Toward Privacy Standards Based on Empirical Studies

Serge Egelman and Erika McCallister
National Institute of Standards and Technology
100 Bureau Drive
Gaithersburg, MD 20899
{serge.egelman, erika.mccallister}@nist.gov

Abstract
In this paper, we argue that if privacy standards are created to guide "do-not-track" technologies, these standards should be created with the primary stakeholder in mind: the data subject. Previous privacy and security standards have been unsuccessful because implementations were inconsistent, confusing, or not readily apparent to the user. The Fair Information Practice Principles (FIPPs) empower users to make informed decisions about their privacy and should be the basis for any resulting privacy standard. However, research must be conducted to determine best practices for presenting this information to users. We describe one such study that we are currently conducting and what we expect to learn about promoting informed consent with regard to data sharing.

Keywords
Privacy standards, empirical studies, informed consent

Introduction
The US Department of Commerce recently released its Privacy Green Paper [17], which made recommendations for the future of Internet privacy. The Department sought to balance consumer trust with commercial innovation. Among its recommendations was the idea of using enforceable codes of conduct based on a set of Fair Information Practice Principles (FIPPs). The principles of transparency and individual participation are directly relevant to the concept of do-not-track. Transparency means that organizations should notify individuals regarding the collection, use, dissemination, and maintenance of their personal information. Individual participation means that organizations should provide means of consent to individuals regarding the collection, use, dissemination, and maintenance of their personal information, as well as provide a means for access to and correction of personal information. Users must know what information is collected and how it is used to make good decisions about when to use a do-not-track technology.

Background
Studies of user perceptions of privacy have found that Internet users are generally concerned about their privacy when online [1]. However, studies have found that in practice their actions do not reflect their preferences [2, 3, 4]. Part of the problem is that users are often unaware of the types and amounts of information that are shared with affiliated websites, which websites are affiliated, or how they can opt-out of having their information shared [13]. However,

when given effective privacy tools with which they can state their privacy preferences, observed behaviors become better aligned with stated preferences [8, 16]. Such privacy tools are effective because the user is at the center of the design and empirical data on user behavior informs the design decisions [20].

Many current Internet privacy tools do not adequately allow users to make informed decisions because the standards on which they are based have not incorporated empirical data on how to best support users' needs. For instance, studies have shown that many web browser security indicators go unnoticed because the indicators are outside the user's view [18, 19], are inconsistent across vendors and versions [14], and have unclear meanings [9].

The W3C's P3P standard attempted to empower users to make informed privacy decisions by specifying a format in which websites could post machine-readable privacy policies [5], but left it up to software vendors to determine what information to display to users and how it was to be displayed. Despite research showing P3P adoption rates of over 25% on popular websites [7, 10], use of full P3P policies failed to gain traction.[1] This may be due in part to browser-based P3P implementations that were hard to understand and often went unnoticed [6]. The onus of this failure is not necessarily the fault of software developers or designers. P3P is a comprehensive standard in its focus on converting natural language privacy policies into a machine-readable format. However, P3P lacks what has proven to be essential guidance on parsing this detailed information in a way that will allow users to take action. Research has now shown that P3P implementations could be designed to help users make more informed choices (e.g., [8, 11]), but it is likely too late to update the standard at this point, and seems unlikely to gain sufficient traction in the future.

In order to be successful, technical standards used to assist in the implementation of the FIPPs must be objective and based on empirical evidence. Since the FIPPs rely on users being able to provide informed consent for data sharing activities, technical standards need to specify how to effectively communicate privacy information to users. At NIST, we are in the early stages of conducting a study to determine effective interfaces for obtaining informed consent for websites to share data with affiliates. In the next section, we provide an overview of this study in order to show how empirical studies can better inform privacy standards.

