P3P User Agent Accuracy

Lorrie Faith Cranor[*] and Joel R. Reidenberg[†]

This paper is an edited excerpt of the discussion paper "Can user agents accurately represent privacy notices?" presented at TPRC2002. The full paper is available from http://intel.si.umich.edu/tprc/archive-search-abstract.cfm?PaperID=65 .

P3P user agent implementers need to find ways of simplifying the terms and definitions in the P3P specification in order to make their products useable by non-expert end users. As a result implementers often bundle together multiple elements and, then, describe those elements with a single term, with paraphrased definitions, or with jargon replaced by more readily accessible terms.[1] This simplification process is essential to usable product design, but reduces the precision of the P3P terms and may introduce some confusion.

P3P user agents may provide confusing and potentially misleading information when describing settings with insufficient detail or presenting summary information about P3P policies that is inaccurate, imprecise, or incomplete. This is an inherent problem for P3P user agents, particularly for those that have been designed in good faith to represent accurately P3P policies. The problem is also exacerbated because future P3P user agent implementations may be designed purposely to misrepresent privacy policies, or to represent them with a specific bias that arguably results in an inaccurate representation.

Some user agent implementers may wish to convey information that supports their particular agenda. For example, a marketing company might develop a P3P user agent that described telemarketing with positive sounding terms such as "personalized home shopping opportunity," while a privacy watch-dog group might label any site that does telemarketing as an "evil privacy invader." Hopefully, such a strong agenda would be obvious to users, who could decide whether the implementer's position matches their own. More troublesome, perhaps, would be a user agent in which the bias is more difficult to detect and would more likely result in users misunderstanding a web site's privacy policy. Imagine, for example, a user agent that provided a numeric rating—perhaps on a scale of 1 to 10—to describe a web site's P3P policy. The agent would necessarily make subjective judgments when determining the rating for each web site. Unless the strategy for assigning ratings was clearly explained to users, the user's assumptions

[*] Lorrie Faith Cranor <http://lorrie.cranor.org/> is a Principal Technical Staff Member at AT&T Labs-Research. She is chair of the P3P Specification Working Group at the World Wide Web Consortium and one of the creators of the AT&T Privacy Bird software. This paper represents her personal views.

[†] Joel R. Reidenberg <http://reidenberg.home.sprynet.com> is a Professor of Law at Fordham University School of Law. He participated as an invited expert in some of the deliberations of the P3P Working Groups.

[1] *See* Lorrie Faith Cranor, *Web Privacy with P3P* (O'Reilly and Associates, 2002), chapter 14.

about the meaning of a 10 rating versus a 1 rating may not match the user agent's actual behavior.[2]

Some companies have expressed concern about legal issues surrounding P3P policies that will be displayed to a user in a format not under corporate control.[3] To date, however, we have found few documented objections by companies to the way implementers have chosen to display P3P policies in their currently available implementations. Beyond problems with omitted data, and concerns about the readability of some of the phrases, only a few other specific concerns have been raised. One such concern relates to user agent wording. In particular, the issue is whether wording might suggest to users that web sites will routinely collect all of the types of data and use the collected data for all of the purposes mentioned in their P3P policy, or whether the wording suggests only that the data might be collected and might sometimes be used for these purposes.[4] Many of the definitions in the P3P specification contain the word "may" to express the later idea.[5] However, the word "may" is omitted in some of the information displayed in user agent policy summaries. For example, the beta release of the Privacy Bird includes the heading "How your information will be used." Some companies have suggested that this heading be changed to "How your information may be used."

Legal Implications

The accuracy of P3P implementations raises a number of significant legal issues for users, web sites and implementers. At the outset, inaccurate representations by user agents of a web site's privacy policy can undermine the main purpose of the P3P standard, namely the enabling of

[2] Indeed in 2000 a company called Enonymous hired people to rate thousands of web site privacy policies on a scale of 1 to 4. They provided user agent software that informed users about the ratings of the web sites they visited. This rating system proved controversial because the criteria Enonymous used for assigning a rating was not well documented and not applied uniformly by their human raters. A well-known privacy advocate praised the system, and then retracted his statement when he discovered that his own web site mysteriously had been assigned a poor rating. *See* Declan McCullagh, Odd Privacy Ratings Exposed, *Wired News* (12 April 2000) at http://www.wired.com/news/print/0,1294,35587,00.html.

