
PRIVACY CON
FEDERAL TRADE COMMISSION
DC // 1.14.16

Session 2: Consumers' Privacy Expectations

PRIVACYCON
FEDERAL TRADE COMMISSION

Serge Egelman

University of California, Berkeley

Android Permissions Remystified: A Field Study on Contextual Integrity

Co-authors: Primal Wijesekera (University of British Columbia); Arjun Baokar, Ashkan Hosseini, David Wagner (University of California, Berkeley); Konstantin Beznosov (University of British Columbia)

PRIVACYCON

FEDERAL TRADE COMMISSION

helping users make better mobile privacy decisions

Serge Egelman, UC Berkeley / ICSI

android comprehension study

online survey of 308 Android users

laboratory experiment with 24 users

A. P. Felt, E. Ha, S. Egelman, A. Haney, E. Chin, and D. Wagner. *Android Permissions: User Attention, Comprehension, and Behavior.* In Proceedings of the 2012 Symposium on Usable Privacy and Security (SOUPS). *Best Paper Award!*

suggestions

- many were habituated—too many requests
- only prompt when necessary

- many were unaware—too late in the process
- provide information earlier

- understanding requires knowing *all* permissions—too many permissions
- narrow list of possible permissions

impact on status quo

55% of permissions could be granted automatically

- reversible
- low risk

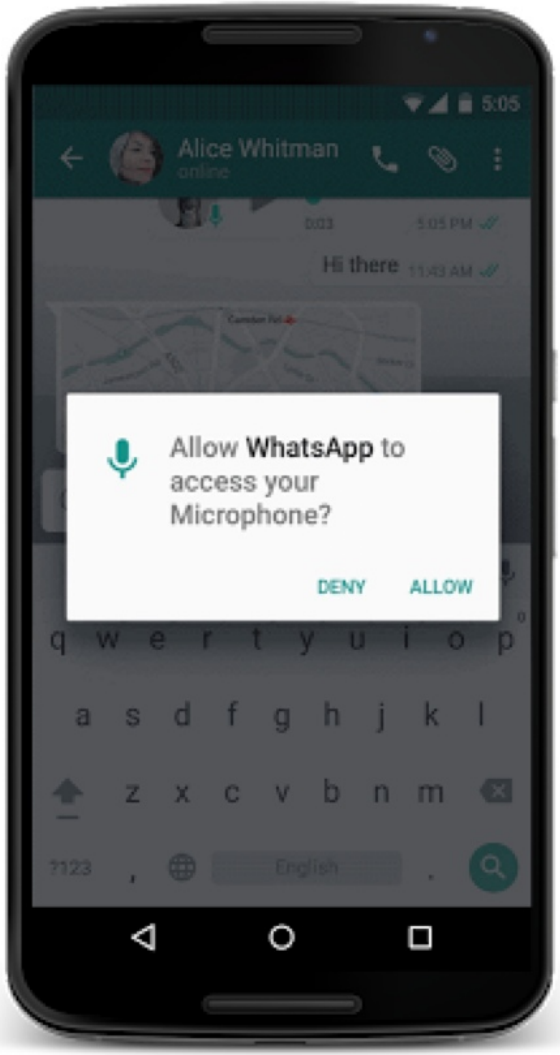
16% could use runtime dialogs

- adds contextual data

caveat: this does *not* reflect frequency of use.

A. P. Felt, S. Egelman, M. Finifter, D. Akhawe, and D. Wagner. *How to Ask for Permission.* Proceedings of the USENIX Workshop on Hot Topics in Computer Security (HotSec), 2012.

things improved



how often are resources accessed *in practice*?

dynamic analysis

- modified Android kernel and gave phones to 36 people
- hooked all API methods invoking permission checks

P. Wijesekera, A. Baokar, A. Hosseini, S. Egelman, D. Wagner, and K. Beznosov. *Android Permissions Remystified: A Field Study on Contextual Integrity.* Proceedings of the 24th USENIX Security Symposium, 2015.

contextual data

timestamp

visibility

screen status

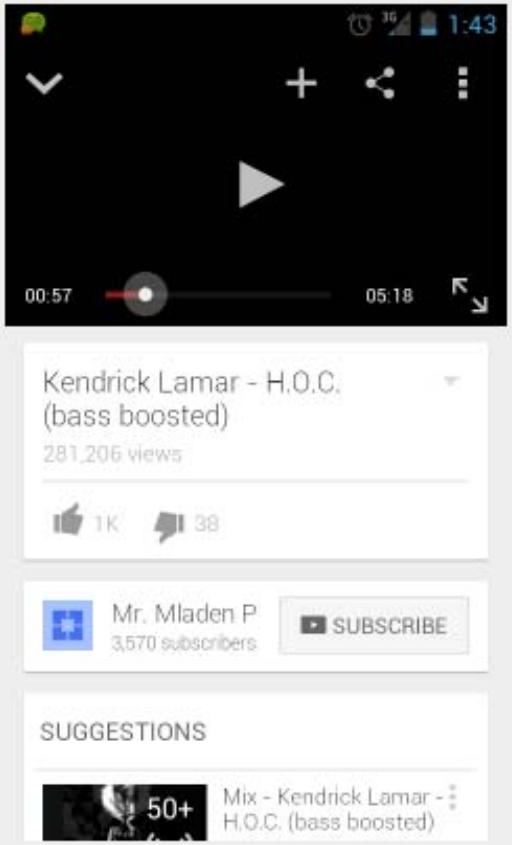
connectivity

location

view

history

Survey



"When this photo was taken, the com.mobilityware.solitaire was Scanning for WiFi"

3. On a scale of 1-5 how much did you expect this app to be accessing this resource?

◯ 1 (Least Expected) ◯ 2 ◯ 3 ◯ 4 ◯ 5 (Most Expected)

4. If you were given the choice, would you have prevented the app from accessing this data?

◯ Yes ◯ No

5. Why?

6. Is it okay for the researchers to view this screenshot?

◯ Yes ◯ No

Next

the results

36 Android smartphone users

6,048 hours of real-world use

27 million permission requests

incorrect mental models

invisible permissions

75.1%

background app (0.70%)
invisible service (14.40%)
screen off (60.00%)

non-indicative indicators

icon is visible for only
0.04% of
accesses to location.

runtime requests?

