



KEEPING THE INTERNET
OPEN • INNOVATIVE • FREE

www.cdt.org

CENTER FOR DEMOCRACY
& TECHNOLOGY

1634 I Street, NW
Suite 1100
Washington, DC 20006

P +1-202-637-9800
F +1-202-637-0968
E info@cdt.org

Looking Back at P3P: Lessons for the Future

Ari Schwartz

November 2009

I. Introduction

A number of people who work on data protection have begun examining the idea of machine-readable statements that can express the privacy practices of a Web site or a third-party intermediary, such as a network advertiser or an analytics company.¹ The theory is that such statements would provide a clear, standardized means of rendering potentially complex privacy policies into a format that could be automatically parsed and instantly acted upon.

The idea is a good one. It harnesses the power of information technology to create a means for transparency and user choice. However, it is hard to overlook the fact that there is already a Web standard to do precisely the same thing, and it hasn't been very successful.

The Platform for Privacy Preferences (P3P) is a standard of the World Wide Web Consortium (W3C), the main standard setting body for the Web. P3P has never been fully implemented as its creators had hoped. While it is in use today and functions in some ways as we thought it might, P3P is unlikely to be broadly adopted or to accomplish all that those pushing for machine-readable policies would like.

This is not meant to suggest that using P3P is passé; or that creating new machine-readable standards based on P3P is a waste of time; or that creating interfaces that could be used for machine-readable policies is a fruitless exercise. In fact, the opposite is true. Machine-readable policies, like other PETs, hold considerable promise and deserve attention. However, to create machine-readable policies that work, we need to learn from how P3P was created and promoted, study its shortcomings, and draw from the immense amount of effort put into the project, where possible.

I worked actively on the P3P standard process and helped to promote its deployment from 1998 – 2003. During that time, we ran into many obstacles as we sought full-scale P3P implementation. This paper is meant to summarize the issues involved and my recommendations (political, economic and ethical) for those who would like to build and promote machine-readable privacy standards in the future.

¹ Ideas on machine-readable policies have been discussed at recent conferences such as the Privacy Bar Camp DC; the NYU privacy legislation symposium; the Engaging Data Forum at MIT and other events that I've attended. As recently as Winter 2009, companies have come to discuss this with the Center for Democracy and Technology (CDT) as if it were a completely new idea.

II. History

P3P has a long and complex history detailed by Carnegie Mellon Professor Lorrie Cranor in her book on privacy and P3P.² I will refrain from repeating this story and instead only focus on parts that are relevant to understanding the hurdles and achievements with P3P.

The theory behind P3P can be traced back to the mid-1990s. Many have claimed credit for the idea of using machine-readable policies for variety of different social purposes. This was just before the birth of XML and there was a realization that metadata would be useful for different purposes but few ideas how to make it a success in a public policy framework. As the privacy debate, in the United States and elsewhere, began to focus on encouraging companies to post human-readable privacy policies and as criticism increased about the complexity of those notices, there was a call to simplify them through standardization. If policies could be narrowed down to the equivalent of a multiple-choice set of options, then they could be made machine-readable.

After discussions about this theory at the Internet Privacy Working Group,³ the idea of P3P was passed to the main standards setting body for the Web, the World Wide Web Consortium (W3C). The W3C was charged with creating a P3P working group that would create the technical standards, the vocabulary, and the data schemas that would be used to make up the multiple choice questions. The W3C started its work on P3P in 1997 and the P3P Specification Working Group was chartered in July 1999.⁴

A. Building and Over-Building

Early on, as it became apparent that there were disparate views within the P3P Specification Working Group, it was decided that a set of “Guiding Principles” should be adopted to structure and inform future work. The principles adopted were as follows:

- *Information Privacy*
Service providers should preserve trust and protect privacy by applying relevant laws and principles of data protection and privacy to their information practices. Including:
 - *Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data*
 - *OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data of 1980*

² Lorrie Cranor, *Web Privacy with P3P*, O’Reilly, Sebastopol, CA, 2002.