[3] In a 15 October 2001 memo to the P3P Specification Working Group, Cheryl Charles, Senior Director of BITS (The Technology Group for The Financial Services Roundtable), expressed the concern "There are potential conflicts between how a P3P implementation characterizes site behavior and a company's own plain language privacy policy, which could appear to lead to charges of bad faith," at http://lists.w3.org/Archives/Public/www-p3p-public-comments/2001Oct/0015.html.

[4] Representatives of several companies have raised this concern in personal conversations with one of the authors.

[5] The word "may" was added to many of the definitions in the P3P specification after one company complained about this problem. They actually proposed that the P3P specification include separate "will" and "may" elements so that companies could distinguish between routine and occasional practices. However, the consensus of the working group was that such a distinction would not be useful to end users. See section 3.3.4 of 18 October 2000 P3P 1.0 Working Draft at http://www.w3.org/TR/2000/WD-P3P-20001018/ ("Note, that the working group discussed at length the possibility of allowing sites to distinguish between purposes they may engage in and purposes they will engage in. The consensus of the working group was that such a distinction is not necessary. However, some members disagreed with this conclusion stating: Yes, no and may all need to be response options in the vocabulary. If no and may are the only options, then the meaning of may is corrupted to equal yes….")

automated decisions based on notice of information practices and consent by the user. The validity of automated actions or agreements between web sites and users for the use of personal information is jeopardized by inaccurate implementations. Yet, even if agreements and automated actions are considered valid, an inaccurate translation of a web site's policy creates ambiguity as to the applicable privacy protections for the personal information. Are the applicable protections those understood by the user based on the user agent's unfaithful translation or those actually disclosed by the web site? From the user's perspective, the inaccurate portrayal of a site's policy might appear as fraud or deception. Should the web site be responsible for this problem or might the implementer bear liability? Web sites are also sensitive that their privacy policies not be portrayed to users in a false light. If an implementation unfairly tarnishes a web site, could the implementer be held liable for defamation?

Although established legal principles offer some guidance for these questions, the novelty of P3P and its user implementations do not fit precisely into the traditional doctrine.

Solutions

In essence, the technological mediation by software agents that is designed to ease the ability of users to understand the privacy practices of web sites risks adding ambiguity, confusion and legal uncertainty. The potential confusion and legal uncertainty demonstrates that the trustworthiness of agents is critical for the success of technical solutions for the protection of privacy. Trustworthiness of agents is in the mutual interests of both consumers and web sites. Several approaches may assure consumers and web sites of the trustworthiness of agents as well as provide implementers with greater legal certainty against potential liability claims.

User agent documentation

The first step in promoting accuracy is comprehensive user agent documentation. User agents without any documentation conceal how the user agent represents web site policies.[6] Transparency will reveal the simplification and choices made by the implementer. Comprehensive documentation enables users and web sites to interpret the meaning of descriptions generated by the agent and to understand how the agent reacts. Transparency also provides an incentive for the implementer to make reasonable choices and creates an opportunity for feedback from sites and users who believe that the choices distort P3P statements.

Transparency will not, however, assure that the simplification and choices faithfully translate P3P statements. Transparency will only expose issues to those sophisticated users and web sites who review the documentation and who examine how particular design choices interact with specific sites, P3P statements and agent configurations.

[6] Software agent researchers acknowledge the importance of agent documentation. *See* Batya Friedman and Helen Nissenbaum, Software Agents and User Autonomy, *Proceedings of Autonomous Agents 97* at http://doi.acm.org/10.1145/267658.267772 ("Sometimes, in order to use the services of an agent as desired, a user must know how the agent goes about its task. When the designer of a software agent does not make information accessible to the user, then the user's autonomy can be undermined.")

Certification

A more complete and direct way to assure that user agents accurately represent web site P3P policies is through certification. Certification benefits users and web sites by assuring that the parties interact on the basis of a shared understanding of the terms for the use of personal information. Certification further benefits web sites by the assurance that their policies will not be misrepresented to users. Implementers can benefit by the assurance that the design choices do not potentially cause the defamation of web sites.

Certification, though, raises its own set of issues. The establishment of criteria to evaluate accuracy is likely to be difficult. While there are some things that P3P user agents might do that would be readily identified as inaccurate or misleading, there is a large gray area in which user agents might present factual information side-by-side with subjective judgments that may be deemed misleading by some people, but not others. The fact that consumers and businesses often have different views about what constitutes a good privacy policy increases the likelihood that developing a standard for user agents would be a controversial endeavor.