213 requests per hour!

- location (10,960/day)
- reading SMS data (611/day)
- sending SMS (8/day)
- reading browser history (19/day)

asking each time is infeasible

- …but 80% of participants wanted to block at least 1 request
- on average, they wanted to block 35% of all requests

predicting expectations?

expectations predicted blocking
($r=-0.39$, $p<0.018$)

- decision-making based on <*application,permission*> is only correct ~50% of the time
- increases to ~85% when examining <*application,permission,visibility*>
- privacy is deeply personal

future work

- implementing classifier
- constructing ecosystem
  - “hard” vs. “soft” policy
  - soft policy:
    - similar users
    - prompts
    - other behaviors

conclusion

human attention is a finite resource

focus attention on unexpected data uses

Serge Egelman
egelman@cs.berkeley.edu

Ashwini Rao

Carnegie Mellon University

Expecting the Unexpected: Understanding Mismatched Privacy Expectations Online

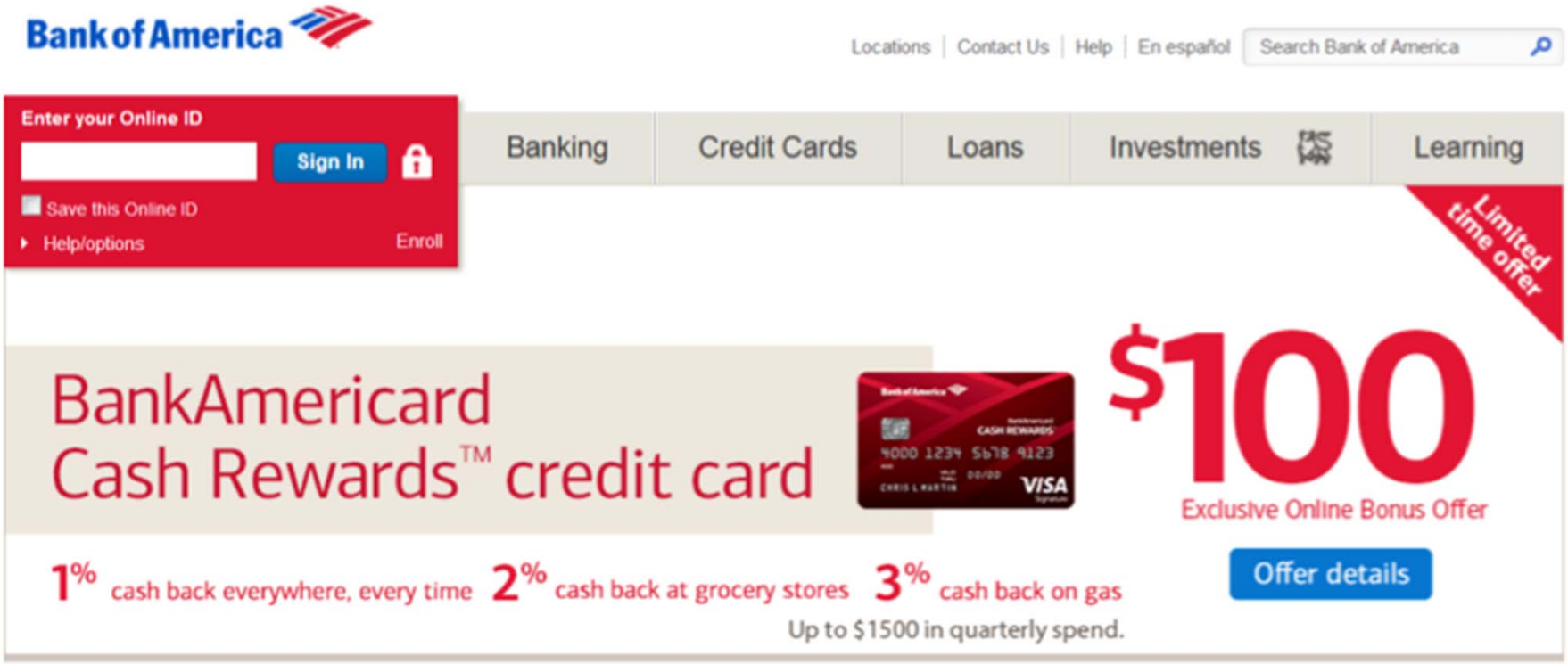
Co-authors: Florian Schaub, Norman Sadeh, Alessandro Acquisti, Ruogu Kang (Carnegie Mellon University)

Expecting the Unexpected: Understanding Mismatched Privacy Expectations Online

Ashwini Rao
School of Computer Science
Carnegie Mellon University

Joint work with Florian Schaub, Norman Sadeh, Alessandro Acquisti and Ruogu Kang

MOTIVATION



What data does Bankofamerica.com collect?

Bank of America

Last updated July 24, 2014

This U.S. Online Privacy Notice (Notice) applies to this Bank of A
application) and any Bank of America U.S. affiliate or subsidiary c
Site, and, collectively, Sites). The term "Bank of America" or "we'
and non-banking U.S. affiliates or subsidiaries of Bank of America
describes how Sites may collect, use and share information from
may be collected and used for advertising purposes.

Bank of America provides other online interfaces not covered by t
from one of these sites, please review the online privacy practices
information may be collected, used and shared.

For U.S. account holders and visitors to this Site, we will use and
about you in accordance with the Bank of America U.S. Consume
use and sharing of information. For Non-U.S. account holders util
account information in accordance with the privacy disclosure tha
security rules applicable to the Bank of America affiliate or subsid

Additional information on our Privacy & Security practices may be
Asked Questions (FAQs). Although the additional information is p
of this Notice control, and by using the Site, you agree to the term

Collecting and Using Information

Personal Information We Collect Online

Personal Information means personally identifiable information su
surveys, applications or other online fields including name, postal

Privacy policies are long and difficult to understand[1]

How can we help users understand online data practices?

