Beyond Concern: Understanding Net Users' Attitudes About Online Privacy

17 August 1999

Lorrie Faith Cranor
AT&T Labs-Research
Shannon Laboratory
Florham Park, NJ 07932
lorrie@acm.org
http://www.research.att.com/~lorrie/

Joseph Reagle *
World Wide Web Consortium
Massachusetts Institute of Technology
Cambridge, MA 02139
reagle@w3.org
http://www.w3.org/People/Reagle/Overview.html

Mark S. Ackerman **
Information and Computer Science
University of California, Irvine
Irvine, CA 92697
ackerman@ics.uci.edu
http://www.ics.uci.edu/~ackerman/

* Some of this work was conducted while the author was a Resident Fellow at the Harvard Law School Berkman Center for Internet and Society.

** Some of this work was conducted while the author was a visiting researcher at AT&T Labs-Research.

http://www.research.att.com/projects/privacystudy/

Introduction

Over the past decade, numerous surveys conducted around the world have found consistently high levels of concern about privacy. The more recent studies have found that this concern is as prevalent in the online environment as it is for physical-world interactions. For example, Westin (Harris 1998) found 81% of Net users are concerned about threats to their privacy while online. While many studies have measured the magnitude of privacy concerns, it is still critical to study the concern in detail, especially for the online environment. As Hine and Eve (1998) point out:

> Despite this wide range of interests in privacy as a topic, we have little idea of the ways in which people in their ordinary lives conceive of privacy and their reactions to the collection and use of personal information (Hine and Eve 1998, 253)

With this study, we have tried to better understand the nature of online privacy concerns; we look beyond the fact that people are concerned and attempt to understand what aspects of the problem they are most concerned about. We hope our results will help inform both policy decisions as well as the development of technology tools that can assist Internet users in protecting their privacy.

This insight should be helpful to ongoing privacy activities. Efforts such as the World Wide Web Consortium's Platform for Privacy Preferences (P3P) specification and self-regulatory efforts such as TRUSTe and BBBOnline make numerous assumptions about how users perceive privacy. The P3P specification will lead to interoperable client and service programs that represent site privacy practices in ways that can be understood and processed automatically on behalf of the user: aiding the user in finding sites and practices she finds most appropriate (Reagle and Cranor 1999). Trust label programs promote guidelines about privacy disclosures and associate a trusted and branded icon with sites that follow those guidelines (Benassi 1999). Consequently, a better understanding of privacy concerns will lead to designs that best meet users' needs. In particular, we hope to gain an understanding that will inform the development of P3P agents and vocabulary — the set of privacy disclosures that can be understood by a user agent.

This is a report of our preliminary findings concerning Internet users' attitudes about privacy. We report our findings on general attitudes about online privacy, attitudes about specific current and anticipated online information practices, and attitudes about privacy regulation and self-regulation. We then describe the major factors that motivate concern about privacy, and discuss some technical and policy implications of our findings.

Survey Methodology

Survey Development

During the summer of 1998 we developed a series of survey questions designed to provide insight into Internet users' attitudes about privacy. We were interested in several privacy issues:

- We wanted to know how people would respond to situations where personal information is collected. In our pre-study, we determined that it was important to ask participants about their concerns through specific online scenarios. Therefore, in addition to the closed form survey questions, we also asked for their reasoning through open-ended questions.
- We wanted to know the sensitivity of particular privacy practices relevant to the design of the P3P vocabulary and P3P user agents. We were interested in testing the design of P3P and in creating better privacy user interfaces. Again, we probed for the reasons behind the respondents' sensitivities through open-ended questions in addition to standard-form survey questions.
- We wanted to determine participants' general attitudes and demographics. We largely used questions that had appeared on other surveys so we could match our sample against others.

We developed our survey and pre-tested it with non-technical employees and summer students at AT&T Labs, as well as with two classes at Harvard and MIT.

Sample Characteristics and Response Rate

Prospective survey participants were selected from the Digital Research, Inc. (DRI) Family Panel. The DRI Family Panel is a group of Internet users that evaluates products and responds to surveys for *FamilyPC* magazine. Approximately one-third of the panel members are *FamilyPC* subscribers, and most of the panel members who are not subscribers joined the panel after visiting the FamilyPC Web site.

Invitations to complete a Web-based survey were emailed to 1,500 Family Panel members (selected randomly, but weighted so that approximately 20% were sent to members outside the US), resulting in 523 surveys completed between November 6 and November 23, 1998 — a response rate of 35%. Code numbers were used to ensure that each respondent filled out the survey only once, and a sweepstakes was offered to encourage participation.

Out of the total sample, 405 completed surveys were from the United States, 88 were from Canada, and 30 were from other countries. We report on only the United States participants here. We eliminated surveys from respondents who did not answer at least two of our demographic questions, leaving us with 381 respondents in our US sample.

We did not obtain a statistically representative sample of United States citizens. However, our sample holds similar attitudes about privacy as the 460 Internet users in Westin's April 1998 sample, with our sample tending towards slightly more concern about privacy. For example, 87% of our US sample and 81% of the Net users in Westin's sample were somewhat or very concerned about threats to their personal privacy while online.

Our US sample differed from a nationally representative sample in some demographic areas. Most significantly our sample was more educated and had more Internet experience than nationally representative samples of Internet users, such as Westin's April 1998 sample or the IntelliQuest (http://www.intelliquest.com/) third-quarter 1998 sample. While 37% of Westin's sample and 36% of the IntelliQuest sample reportedly held college and/or postgraduate degrees, 48% of our sample reported such degrees.

