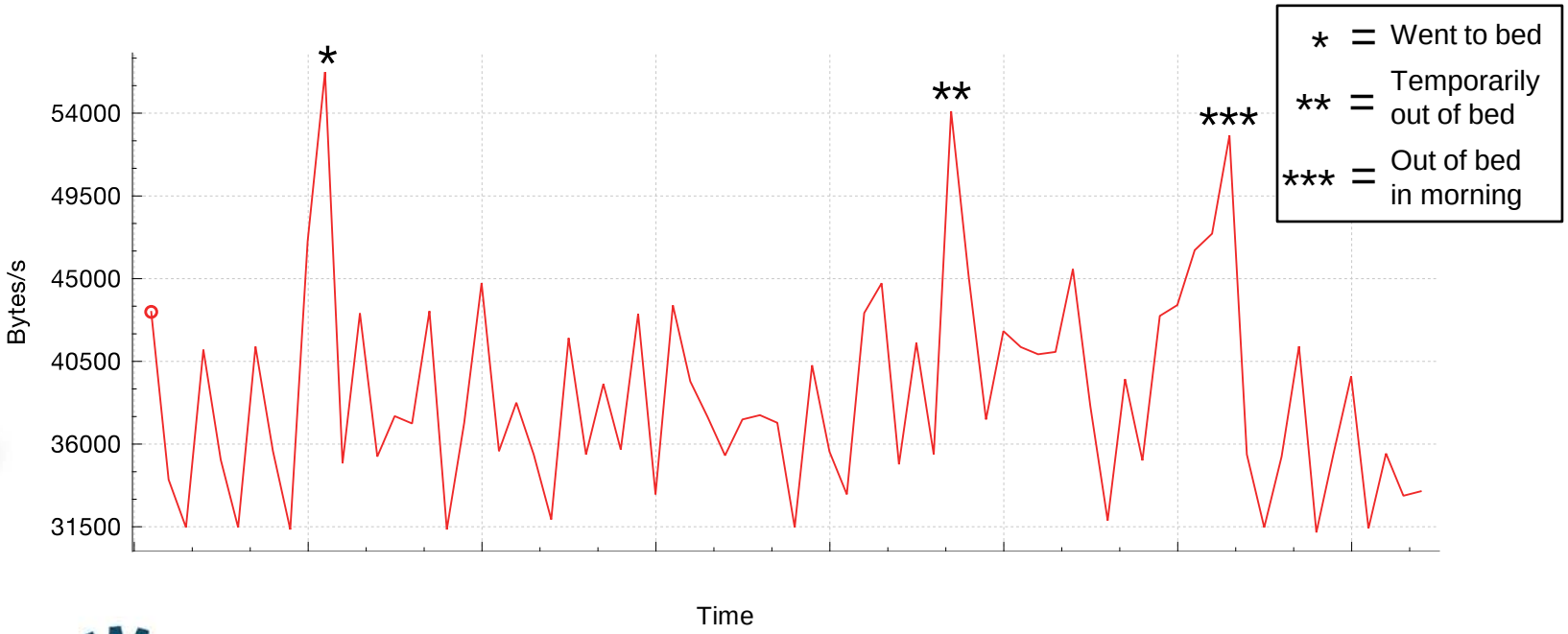
2. Identify device

Sleep monitor



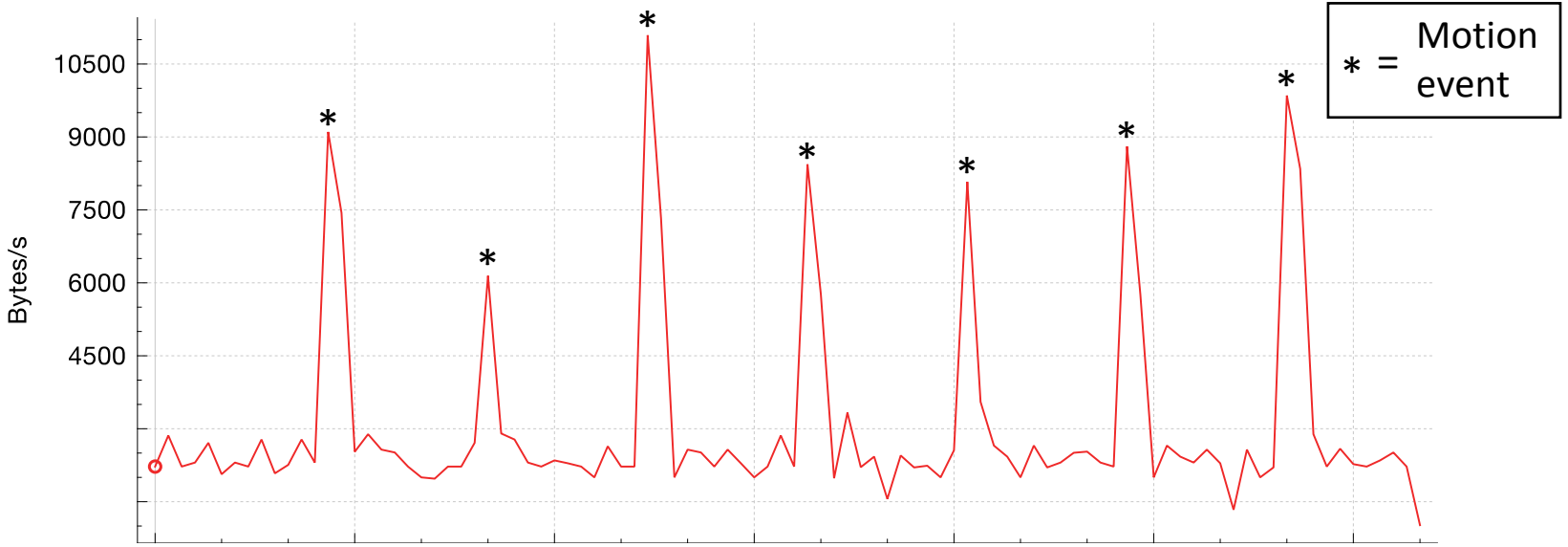
3. Infer behaviors from traffic patterns

