Enacting Privacy in Internet Standards[*]

Nick Doty; UC Berkeley, School of Information

April 7, 2014

Privacy on the Web and Engineering Ethics

The functionality of the Internet and the Web are determined in large part by the standards that allow for interoperable implementations; as a result, the privacy and security of our online interactions are greatly impacted by the work done within standard-setting organizations, like the Internet Engineering Task Force (IETF) and the World Wide Web Consortium (W3C). This is just one instance of a phenomenon variously referred to as *values-in-design* (Nissenbaum 1998) or *technological delegation* (**???**): basic matters of public policy importance can be determined (or regulated: Lessig 1999; Langdon Winner 1980) by software architecture, in much the way that urban architecture has (Caro 1975).

Software engineering shares with other types of engineering an impulse to "build", "make" or "create". That impulse can create an ethic to do something, to build something in part exactly because one *can* do so.[1] To solve a difficult problem, even without a particularly remunerative or societally valuable outcome is often considered sufficiently motivating reward. A common method of recruiting software engineers is extolling the set of "hard problems" to work on and the opportunity to do so. We can see both exploratory motivations (a la climbing a mountain "because it's there") and a motivating sense of independence (showing that you can do it on your own, through use of technology) here.

At the same time, the Internet/high-technology field faces a challenge in response: just because you can do something, *should* you? Related: given the privilege of those few who can make potentially enormous differences through the creation and use of technology, are engineers doing the best to live up to that opportunity? (That is, are they satisfying the responsibility that comes with great power (**???**)?) This became a particularly common question in the responses to the suicide of Aaron Swartz (aaronsw) and in the local debates about tech company social responsibility, evictions and housing in San Francisco. These ethical questions

[*] An extremely rough, vodka- and caffeine-infused first draft, towards a prospectus. For blindingly obvious reasons, please do not cite or distribute.

[1] Nick Doty, because you can is reason enough to do something, via email, August 2013.

of software engineering are deeply tied to what we might call the engineering ethic. Given the outsized role that Internet engineering and the choices of many individual software engineers have for values such as privacy, I will direct my research to understanding how privacy is or is not supported by those who develop the Internet and the Web.

The Standard-Setting Process

While software engineering can be particularly attractive to an independent mindset, the Internet and Web are, perhaps uniquely, not things that can be built on one's own. These technologies require, as a matter of course, *interoperable implementations* to prove new technology. The great successes of these technologies are attributed to their *decentralized* nature: Web servers and clients; email servers and clients; chat servers and clients can all be built independently, by different firms using different software stacks running on different operating systems and boasting different features and user interfaces, and still work together.

As an apparently necessary consequence, Internet and Web standards have also been developed using a *consensus* model, rather than through some mandatory or legal mechanism, where adoption of the standard is a (perhaps, *the*) key criterion (Cargill 1989). Dependence on voluntary adoption also lead these organizations towards the *multistakeholder* model: discussions involve potentially all those stakeholder groups who will affect (or, to some extent, those affected by) the implementation, adoption and deployment of a standard. Multistakeholderism is the historical and preferred paradigm for Internet governance and, despite rampant confusion over the term, standard-setting organizations such as IETF and W3C are often considered part of the Internet governance ecosystem because of the impact these discussions have on the implementation of Internet technologies.

Of course, this simplifying story is not entirely correct. While the Internet and the Web *do* distinguish themselves by their decentralized architectures and the consensus standard-setting process *is* markedly distinct from *de jure* standard-setting or single-firm technical design, the development of standards may not be as entirely "flat" or "open" a process as it's imagined. Even in those processes that might satisfy an extremely strict set of procedural requirements (up to Habermas' standards, some have argued (Froomkin 2003)), large firms with large customer install bases, significant financial interests and numerous employed experts inevitably play a larger role. As I and others have noted (Doty and Gupta 2013; Guston and Sarewitz 2006, in particular, Patrick Feng; Waz and Weiser 2012), simple logistical difficulties (the resources need to follow and understand an ongoing standard-setting process) or lack of available organizational expertise make participation a challenge. And while consensus and the technical merits of arguments might be laudable goals for decision-making, feminist critiques suggest that "meritocracy" may mask other factors, including socioeconomic privilege or other types of political or economic power (**???**). Finally, for better or for worse, standard-setting organizations may see a greater influence from those individuals who have historically played a larger role in their development.

