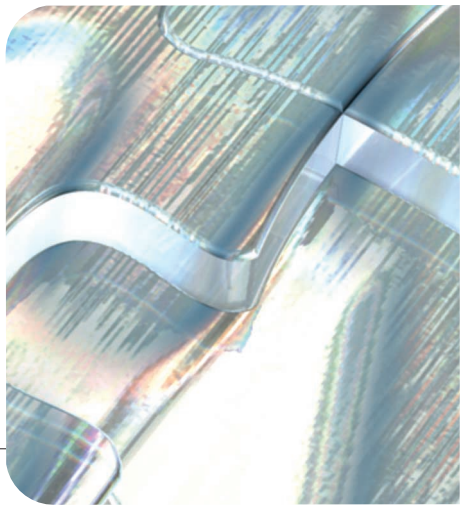
P3P: Making Privacy Policies More Useful

The World Wide Web Consortium's platform for privacy preferences (P3P) lets Web sites convey their privacy policies in a computer-readable format. Although not yet widely adopted, P3P promises to make Web site privacy policies more accessible to users.



Lorrie Faith Cranor
AT&T Labs-Research

Many Web sites now post privacy policies. Consumer and industry groups encourage such policies, and some jurisdictions even require them by law. Essentially, they aim to tell Web site visitors how sites will use their personal information and what privacy choices are available to them. Thus, individuals should be able to gather the information they need to use privacy-related options at the sites they visit.

Unfortunately, although many sites post privacy policies, few visitors read them. Numerous studies have shown that consumers find privacy policies time-consuming to read and difficult to understand, and readability experts have found that comprehending privacy policies typically requires college-level reading skills.[1–3] In addition, privacy policies have no standardized format, making it difficult to compare them. Consumers who do read these policies are also frustrated by the fact that they can change unexpectedly.

In April 2002, the World Wide Web Consortium published the platform for privacy preferences,[4] which specifies a standard computer-readable format for Web site privacy policies. P3P-enabled Web browsers read policies published in P3P format and compare them with user-specified privacy settings. Thus, users can rely on their agents to read and evaluate privacy policies on their behalf. Furthermore, the standardized multiple-choice format of P3P policies facilitates direct comparisons between policies and the automatic generation of standard-format human-readable privacy notices.

Both Microsoft Internet Explorer 6 and Netscape Navigator 7 have built-in P3P functionality, and several P3P user agents and Web tools exist, but P3P adoption has been slow. Although the first generation of P3P user agents has limited functionality and suffers from several user interface problems, early adopters have found these tools useful, and usability test subjects found it significantly easier to answer questions about a Web site's privacy policy using a P3P user agent than by reading the policy.[5,6] As P3P gains wider adoption and better P3P user agents become available, Web site privacy policies might finally become useful.

How P3P works

The P3P 1.0 specification defines a standard way of encoding Web site privacy policies in an XML format, as well as mechanisms for locating and transporting P3P policies.

P3P policies have eight major components (described in the "P3P major components" sidebar), most of which contain multiple subcomponents and attributes. XML elements represent each component. For example, a *purpose* element represents collected data use. The specification defines 11 purpose subelements, each representing a data use. In addition, each purpose subelement has a *required* attribute that indicates whether the data can be used for this purpose all the time, on an opt-in basis, or on an opt-out basis.

A P3P *statement* comprises the purpose, data, recipients, retention, and consequence elements. A P3P policy contains one or more statements. Sites use the statement structure to indicate types of data that are treated similarly. For example, a site might have one statement to describe the information it stores in log files and another to describe the information it collects from individuals who make purchases at the site. Figure 1 shows a P3P policy for a site with a relatively simple privacy policy with only one statement.

P3P major components

The Platform for privacy preferences specification has eight major components, most with multiple subcomponents and attributes:

- *Entity* lists contact information for the business, organization, or person who owns the site.
- *Access* states whether individuals can find out what personal data a site keeps about them in its databases (six types of access policies exist).
- *Disputes* describes how to resolve privacy-related disputes with the site (customer-service desk, privacy seals, relevant privacy laws, and so on); it includes the *remedies* subelement.
- *Data* lists the kinds of data collected (17 data category elements and dozens of specific data elements exist).
- *Purpose* states how collected data is used and whether individuals can opt in or out of any of these uses (11 types of purposes and one *other-purpose* exist; each can take a *required* attribute).
- *Recipient* states whether and under what conditions data can be shared and whether there is an opt in or out (six types of recipient policies exist; each can take a required attribute).
- *Retention* states policies for periodic purging of collected data (five types of retention policies are specified).
- *Consequence* provides a human-readable explanation of site's data practices.