Sleep monitor

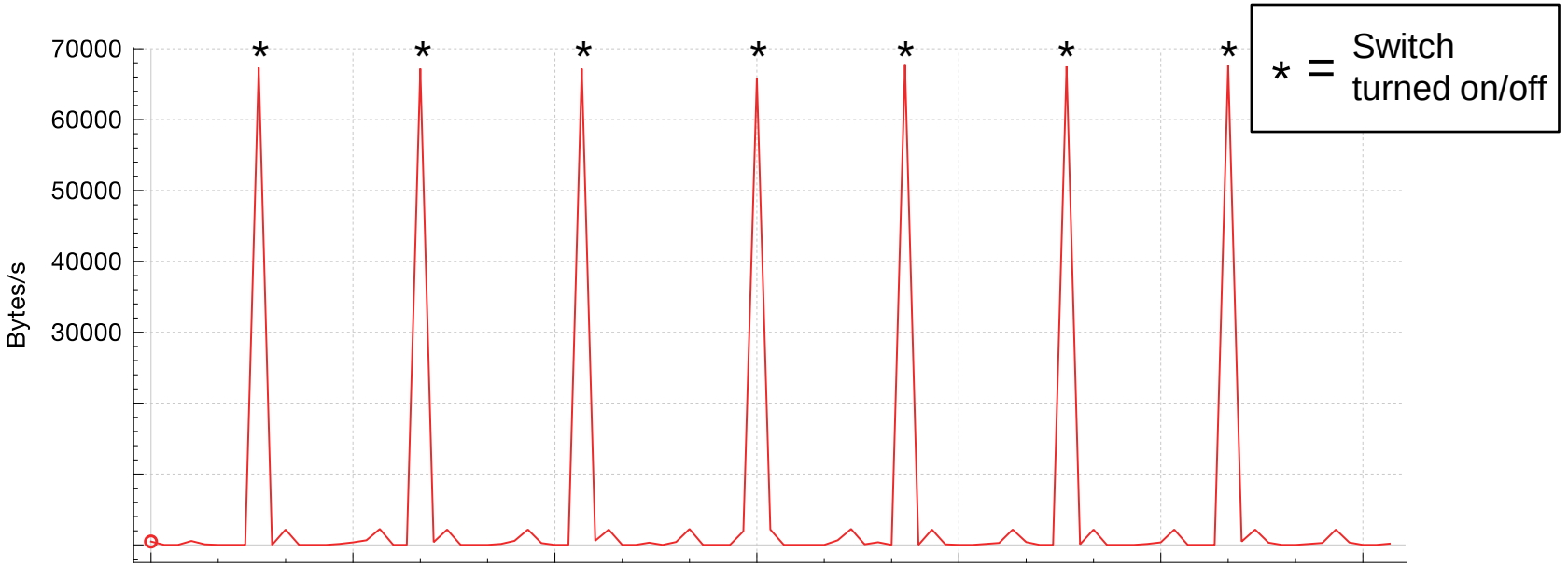


Security camera monitoring

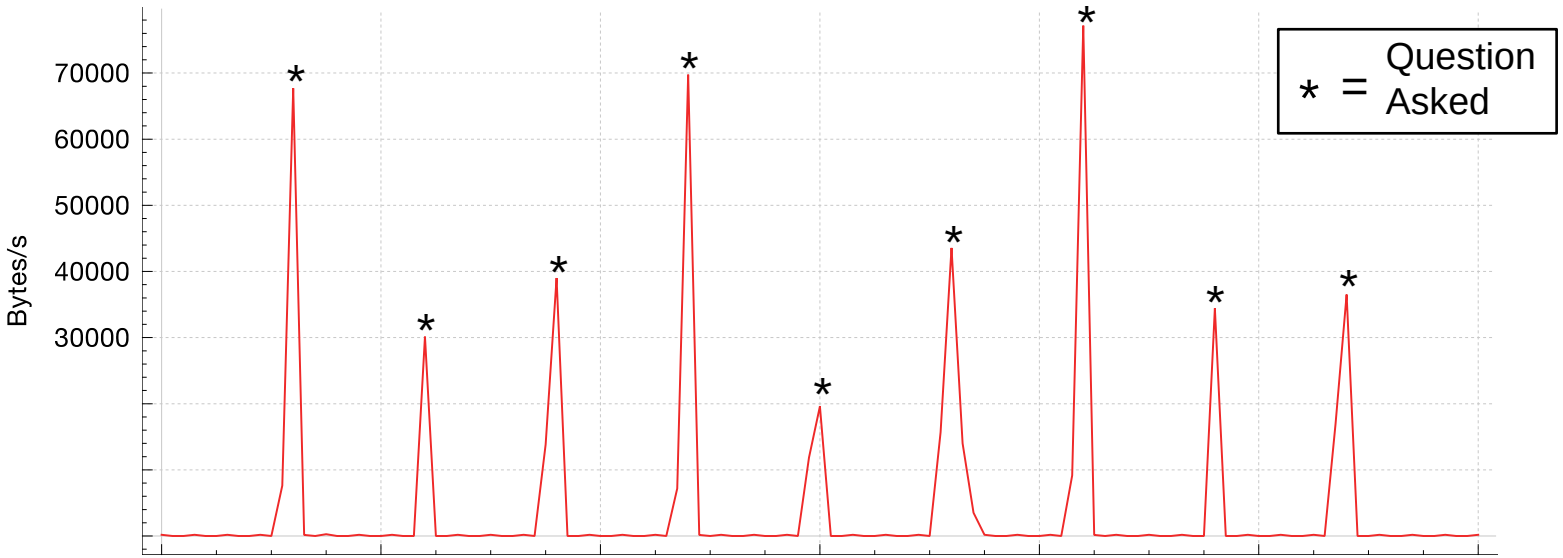
Presence of motion in home



Power to physical appliance



Interactions with personal assistant



Future questions

- How widespread is this privacy risk?

- Can we design privacy preservation methods that address metadata?
  - e.g. probabilistic traffic injection

Conclusion

- Encryption alone does not provide adequate privacy protection for smart home IoT devices

- Link to paper: http://datworkshop.org/papers/dat16-final37.pdf

Cross-App Tracking via Nearby Bluetooth Devices

Aleksandra Korolova

University of Southern California

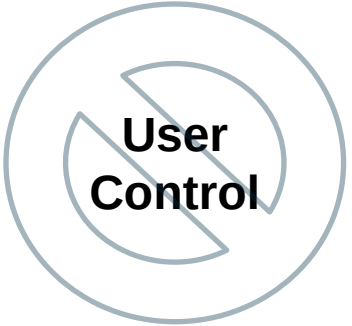
Joint work with Vinod Sharma, USC

Bluetooth-enabled Devices Everywhere

What are the Privacy Implications?

New channel for:

- Profiling the user
- Tracking across apps



The Bluetooth LE Protocol

Devices transmit info	Example
