Binding Privacy Rules to Data: Empowering Users on the Web
W3C Privacy Workshop

July 13-14, 2010

John Morris, Alissa Cooper, and Erica Newland
Center for Democracy & Technology

1. Introduction

This paper describes and discusses a new approach to protecting users' privacy in the online environment – the idea of binding relevant privacy rules to the user data itself, so that recipients of user data are aware of the applicable constraints on their use or retention of the data. This approach is being considered within the W3C in the Device API and Policy Working Group (DAP),[1] and it has been implemented for location data by the IETF's Geopriv Working Group.[2] A goal of this paper is to briefly describe the approach, to explain why it is worth considering, and to discuss a number of arguments that have been advanced in opposition to the proposed approach.

A separate workshop position paper, "Privacy Rulesets: A User-Empowering Approach to Privacy on the Web," suggests a possible method to compactly express bundles of privacy rules, which could be conveyed (as this paper suggests) along with user data. These two papers present ideas that could be considered and adopted independently of each other, but the two papers taken together represent what the authors are urging the W3C DAP WG to adopt. Neither of the papers attempts to define any specific mechanisms for actually conveying rulesets along with user data; such mechanisms could include application header fields, URI parameters, or as parameters in web API functions.

2. Binding Rules to Data

The central feature of the binding rules approach is that user information is always bound or transmitted with applicable privacy rules to ensure that entities that receive the information are informed of how they may use it. By creating a structure to convey the users' preferences along with their information, the likelihood that those preferences will be honored necessarily increases. In particular, no recipient of the information can disavow knowledge of users' preferences for how their information may be used. The binding of privacy rules to user information can convey users' desire for and expectations of privacy, which in turn helps to bolster social and legal systems' protection of those expectations.

Applying and affixing usage rules to information is a well-known way of protecting information. In the copyright space, the Creative Commons[3] model is the most prominent

[1] http://www.w3.org/2009/05/DeviceAPICharter.
[2] http://datatracker.ietf.org/wg/geopriv/charter/.
[3] http://creativecommons.org/.

example, allowing an owner of a work to set four types of rules ("Attribution," "Noncommercial," "No Derivative Works" and "ShareAlike") governing the subsequent use of the work. After the author sets these rules, the rules are conveyed together with the work itself, so that every consumer of the work is aware of the copyright terms.

Classification systems for controlling sensitive documents within an organization are another example. In these systems, when a document is created, it is marked with a classification such as "SECRET" or "PROPRIETARY." Each recipient of the document knows from this marking that the document should only be shared with other people who are authorized to access documents with that marking. Classification markings can also convey other sorts of rules, such as a specification for how long the marking is valid (a declassification date). The United States Department of Defense guidelines for classification[4] provides one example.

None of these examples of binding usage rules to information are self-executing. Unlike some technical strategies (such as encryption), these systems rely on external, non-technical mechanisms (such as laws, contracts, or company rules) to enforce the protection of the information. The proposal here is to create technical requirements to ensure that the applicable rules are always transported with the relevant data, and to leave to regulatory, legal, and market forces the enforcement of those rules.

3. Why Consider a New Approach to Privacy

Traditionally, the extent to which data about individuals enjoys privacy protections on the Internet has largely been decided by the recipients of the data. Internet users may or may not be aware of the privacy practices of the entities with whom they share data. Even if they are aware, they have generally been limited to making a binary choice between sharing data with a particular entity or not sharing it. Internet users have not historically been granted the opportunity to express their own privacy preferences to the recipients of their data and to have those preferences honored.

This paradigm is problematic because the interests of data recipients are often not aligned with the interests of data subjects. While both parties may agree that data should be collected, used, disclosed and retained as necessary to deliver a particular service to the data subject, they may not agree about how the data should otherwise be used or retained. For example, an Internet user may gladly provide his email address on a Web site to receive a newsletter, but he may not want the Web site to share his email address with marketers, whereas the Web site may profit from such sharing. Neither providing the address for both purposes nor deciding not to provide it at all is an optimal option from the Internet user's perspective.

[4] U.S. Department of Defense, "National Industrial Security Program Operating Manual", DoD 5220-22M, January 1995.

Moreover, it is broadly accepted that the privacy-policy-based approach to privacy on the Web has not protected users' privacy.[5] When web sites only disclose what they do with users' information deep in a lengthy document of fine print, no real notice happens, and no real consent is given. As the head of the Bureau of Consumer Protection of the U.S. Federal Trade Commission (FTC) recently said, the old approach to privacy "depended on the fiction that people were meaningfully giving consent" – he made plain that the FTC was trying to determine what a "post-disclosure era" would look like (in other words, what would replace the existing privacy policy regime on the Web).[6] It is a near certainty that – at least in the United States – lawmakers and/or regulators will soon propose new approaches to online privacy.

Thus, there is general consensus that the notice-and-consent regime that the Web (as well as recent standards like the W3C's Geolocation API) relies upon is ineffective, and regulators are looking to determine a new and more privacy-protective regime. Technology and standards designers could choose to continue to rely upon the outgoing regime until regulators compel them to change, but doing so is only likely to delay the inevitable. A more forward-looking approach would be to get a head start on developing new options for users to direct how their data will be used.

4. Arguments Against the Binding Rules Approach

Both in the W3C's Geolocation Working Group and elsewhere, a number of arguments have been raised against using the idea of binding privacy rules to data. This section briefly recaps some of the criticisms and responses to the,m without intending to be an exhaustive discussion of either side of the arguments.

a. The W3C Geolocation Working Group already rejected this approach, and we should not revisit that decision.