Purpose, data, recipients, retention, and consequence are part of P3P's statement structure. Sites use this structure to indicate the types of data they treat similarly.


```
<POLICIES xmlns="http://www.w3.org/2002/01/P3Pv1">
<POLICY discuri="http://p3pbook.com/privacy.html"
        name="policy">
  <ENTITY>
  <DATA-GROUP>
    <DATA ref="#business.contact-info.online.email">privacy@p3pbook.com
    </DATA>
    <DATA ref="#business.contact-info.online.uri">http://p3pbook.com/
    </DATA>
    <DATA ref="#business.name">Web Privacy With P3P</DATA>
  </DATA-GROUP>
  </ENTITY>
  <ACCESS><nonident/></ACCESS>
  <STATEMENT>
    <CONSEQUENCE>We keep standard web server logs.</CONSEQUENCE>
    <PURPOSE><admin/><current/><develop/></PURPOSE>
    <RECIPIENT><ours/></RECIPIENT>
    <RETENTION><indefinitely/></RETENTION>
    <DATA-GROUP>
       <DATA ref="#dynamic.clickstream"/>
       <DATA ref="#dynamic.http"/>
    </DATA-GROUP>
  </STATEMENT>
</POLICY>
</POLICIES>
```

Figure 1. Example P3P policy. This relatively simple policy contains one statement, comprising purpose, data, recipients, retention, and consequence elements.

The P3P 1.0 specification also includes syntax for a P3P *compact policy*—an abbreviated version of an XML P3P policy that describes a Web site's data practices with respect to cookies. Compact policies consist of combinations of three-letter tokens, many of which can be modified by a compact version of the required attribute. Fifty-two such tokens are specified. P3P-enabled Web sites use these optional compact policies to facilitate rapid cookie-blocking decisions.

P3P *policy reference files* are XML-encoded files that indicate the parts of a Web site to which a P3P policy applies. These files specify the location of one or more P3P policies and a URL or set of URLs to which each applies. Most Web sites place their policy reference files at a standard well-known location: /w3c/p3p.xml. A P3P user agent makes an HTTP `GET` request for the file to learn the location of a site's P3P files. After parsing the file, the user agent makes additional `GET` requests to obtain P3P policy files. Because P3P policies generally apply to many (or all) URLs on a site, the user agent

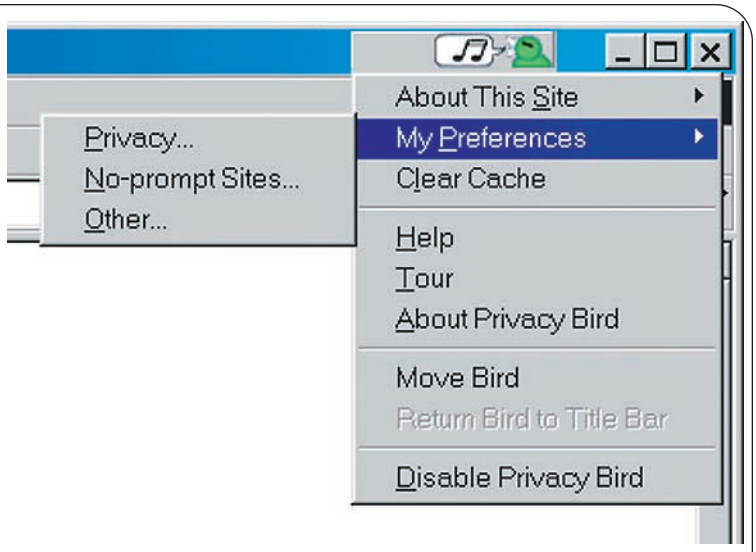


Figure 2. The Privacy Bird's "green bird" icon and My Preferences menu. By turning green, the icon shows that the site's policies match the user's preferences.

doesn't need to fetch the files every time the user requests a new page on a site. By default, P3P files have a lifetime of 24 hours, so if a user returns to a site within one day, the agent doesn't need to fetch them again. Another option is to embed P3P policies in policy reference files, which simplifies site administration and reduces the number of round trips necessary to retrieve P3P files from a site.