Approach: focus on user expectations

[1]McDonald and Cranor. The cost of reading privacy policies. ISJLP 2008.

Users' **expect** websites to engage in certain data practices (collection, sharing etc.)

- Possibly vary by contextual and user characteristics

User expectations may not match actual data practices of online services

Could we generate effective privacy notices by **extracting and highlighting data practices that do not match users' expectations**?

From Policies to Effective Notices

Last updated July 24, 2014

This U.S. Online Privacy Notice (Notice) applies to this Bank of A
application) and any Bank of America U.S. affiliate or subsidiary c
Site, and, collectively, Sites). The term "Bank of America" or "we'
and non-banking U.S. affiliates or subsidiaries of Bank of America
describes how Sites may collect, use and share information from
may be collected and used for advertising purposes.

Bank of America provides other online interfaces not covered by t
from one of these sites, please review the online privacy practices
information may be collected, used and shared.

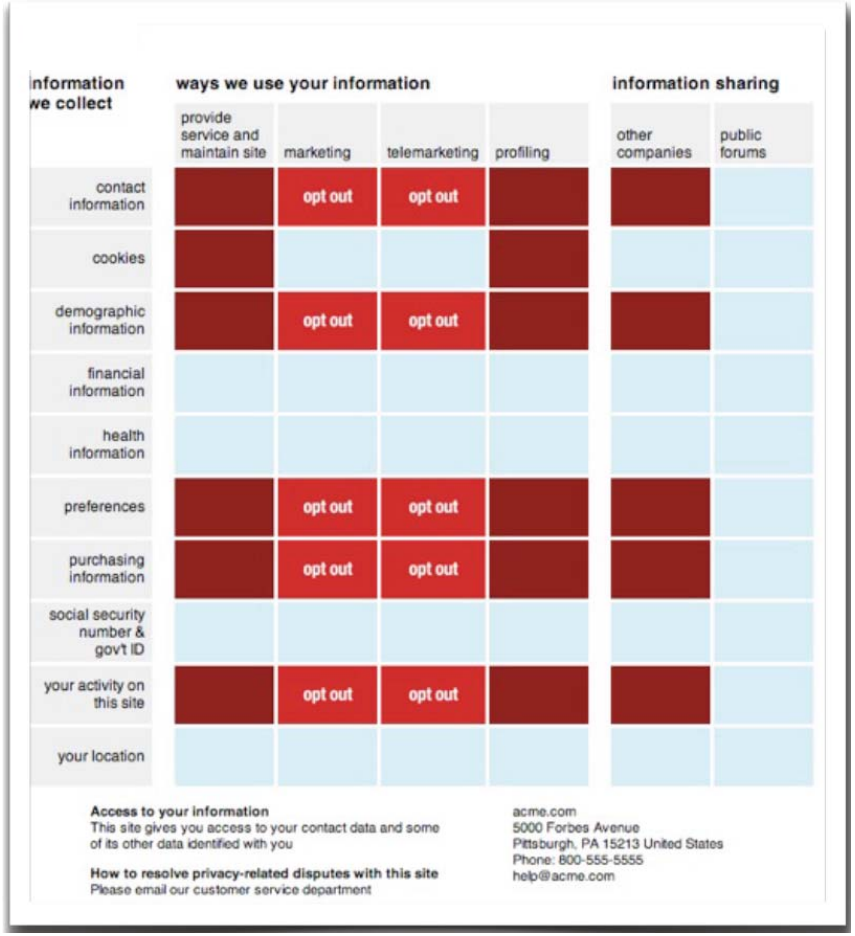
For U.S. account holders and visitors to this Site, we will use and
about you in accordance with the Bank of America U.S. Consume
use and sharing of information. For Non-U.S. account holders util
account information in accordance with the privacy disclosure tha
security rules applicable to the Bank of America affiliate or subsid

Additional information on our Privacy & Security practices may be
Asked Questions (FAQs). Although the additional information is p
of this Notice control, and by using the Site, you agree to the term

Collecting and Using Information

Personal Information We Collect Online

Personal Information means personally identifiable information su
surveys, applications or other online fields including name, postal



information we collect	ways we use your information				information sharing	
	provide service and maintain site	marketing	telemarketing	profiling	other companies	public forums
contact information		opt out	opt out			
cookies						
demographic information		opt out	opt out			
financial information						
health information						
preferences		opt out	opt out			
purchasing information		opt out	opt out			
social security number & gov't ID						
your activity on this site		opt out	opt out			
your location						

