Increasing the Salience of Data Use Opt-outs Online

Namita Nisal
University of Michigan
Ann Arbor, MI, USA
namitan@umich.edu

Florian Schaub
University of Michigan
Ann Arbor, MI, USA
fschaub@umich.edu

Sushain K. Cherivirala
Carnegie Mellon University
Pittsburgh, PA, USA
sushain@cs.cmu.edu

Shomir Wilson
University of Cincinnati
Cincinnati, OH, USA
shomir.wilson@uc.edu

Kanthashree M. Sathyendra
Carnegie Mellon University
Pittsburgh, PA, USA
ksathyen@andrew.cmu.edu

Lorrie Faith Cranor
Carnegie Mellon University
Pittsburgh, PA, USA
lorrie@cs.cmu.edu

Margaret Hagan
Stanford University
Palo Alto, CA, USA
mdhagan@stanford.edu

Norman Sadeh
Carnegie Mellon University
Pittsburgh, PA, USA
sadeh@cs.cmu.edu

Copyright is held by the author/owner. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee. Poster presented at the 13th Symposium on Usable Privacy and Security (SOUPS 2017).

Abstract
Prior work has shown that many individuals are concerned about their privacy but feel resigned with respect to their abilities to manage or protect their privacy. At the same time, individuals may not be aware of privacy choices that are available to them. For instance, many privacy policies contain links that allow users to opt out of certain data use practices, such as the use or sharing of contact information for marketing purposes. We propose a method to make such privacy choices more salient to Internet users. Based on an approach to automatically identify and classify opt-out choices in privacy policies, we designed a browser extension that notifies users about available opt-out choices on the websites they visit. The goal of our work is to not only make users aware of available choices but also make it easier for them to utilize those choices.

Author Keywords
Privacy; usability; privacy policy; opt-out.

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous.

Motivation

Surveys and experiments frequently show that many individuals are concerned about their privacy, but feel helpless and resigned with respect to how they could better protect and manage their privacy online [10, 1]. Most websites, apps and services have privacy policies that inform users about an entity's data collection, use and sharing practices, as well as certain privacy choices, for example the ability to opt-out of the use of contact information for marketing purposes or targeted advertising. However, website privacy policies are long, verbose documents that are often difficult to understand [3, 2, 8]. They are frequently not written for users but rather serve to demonstrate compliance with legal or regulatory notice requirements [8]. Furthermore, reading all privacy policies of websites, app and services one encounters is impractical [5]. As a result, individuals are typically not aware of the privacy choices they could take with respect to a given service and generally struggle to opt-out [4].

As part of the Usable Privacy Policy Project [6],[1] our research focuses on automatically analyzing privacy policies at scale in order to facilitate the provisioning of user-friendly privacy notices and controls [9]. We developed a browser extension that makes opt-out choices described in privacy policies salient for users by leveraging an approach for automatically identifying and extracting privacy choices from privacy policies [7]. Subsequently, we provide an overview of the opt-out choice extraction approach and the browser extension's user experience design. We also discuss further directions for future research.

Opt-out Choice Extraction from Privacy Policies

We treated the problem of extracting choice instances as a binary classification problem where we labeled sentences from a privacy policy as containing a choice instance (positive) or not (negative) based on the OPP-115 corpus of annotated privacy policies [11]. We focus on extracting opt-out choices with hyperlinks from privacy policies, because users can directly act on those choices.

Leveraging multiple features, such as stemmed unigrams and bigrams, relative location in document, and opt-out specific phrases, we used logistic regression classification to identify opt-out choice instances [7]. In a second classification step, we used additional features, such as hyperlink composition, anchor text and similarity between privacy policy and opt-out URL, to develop more fine-grained classifiers that distinguish between different types of opt-outs. For training, we annotated a set of 125 positive instances to assign two additional labels to each of them; these were *Party Offering Choice* (i.e., first party or third party) and *Purpose* (e.g., advertisement, data sharing, communications, analytics or cookies).

Evaluation of our classifiers indicates high accuracy and robustness for opt-out hyperlink identification and classification, with F-1 scores of 0.947 and 0.940 for the two most frequent opt-out link types (opt-outs for first party communication and third party advertising, respectively).

Browser Extension UX Design

We designed a browser extension to present opt-out choices extracted from a website's privacy policy in user-friendly manner that helps individuals understand and act on available opt-out choices. The challenges were to keep the user engaged with the plugin and use simple language that each user would understand, while also reducing the risk of habituation to the browser extension.

Figure 1 shows the browser extension's main panel opened on a newly loaded website. The main panel is displayed

[1] https://www.usableprivacy.org

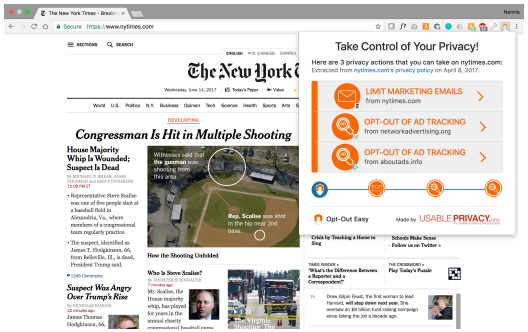


Figure 1: Browser extension showing two available opt-out choices extracted from the website's privacy policy.