Although the Geolocation WG did decide not to use the basic approach suggested here, the primacy focus of that WG was on finalizing and standardizing a pre-existing API, and the WG was not open to building privacy protections into its API. In explaining its rejection of the binding rules approach, the WG specifically left for another day the broader question of how privacy should be handled for devices; the WG chairs wrote:

> "The working group concluded that privacy protection does not belong in the Geolocation API itself, but is better handled as part of a more generic privacy and security framework for device access. The recently formed Device API and Policy Working Group is chartered to develop precisely such a framework (http://www.w3.org/2009/05/DeviceAPICharter)."[7]

---

[5] See, e.g., http://www.nytimes.com/2010/02/28/technology/internet/ 28unbox.html.
[6] http://mediadecoder.blogs.nytimes.com/2010/01/11/ftc-has-internet-gone-beyond-privacy-policies/,
[7] http://lists.w3.org/Archives/Public/public-geolocation/2009Oct/0009.html.

b. The binding rules approach does not protect privacy through technical means (such as encryption).

The binding rules approach does not, by itself, provide technical means through which it can be reasonably guaranteed that users' privacy rules will be honored by recipients of their data. The privacy protections in the approach are largely provided by virtue of the fact that data recipients are informed of relevant privacy rules, and are expected to only use location in accordance with those rules.

By binding privacy rules to users' data, however, the approach provides valuable information about users' privacy preferences, so that non-technical forces such as legal contracts, governmental consumer protection authorities, and marketplace feedback can better enforce those privacy preferences. If a commercial recipient of personal information, for example, violates rules bound to that information, the recipient can, in a growing number of countries, be charged with violating consumer or data protection laws. In the absence of a binding of rules with personal information, consumer protection authorities are less able to protect consumers whose information has been abused.

c. Implementing a rules interface in a user agent would be hard, and users might be confused.

Without question user interfaces are hard. But ultimately, users must be given greater control over their information, and there is likely to be less user confusion if privacy is addressed in the user agent rather than if every individual website or app has to figure out ways to give users that control. Browser makers have previously gone down the path of giving users greater control – over, for example, cookies. When cookies were first introduced on the Web, browsers provided no way for users to control their use.[8] As concerns were raised about potentially privacy-invasive uses of cookies, browser vendors began to add cookie controls into their products, beginning with rudimentary tools and evolving over time to the more sophisticated controls in place today. Today, although many users do not use cookie controls, an increasing number do, and those users' privacy has been significantly enhanced.[9]

The separate but related proposal for "privacy rulesets" may help to reduce user confusion, especially if rulesets can be implemented consistently across browsers. By offering users a finite set of privacy choices, the ruleset approach may enhance consumer understanding.

---

[8] *See Federal Trade Commission Staff Report. Public Workshop on Consumer Privacy on the Global Information Infrastructure, Part III: Enhancing Consumer Protection Online* (Dec. 1996), available at http://www.ftc.gov/reports/privacy/Privacy4.shtm.

[9] http://www.comscore.com/Press_Events/Press_Releases/2007/04/comScore_Cookie_Deletion_Report.

d. Users would blame the browser when websites violate the users' rules.

If a browser provides a user interface about privacy rules and those rules are later violated, there is a risk that the browser gets some blame. There are, however, affirmative steps that can be taken when designing the user interface to mitigate this possibility. The user interface must make clear that it is soliciting rules to be conveyed to the recipient of the information, and that recipient is responsible to follow the rules. Instead of asking, for example, "how long do you want your data to be retained by the receiving website," a user agent could ask about "what time limit on retention do you want to send to the receiving website." By being careful to convey the limits of the browsers' control over later uses of the data, the user agent can reduce the risk that it would be blamed for a rules violation by a receiving entity.

e. Rather than providing incomplete privacy protection, it is better for users to think there is no privacy protection.

In the security context, there may be real risk if users mistake weak protection for adequate protection – they may expose critical data (such as, say, bank account login information) and then suffer catastrophic harm. And there often is an available way to achieve real security, even if it means a delay or inconvenience in performing a transaction.

The privacy context is quite different. The harm is often more incremental, and users are better off if even a subset of recipients of their information honor their privacy rules. Users today are often presented with a "Hobson's choice" with regards to their data: using a service requires implicit acceptance of all future data uses by the service provider, and the only other option is to not use the service at all. Unlike in the security context, users often have no alternative to this "take-it-or-leave-it" approach to privacy, and so users are forced to give up their privacy. Any enhanced privacy protections, even if incomplete, will offer users a substantive improvement over the status quo.

f. We are not sure it will work.

The binding rules approach is new to the applications layer, and there is certainly no guarantee that this framework will succeed. But, one thing is certain: the status quo has failed to protect user privacy in any meaningful way. Doing nothing to change how privacy is handled online will perpetuate that user harm. Alternatively, if the privacy community, the W3C, and the leading browser makers can get behind a new approach to privacy, there are good prospects for success.

* * * * *

By binding privacy rules to data, users can – for the first time – be given some element of control, as well as some legal claim, over how their data may be used. The approach will also better enable regulators to enforce user preference and to take action against the greatest privacy-offenders. It may not be perfect in the short term, but it would be an important step in the right direction and an important foundation for future user protection.