³ The Internet Privacy Working Group (IPWG) is a forum of public interest groups, companies and academics convened by the Center for Democracy and Technology (CDT).

⁴ The P3P site has a history of all versions of the specification — <http://www.w3.org/p3p>.

- *US HEW Fair Information Principles of 1971*
- *Notice and Communication*
Service providers should provide timely and effective notices of their information practices, and user agents should provide effective tools for users to access these notices and make decisions based on them.
- *Choice and Control*
Users should be given the ability to make meaningful choices about the collection, use, and disclosure of personal information. Users should retain control over their personal information and decide the conditions under which they will share it.
- *Fairness and Integrity*
Service providers should treat users and their personal information with fairness and integrity.
- *Security*
While P3P itself does not include security mechanisms, it is intended to be used in conjunction with security tools. Users' personal information should always be protected with reasonable security safeguards in keeping with the sensitivity of the information.⁵

These principles helped resolve questions that arose about the intent of the standard.

Despite having this road map, the P3P specification changed dramatically over time. Pieces were added and then taken away. Professor Cranor has aptly compared the process to out-of-control construction on a kitchen that at first only needs a small new appliance (a toaster) but ends up with a plan for new cabinets, floors and lighting. Controversial ideas for negotiation, automated data transfer and others were added. Fortunately, discussions about the complications introduced by these additions — as well as the significant work required just to finish the vocabulary alone — led the group to cut back on all of these ideas and to more or less return to the original plan. However, a lot of time and effort was wasted debating these large-scale additions to the specification.

B. Caught Up in the Politics of Privacy

P3P had many critics when it was first created. At first, most of the concern came from some influential privacy advocates who believed that P3P was merely a ruse to stop greater regulation of the online industry. Later, concern came from traditional industry members that either did not want to have to implement P3P or that saw P3P was too transparent and therefore a threat to existing business models that consumers would disapprove of once they realized how their information was being used.

1. Criticized by some privacy advocates as an industry subterfuge

⁵ <http://www.w3.org/TR/NOTE-P3P10-principles>.

Early in its development, critics of P3P raised concerns that the standard was intended to stave off consumer privacy legislation in the United States and to allow companies to evade current law in the European Union.

The early decision to tie an automated data transfer standard, known as the Open Profiling System (OPS), to P3P was particularly damaging. A preliminary assessment from the Article 29 Working Party in the EU, written in July 1998 before the Specification Group was even formed, raised concerns about several issues including a fear that OPS would be used to negotiate away privacy protections girded by law.⁶

The legitimate concern that companies would use OPS to limit user choice was raised again and again, even after OPS was completely removed from the specification. When P3P was defended as merely one piece of a broader set of solutions in technology and law, many critics were still concerned. As librarian and activist Karen Coyle said in 1999: “Many people will not understand that ‘privacy practices’ are not the same as ‘privacy.’ P3P therefore allows sites to create an air of privacy while they gather personal data.”⁷

CDT worked with the Ontario Privacy Commissioner, Ann Cavoukian, and her staff to lay out the reasons why, once OPS was removed, a correctly implemented P3P actually could strengthen privacy. In 2000, we published a paper⁸ plainly stating that P3P was not a panacea for privacy. We emphasized that neither P3P nor any other privacy enhancing technology (PET) can solve all privacy issues. Instead, we argued, P3P needs to be used in concert with effective legislation, policy oversight and other privacy enhancing tools. We spelled out four ways in which P3P could help protect privacy:

1. *Countries with data protection and privacy laws and others seeking to police compliance with privacy standards could find the automated ability to assess a businesses' privacy statement useful in their broader oversight and compliance program* – Searching and gathering privacy policies could be simplified through P3P as P3P would allow these policies to be collected and analyzed in a standard machine-readable format. Governments and organizations would be able to simply search through P3P statements to find companies whose notice does not meet privacy standards in various areas. In the current version of P3P, companies could even point to regulatory bodies that oversee them to help route privacy complaints.
2. *Users could more easily read privacy statements before entering Web sites* – Privacy notices are frequently written in complicated legalese. P3P implementations could allow users to assess privacy statements prior to visiting a site, and allow users to screen and search for sites that offer certain privacy

⁶ http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/1998/wp11_en.pdf.