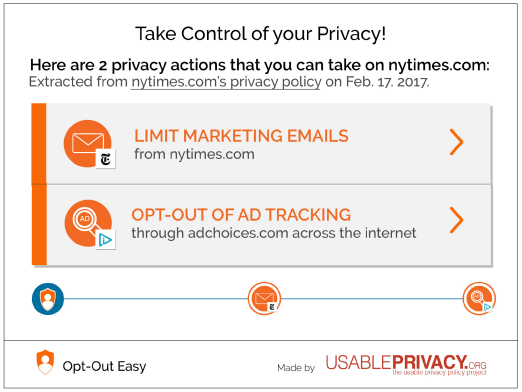


Figure 2: Initial view of the browser extension. Newly extracted opt-out choices are shown in orange.



Figure 3: The user has interacted with both opt-outs already. The extension also displays an overview of websites for which the user has interacted with opt-out choices.

when the user clicks on the extension's icon in the browser. The browser extension lists the opt-out choices for the current website as extracted from the website's privacy policy. All newly extracted choices are shown in orange (see Figure 2). The browser extension uses favicons and text to indicate whether a certain opt-out pertains to the first party, i.e., the website, or a third party, such as a targeted advertising opt-out offered by the advertisement industry. Note, however, that both opt-outs were extracted from the first-party website's privacy policy. Clicking on a choice option takes the user directly to the respective opt-out page, which opens in a new tab.

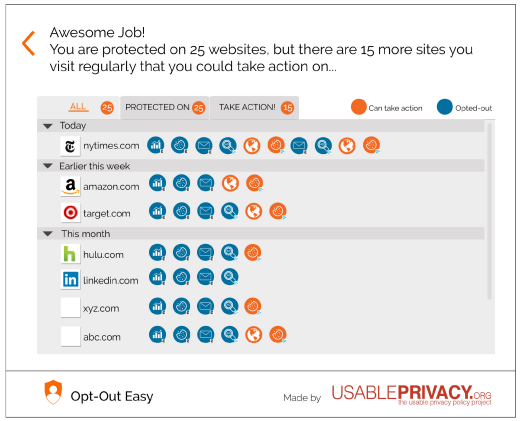


Figure 4: Detailed overview of available and acted on opt-outs for recently visited websites.

Clicking on an opt-out option changes the color of that option, and a timestamp indicates the date visited, as shown in Figure 3. The timeline below the opt-out choices suggests a sequence of actions the user might take and encourages the user to engage with multiple choices. Once the user has interacted with websites' opt-outs, the extension displays an overview grid of those websites to give the user a sense of their level of engagement with available choices. Clicking on the overview grid opens a detailed overview of available and acted on opt-outs for recently visited websites (see Figure 4).

Future Work

The browser extension has been implemented as a Chrome extension. We are currently evaluating the browser extension with users, to further refine it in preparation for a public release. We are also exploring opportunities for maintaining user engagement, for instance by gamifying the opt-out experience. Ultimately, we plan to extend this approach into a solution that helps users keep track of and manage their privacy choices online.

Acknowledgements

This research has been funded by the National Science Foundation under grant agreement CNS-1330596.

REFERENCES

1. Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. 2015. Privacy and human behavior in the age of information. *Science* 347, 6221 (2015), 509–514.
2. Fred H Cate. 2010. The limits of notice and choice. *IEEE Security & Privacy* 8, 2 (2010), 59–62.
3. Lorrie Faith Cranor. 2012. Necessary but not sufficient: Standardized mechanisms for privacy notice and choice. *J. on Telecomm. & High Tech. L.* 10 (2012), 273.
4. Pedro Leon, Blase Ur, Richard Shay, Yang Wang, Rebecca Balebako, and Lorrie Cranor. 2012. Why Johnny can't opt out: a usability evaluation of tools to limit online behavioral advertising. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 589–598.
5. Aleecia M. McDonald and Lorrie F. Cranor. 2008. The Cost of Reading Privacy Policies. *I/S: A Journal of Law and Policy for the Information Society* 4, 3 (2008), 540–565.
6. Norman Sadeh, Alessandro Acquisti, Travis Breaux, Lorrie Cranor, Aleecia McDonald, Joel Reidenberg, Noah Smith, Fei Liu, Cameron Russel, Florian Schaub, and others. 2013. *The Usable Privacy Policy Project:*

Combining Crowdsourcing, Machine Learning and Natural Language Processing to Semi-Automatically Answer Those Privacy Questions Users Care About. Technical Report. Technical Report CMU-ISR-13-119, Carnegie Mellon University.

7. Kanthashree Mysore Sathyendra, Florian Schaub, Shomir Wilson, and Norman Sadeh. 2016. Automatic Extraction of Opt-Out Choices from Privacy Policies. In *2016 AAAI Fall Symposium Series*.
8. F. Schaub, R. Balebako, and L. F. Cranor. 2017. Designing Effective Privacy Notices and Controls. *IEEE Internet Computing* 21, 3 (May 2017), 70–77.
9. Florian Schaub, Rebecca Balebako, Adam L Durity, and Lorrie Faith Cranor. 2015. A Design Space for Effective Privacy Notices. In *SOUPS '15: Symposium on Usable Privacy and Security*. USENIX.
10. J. Turow, M. Hennessy, and N. Draper. 2015. *The tradeoff fallacy: How marketers are misrepresenting american consumers and opening them up to exploitation.* Technical Report. Annenberg School for Communication, University of Pennsylvania.
11. 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, and N Sadeh. 2016. The Creation and Analysis of a Website Privacy Policy Corpus. In *Annual Meeting of the Association for Computational Linguistics, Aug 2016*. ACL.