Research questions

Given these properties of consensus standard-setting, what factors of the standard-setting process affect online privacy?

(1) How does the multistakeholder model determine the *legitimacy* of standards that affect privacy?
(2) What role do the organizational structures (of standard-setting organizations and their participants, or the member organizations) play in determining privacy outcomes?
(3) What expertise relevant to privacy is present in technical standard-setting organizations?

Research plan

For research purposes, the relative openness and diversity of participants in technical standard-setting contexts provide for a rich corpus to study. Communications are extensively documented (for legal, ethical and pragmatic reasons); many discussions are public; and in-depth debates between participants with radically different viewpoints (and representing different stakeholders) are common.

I plan to pursue an ethnographic approach: qualitative methods including semi-structured interviews; automated and manual text analysis; and fieldnotes from observing and participating in these processes. Ongoing interviews of participants in W3C and IETF standards activities — engineers, laywers and businesspeople who volunteer their time to develop standards but also have roles within tech companies or other organizations — provide a source for the variety of participants, their expertise and their own perspectives about the standardization process, legitimacy and the impact on privacy. Surveys and background research of participant and attendee lists can help in analyzing the make-up of participants and their educational/professional backgrounds. As a participant observer (**???**), I also aim to include my own perspective, including the challenges of working in a diverse multistakeholder setting.

Privacy in Standards

The standards and interoperable implementations of the Web, rather than defining an entire software system, make up a platform on top of which email, web sites, web applications of all kinds are built by innumerous differente vendors. As a result, enacting privacy from within standards has different challenges and sees different approaches than software design in general. Standards development sees constant questions of scoping: what is defined within the standard (necessary or desirable to be standardized) and what is left open to varying implementations? Related, standard-setting involves questions of which "layer" will handle a particular responsibility: is a particular communication embedded at a low layer for all Internet communications or specific to a particular higher-level application? This provides an inherent tension for privacy: as a value defined often through user awareness and understanding of data practices, privacy can be heavily influenced by the particular presentation or user experience. However, standardization of user interface has generally been strongly resisted in standardization, exactly because the variation of UI and UX is a key area of variation and competition between implementers.

Perhaps directly as a consequence of this tension, prominent privacy standardization efforts in Web standards have tended to be basic communication protocols. The Platform for Privacy Preferences (P3P) enabled a machine-readable language for describing a site's privacy policy so that browsers could present privacy policy information to user in some automated fashion. Do Not Track (DNT) is a simple mechanism for a user to express a preference not to be tracked online and automatically communicate that preference in all Web communications. Unsurprisingly, both standards have faced challenges in specification and adoption in part because of disputes over the exact presentation to the end user.

Recent surveillance revelations — in particular, revelations in September 2013 that the National Security Agency had influenced standard-setting organizations to intentionally introduce weaknesses into the cryptographic standards relied on for Internet confidentiality — prompted widespread criticism and sometimes emotional responses — "we were naive never again" (Thomson 2014) — from the engineering community. But some responses from the Internet engineering community are more pragmatically-focused, as seen in renewed, consensus[2] efforts to encrypt virtually all Web communications. That frustration from the engineering community, and the pragmatic response to assert control (through end-to-end encryption, say), may be a theme of the engineering ethic and of the community's way of handling privacy.

Research questions

(4) Where and when do Internet and Web standardization participants attempt to enact privacy?

(5) How do participants' views and preferences around privacy and security translate into technical or other responses?

[2] Russ Housely, reporting "hums" via email, November 2013.

Research plan

As above, ethnographic study of standard-setting participants (as described: through fieldnotes, semi-structured interviewing, text analysis of mailing list archives) is appropriate to get at the participants' own views and motivations. However, in addition, to understand the technical decision-making regarding questions of scope and layering for question (4), I will also review and compare the technical documents (RFCs, W3C Recommendations, or even browser feature implementations).

Tools

Given the challenges — procedural and substantive — to enacting privacy on the Internet and the Web through standards, how might we do better? This project plans to look at two different classes of tools: privacy review guidelines and privacy design patterns.

