Carnegie Mellon University

Institute for Software Research
School of Computer Science
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213-3890

February 18, 2011

Federal Trade Commission
600 Pennsylvania Avenue, NW
Washington, DC 20580

RE: COMMENTS ON FEDERAL TRADE COMMISSION PRELIMINARY STAFF REPORT
Protecting Consumer Privacy in an Era of Rapid Change: A Proposed Framework for Businesses and Policymakers

Dear FTC Staff:

With growing recognition that website privacy policies are failing consumers, numerous suggestions[1] are emerging for technical mechanisms that would provide privacy notices in machine-readable form, allowing web browsers, mobile devices, and other tools to act on them automatically and distill them into simple icons for end users. Other proposals are focused on allowing users to signal to websites, through their web browsers, that they do not wish to be tracked.[2] These proposals may at first seem like fresh ideas that allow us to move beyond impenetrable privacy policies as the primary mechanisms of notice and choice. Facilitating transparency and control through easily recognizable symbols and privacy controls that need be set only once are laudable goals. However, in many ways, the conversations around these new proposals are reminiscent of those that took place 15 years ago that led to the development of the Platform for Privacy Preferences (P3P) standard[3] and several privacy seal programs.[4]

---

[1] Here are just a few recent examples: TRUSTe has stated an intention to support efforts to develop XML privacy policies http://www.truste.com/blog/?p=879 . Mozilla has launched a privacy icons project https://wiki.mozilla.org/Drumbeat/Challenges/Privacy_Icons . The Interactive Advertising Bureau (IAB) CLEAR Ad Notice project plans to integrate XML privacy notices http://www.iab.net/clear .

[2] http://donottrack.us

[3] For a more complete history of P3P see chapter 4 of Lorrie Faith Cranor, *Web Privacy with P3P,* O'Reilly, 2002. For another account of the history and a discussion of related policy issues see: Harry Hochheiser, The Platform for Privacy Preferences as a social protocol, *ACM Transactions on Internet Technology,* 2(4), 2002. For a more recent account see also: Ari Schwartz, Looking Back at P3P: Lessons for the Future, November 2009, http://www.cdt.org/files/pdfs/P3P_Retro_Final_0.pdf

[4] One of the early proposals for the eTRUST privacy seal program (later changed to TRUSTe) involved the seal provider offering three different levels of trust marks to describe three different types of data sharing practices. Each mark would have its own icon. Esther Dyson describes this in: Labels and Disclosure Part II: Privacy. *Release 1.0.* 19 February 1997. http://cdn.oreilly.com/radar/r1/02-97.pdf

institute for SOFTWARE RESEARCH

Lorrie Faith Cranor
Associate Professor
Computer Science and Engineering & Public Policy
(412) 268-7534 • lorrie@cs.cmu.edu • http://lorrie.cranor.org/

P3P was developed with the idea that your web browser should be able to read privacy policies for you and take actions on your behalf without interfering with your web browsing experience. After nearly two years of informal discussions, in 1997 W3C launched a five-year process that led to the publication of the P3P 1.0 specification in 2002. The original idea for P3P involved a protocol in which web browsers would negotiate with websites over privacy on behalf of their users. The negotiation protocol proved problematic for a variety of reasons and was dropped before the specification was released. A subsequent P3P 1.1 effort produced a working draft that included additional P3P vocabulary elements, backwards-compatible syntax changes, and plain-language definitions of P3P elements. However, the P3P working group was closed in 2006 due to lack of industry participation, and P3P 1.1 was never finalized.

P3P 1.0 provides an XML format for website privacy policies, and a protocol for locating and retrieving these policies and associating them with online resources. The XML format encodes the P3P "vocabulary," a privacy taxonomy that was the subject of much debate and disagreement during the P3P development process. The P3P protocol is fairly simple, designed so that no special software would be required for web servers to comply with P3P. Websites can become P3P-enabled simply by placing P3P files at designated locations on their servers. Most of the complexity associated with the P3P protocol centers around performance optimizations designed to reduce the number of P3P requests that user agents must make to locate and fetch up-to-date P3P policies.

The P3P 1.0 specification also describes a P3P "compact policy" format for providing a summary of the privacy policy for cookies that can be transferred in an HTTP header. The compact policy was intended as a supplement to a full P3P policy, designed to allow browsers to evaluate quickly the policies associated with cookies. The P3P specification requires sites using compact policies to provide accompanying full P3P policies.