Although the vast majority of Web sites use the well-known location, P3P 1.0 supports two additional mechanisms for locating policy reference files. Webmasters can place the files at arbitrary locations on their sites and reference them through links embedded in HTML content or in special P3P HTTP headers. P3P HTTP headers can also transport P3P compact policies.

A separate W3C specification—A P3P Preference Exchange Language (APPEL)—provides syntax for encoding user preferences about privacy policies.[7] APPEL is a rule-based language encoded in XML. P3P user agents can compare APPEL-encoded preferences with a P3P policy to determine whether a site's policy matches a user's preferences. However, P3P user agents are not required to use APPEL, which is not an official W3C recommendation and is considered somewhat experimental. AT&T's Privacy Bird and several other P3P software implementations use APPEL, however.

P3P software and services

A variety of software tools and services to support P3P exist, including P3P user agents, P3P editors and validators, and services that help P3P-enabled Web sites.

User agents

Because both Microsoft IE6 and Netscape Navigator 7 Web browsers include basic P3P functionality, most Windows users are probably already using P3P user agents. However, whenever I speak to an audience about privacy, I often ask for a show of hands to see how many people using these browsers are aware of the P3P features. Although typically most audience members use the browsers, very few know of the P3P features. Thus, most of them are apparently using P3P features at their default settings without customizing them to reflect their personal privacy preferences.

IE6 automatically checks the HTTP headers sent with cookies for P3P compact policies. Under its default setting, IE6 blocks cookies without compact policies set by a third-party Web site—that is, if the cookies are associated with an advertisement or other content embedded in a Web page served from a domain different from that of the page in which it is embedded. IE6 blocks or restricts other cookies depending on the compact policy and the user's cookie settings. A small icon featuring a picture of an eye with a do-not-enter sign appears in the lower right corner of the browser window when a cookie is blocked or restricted. Although users might not notice this icon and might be unaware of the P3P feature, Web sites that set third-party cookies are increasingly aware of it. When blocked cookies start interfering with their sites' functionality, many Webmasters quickly add P3P and compact policies to their sites.

IE6 also offers a *privacy report* feature that users can select from the browser's view menu. Selecting this feature causes the browser to check for a site's full P3P policy. If the browser can fetch the policy, it parses the XML and displays a human-readable representation of the policy.

Navigator 7 P3P features are similar to IE6's, but it uses a slightly different cookie interface and default settings. Netscape can also generate a human-readable version of a site's P3P policy. A Netscape P3P policy representation is shorter and uses sentence fragments and bulleted lists whereas IE6 uses complete sentences and paragraphs.

AT&T's Privacy Bird is an IE5/IE6 add-on freely available at http://privacybird.com. Once installed, a bird icon appears on the right side of the title bar, as Figure 2 shows. Privacy Bird checks for P3P policies for all content in a page at every site a user visits and compares them with the user's privacy preference settings, configured through a menu accessed by clicking on the bird. When a site's policies match a user's privacy preferences, the bird icon turns green; when they don't match, the icon turns red. When a site isn't P3P-enabled, the icon turns yellow. Symbols in the bird's song "bubble" also help distinguish the three icons. Moreover, users can configure Privacy Bird to play distinctive sounds corresponding to each icon.

Privacy Bird can also generate and display a human-

readable version of a site's P3P policy. Like the Netscape version, it uses short phrases and bulleted lists.

Privacy Bird offers more configuration options than IE6 and Netscape 7, and it allows users to import APPEL preference files. However, Privacy Bird (version beta 1.2) has no cookie-blocking capabilities.

My colleagues and I conducted a laboratory study involving 12 experienced IE users who had never used Privacy Bird or IE6's P3P features. After training them to use these tools, we asked our subjects to visit Web sites and answer questions about their privacy policies using three techniques: Privacy Bird, IE6's P3P features, and the human-readable policies.

We observed the subjects as they performed these tasks and questioned them about their experiences after they finished. They reported that finding information was significantly easier using a P3P user agent than by reading Web site privacy policies. Of the two P3P user agents they tried, our subjects found Privacy Bird to be the most useful and easiest to use to find and understand privacy policy information.[5]

Future P3P user agents might be built into electronic wallets, search engines, and other tools. For example, my colleagues and I are currently developing an experimental P3P-enabled search engine that lets users sort their search results so sites that match both their search criteria and their privacy preferences appear first. Tools such as this will let users more easily compare privacy policies at similar Web sites and identify sites with acceptable policies.