Furthermore 77% of our sample reported that they make online purchases compared with 23% of Westin's sample and 20% of the IntelliQuest sample. Finally, fifty-one percent of our sample reported household incomes greater than $50,000, compared to 43% of Westin's sample and 55% of the IntelliQuest sample. The higher education and income levels coupled with increased number of online purchasers in our sample is consistent with Westin's (1998, 40) finding that online purchasers are more educated and affluent than other members of the public.

Our sample is certainly not statistically representative of US Internet users. However, our users are heavy Internet users — 65% report using the Internet several times a day — and quite possibly lead innovators. We base this on the above statistics, their self-selection in an opinion-formation group, and much of our qualitative data. As such, we believe that this sample is important for understanding the future Internet user population. As more people start using the Internet and gaining experience with email, the World Wide Web, and electronic commerce, we would expect their attitudes about privacy, if not their online behavior, to more closely match those of our sample.

The table below summarizes the demographic and attitudinal differences between our US sample and the Net users in Westin's April 1998 sample.

	Our US Sample	Westin's Sample
Unsolicited commercial email is very serious	52%	48%
Web sites collecting personal information from children is very serious	93%	85%
Web sites collecting email addresses from visitors without consent to compile email marketing lists is very serious	80%	70%
Tracking Web sites people visit and using that information improperly is very serious	87%	72%
Very or somewhat concerned about threats to personal privacy while online	87%	81%
Have personally been the victim of an online privacy invasion	19%	6%
College and/or post graduate degree	48%	37%
Send or receive email	100%	80%
Visit World Wide Web sites	100%	81%
Have made online purchases	77%	23%

Figure 1: Comparisons with Internet Users in Westin's April 1998 Sample

In the following sections, we present the findings from our survey. We have separated this analysis into three sections, the respondents' general attitudes about privacy, their attitudes about current and anticipated online practices, and their attitudes about privacy regulation. We present each in turn.

General Attitudes about Online Privacy

Overall, our respondents registered a high level of concern about privacy in general and on the Internet. Only 13% of respondents reported they were "not very" or "not at all" concerned. Nonetheless, while the vast majority of our respondents were concerned about privacy, their reactions to scenarios involving online data collection were extremely varied. Some reported that they would rarely be willing to provide personal data online, others showed some willingness to provide data depending on the situation, and others

were quite willing to provide data — regardless of whether or not they reported a high level of concern about privacy. Thus it seems unlikely that a one-size-fits all approach to online privacy is likely to succeed.

In order to understand our respondents' attitudes, we used standard multivariate clustering techniques to divide our respondents into three clusters, similar to the clusters Westin (Harris 1991) found in his privacy survey results. Based on general attitudes about privacy as well as their responses to specific scenarios, the clustering methods classified 17% of our respondents as **privacy fundamentalists**, 56% as members of the **pragmatic majority**, and 27% as **marginally concerned**. We will present each group more fully as we discuss their data below, but some general characteristics are important to note.

- The privacy fundamentalists were extremely concerned about any use of their data and generally unwilling to provide their data to Web sites, even when privacy protection measures were in place. They were twice as likely as the other groups to report having been a victim of an invasion of privacy on the Internet. About a third of the fundamentalists refused to answer our survey question about their household income (as compared with 14% of the pragmatists and 3% of the marginally concerned).
- The pragmatists were also concerned about data use, but less so than the fundamentalists. They often had specific concerns and particular tactics for addressing them. For example, the concerns of pragmatists were often significantly reduced by the presence of privacy protection measures such as privacy laws or privacy policies on Web sites.
- The marginally concerned were generally willing to provide data to Web sites under almost any condition, although they often expressed a mild general concern about privacy. Nonetheless, under some conditions, the marginally concerned seemed to value their privacy. For example, they highly rated the ability to have themselves removed from marketing mailing lists.

Demographic differences

Westin (Harris 1998) and others have found demographic differences, although weak, among groups with different levels of concern about online privacy. For example, Westin found that 87% of female Internet users were very concerned about threats to their personal privacy while only 76% of male Internet users were very concerned. Furthermore, he found that women registered higher levels of concern about every privacy-related issue they were questioned about. Although we found no statistically significant differences based on gender or other demographics within our sample, the trends in our data were consistent with Westin's findings.

Attitudes about Current and Anticipated Online Information Practices

Our survey included 14 questions that explored four different scenarios in which the user is asked to provide personal information to Web sites. We asked our respondents whether they would type in the requested information in each situation. We also asked our respondents how comfortable they generally feel providing each of 12 specific pieces of information to Web sites, and we asked for feedback on tools for protecting online privacy. Based on the responses to these questions we made a number of observations about current and anticipated online information practices.

Internet users are more likely to provide information when they are not identified

We presented respondents with two scenarios in which the first part of each scenario described a situation in which a Web site requested only information that was not personally identifiable. The second part of each scenario described the same situation, but this time the Web site also asked for personally identifiable information. In both cases respondents were much less willing to provide information when personally identifiable information was requested.

In a scenario involving a banking Web site, 58% of respondents said they would provide information about their income, investments, and investment goals in order to receive

customized investment advice. However only 35% said they would also supply their name and address so that they could receive an investment guide booklet by mail.

In a scenario about a news, weather, and sports Web site, 84% of respondents said they would provide their zip code and answer questions about their interests in order to receive customized information. But only 49% said they would provide information if they were also required to provide their name.