P3P user agent tools have been integrated into the Microsoft Internet Explorer 6, 7, and 8 web browsers, as well as Netscape 7. P3P was never implemented for Firefox,[5] Safari, or Chrome, although a number of prototype plug-ins and extensions have been developed. In addition, a variety of P3P authoring tools[6] have been developed as well as prototype P3P user agents.[7] In our studies at Carnegie Mellon University, we have demonstrated that the information provided by our prototype P3P user agent tools informs consumers about privacy and impacts purchasing decisions.[8]

---

[5] It appears that some Mozilla developers were planning to implement P3P support at one point, and had begun writing the code. However, this project appears to have been abandoned. http://www-archive.mozilla.org/projects/p3p/ .

[6] One of the most popular P3P authoring tools is the P3P Policy Editor distributed for free by IBM http://www.alphaworks.ibm.com/tech/p3peditor

[7] I have been involved in the development of an IE browser helper object called Privacy Bird http://privacybird.org and a P3P-enabled search engine called Privacy Finder http://privacyfinder.org . Privacy Finder demonstrates the use of P3P to help users select privacy-protective sites from among search results. It also integrates a privacy "nutrition label" generated automatically from P3P policies http://cups.cs.cmu.edu/privacyLabel/ .

[8] S. Egelman, J. Tsai, L. Cranor, and A. Acquisti. 2009. Timing Is Everything? The Effects of Timing and Placement of Online Privacy Indicators. CHI '09: Proceedings of the SIGCHI conference on Human Factors in Computing Systems. http://www.guanotronic.com/~serge/papers/chi09a.pdf; P.G. Kelley, L.J. Cesca, J. Bresee, and L.F. Cranor. Standardizing Privacy Notices: An Online Study of the Nutrition Label Approach. CHI2010. http://www.cylab.cmu.edu/research/techreports/tr_cylab09014.html

The Internet Explorer P3P implementation is probably the most widely used P3P tool given the widespread use of IE. However, it appears that most IE users are completely unaware of P3P. P3P functionality is associated with three user interface components in IE, although the interface does not explicitly mention P3P in any of those places. First, in the View menu, users have the option of viewing a "Privacy Report." Clicking on this option causes IE to check whether a website has a full P3P policy. If IE finds a P3P policy, it fetches it and translates the XML code into English (or the appropriate language for that version of IE), using the technical definitions of P3P elements found in the P3P 1.0 specification. Second, the default cookie setting in IE (the medium setting) bases third-party cookie-blocking decisions on P3P compact policies. Third-party cookies without compact policies are blocked. IE analyzes any compact policies it finds associated with cookies and determines whether or not the policies are "satisfactory." Those third-party cookies found to have unsatisfactory compact policies are blocked. Finally, when IE blocks cookies it places a small icon in the bottom right area of the browser chrome that looks like a do-not-enter sign overlapping an eye. Most users do not seem to have any idea what the icon means. However, those who click on the icon are shown a list of blocked cookies. Users can click on a link for each blocked cookie to display a privacy report if there is a full P3P policy associated with it.

Microsoft's decision to base third-party cookie-blocking decisions in Internet Explorer on P3P compact policies led to widespread adoption of P3P among advertising networks and other companies making substantial use of third-party cookies. P3P was adopted by about a third of the most popular websites, but never saw widespread adoption beyond popular sites and those that use third-party cookies.[9]

Arguably, the largest barrier to P3P adoption has not been problems with the P3P vocabulary or difficulties with the technical mechanisms, but rather lack of incentives to adopt. As Dyson observed in 1997, "Industry disclosure schemes often founder without strong government/public pressure. Otherwise, companies are simply too busy to adopt them, and customers don't factor the information disclosed into their buying habits."[10] By the time the P3P specification was released in 2002, government pressure had subsided and industry had largely lost interest in P3P.

Besides a set of companies who adopted P3P because they were positioning themselves as privacy leaders, most of the adopters decided to implement P3P in order to prevent IE6 from blocking their cookies. However, over time we began to observe that many of these companies did not appear to be making serious efforts to implement P3P, and instead were offering minimal policies designed to prevent IE cookie blocking. Initial signs of this were the large number of sites that implemented P3P compact policies without the corresponding full policies, and an unexpectedly large number of sites that had syntax errors in their P3P policies.[11]

More recently, our research has found that a large fraction of sites adopting P3P compact policies have misrepresented their privacy practices, most likely in an effort to prevent IE from blocking their cookies. [12] We collected CPs from 33,139 websites and used automated techniques to detect syntax

[9] L. F. Cranor, S. Egelman, S. Sheng, A. M. McDonald, and A. Chowdhury. P3P deployment on websites. *Electronic Commerce Research and Applications,* 7(3):274{293, 2008. http://lorrie.cranor.org/pubs/p3p-deployment.html

[10] Dyson 1997.