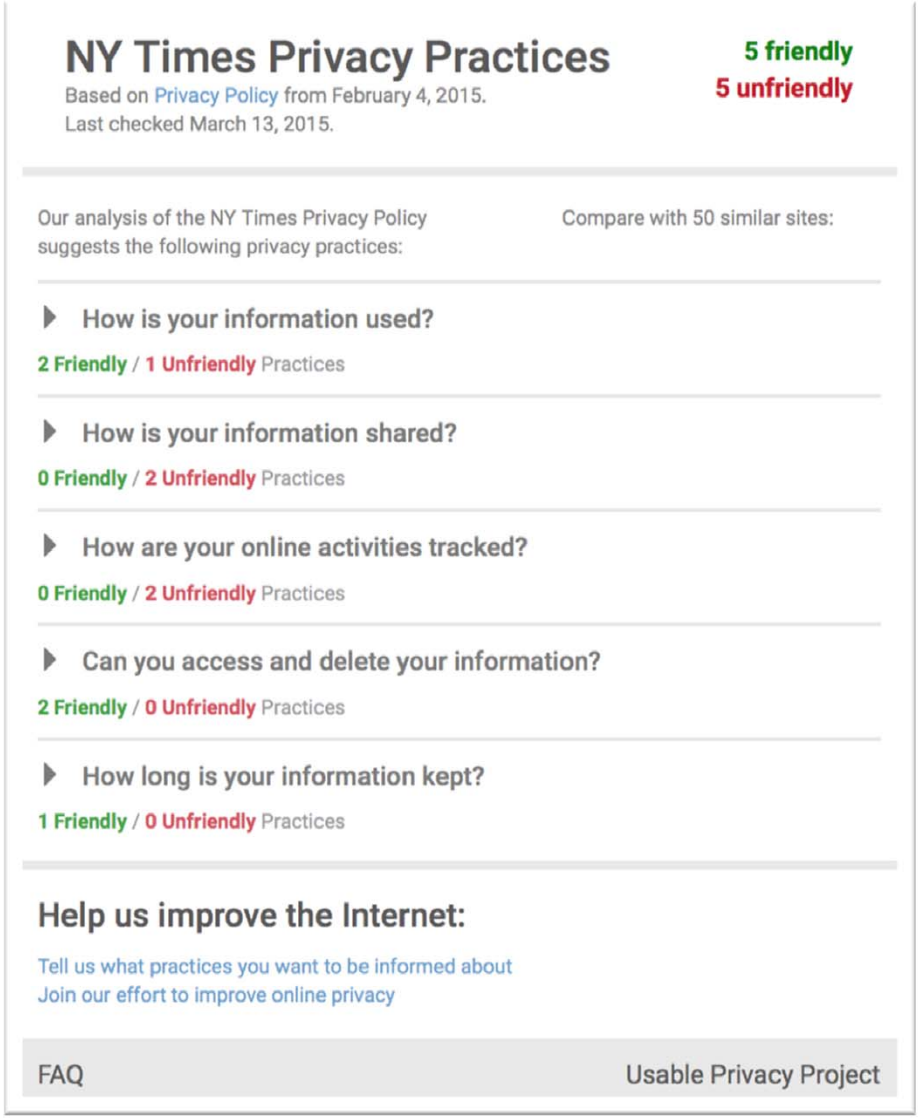
Access to your information
This site gives you access to your contact data and some of its other data identified with you

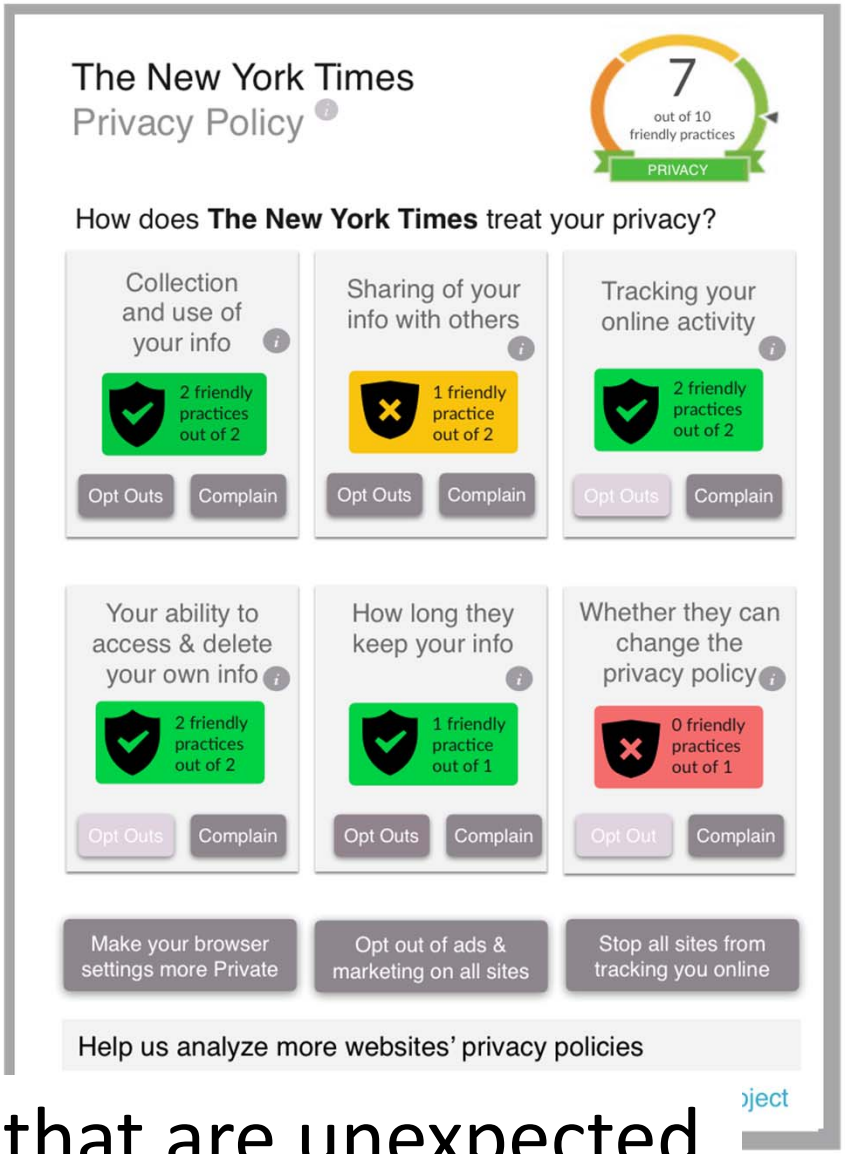
How to resolve privacy-related disputes with this site
Please email our customer service department

acme.com
5000 Forbes Avenue
Pittsburgh, PA 15213 United States
Phone: 800-555-5555
help@acme.com

Focus on expectations is complementary to visual formats

Kelley et al. A "Nutrition Label" for Privacy. SOUPS 2009.





Highlight/display practices that are unexpected

Screens designed by Leon et al, & Margaret and Kursat

RESEARCH QUESTIONS

Main Research Questions

- How do we define “expectation”?
- How do we measure expectations and mismatches in expectations?