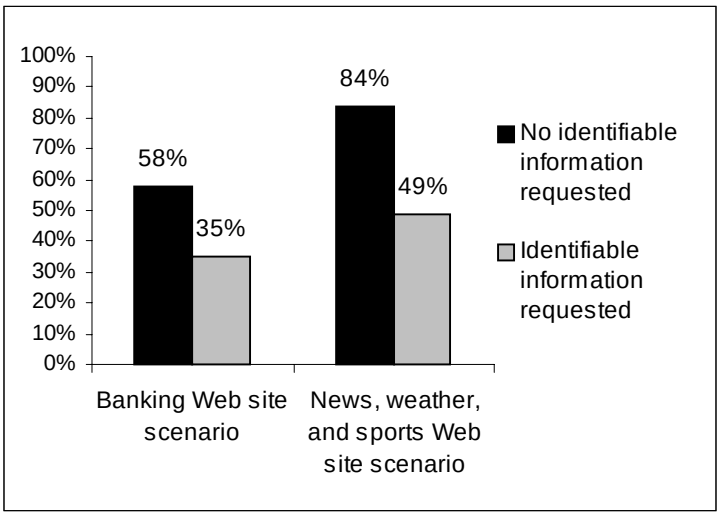

Figure 2: Respondents who probably or definitely would provide requested information

Some types of data are more sensitive than others

As mentioned above, we asked respondents how comfortable they feel providing each of 12 specific pieces of information to Web sites. We also asked them how comfortable they would be if a child in their care between the ages of 8 and 12 were asked to provide this information.

We found significant differences in comfort level across the various types of information. Not surprisingly, the vast majority of respondents said they were always or usually comfortable providing information about their own preferences, including favorite television show (82%) and favorite snack food (80%). A large number also said they were always or usually comfortable providing their email address (76%), age (69%), or

information about their computer (63%). About half said they were always or usually comfortable providing their full name (54%) or their postal address (44%). Few said they were always or usually comfortable providing information about their health (18%) or income (17%), or phone number (11%). None of the respondents said they were always comfortable providing their credit card number or social security number, and only a very small number said they would usually feel comfortable providing credit card number (3%) or social security number (1%).

Respondents were consistently less comfortable allowing a child to provide each of these types of information than they would be providing it themselves, with the biggest differences reported in the number of respondents who said they were always or usually comfortable with a child providing email address (16%) and age (14%).

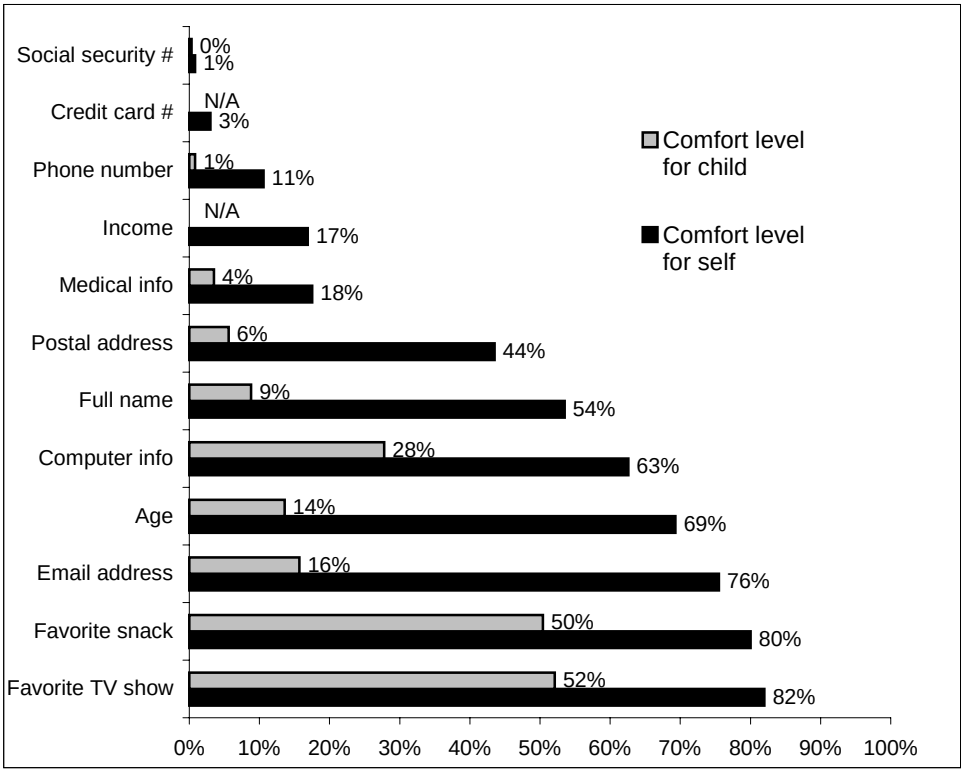

Figure 3: Respondents who are always or usually comfortable providing information

While each of our clusters reported different levels of comfort, the relative sensitivity to each type of data was consistent across clusters. That is the members of each cluster held similar views about which types of data were the most and least sensitive.

It is interesting to note the differences in sensitivity to seemingly similar kinds of data. For example, while postal mail address, phone number, and email address can all be used to contact someone, most of our respondents said they would never or rarely feel comfortable providing their phone number but would usually or always feel comfortable providing their email address. The comfort level for postal mail address fell somewhere in between. We suspect this may have to do with different levels of annoyance related to unsolicited communications in each medium as well as the availability of coping strategies to deal with this annoyance (Culnan 1993). For example, Westin (Harris 1991) found people were much more likely to describe marketing solicitations as an invasion of privacy when the solicitation was conducted via phone calls than when it was conducted via postal mail.

We also suspect that awareness of problems associated with divulging different types of information may affect the level of concern. Publicity surrounding identity theft and credit card fraud may have raised awareness about the dangers of social security numbers and credit card numbers falling into the wrong hands. But there has been less publicity about the dangers associated with disclosure of medical records. This may account for the fact that the concern reported about credit cards and social security numbers is significantly higher than that for medical records — which could be argued to be just as sensitive.