⁷ Karen Coyle, “P3P: Pretty Poor Privacy?” <http://www.kcoyle.net/p3p.html>.

⁸ P3P and Privacy: An Update for the Privacy Community, Ann Cavoukian, Mike Gurski, Deirdre Mulligan and Ari Schwartz, March 28, 2000
<http://www.cdt.org/privacy/pet/p3pprivacy.shtml>.

protections.

3. *P3P could cut through the legalese* – A company's P3P statement cannot use difficult to understand or unclear language. The standardization and simplification of privacy assertions into statements simple enough to be automated will allow users to have a clear sense of who does what with their information.

4. *Enterprising companies or individuals could develop more accurate means of rating and blocking sites that do not meet certain privacy standards or allow individuals to set these standards for themselves* – Creating the tools and knowledge that support products to rate and vet Web sites is difficult and time consuming. By providing an open standard, P3P could enhance the transparency, accuracy and detail of existing products, and could encourage an influx of new privacy enhancing products and services.

2. Cited by some industry advocates as a substitute for legislation

Unfortunately, several libertarian commentators and US politicians – even though they were not working on the specification— actively promoted P3P as a stand-alone solution, thereby reviving the concerns that CDT and the Ontario Privacy Commissioner had attempted to dispel. For example, in testimony before the Senate, the chairman of the US Federal Trade Commission (FTC) cited a Progress and Freedom Foundation report that suggested that 23 percent of Web sites had implemented P3P as a reason not to implement privacy legislation.⁹ The chairman neglected to mention that the report did not look into whether the P3P policies were compliant with the P3P standard, which many were not, and did not assess whether the policies actually offered privacy protections commensurate with either the European Data Protection Directive or the proposed standard in the bill that he was arguing against.

It is interesting to compare this reaction to that of the European officials who looked at P3P at the same time and correctly saw both its value and its limitations:

Among European Privacy Protection Commissioners the consensus grows: P3P is useful for online privacy, but not sufficient on its own because P3P offers only a basic standard for privacy protection. Under any circumstances, additional effective privacy monitoring and precise laws to protect Internet users are required.¹⁰

3. Criticized by some in industry for providing consumers too much transparency

⁹ Statement for the record of FTC Chairman Timothy Muris, S. Hrg. 107-1150, Hearing before the Senate Energy and Commerce Committee on S. 2201, the Online Personal Privacy Act, April 25, 2002. p. 11.

¹⁰ Independent Centre for Privacy Protection Schleswig-Holstein, Press Release on P3P, August 29, 2000 – http://www.datenschutzzentrum.de/somak/somak00/p3pe_pm.htm.

After the Specification Group started its work, many companies became increasingly concerned that P3P would empower users with too complete an understanding of how they were being tracked by companies. Most of these companies would only discuss these ideas behind closed doors, but at least one of the companies' analyses was made public.

Two Citibank employees published a paper expressing "concern that P3P would let ordinary users see, in full gory detail, how their personal information might be misused by less trusted or responsible web site operators."¹¹ This criticism from industry came up frequently in the P3P Working Groups. While a majority of the Working Group remained committed to the guiding principle of transparency, different companies ended up making different choices about how much they really wanted to be transparent with consumers. Two examples:

- A number of company argued that instead of only offering binary responses within the categories for types and uses of data, P3P should contain 3 options — Yes (we collect this data), No (we do not collect this data), Maybe (we may collect this data). The majority of the group felt strongly the binary yes/no option was important for transparency and that "Maybe" had to be treated as a "Yes" to be understood by consumers. One company, which had spent dozens of hours and thousands of dollars following the P3P process, was extremely insistent on this point and, in the end, never implemented P3P.¹²
- When P3P was finally implemented, a company that had worked on the specification complained, behind closed doors, that implementing the full specification would make them look bad and could stop users from accessing some of their sites. After realizing that implementing only part of the specification might leave them open to a charge of deceptive practices in the US and Europe, the company did implement a policy that was compliant with the specification.

