Comments of Nick Doty on *Commercial Data Privacy and Innovation in the Internet Economy: A Dynamic Policy Framework*, Docket #101214614-0614-01

The Internet Policy Task Force has completed a commendably broad review of the diverse and manifest privacy issues and their effects on commerce and outlined several possible options to address them. I appreciate the opportunity to offer comments. In response to the task force's questions for further discussion I've outlined some steps the Department might take to contribute to the *technical* solutions proposed in the green paper. My comments are based on both the history of, and ongoing debate over, interactions between Web technology and policy.

The advantages of machine-readable policy expressions

2.e. Should there be a requirement to publish PIAs in a standardized and/or machine-readable format?

In short, yes, absolutely. Making privacy policies (or privacy impact assessments, or other forms of privacy notices) standardized and machine-readable has the potential to supplement and replace unwieldy and unread legalese documents with short, iconic, recognizable, easily-comparable and automatically-actionable statements of what's important to the consumer. Machine-readability in particular provides the promise of applying the modern revolution in information organization and retrieval and its gains in efficiency to the challenge of managing one's own personal information in a highly networked world. Specifically, machine-readable privacy policies may:

- allow clients (Web browsers, smartphone applications or other software) to read and highlight the key terms of a privacy policy,
- let different individual consumers specify in advance what policies they are comfortable with or would like to avoid,
- provide advocates (whether it's the Chamber of Commerce or the Electronic Frontier Foundation) the chance to recommend privacy settings,
- enable automated comparison of sites based on their privacy features and thus encourage real competition on privacy terms,
- help aggregate and analyze the disparate privacy statements a user may encounter from multiple parties in a single transaction,
- disseminate new developments in readability or presentation to every privacy policy at once (rather than waiting for all companies and web sites to update their privacy notice presentations) and
- force both large businesses and small web site proprietors to think through and be explicit about their data usage policies.

Privacy policies that are not only standardized but also machine-readable may also contribute answers to other of the task force's open questions. For example, (see *2.g.*), an automated client could easily consume and combine multiple policies for the multiple sources of content (a web site, an ad network, an analytics provider) a consumer will often see on a single Web page. Allowing automated access also allows an automated client to highlight key privacy practices even on devices with small screens, a growing concern given the popularity of smartphone access to Web applications (see

2.h.). In fact, with the growing promise of an "Internet of Things" where a multitude of devices from cars to tennis shoes have access to the Internet, machine-readable policies would let users continue to have control over use of their data even without the capability to read online text. Finally, machine-readable privacy statements may also provide one answer to the question of how purpose specifications can be implemented (*2.m.*); P3P, for example, contained a mandatory `<PURPOSE>` element.[1]

Challenges to the use and adoption of machine-readable policy

But the Department should also be aware that these ideas have been considered before and the development of policy expression languages and technology (though difficult) is not the only or even primary blocker of their adoption. The history of the Platform for Privacy Preferences Project (P3P) is the most common and relevant example: despite becoming a well-defined technical standard and reaching multi-stakeholder agreement on the legal terms used, P3P was never broadly implemented by Web browser vendors and therefore rarely taken advantage of by consumers.

Complexity — both in the user interface for the consumer and the language complexity for the developer — has certainly been one inhibitor. I agree with the W3C comment that more research is necessary on building easy-to-understand interfaces to translate machine-readable policies and consumer privacy preferences for the end user. But adoption on the Web also depends on the complexity for the designer or developer of a Web site: not all Web sites are run by large corporations or developed by sophisticated engineers and the adoption of standards often depends on straightforward implementation (as in the case of the original HTML standard in contrast to competing hypertext systems). Opposition to recent machine-readable privacy preference proposals has also included concern over complexity for the Web developer.[2] To that end, we have published a brief technical proposal for simple negotiation of policy disclosures at a recent W3C workshop.[3]

In fact, recently there has been a spate of proposed lightweight policy expression technologies:

- Mozilla has experimented with a set of Privacy Icons (with machine-readable backing) to highlight several key facets of a site-wide privacy policy.[4]
- Privicons, a very simple syntax (a mix of human-writable and machine-readable) for specifying distribution limits on email, has been suggested by researchers at Stanford and elsewhere.[5]