DEFINING EXPECTATION

Types of Expectations

Research in non-privacy domains e.g. consumer psychology[1,2] shows that **users can have different types of expectations**

-e.g. “Desired,” “Minimally Tolerable”

Privacy research has not focused on multiple types of expectations that users can have

[1]Miller J. A. Studying satisfaction … Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction 1977
[2]Swan J. E. and Trawick I. F. Satisfaction related to predictive vs. desired expectations. Refining Concepts and Measures of Consumer Satisfaction and Complaining Behavior 1980

Expectations
(what is likely)

vs

Expectations
(what is desired)

Actual
practices

MEASURING EXPECTATIONS

Identifying Mismatched Expectations

- Present users with actual websites
- Ask participants what they **assume** the website does ("will" or likely expectation)
- Extract practices disclosed in website privacy policies
- **Compare people's expectations with actual practices & identify mismatches**

Organizing Websites & Participants

Website characteristic	
Type	Finance
	Health
	Dictionary
Popularity	More
	Less
Ownership	Private
	Government

User characteristic
Demographic: age, gender, education, occupation computer background, state of residence
Privacy protective behavior
Familiarity with privacy concepts and tools
Knowledge of privacy concepts and tools
Negative online experience
Online privacy concern
Experience with website: amount of recent use, has account, familiarity, trust

Data Practices Considered

Action	Scenario	Information type
Collection	With account	Contact
		Financial
		Health
		Current location
	Without account	Contact
		Financial
		Health
		Current location
Sharing	For core purpose	Contact
		Financial
		Health
		Current location
	For other purpose	Contact
		Financial
		Health
		Current location
Deletion	–	Personal data

Example Scenario Description

*"Imagine that you are browsing [website name] website. You **do not have a user account** on [website name], that is, you have not registered or created an account on the website"*

*"What is the **likelihood that [website name] would** collect your information in this scenario? …"*

		Likely	Somewhat likely	Somewhat unlikely	Unlikely
Collects your **Contact information**	Email address	○	○	○	○
	Postal address	○	○	○	○
	Phone number	○	○	○	○
	Other Please specify	○	○	○	○

Desired vs. Likelihood Expectation

"Do you think that [website] ***should or should not be allowed to collect*** *your information in this scenario? …"*

Vs.

"*What is the* ***likelihood that [website] would collect*** *your information in this scenario? …"*

Study Deployment

- Between-subjects study
  - Total 16 websites
  - Each participant randomly assigned to one website; 15 per website
  - Total 240 participants recruited from Mechanical Turk crowdsourcing platform
  - Study piloted via interviews and then deployed as online survey

Extracting Data Practices from Policies

- Annotation techniques
  - Manually using experts or crowd-workers
  - Semi-automatically by combination crowd-sourcing, machine learning and NLP e.g. Usable Privacy Policy project[1]
- Annotations indicate if a website is ***clear*** (Yes, engages; No, does not engage), ***unclear*** or ***does not address*** a data practice in it's policy

[1]Sadeh et al. The Usable Privacy Policy Project: Combining Crowdsourcing, Machine Learning and Natural Language Processing to Semi-Automatically Answer Those Privacy Questions Users Care About. Tech. report CMU-ISR-13-119. 2013

Different Types of Mismatches

Yes (website)	vs	**No** (website)
No (user)		**Yes** (user)

- Website shares data, but user doesn't think so
- user may use website and give up data unknowingly

- Website doesn't share data, but user thinks so
- user may not use website and lose utility

Different types of mismatches could impact user data privacy differently

RESULTS

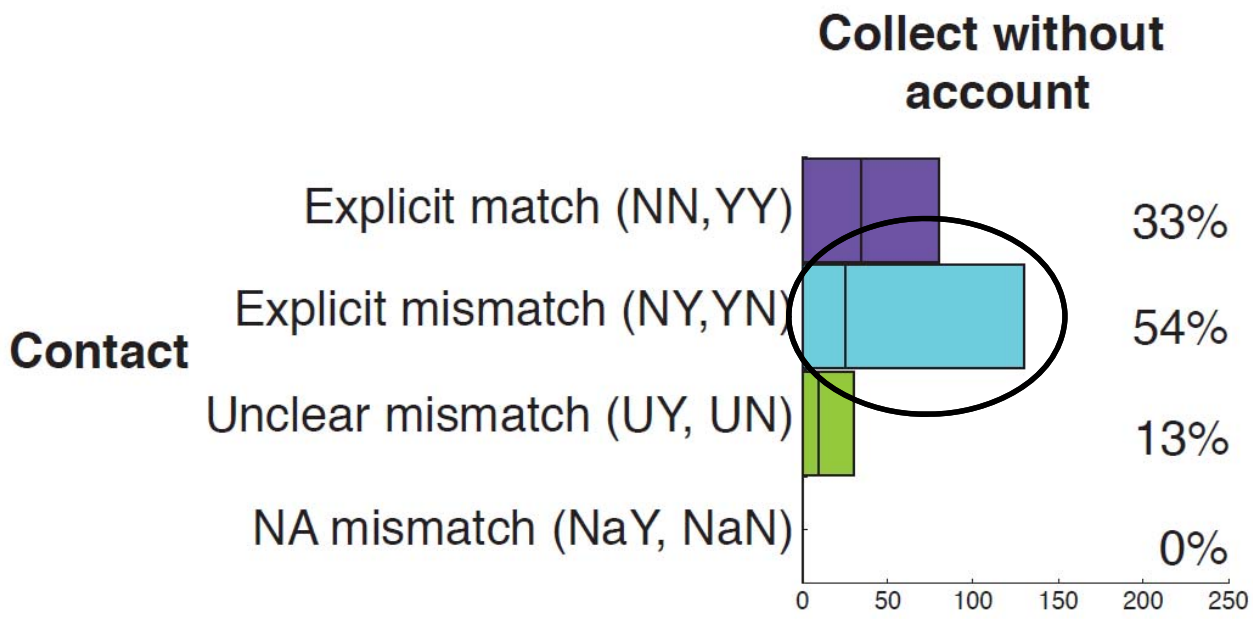
Impact of Website Characteristics

- Only **website type** had statistically significant impact on user expectations
  - Popularity and ownership did not
  - Type impacts expectations only for **financial and health information**, and not contact and current location information