There have also been many positive stories about companies that instituted new privacy-friendly policies when confronted with having to implement P3P. The transparency that P3P offers clearly had an impact on companies when they confronted the realization that P3P would make their privacy policies much more public.

¹¹ Kenneth Lee and Gabriel Speyer, "White Paper: Platform for Privacy Preferences Project (P3P) and Citibank" http://www.w3.org/P3P/Lee_Speyer.html.

¹² This point is reflected in early public comments from BITS, The Technology Group for The Financial Services Roundtable available at <http://www.w3.org/2002/p3p-ws/pp/bits.pdf>. And more strongly at <http://www.bitsinfo.org/downloads/Comment%20letters/W3CCommentLetter.pdf> — where BITS made clear that their specific goal was to try to make P3P statements as confusing as written statements are on the Web: "[O]ne of the most significant decisions of the P3P Working Group was not to enable use of the word "may" within the P3P nomenclature. We believe that the P3P nomenclature should enable verbatim translation of existing plain language policies, and that failure to incorporate that capability will materially affect the speed with which this standard is adopted in the marketplace."

C. Web sites build to the implementation, not the specification

Throughout 1998 and 1999, there was a lot of discussion about whether P3P had a “chicken and egg” problem. The concern was that P3P policies wouldn’t be created until there was implementation in a widely used consumer product such as a Web browser, but the browser implementation wouldn’t do anything until there were policies online. There was an effort to get many sites compliant, but until consumer products existed those efforts were not very successful.

In October 2000, after the second working draft of the P3P specification was released, several consumer products were created. Most notably, Microsoft built P3P capabilities into Internet Explorer 6. However, those features mostly focused on utilizing an optional part of the P3P specification called the “compact policy.” The compact policy takes all of the categories of information and all of the purposes for which they were used and ties them together, losing much of the subtlety that P3P full policies promised, but gaining an ability to read the policies more quickly. Internet Explorer 6 also put the strongest defaults on the use of “third-party cookies,” a term that is not even in the P3P specification. Microsoft decided not to utilize the main source of metadata — the full P3P policy as opposed to the compact policy — from P3P policies to help consumers control the release of their personal information based on what is actually happening to that data rather than an abstract summary offered by the compact policy. Because of these decisions, the P3P compact policies are in widespread use among companies that place third-party cookies demonstrating the power of a single implementation in the browser.

Unfortunately, there are still no good tools that make use of the metadata and this is why the main portion of the P3P specification is only used by a minority of Web sites today.

III. Recommendations for the Future

When thought of as an important experiment in categorizing privacy practices, P3P has been a qualified success. On the other hand, if the goal of P3P was either to protect the privacy of users on its own or, for the Internet industry, to stave off the threat of regulation, P3P should be viewed as an abject failure.

However, as a standard that works in conjunction with “additional effective privacy monitoring and precise laws to protect Internet users” (as the Independent Centre for Privacy Protection Schleswig-Holstein) current P3P implementations are a minor success and an indicator of what can still be accomplished with machine-readable policies.

Also, as a case study in the pitfalls and potentials of efforts to develop PETs, P3P is undeniably valuable. As new metadata standards for privacy are created based on P3P and as other PETs are explored, there are several lessons to learn from the P3P project experience:

A. Keep it simple

P3P is far too complex as it stands today.