[11] Cranor et al. 2008.

[12] P.G. Leon, L.F. Cranor, A.M. McDonald, and R. McGuire. Token Attempt: The Misrepresentation of Website Privacy Policies through the Misuse of P3P Compact Policy Tokens. WPES 2010. http://www.cylab.cmu.edu/research/techreports/2010/tr_cylab10014.html

errors and inconsistencies in 11,176 of them, including 134 TRUSTe-certified websites and 21 of the top 100 most-visited sites. Upon further investigation, we discovered thousands of sites that had the identical erroneous policies and traced these policies to a Microsoft support website[13] and several blog posts that recommended posting these policies to prevent cookie blocking.

We also found sites that posted compact policies that bore little resemblance to proper compact policy syntax and were clearly meant to circumvent IE. For example Amazon posted a compact policy containing the single made-up token "AMZN" and Facebook posted a compact policy containing only tokens for the disputes and remedies elements (no data categories, purpose, recipients, access, or retention tokens, which are required for a valid policy). During a preliminary study in 2009 we observed that the Facebook compact policy contained the single made-up token "HONK."

We also discovered that when IE analyses compact policies to determine whether they are satisfactory, it simply looks for combinations of tokens that appear on a list of unsatisfactory tokens. IE apparently does not test the compact policy to determine whether it is syntactically valid. As a result, compact policies that consist entirely of made-up tokens will never be flagged as unsatisfactory.

After our paper was published, we noticed that Facebook updated it's P3P compact policy to the invalid policy:

> P3P:CP="Facebook does not have a P3P policy. Learn why here: http://fb.me/p3p"

The link in the compact policy goes to a page that explains:[14]

> The organization that established P3P, the World Wide Web Consortium, suspended its work on this standard several years ago because most modern web browsers do not fully support P3P. As a result, the P3P standard is now out of date and does not reflect technologies that are currently in use on the web, so most websites currently do not have P3P policies.

While Facebook is now trying to be more upfront about their bogus P3P compact policy, the fact remains that they have put a compact policy in place to circumvent IE's cookie blocking mechanism, on which many consumers rely.

Amazon took a different approach, and changed its compact policy to a policy that appears to be syntactically valid. However, the corresponding full P3P policy is not valid, and contains the following text:[15]

> Because some browsers require a P3P policy, we have created a compact P3P policy that outlines some, but not all, of the details of our privacy practices. The compact policy relates primarily to our use of HTML cookies and personally identifiable information associated with such cookies. However, we have not included a full P3P policy because the binary limitations of the required XML code currently do not fully and adequately express our policies and practices.

---

[13] Microsoft Support. Session variables are lost if you use FRAMESET in Internet Explorer 6, April 2006. http://support.microsoft.com/kb/323752 (this page was removed shortly after our paper was published in September 2010).

14 Excerpted from https://www.facebook.com/help/?topic=p3p visited on February 8, 2011.

[15] Excerpted from Amazon.com full P3P policy posted at http://www.amazon.com/w3c/p3p_full.xml and visited on February 8, 2011.

Instead, we ask that you read our full Privacy Notice at www.amazon.com/privacy.

While it is not entirely clear what Amazon's concern is, their reference to "binary limitations" suggests that they are uncomfortable selecting from among some of the multiple-choice P3P vocabulary elements. If Amazon's P3P compact policy matches their actual data practices with respect to cookies, then they are no longer circumventing the IE cookie-blocking mechanism. However, they are not fully complying with P3P either.

The lack of overall P3P compliance demonstrates the ineffectiveness of P3P as a self-regulatory program. After we found that TRUSTe websites with compact policies were just as likely to have errors as the other websites with compact policies that we surveyed, TRUSTe acknowledged that P3P compliance was not part of their routine review process and they did not expect to make it part of their process.[16] Indeed, it seems the industry has all but given up on P3P, but cannot abandon it completely as long as Microsoft keeps using it as part of their cookie-blocking filter in Internet Explorer.