Many factors are important in decisions about information disclosure

Web site privacy policies include a wide range of privacy practice details. A number of efforts have tried to find ways of highlighting critical points of these policies for users. For example, initially the TRUSTe privacy seal program offered three seals that varied according to policies on sharing information with other parties. The P3P specification includes a vocabulary for encoding these practices in a standard way. Even so, it is unclear how to best (1) display these practices in a way that users can quickly evaluate

the practices and (2) design a user interface that permits users to configure an automated tool for evaluating those practices. Consequently, we asked respondents "If you could configure your Web browser to look for privacy policies and privacy seals of approval on Web sites and let you know when you were visiting a site whose privacy practices might not be acceptable to you, which criteria would be most important to you?" We also asked respondents to rate each of 10 criteria as very important, somewhat important, or not important.

Our respondents rated the sharing of their information with other companies and organizations as the most important factor. Ninety-six percent of respondents said this factor was very or somewhat important, including 79% who said it was very important.

Three other criteria emerged as highly important factors: (1) whether information is used in an identifiable way, (2) the kind of information collected, and (3) the purpose for which the information is collected. All of these criteria were rated as very important by at least 69% of respondents and had the same level of importance statistically.

These top criteria are consistent with the findings of other surveys. For example, the GVU survey (1998) asked respondents about seven factors that might influence whether they would give demographic information to a Web site. The factors most often selected by respondents were "if a statement was provided regarding how the information was going to be used," "if a statement was provided regarding what information was being collected," and "if the data would only be used in aggregate form." Providing data in exchange for access to Web pages, product discounts, value-added service, or other terms and conditions were less popular options. The top reason respondents gave for not filling out online registration forms at sites was "information is not provided on how the data is going to be used."

We found three additional criteria that were also very important factors: (1) whether a site is run by a trusted company or organization, (2) whether a site will allow people to find out what information about them is stored in their databases, and (3) whether the site will remove someone from their mailing lists upon request. These criteria were rated as very

important by at least 62% of respondents and had the same level of importance statistically. Interestingly, while none of these criteria were among the top factors for our privacy fundamentalist or pragmatic majority clusters, whether the site will remove someone from their mailing lists upon request was the number one most important factor for our marginally concerned cluster.

The remaining three criteria were rated as important, but considerably less so than the other factors. Not many people rated as very important whether a site posts a privacy policy (49%), whether a site has a privacy seal of approval (39%) and whether a site discloses a data retention policy (32%). These three factors were the least important factors for all three clusters of respondents.

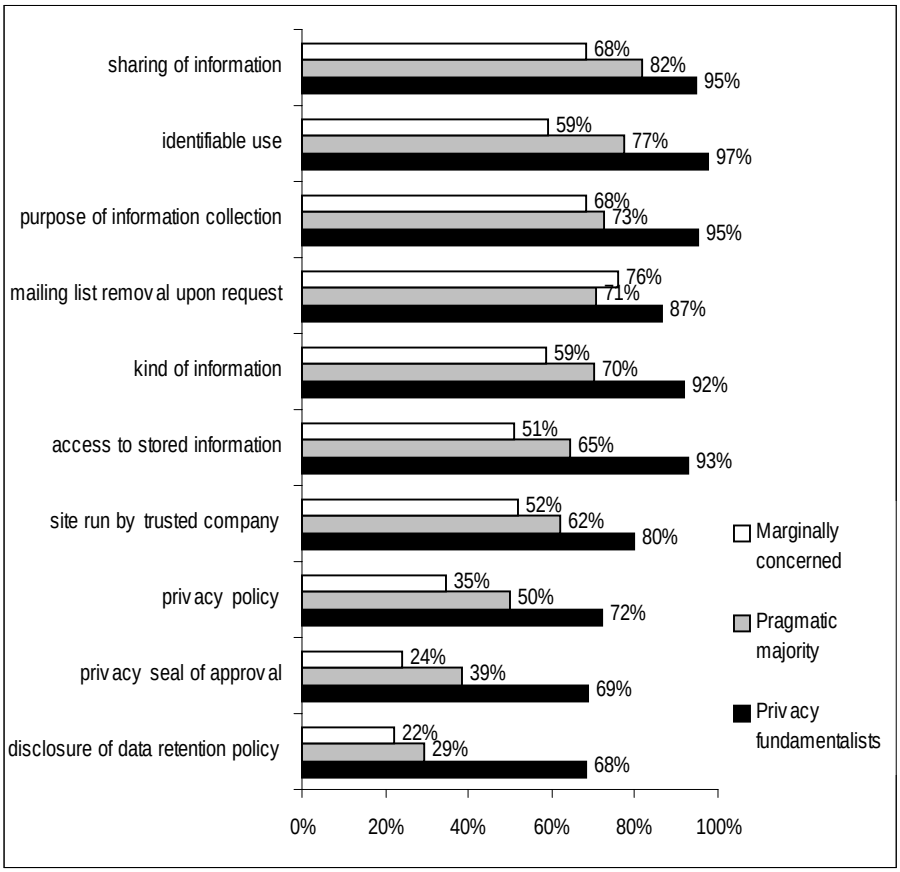

Figure 4: Respondents who consider factor very important

The lack of enthusiasm for knowing whether or not a site posts a privacy policy suggests that it is not enough for people to know *whether* a privacy policy is present — it is more important to know what the policy states. The lack of interest in knowing whether a site has a privacy seal of approval may be indicative of a lack of understanding of privacy seals (which will be discussed later).

The lack of concern for knowing whether a site discloses a data retention policy appears to be due to a distrust that companies will actually remove people from their databases and a belief that it will be impossible to remove information from all the databases it may have propagated to. Typical comments from our respondents were skeptical: “It doesn’t take long for this information to get spread around and a lot of this might have already been done,” “too late: the damage would already be done,” “who knows where they would sell my address to in the mean time,” “once you get on a mailing list, you’re on many mailing lists,” and “maybe they wouldn’t take me off. How would I know?”