Study Design and Goals

We have designed a study to examine how participants' data-sharing decisions change based on the presence of salient information describing the data to be shared. Specifically, we are examining a popular single sign-on (SSO) interface to examine whether participants make different decisions about whether to use SSO based on how the data being shared is described. Many websites are opting to support various SSO platforms because in addition to simplifying their authentication implementations, it allows these websites to collect more data about their users and their users' habits. Providing informed consent in this situation goes to the very heart of "transparency." While our study specifically examines privacy trade-offs when using an

[1] The P3P compact policy is widely used today, but use of the full XML policy never reached the level that its creators expected [15].

SSO implementation, the results should be generalizable to the design of any dialog used to solicit informed consent to collect or share personal information from users. Specifically, we expect these results to be relevant to the design of a granular do-not-track interface that allows users to opt-in or opt-out of tracking based on the requesting entity and the data requested.

The particular SSO implementation that we are studying was developed by a social networking website. In addition to offering affiliated websites the ability to authenticate users, the social networking site also provides the affiliated websites with personal information from users' profiles. This data may be used for marketing, user profiling, or other unknown reasons. However, the user must first consent to sharing this data. Figure 1 depicts a screenshot of the original consent dialog.[2] As can be seen, the dialog ostensibly supports some of the FIPPs by providing a list of data being requested and the name of the organization requesting it. In this particular example, the dialog is requesting:

- Name
- Profile picture
- Gender
- Networks
- User ID
- List of friends
- Any other public profile information

[2] Certain commercial products are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the software identified is necessarily the best available for the purpose.

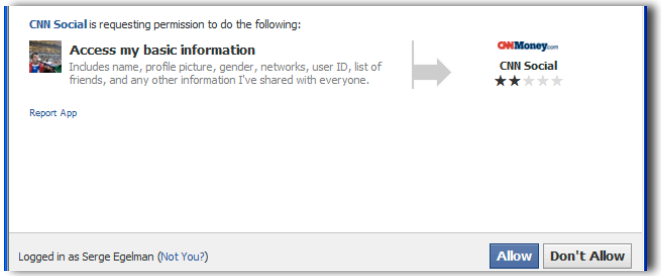

Figure 1: Screenshot of the consent dialog for sharing profile data with participating websites.

Figure 2: Screenshot of consent dialog in our experimental condition. Here, data to be shared is displayed verbatim.

We are examining whether the information is presented clearly enough to facilitate informed consent. Specifically, we have created an experimental condition (Figure 2) wherein users instead see their data verbatim, in addition to the descriptions of that data. If we find that users in the experimental condition were significantly more or less likely to use the SSO option to authenticate to the same websites as users in the control condition, then it is likely because they better understood what data the websites were requesting. This will yield important guidance on how to better request informed consent from users such that it is truly informed.

Conclusion

Our study aims to answer several questions about how to more effectively support the FIPPs through better user interaction design. At the workshop, we hope to present a larger set of questions related to the concept of do-not-track and how empirical studies will better inform a technical standard. Some related standards proposals focus on the binary decision of track or no-track [12]. We propose the use of empirical research to create objective and usable standards that balance user privacy preferences, tailoring to users' needs, and the commercial innovation that can be gained through sharing user data. Do-not-track is not a binary question; it should be more granular by focusing on the ways of conveying information to the user. Do-not-track cannot be effective without the transparency created through granularity. Empirical research, such as our current study, should be used as input for any potential do-not-track standards to improve the usability, effectiveness, and adoption.

References

[1] Ackerman, M. S., Cranor, L. F., and Reagle, J. 1999. Privacy in e-commerce: examining user scenarios and privacy preferences. In *Proceedings of the 1st ACM conference on Electronic commerce* (EC '99). ACM, New York, NY, USA, 1-8.

[2] Acquisti, A. 2004. Privacy in electronic commerce and the economics of immediate gratification. In *Proceedings of the 5th ACM conference on Electronic commerce* (EC '04). ACM, New York, NY, USA, 21-29.

[3] Acquisti, A. and Grossklags, J. 2005. Privacy and Rationality in Individual Decision Making. *IEEE Security and Privacy* 3, 1 (January 2005), 26-33.

[4] Berendt, B., Gunther, O., and Spiekermann, S. 2005. Privacy in e-commerce: stated preferences vs. actual behavior. *Commun. ACM* 48, 4 (April 2005), 101-106.