MAC address	A4:77:32:4E:43:92
Name	Joe’s Chromecast
Manufacturer data	16 bit id
Service UUIDs	0000fea0-0000-1000-8000-00805f9b34fb

Frequency: Every 20ms-10s
Range: up to 100m

App Developers

Can request info of all nearby devices from OS

- Unlimited frequency
- User permission not needed in iOS & Android 5
  - iOS modifies MAC addresses
  - Android 6 requires location permission

How does BLE enable Profiling?

Dan Jo's FitBit

Kevo

[TV] Samsung 9 Series (65)

mamaRoo

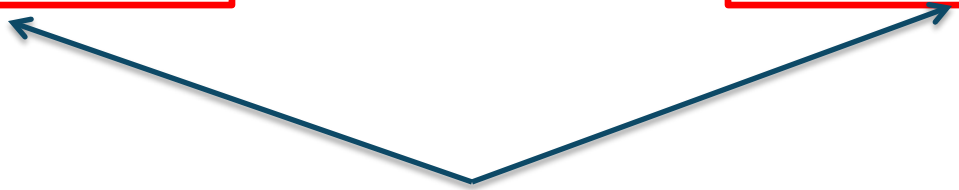
How does BLE enable Tracking?

APP 1

USER 1	USER 2
AA:BB:CC:DD:EE	FF:FF:FF:FF:FF
A1:B1:C1:D1:E1	55:55:55:55:55
AA:AA:AA:AA:AA	BB:BB:BB:BB:BB

APP 2

USER X	USER Y
AA:BB:CC:DD:EE	CC:CC:CC:CC:CC
A1:B1:C1:D1:E1	66:66:66:66:66
AA:AA:AA:AA:AA	DB:DB:DB:DB:DB



Likely the same user

Our Study

70 volunteers collect nearby Bluetooth-enabled device data every 10 minutes for 1 week in June 2016

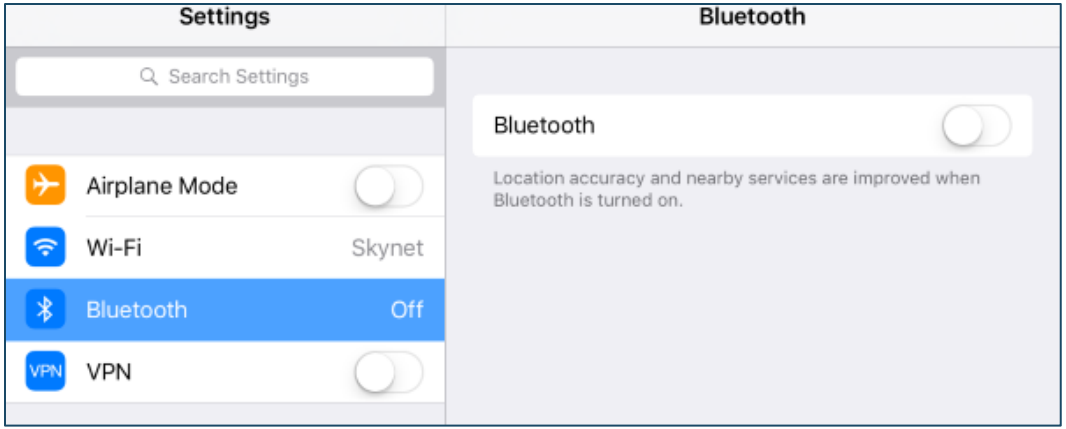
Findings:

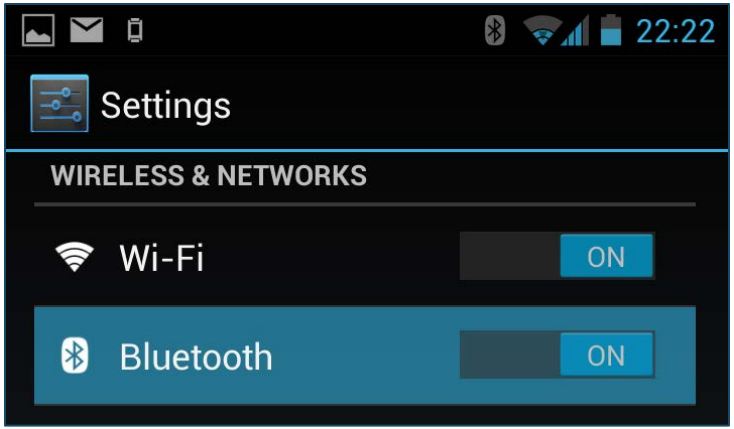
- **1,000+** distinct device names
- **87%** uniquely identifiable by apps used every 5 hours
- **60%** uniquely identifiable by apps used once a day

Is BLE-based Profiling & Tracking Happening?

- Hard to tell, which is a problem in itself!
- What we know:
  - increasing # of apps declaring Bluetooth use
  - some apps access Bluetooth quite often
  - nothing stops them from using obtained info for profiling & tracking

Controls Available to Individuals





Device-level Bluetooth ON / OFF

- Hardly a meaningful choice long-term

Conclusions – Changes are Needed

I. Profiling & Tracking using nearby BLE devices feasible

I. Can increase privacy without hurting functionality, but changes needed by:
   - Bluetooth Special Interest Group
   - Apple, Google
   - Device manufacturers

A role for FTC?

Cross-App Tracking via Nearby Bluetooth Devices

Aleksandra Korolova

University of Southern California

Study details: http://bit.ly/BluetoothPrivacy

Joint work with Vinod Sharma, USC

PRIVACYCON

Rating Products & Services for Privacy, Security & Data Practices

Maria Rerecich
Consumer Reports

This research was funded by the Ford Foundation

PRIVACYCON

Testing Philosophy

- Comparative Testing -> Informed choice
- Defined protocols and procedures
- Consumer-focused
- Companies compete
- Products improve
- Industry standards and regulations advance

Developing a Standard: Why Now?

- Shift from hardware-only to software-enabled and connected, i.e. Internet of Things
- Uncharted territory for consumers
- No easy way to navigate the privacy and security of this new digital world
- Need for a consistent, accessible standard to measure these products

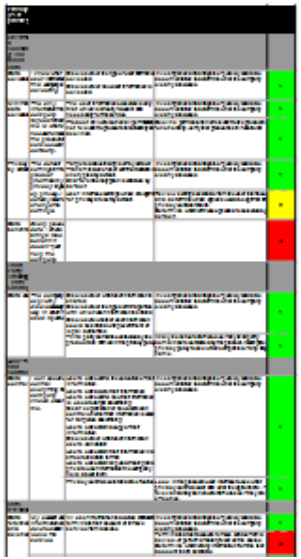
Organizing Principles

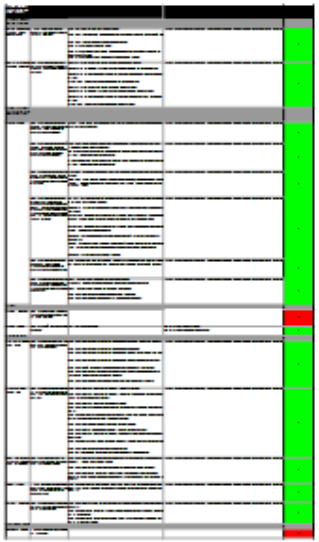
- Security – “Is it safe?”
- Privacy – “Is it private?”
- Governance and Compliance – “Are the policies strong for consumers?”
- Ownership – “Is it mine?”

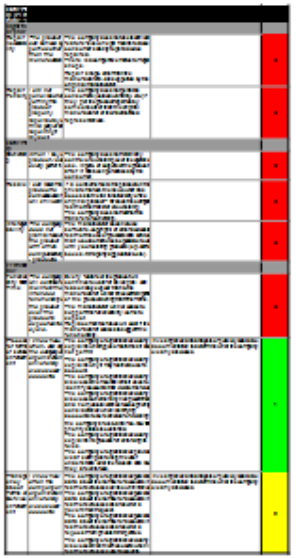
Example

- Criterion: The product is protected from known software vulnerabilities that present a danger from attackers.
- Indicator: The software is secure against known bugs and types of attacks.
- Procedure Overview: Launch activities from user interface. Monitor communication to/from device.