Likewise, one of our scenario questions asked respondents whether they would be more or less likely to provide data to a Web site if it had a privacy policy that explained that their information would be removed from the site’s database if they did not return to the site for three months. Seventy-eight percent of respondents said that such a retention policy would not influence them in any way. Five percent said they would be less likely to provide information in that case (their comments suggested they viewed having their information removed from the database as an inconvenience should they return to the site after three months), and 17% said that such a retention policy would make them more likely to provide information. But other factors such as the existence of privacy policies, privacy seals, and privacy laws appeared to be much more influential than retention policies.

Acceptance of the use of persistent identifiers varies according to their purpose

Some Internet users are concerned that their online activities may be tracked over time. This can be accomplished using persistent identifiers stored on a users computer. These are often referred to as cookies. When asked about Web cookies, 52% of our respondents

indicated they were concerned about them (and another 12% said they were uncertain about what a cookie is). Of those who knew what cookies were, 56% said they had changed their cookie settings to something other than accepting all cookies without warning.

Comments to our free response questions suggest considerable confusion about cookies among our respondents. For example many respondents seemed to believe that cookies could cause identifying information about them to be sent automatically to Web sites. One respondent wrote, “cookies can determine my identity from visiting the site,” and another wrote “I may have a false sense of security but I understand that as long as I accept ‘no cookies’ the site managers cannot access my email address and other personal information.” Others understood that cookies need not be used to extract personal information from them, but did not seem to understand that cookies could be used to track their behavior. One respondent wrote, “A cookie can only provide information I have already given, so what is the harm?” Still another was simply confused: “I am not quite sure what cookie is, but I have an idea.”

We also included three scenario questions in which we described the use of persistent user identification numbers that browsers could automatically send back to Web sites on return visits. While the behavior we described could be implemented using cookies, we did not refer to cookies in these questions. In a scenario in which a site uses a persistent identifier to provide a customized service, 78% of respondents said they would definitely or probably agree to the site assigning them such an identifier. When we indicated the identifier would be used to provide customized advertising, 60% of respondents said they would definitely or probably agree to the site assigning them an identifier. But when we indicated that the identifier would be used to provide customized advertising across many Web sites, only 44% of respondents said they would definitely or probably agree to using such an identifier. We found similar trends across all three clusters of respondents. Thus it appears that most of our respondents are not opposed to the use of persistent identifiers or state management mechanisms such as cookies; however, many have misconceptions about these technologies and concerns about some of their uses.

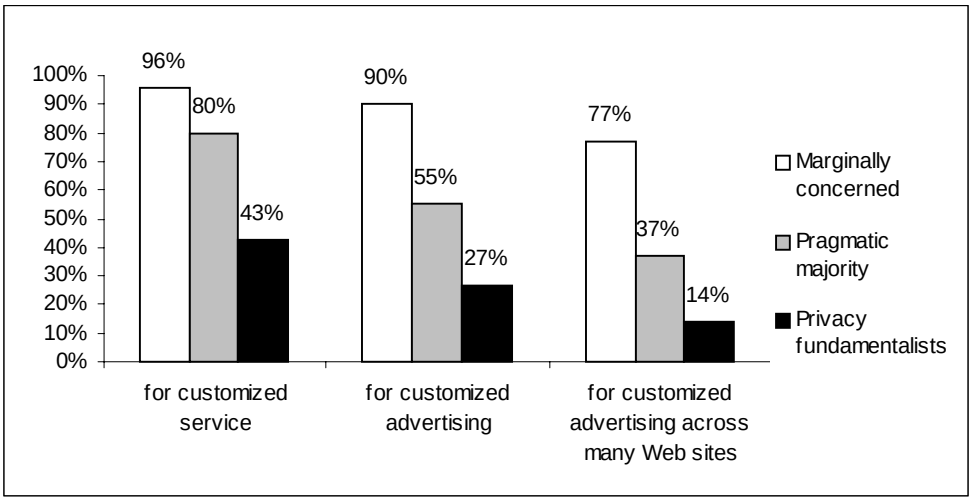

Figure 5: Respondents who would probably or definitely agree to a site assigning persistent identifier

Internet users dislike automatic data transfer

In the survey, we also described a number of browser features that would make it easier to provide information to Web sites and asked respondents which features they would use. We found that while our respondents said they are interested in tools that make using the Web more convenient, most do not want these tools to transfer information about them to Web sites automatically.

The most popular feature we described was an "auto-fill" button that users could click on their browsers to have information they had already provided to another Web site automatically filled in to the appropriate fields in a Web form. Sixty-one percent of our respondents said they would be interested in such a feature, while 51% said they would be interested in a similar feature that would automatically fill out forms at sites that have the same privacy policies as other sites the user had provided information to (no button click would be necessary to activate the auto-fill). Both of these features would require a user to click a submit button before any information was actually transferred to a Web site. Thirty-nine percent of respondents said they would be interested in a feature that automatically sent information they had provided to a Web site back on a return visit.

However, there was little interest in two features that would automatically send information to Web sites without any user intervention: a feature that notified the user that it had sent the information was of interest to 14% of respondents, and a feature that provided no indication that it had transferred data was of interest to only 6%. Thus 86% of our respondents reported no interest in features that would automatically transfer their data to Web sites without any user intervention.