[5] Cranor, L. 2002. Web Privacy with P3P. O'Reilly & Associates, Sebastopol, CA.

[6] Cranor, L. F., Guduru, P., and Arjula, M. 2006. User interfaces for privacy agents. *ACM Trans. Comput.-Hum. Interact.* 13, 2 (June 2006), 135-178.

[7] Egelman, S., Cranor, L. F., and Chowdhury, A. 2006. An analysis of P3P-enabled web sites among top-20 search results. In *Proceedings of the 8th International Conference on Electronic Commerce.* (ICEC '06). ACM, New York, NY, USA, 197-207.

[8] Egelman, S., Tsai, J., Cranor, L. F., and Acquisti, A. 2009. Timing is everything?: the effects of timing and placement of online privacy indicators. In *Proceedings of the 27th international conference on Human factors in computing systems* (CHI '09). ACM, New York, NY, USA, 319-328.

[9] Friedman, B., Hurley, D., Howe, D. C., Felten, E., and Nissenbaum, H. 2002. Users' conceptions of web security: a comparative study. In *CHI '02 extended abstracts on Human factors in computing systems* (CHI EA '02). ACM, New York, NY, USA, 746-747.

[10] Jensen, C., Sarkar, C., Jensen, C., and Potts, C. 2007. Tracking website data-collection and privacy practices with the iWatch web crawler. In *Proceedings of the 3rd Symposium on Usable Privacy and Security* (SOUPS '07). ACM, New York, NY, USA, 29-40.

[11] Kelley, P. G., Bresee, J., Cranor, L. F., and Reeder, R. W. 2009. A "nutrition label" for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security* (SOUPS '09). ACM, New York, NY, USA, , Article 4 , 12 pages.

[12] Mayer, J., Narayanan, A., and Stamm, S. "Do Not Track: A Universal Third-Party Web Tracking Opt Out." IETF draft document. Accessed: March 18, 2011. http://tools.ietf.org/html/draft-mayer-do-not-track-00.

[13] McDonald, A. M. 2010. Cookie confusion: do browser interfaces undermine understanding?. In *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems* (CHI EA '10). ACM, New York, NY, USA, 4393-4398.

[14] Schultze, S. "Web Browser Security User Interfaces: Hard to Get Right and Increasingly Inconsistent." Accessed: March 17, 2011. http://www.freedom-to-tinker.com/blog/sjs/web-browser-security-user-interfaces-hard-get-right-and-increasingly-inconsistent.

[15] Schwartz, A. "Looking Back at P3P: Lessons for the Future." November 11, 2009. http://www.cdt.org/paper/looking-back-p3p-lessons-future.

[16] Tsai, J., Egelman, S., Cranor, L., and Acquisti, A. The effect of online privacy information on purchasing behavior: An experimental study. In Proceedings of the 2007 Workshop on the Economics of Information Security (WEIS'07) (Pittsburgh, PA, USA, 2007).

[17] US Department of Commerce. "Commercial Data Privacy and Innovation in The Internet Economy: A Dynamic Policy Framework." December 2010. http://www.commerce.gov/sites/default/files/documents/2010/december/iptf-privacy-green-paper.pdf.

[18] Whalen, T. and Inkpen, K. M. 2005. Gathering evidence: use of visual security cues in web browsers. In *Proceedings of Graphics Interface 2005* (GI '05). Canadian Human-Computer Communications Society, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 137-144.

[19] Wu, M., Miller, R. C., and Garfinkel, S. L. 2006. Do security toolbars actually prevent phishing attacks?. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (CHI '06), Rebecca Grinter, Thomas Rodden, Paul Aoki, Ed Cutrell, Robin Jeffries, and Gary Olson (Eds.). ACM, New York, NY, USA, 601-610.

[20] Zurko, M. E. and Simon, R. T. 1996. User-centered security. In *Proceedings of the 1996 New Security Paradigms Workshop* (NSPW '96). ACM, New York, NY, USA, 27-33.