A Work in Progress







Next Steps

- Create an openly sourced digital standard that can be used to hold manufacturers and providers accountable for how they manage consumers' privacy, security, and data.
- To provide any comments or feedback, please email us at externalrelations@cr.consumer.org

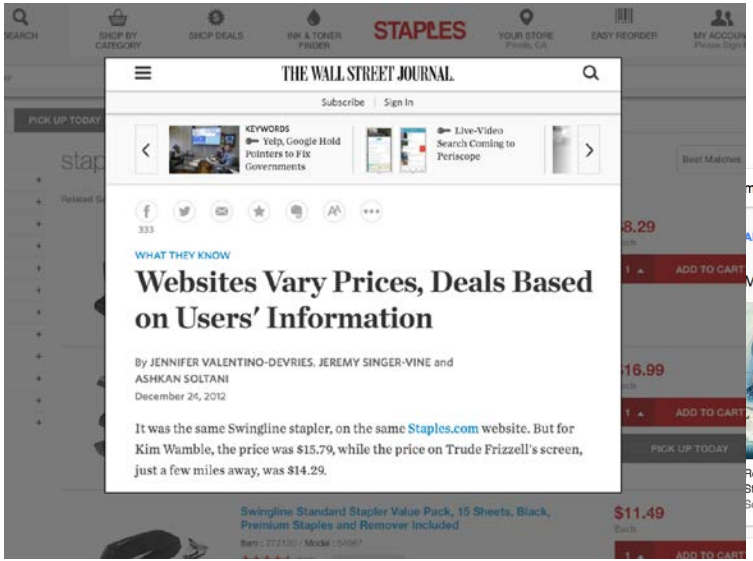
A User-Centered Perspective on Algorithmic Personalization

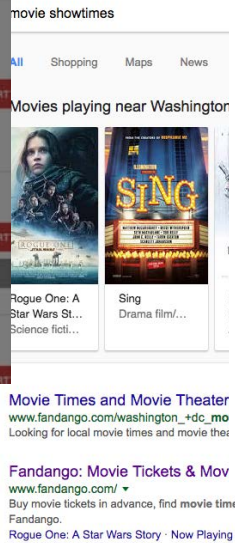
Alethea Lange, Rena Coen, Emily Paul, Pavel Vanegas, & G.S. Hans
Center for Democracy & Technology
UC Berkeley School of Information

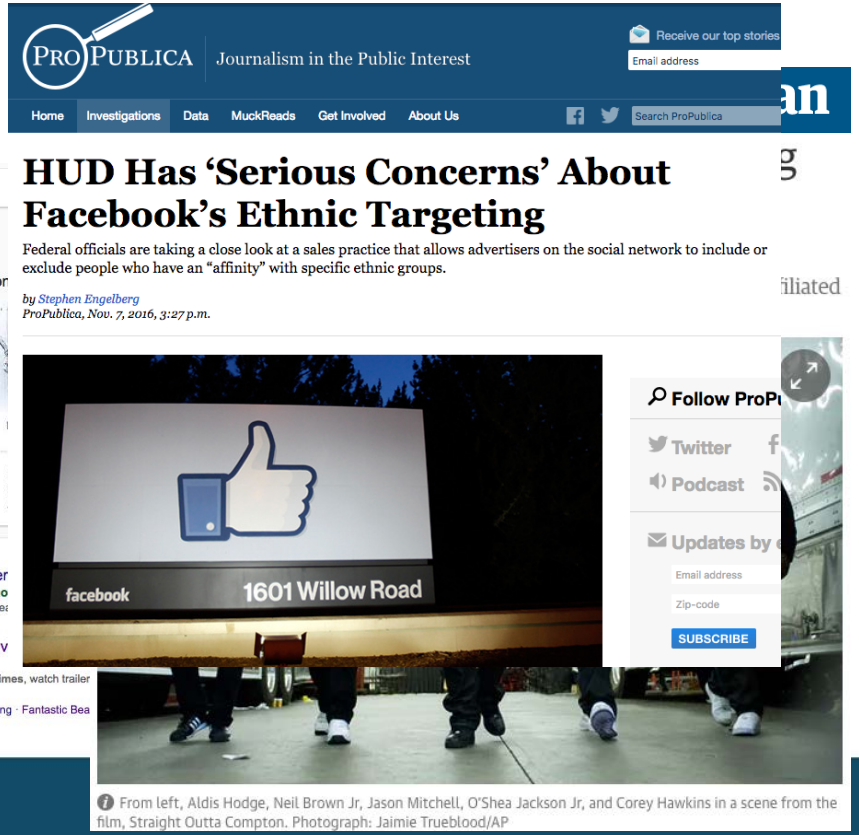
This research was funded by the Berkeley Center for Technology, Society, and Policy and the Berkeley Center for Long-Term Cybersecurity

PRIVACYCON

Personalization Practices







PRIVACYCON

Research Design

DOMAIN	DATA TYPE	SOURCE
Targeted Advertising	Personal Information	Provided
Filtered Search Results	City/Town of Residence	Accurately Inferred
Differential Retail Pricing	Gender	Inaccurately Inferred
	Household Income	
	Race	

CONTEXT

INFERENCE & ACCURACY

PRIVACYCON

Research Design

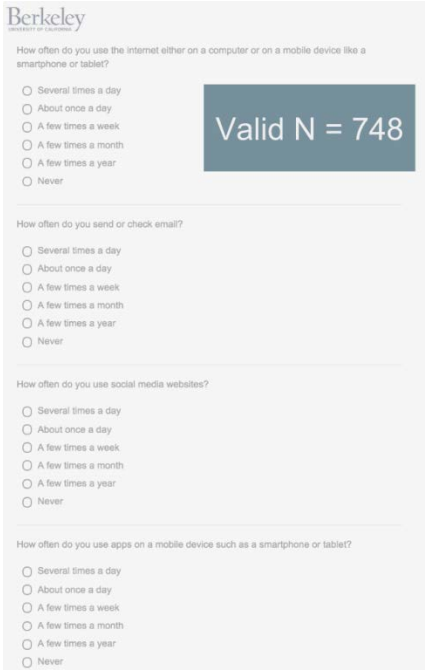
DOMAIN	DATA TYPE	SOURCE
Targeted Advertising	Personal Information	Provided
Filtered Search Results	City/Town of Residence	Accurately
Differential Retail Pricing	Gender	Inferred
	Household Income	Inaccurately
	Race	Inferred

You are reading an article on a website. The ad is shown to you based on your race which was inferred from the webpages you visit and is accurate.