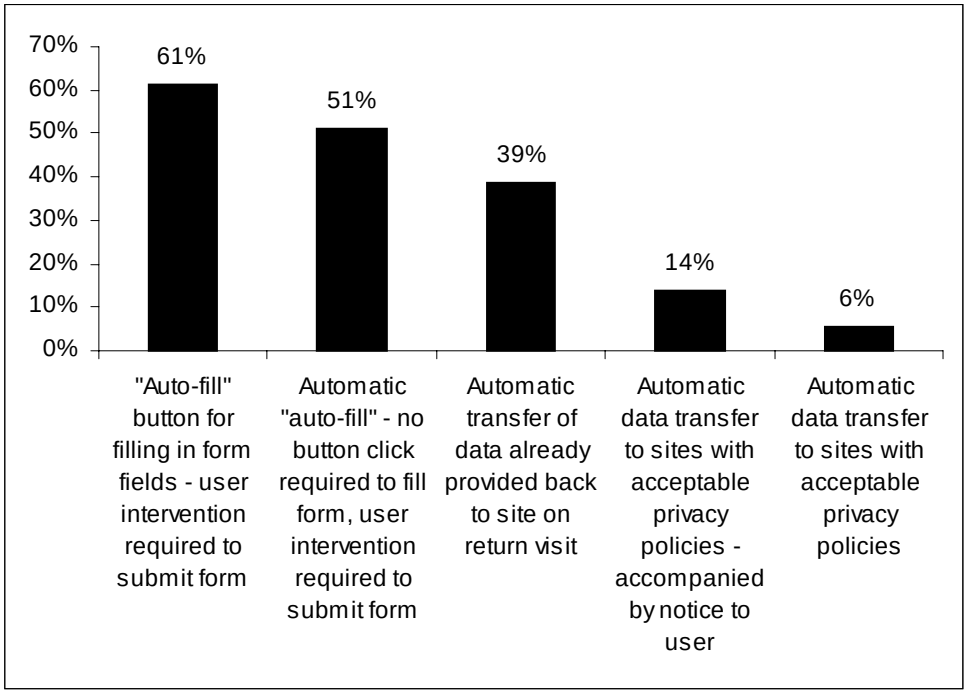

Figure 5: Respondents who would use proposed browser features

Respondents in our privacy fundamentalist cluster had much less interest in any of the described features than the members of the other clusters — only about one-fourth of the privacy fundamentalists were interested in any of the features. However, even the marginally concerned cluster members had little interest in features that would automatically transfer their data to Web sites without any user intervention — only 12% of the marginally concerned were interested in a feature that transferred data without notification.

These findings are consistent with other surveys that found Web users value privacy over convenience. For example, on the GVU survey (1998) 78% of respondents said privacy is more important to them than convenience. Our findings demonstrate how this concern plays out over specific technical features.

Our respondents provided strong comments about automatic data transfer. A large number of respondents made comments about wanting to remain in control over their information and stating that they had no desire for automatic data transfer. Some respondents were concerned with the perils of automatic data transfer in general. For example, one respondent noted that "I want to be in charge of all information sent to other companies. Just because they are similar, doesn't mean I [want] my information shared with them." Another noted the need for updating personal information: "To be able to update or correct the previous info is a good thing." However, most comments revolved around the respondents' desire to maintain control of the process. For example: "Auto[matic] features save time. …However, I do like to know when information about me is being transmitted," "I want to be in control of what is done. This way I know what was done," and "I don't want anything sent automatically. I want to check out everything I am applying for."

Internet users dislike unsolicited communications

On several questions, respondents displayed a desire not to receive unsolicited communications resulting from the provision of information to Web sites. For example, after describing a scenario in which a Web site would offer visitors free pamphlets and coupons, we asked respondents whether they would be more or less likely to provide information to the same Web site with a new condition. Specifically we described a site that had a privacy policy that permitted the site to send periodic updates on products *and* to share identifiable information with other companies that sold products of potential interest. Sixty-one percent of respondents who said they would provide their information to receive pamphlets and coupons said they would be less likely to provide that information if it would be shared for future marketing. However, nearly half of those

respondents said they would be more likely to provide the information if the site offered a way to get off their mailing list in the future.

The reasons for this were obvious in the written comments. As one respondent noted, “I already get too much junk mail.” Others expressed concerns about unsolicited marketing: “I would not want to have telemarketers, email messages, direct mail, etc. coming as I get too much of that anyway.” and “I don’t mind receiving literature that I request, but I DO NOT like to receive unsolicited mail, e-mail or phone calls.”

While respondents indicated a clear dislike for unsolicited communications, they were less concerned (but not unconcerned) about unsolicited email. As discussed earlier, respondents were more comfortable providing their email address than they were their postal address or their phone number. Furthermore, they expressed less concern about unsolicited email and about Web sites collecting email addresses for marketing lists than they did about Web sites collecting personal information from children, or someone tracking what Web sites people visit and using that information improperly.

In the previous sections, we have presented findings about the respondents’ attitudes about current information practices. We found a number of “hot” issues, such as whether they can be identified and the sensitivity of the data items to the individual. We also found a number of important differences in how our privacy clusters (privacy fundamentalists, privacy pragmatics, and privacy unconcerned) weighed these criteria. We also presented findings suggesting that there are some surprising similarities: People do not like unsolicited communications, they can be tolerant of persistent identifiers, and they dislike automatic transfer, although the degree of preference varies among the respondent clusters.

In this next section, we present our results about respondents’ views of privacy regulation.

Attitudes about Privacy Regulation and Self-Regulation

A joint program of privacy policies and privacy seals seemingly provides a comparable level of user confidence as that provided by privacy laws

We described a scenario in which a Web site with interesting information related to a hobby asks for a visitor's name and postal address in order to provide free pamphlets and coupons. Seventy-three percent of respondents said they definitely or probably would provide that information under those circumstances. We then asked whether they would be more or less likely to provide the information:

1. if there was a law that prevented the site from using the information for any purpose other than processing the request
2. if the site had a privacy policy that said the information would be used only to process the request, and
3. if the site had both a privacy policy and a seal of approval from a well-known organization such as the Better Business Bureau or the AAA.

