# The Usable Privacy Policy Project

N. Sadeh\*, A. Acquisti\*, L.F. Cranor\*, T.D. Breaux\*, J.R. Reidenberg<sup>†</sup>, N. Smith<sup>‡</sup>, B. van Schewick<sup>§</sup> \*Carnegie Mellon University, †Fordham University, ‡University of Washington, §Stanford University

https://www.usableprivacy.org/



Norman Sadeh (Lead PI)

## Motivation

## **Notice & choice**

- Websites' privacy policies are the de facto standard for notice & choice online
- Privacy policies provide notice about data practices, e.g., data collection, use, sharing, retention, user access

## **Problem: No one reads** privacy policies

They are oftentimes long and complex

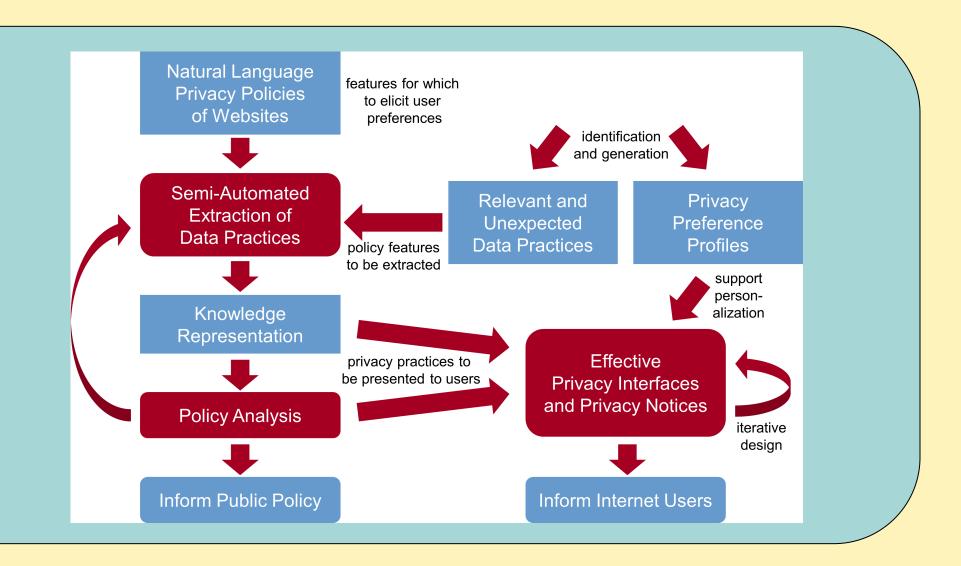
NY Times nytimes.com		D Take a tour
Privacy Practices	Privacy Policy	
Click a category to explore individual practice states	Image: NY Times Privacy Policy from Jun 10, 2015         Reading Level: College (G           107         241 privacy practice statements in total	rade 13)
Third Party Sharing/Collection @	3 Device Information. We may collect non-personal information about the	
User Choice/Control 📀	<ul> <li>computer, mobile device or other device you use to access the NYT Service such as IP address, geolocation information, unique device identifiers, b</li> </ul>	ces,
User Access, Edit and Deletion 🕜	3 type, browser language and other transactional information.	
Data Security 🕜	Cookies, Beacons, Local Storage and Other Similar Technologies. We use           3         "cookies," Web beacons, HTML5 local storage and other similar technologies.	
Data Retention 🕜	<ul> <li>These technologies allow us to manage access to and use of the Services,</li> <li>recognize you and provide personalization, and help us understand how</li> </ul>	people
International and Specific Audiences 🕜	<ul> <li>use the NYT Services. You will not be able to access certain areas of our websites, including NYTimes.com, if your computer does not accept cool from us. We do not respond to browser-based "do not track" signals. For</li> </ul>	
Do Not Track 🛛	<ul> <li>detailed information about our use of cookies and local storage, and how to manage Frequently Asked Questions About Cookies and Similar Technologies in our Help se</li> </ul>	them, see
Policy Change 🕜	We may transmit non-personally identifiable website usage information	
Wensite Categories	other sites. For more information about our third-party ad server, or to	
Website Categories Click a category to browse other sites within that cat Reference Regional Science Society Spo Arts Computers Health Kids and Teens	tegory. your choices about not having this non-personal information used to tar to you, please click here.	get ads r and t is

- Few users read them
- Efforts to improve notice & choice  $\bullet$ lack industry support (e.g., P3P and DNT are stalled)

	Identifiability	not-selected -
nformation You Give Us: We receive and store any information you enter on our Web ite or give us in any other way. Click here to see examples of what we collect. You	Personal Information Type *	Generic personal information -
can choose not to provide certain information, but then you might not be able to take advantage of many of our features. We use the information that you provide for such	Purpose *	Personalization/Customization -
purposes as responding to your requests, <b>customizing future browsing for you</b> , improving our site, and communicating with you.	User Type	not-selected -
improving our exe, and communicating with you.	Choice Type	Don't use service/feature -
lease write your comments for this paragraph	Choice Scope	not-selected -
	<ul> <li>References another place in the policy</li> </ul>	
	Save	
Practices of this paragraph		
First Party Collection/Use		
Does Explicit Collect on website not-selected Generic personal information Basic service/feature not-selected Don't use service/feature not-selected Clone Delete		
Third Party Sharing/Collection		
User Choice/Control		
Liser Access Edit and Deletion		

### Approach

- **Crowdsourcing, Machine Learning and NLP to semi-**automatically annotate privacy poliies
- **Privacy policy analysis**
- **Effective privacy user interfaces incl. privacy** preference modeling
- **Comparison of privacy policy content to** actually occurring practices)



#### Extract data practices from privacy policies

- Combining crowdsourcing, NLP & ML
- Fine-grained annotation scheme

#### Effective privacy user interfaces

- Analyze usability issues
- Provide relevant
- Actionable information & choices
- Support privacy ullet

#### Privacy preference modeling

- Focus on relevant and unexpected practices
- Understand cognitive biases
- Create privacy preference profiles to support personalization

#### Privacy policy analysis

- Inconsistencies in policies
- Ambiguity and vagueness in policy
- Relative and sectoral comparison

information

decision making

## Comparison of privacy policy content to actually occurring practices Analyze potential non-compliance

- Predict from metadata (e.g., from app category)  $\bullet$
- Large-scale through automation (e.g. 18,000 mobile apps and work with Cal AG)

#### **Cooperation with Stakeholders**

- Develop tools for regulators and tech providers
- Help software developers address potential privacy issues before they become prevalent

Interested in meeting the PIs? Attach post-it note below!



National Science Foundation HERE DISCOVERIES BEGIN

NSF Secure and Trustworthy Cyberspace Inaugural Principal Investigator Meeting Nov. 27 - 29<sup>th</sup> 2012 National Harbor, MD