Survey Instrument

qualtrics

amazon mechanical turk



Race	N	%
American Indian or Alaska Native	10	1.34%
Asian	70	9.36%
Black or African American	62	8.29%
Hispanic or Latino	50	6.68%
Native Hawaiian or Other Pacific Islander	3	0.40%
White	597	79.81%
Other	9	1.20%

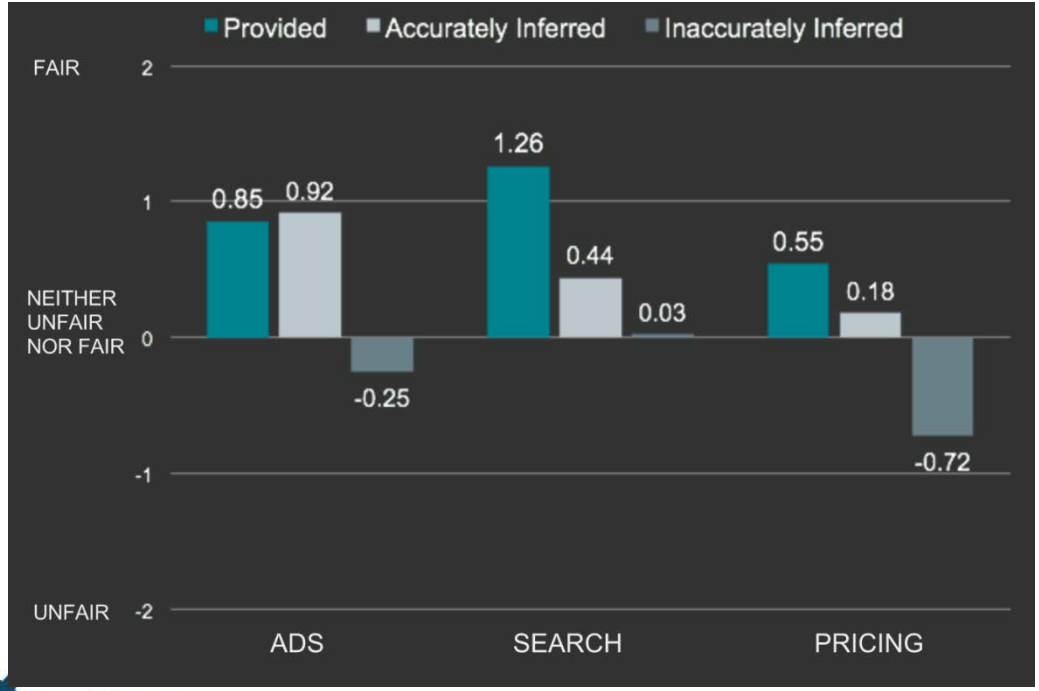
Gender	N	%
Male	347	46.39%
Female	392	52.41%
Female-to-Male transgender	1	0.13%
Male-to-Female transgender	0	0.00%
Genderqueer	5	0.67%
Other	0	0.00%
Prefer not to answer	3	0.40%

Household Income	N	%
Less than $20,000	118	15.78%
$20,000 - $39,999	211	28.21%
$40,000 - $59,999	175	23.40%
$60,000 - $79,999	106	14.17%
$80,000 - $99,999	60	8.02%
$100,000 +	77	10.29%

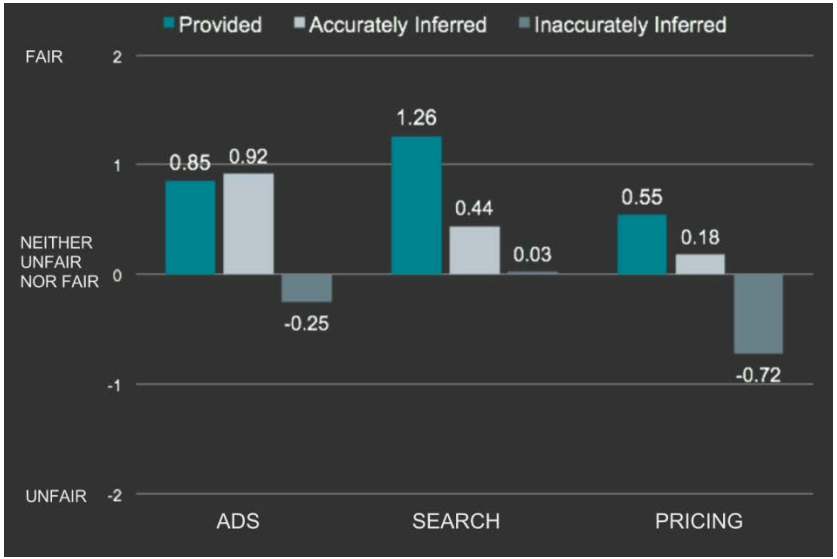
Age	N	%
18-33	406	54.28%
34-49	235	31.42%
50-68	95	12.70%
69 and older	11	1.47%