Impact of User Characteristics

User characteristic (IV)	User expectation (DV)	Model R^2
Privacy knowledge	Collect health info without account	0.10
Privacy concern	Collect location info with account	0.13
	Share contact info for core purpose	0.09
	Share location info for core purpose	0.08
Age	Allow deletion	0.13
Trust in website	Share location info for core purpose	0.08
	Share financial info for other purpose	0.07
	Share health info for other purpose	0.05
	Allow deletion	0.13
Recent use	Collect location info with account	0.13
	Share contact info for core purpose	0.09
	Allow deletion	0.13

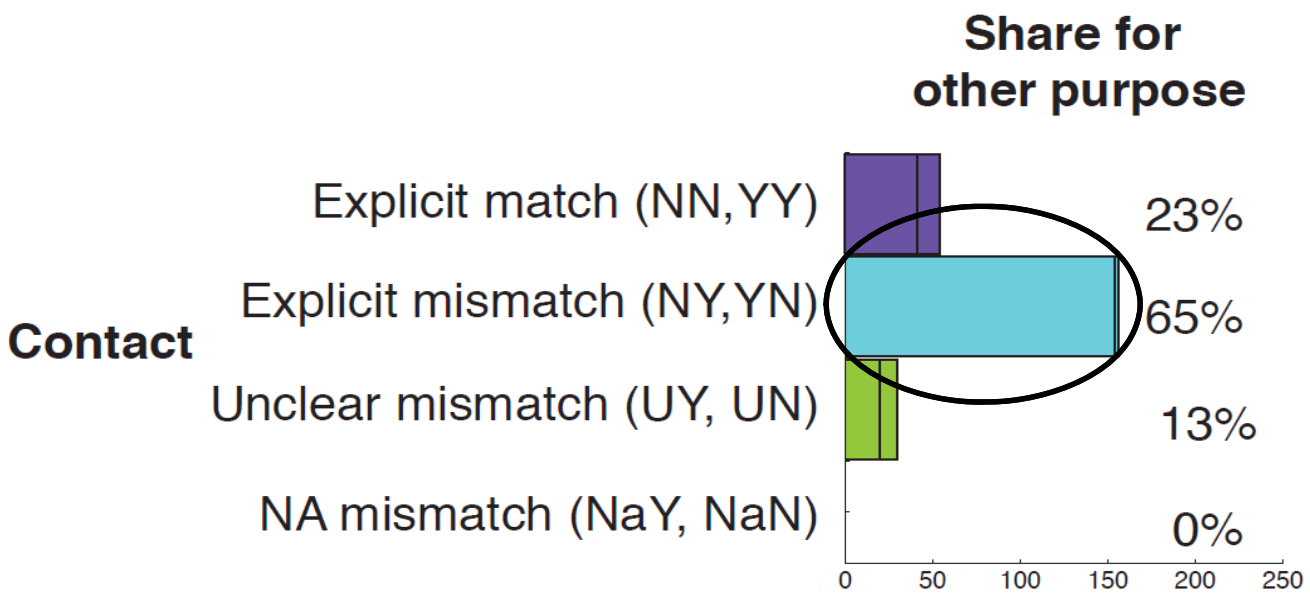
E.g. of Mismatch in Collection DP



Yes—No
(Website – user)
mismatch

Explicit match or mismatch occurs when website is clear about its data practice

E.g. of Mismatch in Sharing DP



No—Yes
(Website – user)
mismatch

Explicit match or mismatch occurs when website is clear about its data practice

Mismatches in Deletion DP

Deletion	% Users expect	% Websites permit
Yes -- full	**32%**	19%
Yes -- partial	**48%**	12%
No	20%	19%

Users expect websites to permit deletion, but websites do not

E.g. of Other Types of Mismatches

- Website specific mismatch
  - users do not expect banking websites to collect health information
  - Banking websites generally do not collect health information, but BankofAmerica website does

DISCUSSION

Potential for "Shorter" Privacy Notices

Display in notice	# practices	% reduction
All practices	17	-
Mismatched practices only	11	35%
Mismatched Yes—No practices only	5	**70%**

Potential reduction in information that users have to process for BOA privacy notice

Future Work

- Analyzing desires vs. likelihood vs. actual practices
- Consider additional data practices of interest to users e.g. tracking
- Test effectiveness of plug-in/notices that highlight mismatches

Heather Shoenberger

University of Oregon

Jasmine McNealy

University of Florida

Offline v. Online: Re-Examining the Reasonable Consumer Standard in the Digital Context

PRIVACYCON

FEDERAL TRADE COMMISSION

OFFLINE V. ONLINE:

RE-EXAMINING THE REASONABLE CONSUMER STANDARD IN THE DIGITAL CONTEXT (AN OVERVIEW)

Heather Shoenberger, University of Oregon

Jasmine McNealy, University of Florida

Methods

Interview

- 30 participants
- 20 women; 10 men;
- Average age: 26;
- 20 White, 5 Hispanic, 5 African-American

Survey

- 871 participants
- 415 men; 454 women;
- 35.9
- 4 Hispanic, 77 African-American, 657 White, 59 Asian, 19 Other.

- **Social Trust**: 6-item scale.
  - "How much do you trust the following institutions or persons in terms of how well they fulfill their responsibilities in collecting and handling consumer data collected online?"
  - The government; Individual advertisers. $\alpha = .86$
- **Control efficacy**:
  - 4-item scale.
  - "I can use online privacy tools to remain anonymous online." $\alpha = .61$

Main Dependent Variables:

- **Always Click Yes**: I always just click “yes” without reading terms of agreement (apps, websites) r=.83**

- **Privacy Concern**: 3-item scale. Data companies collect about me might be used in ways that make me feel uncomfortable $\alpha = .83$

What are consumers' privacy expectations online versus offline?