While 28% of respondents who were uncertain or said they would not provide the information said they would be more likely to provide the information if the site had a privacy policy, 48% said they would be more likely if there was a relevant law, and 58% said they would be more likely if the site had both a privacy policy and a seal of approval. This suggests that the comfort gained from a joint program of privacy policies and privacy seals may be at least as much as that gained from a privacy law. Note that our scenario involved a Web site that requested name and postal address information; it is unclear whether we would find the same results in a scenario involving more sensitive information.

People do not seem to understand privacy seal programs

While a large number of respondents said they would be more likely to provide information in a scenario where a Web site had a privacy policy and a seal of approval, privacy seals were among the least important criteria for determining whether or not to provide information to Web sites. There are several potential reasons for this. It is

important to note that in the scenario question we mentioned two specific well-known organizations as possible seal providers: the Better Business Bureau and AAA. (Note that neither of these organizations actually offered online privacy seals at the time of our survey; however, both have seals that are well known in other domains. Our intention was not to evaluate a particular seal program.) We made no such mention in the criteria question. Efforts to date to raise consumer awareness of privacy seal programs have been minimal, and it is likely that respondents are not sufficiently familiar with the concept of online privacy seal programs that they consider them meaningful unless they are linked to a familiar trusted organization. If that is the case, it indicates that the proponents of such programs need to do more to educate consumers.

A number of respondents commented that they would like to know whether or not sites actually follow their privacy policies, suggesting that they were unaware that seals can help provide assurance that policies are followed. Some commentators did understand this, but wanted to know how they could be assured that the seal was authentic.

A number of respondents commented that they would trust a seal from a trusted third-party. In the question, one of the third-party examples was the Better Business Bureau. Typical comments included: “I trust the BBB, etc.,” “The BBB is very thorough in their recommendation of sites,” and “[the] Better business Bureau has a reputation of protecting the consumer and being responsible for investigating companies, thus, I trust their seal.” However there were some skeptics, who suggest how easy it might be to lose consumers’ trust. “The Better Business Bureau doesn’t have any REAL POWER [sic] over a business....” wrote one respondent. In general, however, seals were perceived as a positive influence in maintaining privacy, as one respondent explained: “This affiliation gives them more credibility and believability as far as I’m concerned. I would check with the reference before I gave the information if it was information that was more personal.”

Technical Implications

As the software engineering community attempts to implement P3P or similar privacy protocols, one of the major issues will be the design of easy-to-use systems for end-users.

Users would likely benefit from systems that assist them in identifying situations where a site's privacy practices is counter to their interest and assisting them in reaching agreement and exchanging data where acceptable to the user.

However, a user interface must not only present an extremely complex information and decision space, it must do so seamlessly and without a distracting interface (Ackerman and Cranor 1999). If a person wishes to control what information she presents to whom, this results in an enormous information space (i.e. each datum a person has about herself against each person or organizational entity with which she comes into contact). Moreover, the space is actually more complex, since there are additional dimensions to information dissemination, as noted in the P3P specification (e.g., purpose, access). Obviously, a matrix-style user interface for private information over each of its ten dimensions would be overwhelming for most users. However, properly designed and abstracted interfaces or borrowed settings (Cranor and Reagle 1997) may help.

One of our goals for this survey was to investigate consumer-driven design issues in privacy protocols and their user clients. We found several items of interest in considering the feasibility of P3P or any other privacy protocol:

- The cluster of privacy fundamentalists and marginally concerned may find extremely simplified interfaces to be adequate for their purposes. For example, a privacy fundamentalist may only want to release information under a small number of circumstances, such as when sites use information only for completing a purchasing transaction. A marginally concerned user would only need to specify those few (already constrained) instances in which she would not permit information collection practices. However, the pragmatists (who are the majority of users) will require more sophisticated and varied interface mechanisms to be most at ease. This cluster of users employs many strategies across a wide range of finely weighed situations. It is unlikely that a highly simplified interface will satisfy them.
- Automatic transfer of data and computerized negotiations with sites are unlikely to be interesting to most consumers.

- Designers should permit users to have differing views of — or ways of looking at — their information. For instance, while it makes sense to include phone number in a contact information category, our respondents considered it to be more sensitive than postal information. Consequently, a user should be able to enter contact information on one page, but be able to drag those pieces of information to different sensitivity buckets or to simply manipulate information as grouped by sensitivity.
- Additional augmentative assistance to consumers will be useful. Many of our respondents expressed confusion over potential risks and rewards for their dissemination of personal information. Having agents that help users (e.g., that provide warnings based on third-party databases of rogue sites) could well be helpful instead of placing the full burden on users themselves.
- Finally, technical mechanisms clearly have limitations. Our respondents were very aware (and vocal) about these limitations.

Policy and Business Implications

What do our results say to those concerned with public policy? Our findings show that users are indeed concerned about privacy. Do our results argue that present day laws and self-regulatory programs are mitigating that concern? Not necessarily.

Our results do permit us to compare assumptions made about Internet users' approaches to privacy with the responses of actual users. For instance, present day US public policy does make a distinction between children and adults, and this seems well founded on the basis of our results. We also found that our respondents cared a great deal about the perceived trustworthiness of the data collecting organization, the purpose of the data collection, and its redistribution policies. Proposed policy solutions need to squarely address each of these topics. Seemingly, much of the discomfort with the Web today results from not knowing, or not trusting the information practices of a site. As Hine and Eve point out:

> Our research showed that, in the absence of straightforward explanations on the purposes of data collection, people were able to produce their own

> versions of the organization's motivation that were unlikely to be favorable. Clear and readily available explanations might alleviate some of the unfavorable speculation (Hine and Eve 1998, 261).