Median Age = 33

City or Town of Residence



City or Town of Residence

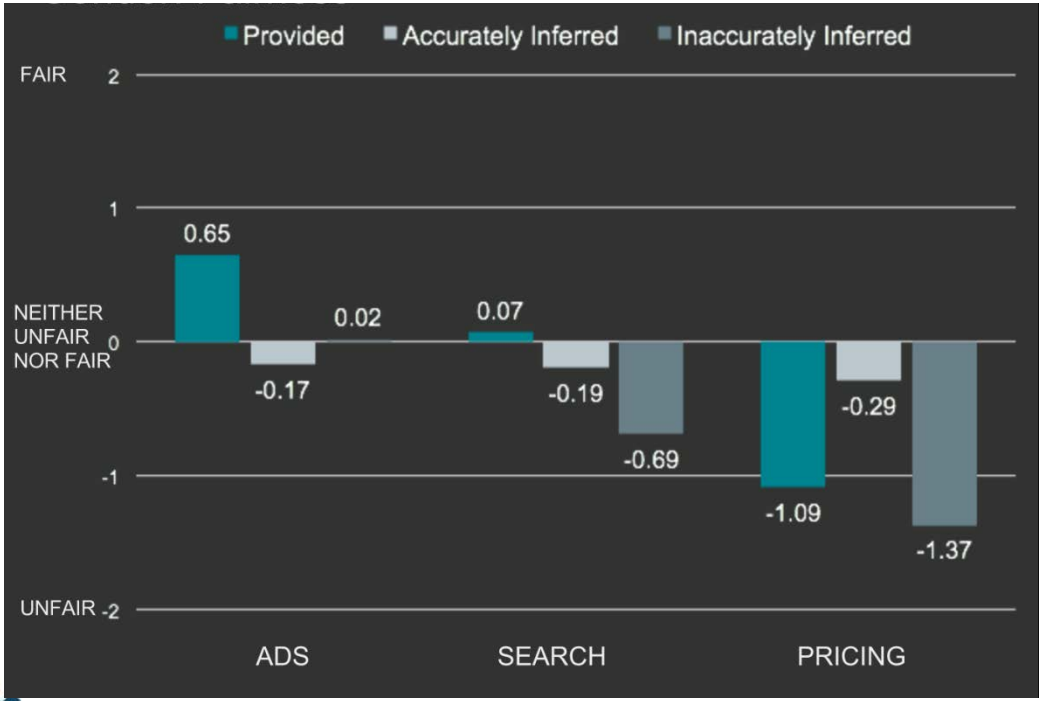


"Getting an **ad** about a product or store in my town is perfectly acceptable and beneficial to me."

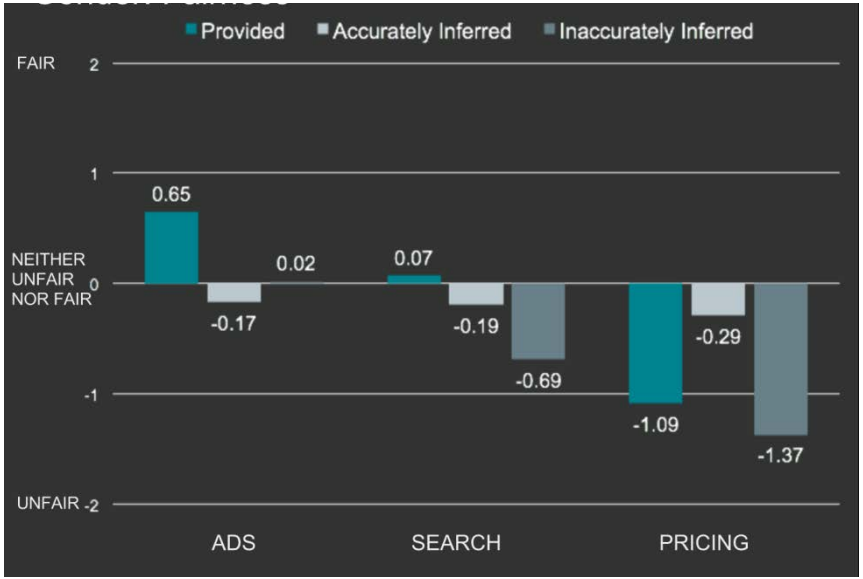
"I would find this acceptable for some things. If I was **searching** for a book or a movie plot . . . it wouldn't be helpful, but if I wanted the weather or a nearby restaurant, it would save time."

"My initial reaction is that **price** discrimination by geography is unfair. . . . After further thought, this effectively mimics the physical world where prices in areas with higher costs and typically higher salaries tend to be more."