Our finding that companies are adopting P3P in order to misrepresent their privacy practices to Internet Explorer's cookie blocking feature – one of the most commonly-used tools that consumers have for protecting their online privacy – raises serious concerns. This would seem to be an area where the US Federal Trade Commission could exert their enforcement authority.

There is currently a lot of interest in privacy nutrition labels, machine-readable privacy policies, and privacy icons. Based on my past work and observations I offer the following recommendations.

Incentives for adoption and mechanisms for enforcement are essential. We are unlikely to see widespread adoption of a privacy policy standard if we do not address the most significant barrier to adoption: lack of incentives. If a new protocol were built into web browsers, search engines, mobile application platforms, and other tools in a meaningful way such that there was an advantage to adopting the protocol, we would see wider adoption. However, in such a scenario, there would also be significant incentives for companies to game the system and misrepresent their policies, so enforcement would be critical. Incentives could also come in the form of regulations that require adoption either for all companies, or for companies that make secondary use of personal data. Before we go too far down the road of developing new machine-readable privacy notices (whether comprehensive website notices like P3P, icon sets, or notices for mobile applications, do-not-track, or other anything else), it is essential to make sure adequate incentives will be put in place for them to be adopted, and that adequate enforcement mechanisms exist.

Standardization benefits consumers. There seems to be a clear advantage of standardized notices for consumers. Standardized notices facility comparisons and allow consumers to become familiar with terminology and where to look to find particular types of information. However, to be effective, standardized notices need to have fairly rigid requirements so that their elements are directly comparable. An earlier attempt at standardized privacy notices, the layered notice developed by The Center for Information Policy Leadership,[17] introduced some good ideas, but allowed so much flexibility that companies ended up using it in fairly inconsistent ways. While there may be a place for

[16] TRUSTe president Frances Maier mentions this in her comments to a blog post on September 14, 2010: http://www.pogowasright.org/?p=13959&cpage=1#comment-257

[17]The Center for Information Policy Leadership. Ten steps to develop a multilayered privacy notice. March 15, 2007. http://www.hunton.com/files/tbl_s47Details/FileUpload265/1405/Ten_Steps_whitepaper.pdf

companies to customize standardized formats to provide specific details, the overall format needs to be fairly uniform.

Machine-readable policies allow for automation. As online interactions get more complicated, it becomes increasingly difficult for users to understand what parties are involved let alone sort through each of their privacy policies. Thus machine-readable policies play an important role because they allow web browsers and other tools to consume policy data automatically and take actions on a user's behalf (blocking cookies and other forms of profiling, warning users, etc.). Machine-readable policies also have benefits for business-to-business interactions, because they allow businesses to more easily determine the policies associated with their service providers and advertising agents.

Layers allow for both simple and detailed views. An extremely simple privacy notice, perhaps in the form of an icon, is likely to appeal to most consumers. On the other hand, some consumers and privacy experts will want to see more detailed disclosures, and in some cases detailed disclosures are required for legal purposes. A layered approach to privacy notices would make very simple notices readily available with links to more detailed notices.

Standard policy types could simplify privacy decision-making. One way of distilling complicated privacy policies down to a small number of icons (similar to the Creative Commons approach) is to identify the most important practices that consumers are likely to want to know about and developing a small number of policy templates that incorporate these practices. For example, a type I policy might commit to not collecting sensitive categories of information and not sharing personal data except with a company's agents, while a type II policy might allow collection of sensitive information but still commit to not sharing them, a type III policy might share non-identified information for behavioral advertising, and so on. Companies would choose which policy type to commit to. They could advertise their policy type with an associated standard icon, while also providing a more detailed policy. Users would be able to quickly determine the policy for the companies they interact with. In addition, the establishment of a clear set of policy types would likely encourage companies to improve their privacy practices so that they could associate themselves with a more privacy-friendly policy type.

The P3P vocabulary offers a good starting point for future privacy vocabularies. The P3P vocabulary has been criticized for its complexity and for its lack of expressiveness. There are clearly some areas that could use some fine-tuning, but after about a decade of use, it seems that overall the P3P vocabulary seems to do a pretty reasonable job. As long as companies are forced to use a fixed vocabulary with multiple-choice fields to express their policies, there are likely to be complaints that any vocabulary is not expressive enough. I recommend collecting very specific examples of problems companies are having expressing their policies in P3P, and using these to help frame discussions about where changes are needed.

Sincerely,

Lorrie Faith Cranor
Associate Professor
Carnegie Mellon University