For example, the standard includes 17 categories for data-type and 12 categories for data-use that Web sites can include in their meta-data; four of the data-use categories cover different types of profiling. There are many legitimate reasons that these categories exist,¹³ but the sheer number leads to far too many combinations and is overwhelming both for programmers and for Webmasters who would otherwise be interested in implementing P3P. Compliance is not difficult for a Web site with a clear and simple privacy policy, but many companies just aren't willing to put in the effort to understand all of the categories and purposes.¹⁴

In their book *Nudge: Improving Decisions about Health, Wealth and Happiness*,¹⁵ Richard Thaler and Cass Sunstein discuss the problems created by overwhelming choice and how to create workable options for individuals in policy and technology. Anyone interested in creating the next round of metadata technologies must read *Nudge* and consider how its recommendations on setting options and defaults would work in the particular context. My reading is that there should be no more than four options and the default should be set higher than average practices on the Internet today.

B. There is no “chicken and egg” problem: build the interface to use metadata first

Too much time and effort was spent trying to convince Web site operators of the value of implementing P3P on their Web sites. Either the market will work or direct regulation will dictate the value for the companies, or the idea will fail, but in no case is it possible for the developers of a concept like P3P to create critical mass of acceptance among Web sites – there are simply too many Web sites to convince to gain that critical mass. The evidence from the relatively successful implementation of P3P for cookies in Internet Explorer demonstrates the value of working with browser makers or with developers in other spaces that have ready access to direct user interfaces (as opposed to add-on tools) to implement solutions that utilize the metadata in ways that clearly benefit consumers. After these solutions are in place, companies will be forced to implement by the economics of having sites blocked or tagged.

C. Manage expectations: companies shouldn't use a metadata solution to argue for less regulation

¹³ In one telling example, when the White House was implementing P3P, officials there found that the specification did not have an option allowing them to express that they were required by law to store information for historical purposes. It was decided that many governments would have this same issue, so a “historical” purpose was added.

¹⁴ Professor Cranor suggests that both categories should be cut down to eight, which would be more manageable for programmers, but would still need to be cut down further by the programmers to be successful. She and her students use P3P policies to automatically generate a privacy “nutrition label” in the form of a table with 10 rows and 6 columns. This format hides some of the complexity of P3P by representing multiple P3P elements in a single row or column. <http://cups.cs.cmu.edu/privacyLabel/>

¹⁵ Richard Thaler and Cass Sunstein, *Nudge: Improving Decisions about Health, Wealth and Happiness*, Penguin, 2009.

If you are looking for a way to prevent over-burdensome privacy legislation or regulation and you believe that metadata tools are a means to accomplish this, you need to think again. Too many companies and trade associations spent more time arguing for the benefits of P3P in Washington and Brussels and too few spent effort building P3P into products. Perhaps at some point, widespread and effective use of metadata tools will justify a loosening of regulatory requirements, but even after adoption is completely ubiquitous, we would need testing and facts to prove that the technology was in fact effective. And of course, it would be just as easy to add metadata requirements into regulations for transparency as it would be to use them to prevent regulation. In fact, it would be a particularly inexpensive addition compared to the rest of the cost of data protection legislation. I am not advocating this approach as a solution as much as I am trying to point out that the development of PETs and the debates over regulation should take place on largely separate tracks, with participants checking in only to ensure that new regulations match the vocabulary in the metadata. Neither the development of PETs nor the regulatory debate will be well-served by those who engage in the PETs development process mainly to bolster their arguments against legislation.

D. Learn from the work that has been done on P3P

Finally, a lot of good work went into P3P. It is not a dead standard. Those who use third-party cookies regularly are implementing it now more than ever. However, it can be improved and it will need to be modernized in order to reach the original vision where the metadata of the full policy is parsed and used regularly. This could mean revamping P3P or it could mean developing something new. In any case, starting from scratch will only mean running into some of the same hurdles faced by the W3C P3P working groups. The history and work of P3P should be a launching place, not something to throw aside.