[1] The Platform for Privacy Preferences 1.0 Specification, 16 April 2002. http://www.w3.org/TR/P3P/#PURPOSE

[2] See, for example, the debate over including GeoPriv privacy preferences in the W3C Geolocation API: http://www.w3.org/2008/12/08-geolocation-minutes#item04

[3] Nick Doty and Erik Wilde, “Simple Policy Negotiation for Location Disclosure”, W3C Workshop on Privacy and Data Usage Control, October 2010. http://www.w3.org/2010/policy-ws/papers/03-Doty-Wilde-Berkeley.pdf

[4] Aza Raskin, “Privacy Icons: Alpha Release”. http://www.azarask.in/blog/post/privacy-icons/

[5] Privicons. http://privicons.org/

- One version of "Do Not Track" (recently implemented by Mozilla in development versions of Firefox) is a simple HTTP header with a binary, 1-or-0 preference.[6]
- Though not privacy-focused, the Creative Commons project has shown great success and adoption of simple (and in many cases machine-readable) representations of media use licensing.[7]

Encouraging simple technical solutions — and resisting the urge (particularly initially) to represent every complexity — can aid adoption by developers and use by consumers.

Many also resist implementing a machine-readable policy system when the enforceability of such a system is unclear: there is a concern (amongst browser makers, a crucial constituency in this case) that malicious sites will lie in these machine-readable ways and browsers will be responsible for giving users a false sense of security. And those concerns have some basis: a recent study found that many of the P3P compact policies found on the Web today don't accurately represent site policies (or even satisfy the requirements of the standard) but were instead apparently set simply to work around cookie restrictions in Internet Explorer.[8]

Enforcement and regulators as conveners

The Department of Commerce can play a role in assuaging the concern of potentially malicious use of inaccurate machine-readable privacy statements and in doing provide clarity to industry and technical standards bodies considering development and adoption of such systems. Signaling that promises made in a machine-readable format are affirmative representations and will be enforced by the Federal Trade Commission under the "deception" clause of Section 5 could give these fledgling formats a much-needed boost. Enforcement actions against P3P compact policy violations would be one step, but simple guidance from FTC commissioners or the Internet Policy Task Force could also go a long way. Engineers often resist technical measures that are not self-enforcing, based on the very reasonable concern that the response from the market or regulators may not enforce such promises; proactive statements from the Department of Commerce and the Federal Trade Commission may mitigate this inhibitor to adoption of privacy-enhancing technologies.

I also support the report's suggestion of "regulators as conveners." A new Privacy Policy Office could provide considerable value just by moderating the many stakeholders (consumers, browser vendors, service providers, privacy advocates, academics, et al.) involved. I would encourage the PPO to engage with technical standards bodies (the W3C, IETF, OASIS and others) as well as with trade groups of advertisers and service providers. These groups certainly satisfy the requirement for an "open, multi-stakeholder process" and working with the technical community early on, rather than expecting machine-readable policy expression systems to follow the creation of codes of

[6] For one description, see http://donottrack.us/

[7] http://creativecommons.org/

[8] Leon, P. G., Cranor, L. F., McDonald, A. M., and McGuire, R. “Token Attempt: The Misrepresentation of Website Privacy Policies through the Misuse of P3P Compact Policy Tokens”, September 10, 2010. http://www.cylab.cmu.edu/research/techreports/2010/tr_cylab10014.html

conduct, will result in technical systems that better fit the architecture of the Internet and the Web.

I note that NIST is currently requesting comment on the nature and extent of government involvement in technical standard setting;[9] DoC involvement in the W3C or IETF on defining machine-readable policy expressions for commercial data privacy could be a model example.

I hope these comments provide some insight into the Department's potential role in promoting both innovation and privacy through a combination of technical and policy tools. I look forward to continuing work in this area.

Sincerely,

Nick Doty

Lecturer / Researcher
UC Berkeley, School of Information
npdoty@ischool.berkeley.edu
http://npdoty.name

These comments are my own and do not represent the School of Information or the University of California more broadly.

[9] "Effectiveness of Federal Agency Participation in Standardization in Select Technology Sectors for National Science and Technology Council's Sub-Committee on Standardization" http://federalregister.gov/a/2010-30864