Privacy Reviews

A privacy review is essentially a procedural tool: at some point in the process of developing a piece of code, a software system or an Internet standard, the object is reviewed with a particular eye toward identifying privacy issues and determining mitigations where appropriate. Many corporations and governments have embraced the approach of *privacy impact assessments* (see, for example: Bamberger and Mulligan 2008, discussing PIAs in US administrative lawmaking; Commissioner 2010, a guide for PIAs from the Australian government; and recently, Oetzel and Spiekermann 2013) to systematically document the information processed by a software system and commonly how it satisfies a set of fair information practices (Department of Health Education and Welfare 1973).

In the space of Internet standards, applying privacy reviews has not been so clear. There is a history of security reviews and "Security Considerations" document sections in RFCs, with a mix of successes and failures (Rabkin, Doty, and Mulligan 2010). Reviewing for privacy — a value much harder to clearly define — has proved more difficult, though attempts have been made and continue (Braman 2012, **???**). The differences in participation (voluntary, bottom-up), software architecture (a generative platform) and organizational structure (a multistakeholder, boundary organization) of Internet standard-setting make a very different environment for privacy reviews than a traditional privacy impact assessment.

Privacy Design Patterns

As an alternative tool, I have proposed privacy design patterns to help translate principles that support privacy into engineering techniques. Design patterns are abstract solutions to common problems within particular contexts. Historically, design patterns are a construct from architecture and urban planning but are also commonly applied within the software developer community. As a familiar method and one suited to documenting problems and solutions in a variety of contexts, we propose it as a specifically "bottom-up" tool for implementing privacy-by-design in software (Doty and Gupta 2013). Past work in this area includes patterns that address the security, usability, and software engineering aspects of designing for privacy (Romanosky et al. 2006), user interface patterns from Egelman (2009) and ongoing work from Hoepman on privacy design strategies as categorizations of privacy design patterns (2012).

Through a past collaboration, I developed privacypatterns.org: a web site with a small collection of privacy patterns (particularly for location-based services) and a workflow for developing additional patterns as a community. I continue to recruit additional participants to contribute patterns

Research questions

Regarding privacy reviews:

(6) How do the conditions in standard-setting organizations differ from other software design situations and how does that influence the use of privacy reviews?

(7) What have been successful or unsuccessful privacy reviews in Internet standards? What properties of the process/substance of those reviews are distinctive?

And patterns:

(8) When is a privacy design pattern more or less useful for software engineers considering privacy in their technical designs?

(9) How does the patterns concept work in the context of Internet standards? Is the level of abstraction useful? Or is it too abstract / too low-level?

(10) What determines willingness to contribute to a community-developed documentation effort?[3]

Research plan

In exploring the history and current state of privacy reviews at IETF and W3C, I plan to interview those experts (of varying backgrounds) who have participated in reviewing standards for privacy or developed standards with privacy implications.

[3] This might apply to both privacy reviews, privacy patterns and even standards participation; as such, it might be a good point for drawing connections between the different sections / research questions.

In addition, I'll complete a textual analysis of past drafts of standards and the conduct of privacy and security reviews in mailing lists (Doty 2014). Finally, I will reflect on my own efforts to encourage and systematize privacy reviews while employed at W3C, including the challenges of recruiting participation and the distinctive procedural difficulties within standard-setting.

While grounded in its own set of literature and history, the privacy patterns project is intended to be very practical. My efforts here will be to develop the project and help a small community contribute patterns and use them "in the wild" for addressing privacy issues in their software designs. As a research matter, I hope to gather and document those practical experiences: talking to those who have written their own patterns or consulted them during an ongoing technical project. That success-and-failure information can be used to improve the design of patterns themselves and the workflow for developing or using them or for learning when the paradigm is and isn't worth continuing.

Future Work

Through the particular forum and case study of Internet standards, this research attempts to learn how privacy is enacted in certain technical designs and to go further to explore how it could be better supported.

Some open directions:

- What should the **future research agenda** be for the area of Internet privacy and standards?
- What software could I **build** to demonstrate or explore the concepts described here?
- To what other settings/communities (open source software development, for example; or Internet multistakeholder governance institutions) might conclusions from privacy in Internet/Web standards extend?

References

Bamberger, Kenneth A., and Deirdre K. Mulligan. 2008. "Privacy Decisionmaking in Administrative Agencies." *The University of Chicago Law Review* 75 (1) (January): 75–107. http://www.jstor.org/stable/20141901.