The choice of one or more certifying authorities presents further difficult practical and political decisions. One option is to rely on the market and self-regulatory organizations to provide certification. This option necessitates that the organizations providing the certification also be sufficiently trustworthy and that the measurement criteria be valid. An opposing option is to rely on a government agency, like the Federal Trade Commission, to certify the accuracy of P3P products. This alternative may be cumbersome and also requires a valid set of measurement criteria. A combination of these two options may be the most promising way to proceed. A potential model already exists for this combined approach under the Children's Online Privacy Protection Act.[7] COPPA allows groups to submit privacy implementing guidelines for approval by the Federal Trade Commission. The FTC has established criteria for approval of these guidelines.[8] Once approved, the guidelines provide a "safe harbor" establishing compliance with COPPA. For P3P implementations, the FTC may through a public proceeding set out criteria for the accuracy of software agents.[9] Independent organizations would then have a trustworthy benchmark or alternatively, the FTC itself might issue safe harbor notices for compliance with the criteria of accuracy.

Finally, once the criteria and the certifying authority are settled, the certification of a user agent may still be difficult to assess in an objective manner. The divergent perspectives of businesses and consumers may affect the measurement of how a user agent behaves.

User Agent Guidelines

If the establishment of a certifying authority proves infeasible, the development of a set of requirements, guidelines, or best practices for user agents might go a long way towards

[7] 15 U.S.C. §§ 6501-6506

[8] See 16 CFR Part 312

[9] The FTC's jurisdictional authority would be based on its "unfair and deceptive practices" authority under 15 U.S.C. §45(a).

improving user agent reliability. Such guidelines might even be used as part of a self-certification system. For example, the W3C has developed accessibility guidelines that web site developers and user agent implementers often use for self-assessment.[10] Companies sometimes refer to their compliance with these guidelines when advertising their products.

As discussed earlier, the definitions in the P3P specification were not written with end users in mind. Thus, user agent implementers are faced with the choice of presenting users with information that will be difficult for them to understand, or making their own decisions about the translation of the definitions into user-friendly terminology with the risk that the translations will not be completely accurate. P3P user agent guidelines might offer a standard set of user-friendly terms for representing the P3P vocabulary in English (or perhaps several natural languages). User agent implementers would likely welcome the opportunity to adopt standard user-friendly terms and avoid the risks associated with inventing their own. However, without a certification authority, implementers may choose to implement user agent guidelines only partially.

Selecting the most appropriate organization and process to develop P3P user agent guidelines is very important for the eventual credibility of any resulting guidelines. As the organization that developed the P3P specification, W3C seems like a logical candidate. Although W3C is not setup to take on the role of a certifying authority, W3C does have experience developing guidelines. However, to develop good guidelines for P3P user agents will likely require participation by usability experts, as well as require conducting user studies, for example to ensure that proposed user-friendly terminology is actually understandable by end users. This would be a new type of undertaking for the W3C and might require considerable resources to complete.

Outside the P3P development efforts, US industry groups have begun talking about the desirability of offering a “privacy nutrition label” or a standard format “short notice” or “layered notice.” Much of this discussion has been in response to criticism that financial privacy notices are lengthy and difficult to read.[11] One group is also investigating ways of using P3P as part of this effort.[12] This or a similar effort could potentially result in the development of a standard mapping of P3P vocabulary elements into a user-friendly format. Such a mapping could serve as a set of guidelines for P3P user agent implementers. However, the often narrow special interest focus will limit the appeal of resulting guidelines. Similarly, it is also uncertain whether industry alliances will have sufficient representative expertise and the commitment to undertake such an effort.

[10] The World Wide Web Consortium’s Web Accessibility Initiative has developed guidelines for developing web browsers and web content that will be accessible to users with a variety of needs. See http://www.w3.org/WAI/.

[11] For example, The Center for Information Policy Leadership, a division of the Hunton & Williams law firm, has created an industry group to develop a proposed short notices “template.” The Center has conducted focus groups to vet their proposed template with consumers.

[12] *See Privacy Regulation Report*, With Industry Divided Over Layered-Notice Approach, Privacy Group Readies “Template”, 5 August 2002.