If we wish to raise the comfort level, we must ensure users are informed and can trust whatever policies are disclosed. If we wish to be pragmatic, we must focus on those things about which users say they are most concerned. These results provide evidence of what those concerns are among an Internet savvy population. We must echo Milne and Boza's conclusions:

> The data from these studies suggest that a trust-enhancement approach is more effective. Trust can be enhanced by building a reputation for fairness, by communication information sharing policies up front and stressing the relational benefits, and by constantly informing the consumers of the organization's activities to serve them better." (Milne and Boza 1998).

Several important caveats and considerations remain. For example, although some privacy advocates consider disclosures on how long an organization retains data to be an important privacy principle, only 29.9% of our sample considered this very important and 18.6% classified it as not important. Thus, the pursuit of this goal in the policy arena might be treated differently from the more highly rated factors. We are not arguing that the pursuit of the duration of retention principle is useless. Our results do not speak to whether the low expectations about access to information should or could be raised in priority — the data can only argue that access is presently considered to be less important by our respondents.

Additionally, to meet the needs of the varied clusters of people, public policy should support flexibility. Both the technical (P3P) and self-regulatory (TRUSTe, BBBOnline) approaches promote privacy practice disclosure upon which users can make their own decisions. However, while users do care very much about some information practices, our respondents placed relatively little value in the presence of Web site privacy proposals or privacy seals. In light of this, a technical approach like P3P is compelling because it permits flexibility and enables users' preferences to be acted upon without requiring too much of their attention. Otherwise, we may have to rely upon an approach

of (1) continued discomfort and confusion or (2) a one-size-fits-all legal approach. Regardless, we acknowledge that an eventual solution might rely upon elements of legal, self-regulatory, and technical approaches to the problem.

Finally, we must caution that policy based solely on survey results is inadequate. People's self-reported preferences often do not match their real world behavior (Turner and Martin 1984). Indeed, we found notable mismatches in our results. For example, while 39% of respondents said they are very concerned about online privacy, only half the members of that group were classified as privacy fundamentalists based on their responses to our scenario questions. More importantly, one would not want to base public policy solely on user expectations.

Policy making based solely on survey results can be described as *self-deprecating*: if the standard of what constitutes reasonable privacy is based on people's expectations, the standard and expectations are mutually influencing, resulting in a downward trend. his melt-down was reflected in some pre-study interviews: Students felt concern would only be frustrating or futile, since they felt they had few choices. Nonetheless, we believe that present day public policy can only improve with more concrete data about users' actual attitudes and expectations of online privacy — if for no other reason to understand the ways in which people's expectations change over time.

Acknowledgments

We would like to thank the members of the P3P Working Groups and our other colleagues for their ideas about privacy and privacy technologies. We would like to thank especially Rebecca Grant, Bonnie Nardi, and Steve Greenspan for pointing us towards relevant literature and reviewing preliminary drafts of our survey instrument. We would also like to thank Parni Dasu for her assistance with cluster analysis, Bob Cuzner and Bob Domine at DRI, Robin Raskin at *FamilyPC*, Roger Clarke for maintaining an online reference list of privacy surveys (http://www.anu.edu.au/people/Roger.Clarke/DV/Surveys.html), and the students and professors at Harvard and MIT who helped pre-test our survey instrument. Finally, would like to thank AT&T Labs-Research for its generous support of this survey.

References

Ackerman, Mark S. and Lorrie Cranor. Privacy Critics: UI Components to Safeguard Users' Privacy. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'99)*, short papers (v.2.), 258-259.

Benassi, Paola. TRUSTe: an online privacy seal program (February 1999). The platform for privacy preferences. *Communications of the ACM* 42(2):56-59.

Cranor and Reagle (1998). Designing a Social Protocol: Lessons Learned from the Platform for Privacy Preferences Project. In Jeffrey K. MacKie-Mason and David Waterman, eds., *Telephony, the Internet, and the Media*. Mahwah: Lawrence Erlbaum Associates.

Culnan, Mary J. (September 1993). "How did they get my name?": an exploratory investigation of consumer attitudes toward secondary information use. *MIS Quarterly* 17: 341-364.

Georgia Tech Graphics, Visualization & Usability Center (1998). GVU's 10th WWW User Survey. http://www.gvu.gatech.edu/user_surveys

Harris, Louis and Associates and Alan F. Westin (1991). *Harris-Equifax Consumer Privacy Survey 1991.* Atlanta, GA: Equifax Inc.

Harris, Louis and Associates and Alan F. Westin (1994). *Equifax-Harris Consumer Privacy Survey 1994.* Atlanta, GA: Equifax Inc.

Harris, Louis and Associates and Alan F. Westin (1996). *The 1996 Equifax-Harris Consumer Privacy Survey.* Atlanta, GA: Equifax Inc.

Harris, Louis and Associates and Alan F. Westin (June 1998). *E-commerce & Privacy: What Net Users Want*. Hackensack, NJ: Privacy & American Business.

Hine, Christine and Juliet Eve (1998). Privacy in the marketplace. *The Information Society* 14(4):253-262.

Milne, George R. and Maria-Eugenia Boza (September 1998). *Trust and Concern in Consumers' Perceptions of Marketing Information Management Practices.* Marketing Science Institute Working Paper Report No. 98-117.

Pew Research Center for the People & the Press (January 1999). *Online Newcomers More Middle-Brow, Less Work-Oriented : The Internet News Audience Goes Ordinary.* http://www.people-press.org/tech98sum.htm

Raab, Charles D. and Colin J. Bennett (1998). The Distribution of Privacy Risks: Who Needs Protection? *The Information Society* 14(4):253-262.

Reagle, Joseph and Lorrie Faith Cranor (February 1999). The platform for privacy preferences. *Communications of the ACM* 42(2):48-55.

Turner, Charles, and Elizabeth Martin, ed. (1984). Surveying Subjective Phenomena. New York: Russell Sage Foundation.