Braman, Sandra. 2012. "Privacy by Design: Networked Computing, 1969–1979." *New Media & Society* 14 (5) (August): 798–814. doi:10.1177/1461444811426741. http://nms.sagepub.com/content/14/5/798.

Cargill, Carl F. 1989. *Information Technology Standardization: theory, Process, and Organizations*. Newton, {MA}, {USA}: Digital Press.

Caro, Robert A. 1975. *The Power Broker: Robert Moses and the Fall of New York*. New York: Vintage Books. http://www.amazon.com/The-Power-Broker-Robert-Moses/dp/0394720245.

Commissioner, Office of the Australian Information. 2010. “Privacy Impact Assessment Guide Office of the Australian Information Commissioner - OAIC.” http://www.oaic.gov.au/privacy/privacy-resources/privacy-guides/privacy-impact-assessment-guide.

Department of Health Education, and Welfare. 1973. “Records, Computers and the Rights of Citizens.” https://epic.org/privacy/hew1973report/.

Doty, Nick. 2014. “Past, Present and Future of Privacy Reviews in Internet Standard-Setting.”

Doty, Nick, and Mohit Gupta. 2013. “Privacy Design Patterns and Anti-Patterns: Patterns Misapplied and Unintended Consequences.” In *A Turn for the Worse: Trustbusters for User Interfaces Workshop*. http://cups.cs.cmu.edu/soups/2013/trustbusters.html.

Egelman, Serge. 2009. “Trust Me: design Patterns for Constructing Trustworthy Trust Indicators.” PhD thesis, ProQuest. http://books.google.com/books?isbn=1109149131.

Froomkin, A. Michael. 2003. “Habermas@Discourse. Net: Toward a Critical Theory of Cyberspace.” *Harvard Law Review* 116 (3) (January): 749–873. doi:10.2307/1342583. http://www.jstor.org/stable/1342583.

Guston, David H., and Daniel R. Sarewitz. 2006. *Shaping Science and Technology Policy: the Next Generation of Research*. Univ of Wisconsin Press. http://books.google.com/books?id=12kOiesm1T0C.

Hoepman, Jaap-Henk. 2012. “Privacy Design Strategies.” http://arxiv.org/abs/1210.6621.

Langdon Winner. 1980. “Do Artifacts Have Politics?” *Daedalus* 109 (1) (January): 121–136. http://www.jstor.org/stable/20024652.

Lessig, Lawrence. 1999. *Code and Other Laws of Cyberspace*. Basic Books. http://books.google.com/books?id=0l1qLyT88XEC.

Nissenbaum, Helen. 1998. “Values in the Design of Computer Systems.” *Computers and Society* 28 (1): 38–39. http://www.nyu.edu/projects/nissenbaum/papers/society.pdf.

Oetzel, Marie Caroline, and Sarah Spiekermann. 2013. “A Systematic Methodology for Privacy Impact Assessments: a Design Science Approach.” *European Journal of Information Systems* (July). doi:10.1057/ejis.2013.18. http://www.palgrave-journals.com/ejis/journal/vaop/ncurrent/abs/ejis201318a.html.

Rabkin, Ari, Nick Doty, and Deirdre K Mulligan. 2010. “Facilitate, Don’t Mandate.” http://www.iab.org/wp-content/IAB-uploads/2011/03/nick/_doty.pdfhttp://www.iab.org/activities/workshops/internet-privacy-workshop-2010/.

Romanosky, Sasha, Alessandro Acquisti, Jason Hong, Lorrie Cranor, and Batya Friedman. 2006. "Privacy Patterns for Online Interactions." In *PLoP 06 Proceedings of the 2006 Conference on Pattern Languages of Programs*. http://portal.acm.org/citation.cfm?id=1415472.1415486.

Thomson, Martin. 2014. "A Statement." Accessed April 8 06:34:37. http://tools.ietf.org/html/draft-thomson-perpass-statement-01.

Waz, Joe, and Phil Weiser. 2012. "Internet Governance: The Role of Multistakeholder Organizations." *World Wide Web Internet and Web Information Systems*. http://www.silicon-flatirons.org/documents/publications/report/InternetGovernanceRoleofMSHOrgs.pdf.