- Showing photos in person is more "intimate" than posting them online. "I would wait for a friendship to develop (offline) before showing any photos to someone in person."
  – *Interviewee*

- Significant differences between indicated sharing behaviors online and offline where sharing online was more likely.

Always clicking “yes” to digital terms of agreement without further investigation.

Race
***Age (-)**
Education
Gender
HHI

Social Trust
***Control Efficacy (+)**

***Negative Experience (-)**
Peer recommendations
***Convenience (+)**
***Site: poor aesthetics unfamiliar (-)**
***Presence of a policy (+)**

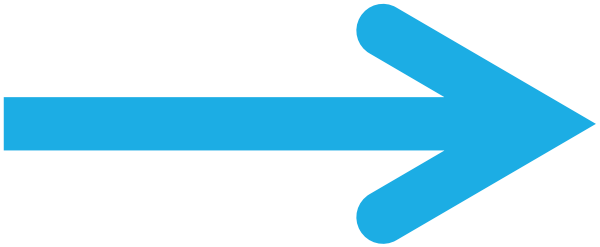
Always clicking “yes”

Privacy Concern

*Age (+)
*Education (+)
Gender
HHI
Race

*Social Trust (-)
*Control Efficacy (-)

*Negative Experience (+)
Peer recommendations
Convenience
* Site: Poor aesthetics, unfamiliar (+)
*Presence of a Policy (-)



Privacy concern

Average or “reasonable” consumer in the digital context

- Convenience and the cues of privacy policies and web design are the biggest predictors in our model for indication of actual behavior.
- If consumers are not reading the policies can there be meaningful control over their data?
- Lower social trust and the cues of privacy policies and web design are an important predictor of privacy concern.
- Focus on building trust to ensure the free flow of data in the digital context?

Suggestions for the FTC and industry in the digital context action items

- Guidelines for those who collect or use consumer data (advertisers, government, news organizations, etc.).
  - Adherence to “average consumer in the digital context’s” expectations of privacy based on type of data collected (photos, location, clickstream data, etc.).
  - Policies that are concise, readable and potentially “designed” for consumer approachability.

Conclusion/Future Research

- Pinpointing consumer expectations of privacy in different data collection scenarios.
- Further data gathering from a more diverse pool of consumers.
- Examine additional contextual variables (e.g., media reports).
- Testing “designed” policies for readable and understanding.
- Creating PSA to notify consumers of new and friendly policies.

Andelka Phillips

University of Oxford

Jan Charbonneau

Centre for Law & Genetics, Faculty of Law, University of Tasmania, Australia

Giving away more than your genome sequence?: Privacy in the Direct-to-Consumer Genetic Testing Space

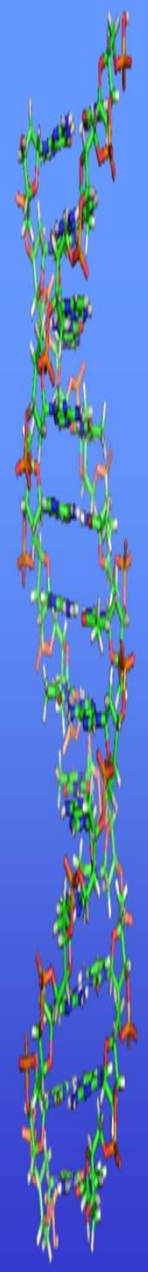
Giving away more than your genome sequence: Privacy in the Direct-to-Consumer Genetic Testing Space

Andelka Phillips, Faculty of Law, University of Oxford
Jan Charbonneau, Centre for Law & Genetics, University of Tasmania

UNIVERSITY OF OXFORD

UNIVERSITY of TASMANIA AUSTRALIA

Genetic Testing: Privacy Concerns



- Characteristics of genetic data
  - Most intimate of personal data: unique identifier of both an individual & their family groups
- Inherently identifiable
  - NOT possible to fully *de-identify* genetic data to make it impossible to *re-identify*
- Irrevocable
  - Once breached, it cannot be changed

Direct-to-Consumer Genetic Testing

- Traditional genetic testing
  - Occurs within each country's healthcare system
  - 'Patient' – enlivens professional/regulatory oversight & established legal duties of care
- Direct-to-consumer genetic testing
  - Commercial transaction
  - Occurs in the marketplace, typically online
  - 'Consumer' – enlivens consumer protection legislation & actions such as contract & negligence

General Public's View: Privacy & DTC

- Australia: GP or DTC?
  - Privacy concerns key constraint (also intention to biobank)
- 'Sharing' in the DTC space
  - Potential to extend beyond consumer-company
- Online panel of 3000 American, Australian and UK respondents (+ Japan)
  - 10% actual consumers; 90% potential consumers

Acknowledgement: DTC research funded by the Australian Research Council Discovery Grant
Personalised Medicine in the Age of Genomic Medicine DP11010069

Privacy & DTC

- Private = not shared; Shared = not private
- Privacy issues arise from sharing
  - Privacy = control over sharing
- Providing permission to share means individuals control personal genetic information
  - Permission = control over privacy

Privacy & DTC Engagement

- If consumers *believe* genetic data will only be shared with permission (*perceived* control)
  - More likely to purchase DTC tests
    - especially if have actually shared with family or online
  - Much more likely to participate in DTC research
    - initially permission-based (non-specific/enduring consent)
    - more likely to have actually shared & more likely to purchase

Sharers are Sharers

- More likely to share DTC results with family (not friends)

- More likely to share with doctors
  - DTC results for 'research, informational & educational use only' – not diagnosis
  - 'It would be 'a very brave' GP who relied on the results of a DTC gene test to manage a patient.' Prof Suther, RCPA

- More likely to share in online health communities & with genetic counselors

Does *perceived* control = *actual* control?