Web site tools

Several companies have developed software and services to help Webmasters P3P-enable their sites (see www.w3.org/P3P/implementations).

One of the most popular P3P tools for Web sites is the P3P Policy Editor, offered as a free download from IBM Alphaworks (www.alphaworks.ibm.com/tech/p3peditor). This tool includes a graphical user interface with which users can visually construct P3P policies. The tool generates XML-encoded P3P policies, as well as a human-readable version of each policy. It also generates policy reference files and compact policies, and explains how IE6 will respond to a particular compact policy. A tutorial on the P3P Policy Editor is available in my book, *Web Privacy with P3P*.[8]

W3C maintains the P3P Validator at www.w3.org/P3P/validator.html. Users type in a URL, and the Validator checks to see whether the site is properly P3P enabled. It checks to make sure P3P files use the correct syntax and are in the proper locations on the site. The Validator is free and easy to use, yet the number of errors I've found in P3P policies suggests that few Webmasters have tried it.

P3P adoption

Several surveys have sought to assess P3P adoption. Checking every Web site in existence for P3P compliance isn't feasible, nor is it a particularly useful metric. My brother's dog has a P3P-enabled Web site, but given the low number of visitors to the site, the fact that it's P3P-enabled isn't very significant. What's more interesting is the fraction of P3P-enabled sites among the most popular sites on the Internet.

My colleagues and I developed software to automatically check a list of Web sites for P3P compliance.[9] In July 2003, we found that 30 percent of the top 100 sites and 23 percent of the top 500 sites were P3P-enabled. Six months earlier, Ernst & Young performed a similar check manually and found that 18 percent of the top 500 sites were P3P-enabled. In general, the most popular sites are the most likely to be P3P-enabled.

Although these numbers demonstrate a slow adoption rate, the adoption levels are substantial for a specification that was published a little over a year earlier. Unfortunately, few researchers have tracked adoption rates for other Web standards, so it's difficult to know how P3P adoption rates compare.

One cause for concern, however, is the large number of sites with P3P errors. Of the 588 P3P-enabled sites we identified during our study, the P3P Validator flagged errors in about a third of them. Many of the errors were relatively minor and didn't interfere with a P3P user agent's evaluation of the site's P3P policy. However, 6 percent of the P3P-enabled sites had more substantial errors that prevented Privacy Bird from evaluating them.

Errors in the implementation of Web-related standards are quite common. For example, over a year after the release of HTTP/1.1, a study found that many Web servers failed various compliance tests.[10] P3P errors arguably have more severe legal and policy-related consequences than errors in the HTTP standard's implementation. While HTTP errors can result in less efficient Web transactions and even occasional server crashing, P3P errors can result in misrepresented privacy policies and misled users.

In addition to monitoring the extent of P3P adoption, our study assessed the types of privacy policies P3P-enabled Web sites were adopting. Our findings uncovered many things:

- The policies of about half the P3P-enabled sites we examined matched the privacy preferences represented by Privacy Bird's medium setting, three-quarters matched its low setting, and less than 20 percent matched its high setting.

Our subjects found Privacy Bird to be the most useful [user agent].

P3P-related URLs

- "A Webmaster's Guide to Troubleshooting P3P": www.oreillynet.com/pub/a/javascript/2002/11/19/p3p.html
- "Help! IE6 Is Blocking My Cookies": www.oreillynet.com/pub/a/javascript/2002/10/04/p3p.html
- Joint Research Center P3P Resource Center: http://p3p.jrc.it/
- P3P information from W3C: www.w3.org/p3p/
- P3P Toolbox: http://p3ptoolbox.org/
- Privacy Features in IE6: http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnpriv/html/ie6privacyfeature.asp

- About half the P3P-enabled sites indicated they might share data with third parties beyond their agents and delivery companies. Of these, about 40 percent report letting users opt in or out of this sharing.
- About two-thirds of the P3P-enabled sites indicated they might contact individuals for marketing purposes, and about one-third indicated this contact might be via telephone. About 70 percent of the sites that said they might contact individuals for marketing claimed to offer opt in or out choices.

We plan to continue monitoring P3P adoption and tracking trends in the types of policies P3P-enabled sites offer.