Gender


Provided
Accurately Inferred
Inaccurately Inferred
FAIR
2
1
NEITHER UNFAIR NOR FAIR
0
-1
UNFAIR
-2
0.65
-0.17
0.02
0.07
-0.19
-0.69
-1.09
-0.29
-1.37
ADS
SEARCH
PRICING

Gender

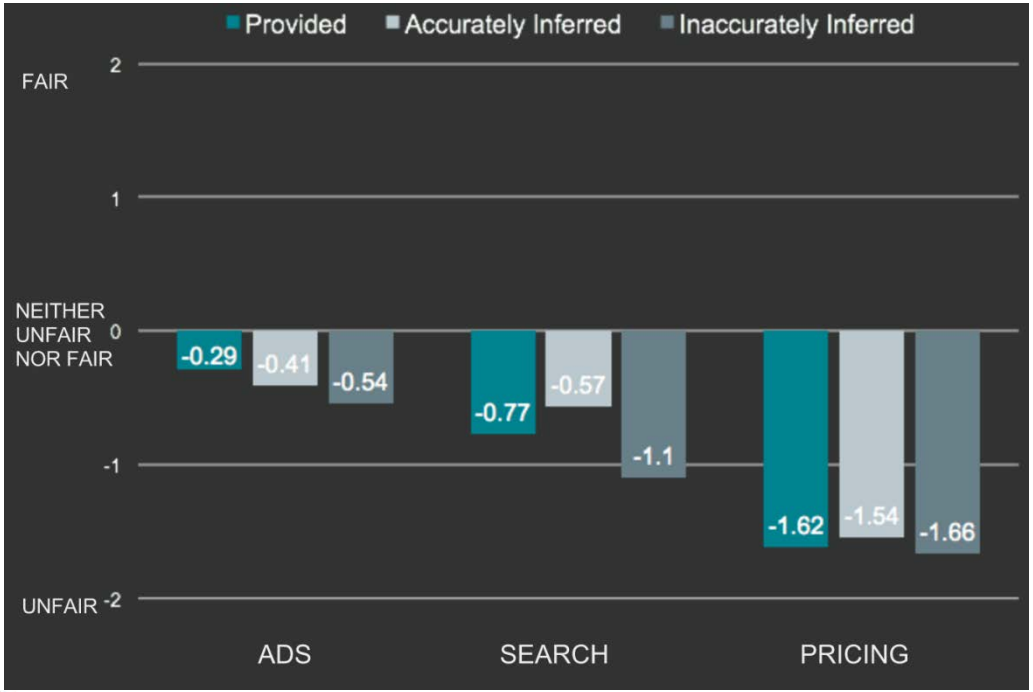


"I don't love the idea of targeted **ads** in general but something based on my gender seems relatively harmless."

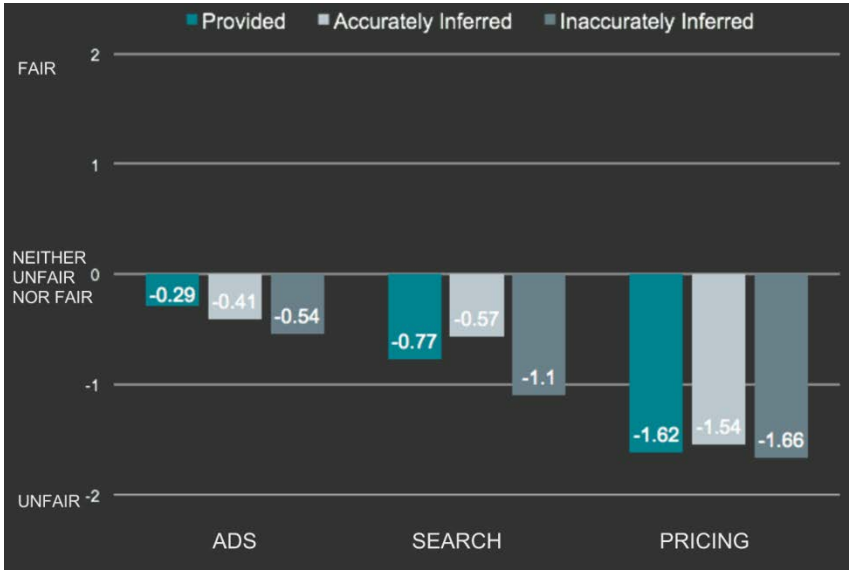
"I don't think my gender provides enough information to improve **[search]** results, and I worry that the filtering will . . . lead to people of different genders having access to different information."

"That seems very unfair. Products should **cost** the same regardless of gender. . . I don't trust a website that would pull something like this."

Race


Provided
Accurately Inferred
Inaccurately Inferred
FAIR
2
1
NEITHER UNFAIR NOR FAIR
0
-1
UNFAIR
-2
-0.29
-0.41
-0.54
-0.77
-0.57
-1.1
-1.62
-1.54
-1.66
ADS
SEARCH
PRICING

Race



"I cannot imagine a scenario where the **[ad]** presented, using this metric, would be anything more than indicative of stereotypes and racism."

"I don't like my **search** results being filtered by my race. I feel like I may be missing out on relevant information that I could use."

"My race shouldn't influence personal product **pricing** regardless of whether the information the company has about me is correct or not"

Key Takeaways

- Reconsider Personalizing in High-Stakes Domains
- Personalization Based on Location Data (City or Town of Residence) is Acceptable Across Contexts
- Data Quality is Meaningful
- Personalized Pricing Should Mirror Offline Practices
- Avoid Personalization Based on Race

Discussion of Session 1

Presenters:

- **Noah Apthorpe** and **Dillon Reisman**, Princeton University
- **Aleksandra Korolova**, University of Southern California
- **Maria Rerecich**, Consumer Reports
- **Alethea Lange**, Center for Democracy & Technology

Moderator:

- **Peder Magee,** Federal Trade Commission

PRIVACYCON

BREAK

PRIVACYCON