- DTC is a commercial transaction
  - Governed by contracts, terms of service & privacy policies (same for online interpretation & sharing sites)

- Australian DTC companies & their privacy policies
  - Privacy policies do NOT comply with *Privacy Act 1988* (Cth) or *Enhanced Privacy Protection Act* (in force 2014)

Click Here Now:
DTC Contracts & Privacy Policies

- Study examined DTC contracts and privacy policies of companies providing tests for health purposes

- These govern:
  - Purchase of genetic tests
  - Use of DTC websites
  - Participation in DTC research

Contracting Online & Consumer Behavior

- When active online we often have ‘inattentional blindness’
- Consumers may not realise they are entering into a contract
- This is particularly relevant to both wrap contracts and privacy policies
  - Consumers often may not even notice, let alone read them

Privacy Risks

- Sharing or sale of sequenced genetic data
- Sharing or sale of other types of personal data
- Possible discrimination on the basis of an individual's genetic makeup

More Privacy Risks

- There is potential for hacking of genetic databases for purposes of:
  - Identity theft
  - Targeted marketing (e.g. pharmaceutical drugs)
  - Discrimination in insurance or employment
  - More remotely, the creation of synthetic DNA

DTC Contracts & Privacy Policies

- Often contracts and privacy policies are not industry specific
- Contracts online more generally often use very similar wording
- Several terms commonly included might be deemed unfair and unenforceable under UK and European Union law

Common Terms

- Consent or agreement with terms **OFTEN DEEMED** through use or viewing of the website or use of services

- Clauses allowing unilateral alteration of terms without notice to consumers

  - Companies could make significant changes to policies on use, storage, sharing, & sale of data without telling consumers.

Significant Clauses

- Clauses stating services are provided for ‘research, informational and educational use only’ &/or ‘recreation’
- Clauses stating company may share data with law enforcement
- Clauses stating company can share with third parties

Need For Reform Of Contracts & Privacy Policies

- Contracts and privacy policies should be drafted so that they can
  - Be easily understood by the consumer
  - Allow for consumers to make informed decisions & have control over their data
    - e.g. could include more opt-ins for specific uses of data
  - Consent should not be deemed through visiting a website

Thank you!

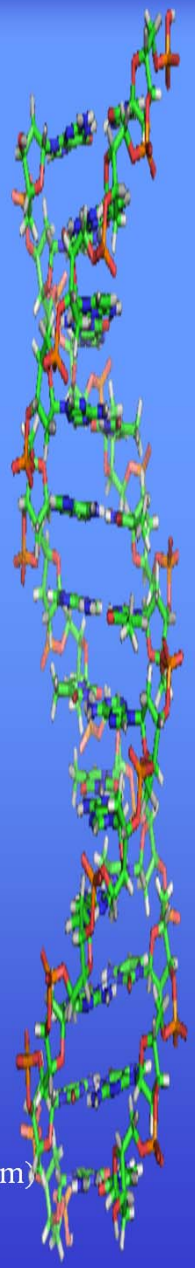
For further information, please contact

Andelka Phillips

andelka.phillips@law.ox.ac.uk

Jan Charbonneau

jan.charbonneau@utas.edu.au



Graphics: DNA available at Google Images, origin not attributed

‘Do Not Access’, ‘Poking holes in genetic privacy’, 16 June 2013 (wwwnhealthtran.com)

‘Private/public key’, ‘Should I get 23andMe DNA Analysis?’, 27 September 2015 (www.hubpages.com)

‘Confidential DNA’, www.councilforresponsiblegenetics.org

‘Privacy policies’, www.runtosucceed.com

Further reading: Genes & Privacy

- Andelka Phillips, ‘Only a Click Away – DTC Genetics for Ancestry, Health, Love… and More?’, *Applied & Translational Genomics* (forthcoming 2016).
- Andelka Phillips & Jan Charbonneau, ‘From the lab to the market’, *Gene Watch* 28(3), 2015.
- Andelka Phillips, ‘Direct-to-consumer genetic tests – more than just a question of health’, *BioNews*, December 2015.
- Andelka Phillips, ‘Genomic Privacy and Direct-to-Consumer Genetics – Big Consumer Genetic Data – What’s in that Contract?’, GenoPri’15 - The 2nd Workshop on Genome Privacy and Security, www.genopri.org/program.html.
- Andelka Phillips, ‘Think Before You Click – Ordering a Genetic Test Online’, *The SciTech Lawyer* 11(2), Winter 2015.

- Christine Critchley, Dianne Nicol & Rebecca McWhirter, ‘Identifying public expectations of genetic biobanks’, *Public Understanding of Science* (forthcoming 2016).

- Margaret Otlowski, ‘Disclosing genetic information to at-risk relatives: new Australian privacy principles, but uniformity still elusive’, *Medical Journal of Australia*, 2015.

- Dianne Nicol, Meredith Hager, Nola Ries & Johnathon Liddicoat, ‘Time to get serious about privacy policies: The special case of genetic privacy’, *Federal Law Review*, 2014.

- Christine Critchley, Dianne Nicol, Margaret Otlowski & Don Chalmers, ‘Public reaction to direct-to-consumer online genetic tests: Comparing attitudes, trust and intentions across commercial and conventional providers’, *Public Understanding of Science*, 2014.

- Dianne Nicol & Meredith Hager, ‘Direct-to-consumer genetic testing – a regulatory nightmare?’, *Medical Journal of Australia* May 2013.

Discussion of Session 2

Discussants:

- **Kristen Anderson,** Federal Trade Commission
- **Alan McQuinn,** Information Technology and Innovation Foundation
- **Darren Stevenson,** University of Michigan and Stanford Law School

Presenters:

- **Serge Egelman,** ICSI/University of California, Berkeley
- **Ashwini Rao,** Carnegie Mellon University
- **Heather Shoenberger,** University of Oregon **& Jasmine McNealy,** University of Florida
- **Andelka M. Phillips,** University of Oxford **& Jan Charbonneau,** University of Tasmania

PRIVACY CON
FEDERAL TRADE COMMISSION