The road ahead

As with most technology, no sooner was P3P 1.0 released than developers and Web site operators began requesting the next version. In spring 2003, W3C launched the P3P 1.1 Specification working group to fix several minor errors in P3P 1.0, add a few new features, and make a few relatively minor changes. As our work progresses, we are attempting to keep our changes backward compatible with P3P 1.0 so P3P-enabled Web sites won't have to change their P3P policies to be readable by P3P 1.1 user agents. We hope to complete our work by summer 2004.

One major new component in P3P 1.1 will be a set of guidelines for P3P user agent implementers, including a set of recommended "plain English" translations of P3P element definitions. The P3P 1.0 specification provides a detailed definition of every XML element that can be included in a P3P policy. Because these definitions were never intended for end users, every user agent implementer has independently decided how to convey information about the elements to users. Consequently, P3P user agents have different ways of displaying the same information, and Web site administrators don't know how their P3P policies will be displayed. Hopefully the working group will agree on a standard set of user-friendly strings, and user agent implementers will adopt them.

Another major component of the P3P 1.1 work is exploring how P3P can be used with other emerging Web technologies such as Web services. In some cases, the P3P working group might recommend interoperability guidelines to developers or other W3C working groups; in others, it might specify new P3P features to facilitate P3P's use with other standards.

We have also received many suggestions for more radical changes to the P3P specification, some of which might be incorporated into a P3P version 2. Although the W3C currently has no definite plans or timetable for this work, many researchers are pursuing projects aimed at exploring more long-term P3P changes. Ideas under consideration include developing a preference language to replace APPEL[11] and adding features that let users consent to a policy or opt in or out of various data practices. Some researchers are also exploring the possibility of a rich negotiation framework that would let user agents negotiate privacy policies with Web sites. Other work focuses on integrating P3P into back-end systems to ensure that promises made in privacy policies are kept in practice.[12,13]

P3P in context

The P3P development process spanned five years and involved dozens of individuals from around the world. Throughout this process, P3P has been somewhat controversial and has received much criticism. Many companies and industry groups have voiced concern about the extent to which P3P would require sites to make disclosures they might not be legally required to make. They've also worried that posting P3P policies might open companies to additional liability.

In addition, some privacy advocates have raised concerns that P3P won't improve privacy protection for individuals. Despite the international composition of the P3P working group, others have claimed that P3P had too much of an American focus.

Part of the reason the P3P development process took so long is that the working group attempted to address as many of these diverse concerns as possible. Harry Hochheiser provides an extensive critique of P3P that references critics' comments and P3P proponents' responses.[14]

One criticism repeatedly leveled against P3P is its failure to address fair information practices. Indeed, P3P was never intended to be a comprehensive privacy "solution" that would address all FIP principles. P3P focuses squarely on increasing the transparency of Web site privacy practices—often referred to as the *notice* principle or *notice and choice*. Early drafts included a protocol that let user agents negotiate with Web sites, allowing Web site visitors to directly exercise choice options. This was eventually dropped in favor of a simpler protocol.

Current P3P user agents provide direct links to pages on P3P-enabled Web sites where users can opt in or out. P3P does not automate the exercising of these options, however. Rather, it standardizes privacy policies to allow

for automated comparisons and consistent display. Thus, individuals can gather information relevant to determining compliance with other FIPs.

Because P3P exposes Web site privacy practices in a way that makes it easier for individuals to understand them and identify sites with objectionable practices, P3P might lead to other privacy improvements, such as reductions in the amount of information collected or secondary uses of that information. Joseph Turow argues that the value of standardized computer-readable privacy policies is so great that Web sites should be required by law to use P3P.[3]

Of course, as critics have argued, other approaches to data privacy protection—such as laws that limit secondary data use—might be more effective than either voluntary or mandatory adoption of P3P. Critics point out that greater transparency doesn't guarantee the availability of meaningful privacy-friendly choices.[14] Many prefer legislative approaches that would force companies to offer privacy-friendly options. There is by no means universal agreement, however, even among privacy advocates, as to what approaches will most effectively increase data privacy protections. It seems unlikely that the US will adopt strict privacy laws anytime soon, and countries that enjoy more legal privacy protections find that enforcing these laws is often problematic. Thus, most P3P proponents advocate a multipronged agenda that includes P3P adoption and other regulatory and self-regulatory approaches.

