# 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/



N. 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

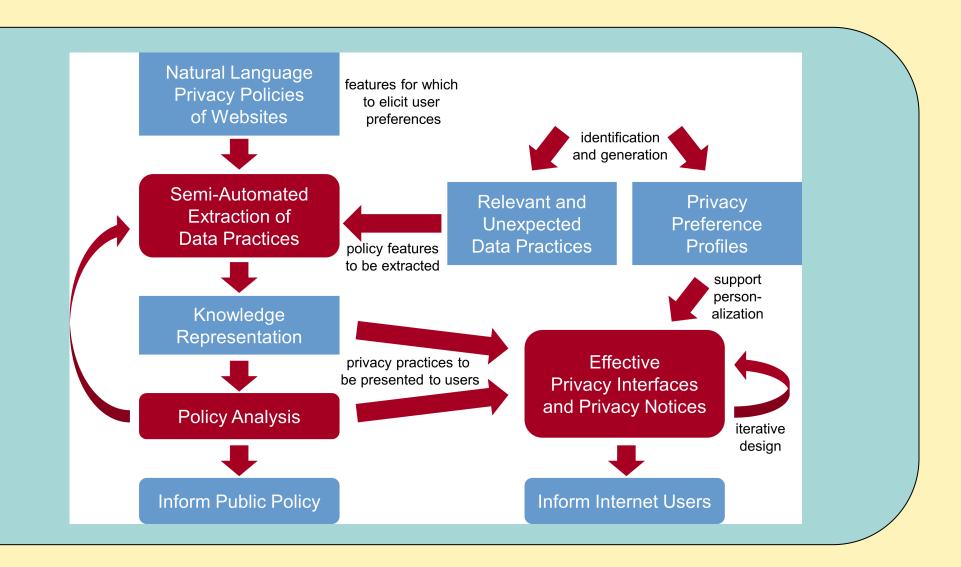
det actegues plot enditedue protecte statementsImage: Collection (Use )Image: Collection (Use ) <th>Privacy Practices</th> <th>Privacy Policy</th>	Privacy Practices	Privacy Policy
First Party Collection (Use )       Image: Additional and Specific Additions (Image: Additional and Other Similar Technologies. We use the Services, see the NTS Services, see the sec the sec transmitter the sec the services, see the sec transmitter the sec the sec transmitter the sec transmitter the sec transmitter the sec the sec transmitter the sec transmitter the sec the sec transmiter the sec the sec transmitter the sec the se		rements.
Upre Choice/Control I Information we finally collect fuols-personal information about the Strekes, specific control Information unique device identifiers, browser, we day and the transactional information and use of the Services, the strekes is proceeding and other similar technologies. These technologies allow us to manage access to and use of the Services, these technologies allow us to manage access to and use of the Services, these technologies allow us to manage access to and use of the Services, these technologies allow us to manage access to and use of the Services, these technologies allow us to manage access to and use of the Services, these technologies allow us to manage access to and use of the Services, these technologies allow us to manage access to and use of the Services, these technologies allow us to manage access to and use of the Services, these technologies allow us to manage access to and use of the Services, these technologies allow us to manage access to and use of the Services, these technologies allow us to manage access to and use of the Services, these technologies allow us to manage access to and use of the Services, these technologies allow us to manage access to and use of the Services, these technologies allow use to accept cookies from us. We do not respond to browser-based do not tracked information and the use of the Services, the use to manage access to and accept cookies from use. We do not respond to browser-based do not tracked information and use of the Services, the use to access to and accept cookies from use to accept cookies from use to accept cookies from use to accept to aclow the set we then the use of the NT Services, the use to the set to a technologies in our Help section.         Werker K & Werker K	First Party Collection/Use 📀	
Uger Choice/Control ©       Image: Choice/Control ©         User Access, Edit and Deletion ©       Image: Choice/Control ©         Data Security ©       Image: Choice/Control ©         Data Security ©       Image: Choice/Control ©         Data Retention ©       Image: Choice/Control ©         Do Not Track ©       Image: Choice/Control ©         Policy Change ©       Image: Choice/Control ©         Web Access (Edit and Deletion ©       Image: Choice/Control ©         Mark Track ©       Image: Choice/Control ©         Policy Change ©       Image: Choice/Control ©         Web Acatego to burves the stervice in the stervice in the NYT Stervice in the N	Third Party Sharing/Collection @	Device Information. We may collect non-personal information about the
War Access, Edit and Deletion Image Access, Edit Access, Edit and Deletion Image Access Accessing Image Access Accessing Image Access Accessing Image Accesses, Image Access, Image Access, Image Access, Image Access, Image A	User Choice/Control 🕜	such as IP address, geolocation information, unique device identifiers, browser
Data Security ①       ①         Data Security ②       ①         Data Retention ②       ②         International and Specific Audiences ②       ②         Do Not Track ②       ③         Policy Change ③       ③         Policy Change ③       ④         Verbasites, Including INTTimes.com, if your computer does not accept cookies from us. We do not respond to browser-based "do not track" signals. For more detailed information about our use of cookies and local storage, and how to manage them, see Frequently Asked Questions About Cookies and local storage, and how to manage them, see Frequently Asked Questions About Cookies and local storage, and how to manage them, see the NUTS services. You will not be about the substitution bout our use of cookies and local storage, and how to manage them, see the auteout to about our use of cookies and local storage, and how to manage them, see the auteout to about our use of cookies and local storage, and how to manage them, see the auteout to about our use of the New York Times when you visit does the sites. For more detailed information about our use of the New York Times when you visit does the sites. For more information mature them so to target about to target about not having this non-personal information automatically and store it in log files. This information mature three your computer or device is accessing the Internet), howoser type, operating system and other usage information about the use of the NYT Services, including a history of the pages you view.	User Access, Edit and Deletion 🚱	