One of P3P's major goals has been to help users learn about Web site privacy policies without having to read lengthy privacy policies at every Web site they visit. As more sites adopt P3P and new P3P user agents emerge, we are starting to realize this goal. We are also beginning to see some anticipated secondary effects. Because sites must choose between multiple-choice options when writing a P3P policy, they sometimes make clearer and more explicit statements in their P3P policies than they made in their human-readable privacy policies. Some sites have improved their privacy policies to look better when displayed by P3P user agents or avoid having their cookies blocked by IE6. Ideally, the increased transparency brought about by P3P will result in more policy improvements and facilitate more informed debate about the effectiveness of regulatory and self-regulatory privacy programs.[15] □

References

1. Harris Interactive, *Privacy Leadership Initiative: Privacy Notices Research Final Results*, Dec. 2001; www.ftc.gov/bcp/workshops/glb/supporting/harris%20results.pdf.
2. W. Rodger, "Privacy Isn't Public Knowledge: Online Policies Spread Confusion with Legal Jargon," *USA Today*, 1 May 2003, p. 3D.
3. J. Turow, *Americans & Online Privacy: The System is Broken*, Annenberg Public Policy Ctr., Univ. of Pennsylvania, June 2003; www.asc.upenn.edu/usr/jturow/internet-privacy-report/36-page-turow-version-9.pdf.
4. L. Cranor et al., *The Platform for Privacy Preferences 1.0 (P3P1.0) Specification*, World Wide Web Consortium Recommendation, April 2002; www.w3.org/TR/P3P/.
5. L. Cranor, P. Guduru, and M. Arjula, "User Interfaces for Privacy Agents," submitted for publication, 2003.
6. L. Cranor, M. Arjula, and P. Guduru, "Use of a P3P User Agent by Early Adopters," *Proc. ACM Workshop on Privacy in the Electronic Society,* ACM Press, 2002, pp. 1–10.
7. L. Cranor, M. Langheinrich, and M. Marchiori, "A P3P Preference Exchange Language 1.0 (APPEL 1.0)," World Wide Web Consortium working draft, April 2002; www.w3.org/TR/P3P-preferences.html.
8. L. Cranor, *Web Privacy with P3P,* O'Reilly & Associates, 2002; http://p3pbook.com/.
9. S. Byers, L. Cranor, and D. Kormann, "Automated Analysis of P3P-Enabled Web Sites," *Proc. 5th Int'l Conf. Electronic Commerce* (ICEC2003), ACM Press, 2003; http://lorrie.cranor.org/pubs/icec03.html.
10. B. Krishnamurthy and M. Arlitt, "PRO-COW: Protocol Compliance on the Web—A Longitudinal Study," *Proc. Usenix Symp. Internet Technologies and Systems* (USITS 2001), Usenix, 2001, pp. 109–122; www.usenix.org/events/usits01/krishnamurthy.html.
11. R. Agrawal et al., "An XPath-based Preference Language for P3P," *Proc. 12th Int'l Conf. World Wide Web,* ACM Press, 2003, pp. 629–639.
12. P. Ashley, C. Powers, and M. Schunter, "From Privacy Promises to Privacy Management: A New Approach for Enforcing Privacy Throughout an Enterprise," *Proc. 2002 Workshop on New Security Paradigms,* ACM Press, 2002, pp. 43–50; http://doi.acm.org/10.1145/844102.844110.
13. G. Karjoth, M. Schunter, and E. Van Herreweghen, "Translating Privacy Practices into Privacy Promises—How to Promise What You Can Keep," *Proc. 4th IEEE Int'l Workshop on Policies for Distributed Systems and Networks,* IEEE Press, 2003, pp. 135–146.
14. H. Hochheiser, "The Platform for Privacy Preference as a Social Protocol: An Examination Within the US Policy Context," *ACM Trans. Internet Technology,* vol. 2, no. 4, 2002, pp. 276–306.
15. L. Cranor and R. Wenning, "Why P3P Is a Good Tool for Consumers and Companies," *GigaLaw.com,* 2002, www.gigalaw.com/articles/2002/cranor-2002-04.html.

Lorrie Faith Cranor *is chair of the platform for privacy preferences project (P3P) specification working group at the World Wide Web Consortium. Her research has focused on a variety of areas in which technology and policy issues interact, including online privacy, electronic voting, and spam. Cranor received a PhD in engineering and policy from Washington University in St. Louis. She is a member of the ACM, the IEEE, and CPSR. For more information about her research, visit her Web site at http://lorrie.cranor.org.*