Data Retention @       •         International and Specific Audiences @       •         Do Not Track @       •         Policy Change @       •         WebSite Categoories (requently Asked Questions About Cookies and Similar Technologies in our Fielp section.       •         WebSite Since Society Sports Are Computer Mean Street Technologies in our Fielp section.       •         Multice Categoories (requently Asked Questions About Cookies and Similar Technologies in our Fielp section.       •         WebSite Since Society Sports Are Computer Mean Street Technologies in our Fielp section.       •         Multice Categoories (requently Asked Questions About Cookies and Similar Technologies in our Fielp section.       •         WebSite Since Society Sports Are Computer Mean Street Field Michael Street Str	Data Security 🕢	3 "cookies," Web beacons, HTML5 local storage and other similar technologies.
International and Specific Audiences ( )       ( )         Do Not Track ( )       ( )         Do Not Track ( )       ( )         Policy Change ( )       ( )         Webstress, including NYTimes.com, if your computer does not accept eookies       ( )         Policy Change ( )       ( )         Vebstress, functional boot our use of cookies and local storage, and how to manage them, see Frequently Asked Questions About Cookies and local storage, and how to manage them, see Frequently Asked Questions About Cookies and Jonat Track ( )         Keis category to browse other sites with a category.       We may transmit non-personally identifiable website usage information to third parties in order to show you advertising for The New York Times when you visit on choices about not having this non-personal information used to target ads to you, please click here.         Margin Tage ( )       ( )	Data Retention 🕑	2 recognize you and provide personalization, and help us understand how people
Do Not Track        Image: Contract Contrect Contract Contract Contract Contract Contract Conternation Cont	International and Specific Audiences 🕢	2 websites, including NYTimes.com, if your computer does not accept cookies
Policy Change <ul> <li>Policy Change </li> <li>We may transmit non-personally identifiable website usage information to third party ad server, or to learn your choices about not having this non-personal information used to target ads to you, please click here.</li> <li>Mandytics, Log Files and Reading History</li> <li>As is true of most web sites, we gather certain information automatically and store it in log files. This information may include Internet protocol (IP) addresses (the region or general location where your computer or device is accessing the Internet), browser type, operating system and other usage information about the use of the NYT Services, including a history of the pages you view.</li> </ul>	Do Not Track 🕢	detailed information about our use of cookies and local storage, and how to manage them, see
Website Categories       parties in order to show you advertising for The New York Times when you visit other sites. For more information about our third-party ad server, or to learn your choices about not having this non-personal information used to target ads to you, please click here.         Regional Science Society Sports       Stores         Arts Computer Health Kids and Teem News       News         As is true of most web sites, we gather certain information automatically and store it in log files. This information may include Internet protocol (IP) addresses (the region or general location where your computer or device is another information about the use of the NYT Services, including a history of the pages you view.         usableprivacy       User Profile       Task       Visualize       Settings       Logout	Policy Change 🙆	
usable <b>privacy</b> User Profile Task Visualize Settings Logout	Click a category to browse other sites within that cat Reference Regional Science Society Spor	category.       to you, please click here.         ports       Analytics, Log Files and Reading History         News       As is true of most web sites, we gather certain information automatically and store it in log files. This information may include Internet protocol (IP) addresses (the region or general location where your computer or device is accessing the Internet), browser type, operating system and other usage information about the use of the NYT Services, including a history of the pages
ment Policy: www.indu.com-privacypolicy-03-2014.csv		addresses (the region or general location where your computer or device is         accessing the Internet), browser type, operating system and other usage         information about the use of the NYT Services, including a history of the pages
First Party Collection/Use	urrent Policy: www.imdb.com-privacypolicy-0	-05-2014.csv

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

	Identifiability	not-selected 🗸
Information You Give Us: We receive and store any information you enter on our Web site 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 are, and communicating with you.	Choice Type	Don't use service/feature -
lease write your comments for this paragraph	Choice Scope	not-selected 🚽
	References another place in the policy	
	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

