AMERICANS & Online Privacy

The System is Broken

A Report from the Annenberg Public Policy Center of the University of Pennsylvania

By Joseph Turow, Ph.D.

Americans and Online Privacy
The System is Broken

By Joseph Turow
June 2003

Americans and Online Privacy
The System is Broken

OVERVIEW

This new national survey reveals that American adults who go online at home misunderstand the very purpose of privacy policies. The study is also the first to provide evidence that the overwhelming majority of U.S. adults who use the internet at home have no clue about data flows—the invisible, cutting edge techniques whereby online organizations extract, manipulate, append, profile and share information about them. Even if they have a sense that sites track them and collect individual bits of their data, they simply don't fathom how those bits can be used. In fact, when presented with a common way that sites currently handle consumers' information, they say they would not accept it. The findings suggest that years into attempts by governments and advocacy groups to educate people about internet privacy, the system is more broken than ever.

- 57% of U.S. adults who use the internet at home believe incorrectly that when a website has a privacy policy, it will not share their personal information with other websites or companies
- 47% of U.S. adults who use the internet at home say website privacy policies are easy to understand. However, 66% of those who are confident about their understanding of privacy policies also believe (incorrectly) that sites with a privacy policy won't share data.
- 59% of adults who use the internet at home know that websites collect information about them even if they don't register. They do not, however, understand that data flows behind their screens invisibly connect seemingly unrelated bits about them. When presented with a common version of the way sites track, extract, and share information to make money from advertising, 85% of adults who go online at home did not agree to accept it on even a valued site. When offered a choice to get content from a valued site with such a policy or pay for the site and not have it collect information, 54% of adults who go online at home said that they would rather leave the web for that content than do either.
- Among the 85% who did not accept the policy, one in two (52%) had earlier said they gave or would likely give the valued site their real name and email address—the very information a site needs to begin creating a personally identifiable dataset about them.
- Despite strong concerns about online information privacy, 64% of these online adults say they have never searched for information about how to protect their information on the web; 40% say that they know "almost nothing" about stopping sites from collecting information about them, and 26% say they know just "a little." Only 9% of American adults who use the internet at home say they know a lot.
- Overwhelmingly, however, they support policies that make learning what online companies know about them straightforward. 86% believe that laws that forces website policies to have a standard format will be effective in helping them protect their information.

- Yet most Americans feel unsure or conflicted about whether key institutions will help them with their information privacy or take it away. Only 13% of American adults who use the web at home trust that the government will help them protect personal information online while not disclosing personal information about them without permission.
- Similarly, only 18% trust their banks and credit card companies and only 18% trust their internet service providers (ISPs) to act that way.
- Parents whose children go online are generally no different on these attitudes, knowledge or actions than the rest of U.S. adults who use the internet at home. Like the others, most parents are concerned, confused, and conflicted about internet privacy.

These are highlights from the most recent Annenberg national survey of internet attitudes and activities. The survey raises questions about the usefulness of trying to educate American consumers in the growing range of tools needed to protect their online information at a time when technologies to extract and manipulate that information are themselves growing and becoming ever-more complex. Our findings instead indicate that consumers want legislation that will help them easily gain access to and control over all information collected about them online. At the end of this report, we therefore suggest that the federal government needs to require online organizations to unambiguously disclose information-collection policies as well as to straightforwardly describe at the start of every online encounter what has and will happen to the specific user's data.

Our examination of online Americans' attitudes, knowledge, and actions regarding their online information was carried out by ICR/International Communication Research for the Annenberg Public Policy Center of the University of Pennylvania.[1] The study was conducted by telephone from February 5 to March 21, 2003 among a nationally representative sample of 1,200 respondents 18 years and older who said they use the internet at home. 516 (43%) of the respondents were parents of a child age 17 or younger.

Our aim was to address two critical public policy questions that had not previously been explored in depth: What level of understanding do Americans have regarding the way organizations handle information about them on the internet? And how much do they trust social institutions to help them control their information online?

[1] Thanks to Tara Jackson, Melissa Herrmann, and Jill Glather and Carol Cassel of ICR for survey and statistical help. Susannah Fox, Robert Hornik, Steve Jones, Mihir Kshirsagar, Deborah Linebarger, Mihaela Popescu, Lee Rainie, and Judith Turow generously listened at various stages of this project and provided useful suggestions. All responsibility for presentation and interpretation of findings rests with the author of this report.

BACKGROUND

An important reason that policy analysts need to know the answer to these questions relates to the absence of U.S. laws to control much of the extraction, manipulation, and sharing of data about people and what they do online. With the exception of certain personal health information,[2] certain types of personal financial information held by certain types of firms[3], and personally identifiable information from children younger than 13 years,[4] online companies have virtually free reign to use individuals' data in the U.S. for business purpose without their knowledge or consent. They can take, utilize and share personally identifiable information—that is, information that they link to individuals' names and addresses. They can also create, package and sell detailed profiles of people whose names they do not know but whose interests and lifestyles they feel they can infer from their web-surfing activities.

Companies continually troll for, and exploit, personally identifiable and non-personally identifiable information on the internet. They often begin by getting the names and email addresses of people who sign up for web sites. They can then associate this basic information with a small text file called a cookie that can record the various activities that the registering individual has carried out online during that session and later sessions. Tracking with cookies is just the beginning, however. By using other technologies such as web bugs, spyware, chat-room analysis and transactional database software, web entities can follow people's email and keyboard activities and serve ads to them even when they are off-line. Moreover, companies can extend their knowledge of personally identifiable individuals by purchasing information about them from list firms off the web and linking the information to their own databases. That added knowledge allows them to send targeted editorial matter or advertising to consumers. More specificity also increases the value of the databases when they are marketed to other interested data-trollers.

Marketers and media firms use consumer information in a broad gamut of ways and with varying concerns for how far the data travel. Some websites unabashedly collect all the information they can about visitors and market them as aggressively as they can to advertisers and other marketers. Though many of these emphasize personally identifiable

---

[2] These regulations relate to Health Insurance Portability and Accountability Act of 1996 (HIPPA). They resulted in the first set of federal privacy rules to protect medical information online and elsewhere. See http://www.consumerprivacyguide.org/law/hipaa.shtml

[3] These "opt-out" regulations relate to the Financial Modernization Act (Graham-Leach-Bliley Act). For an explanation, see the Privacy Rights Clearinghouse site: http://www.privacyrights.org/fs/fs24a-optout.htm

[4] The Children's Online Privacy Protection Act, which went into effect in 2000, requires online services directed at children 12 and under, or which collect information regarding users' age, to give parents notice of their information practices and obtain their consent prior to collecting personal information from children. The Act also requires sites to provide parents with the ability to review and correct information that they collect about their children. See Joseph Turow, *Privacy Policies on Children's Websites: Do They Play By the Rules?* Philadelphia: Annenberg Public Policy Center, 2001. http://www.appcpenn.org/internet/family/

information, not all of them do. Tracking people anonymously can still lead to useful targeting. An important example is the Gator Corporation, which places its tracking files into people's computers when they download free software such as the KaZaA music-sharing program.

The company claims to be in 35 million computers and says that once there, "The Gator Corporation has the ability to ride along with consumers as they surf the Web. That allows us to display targeted ads based on actual behavior and deliver incredible insights."[5] A pitch to potential clients continues:

> Here's an example: Gator knows this consumer is a new parent based on their real-time and historical online behavior—looking for information on childbirth, looking for baby names, shopping for baby products. . . .[6]
>
> Let's say you sell baby food. We know which consumers are displaying behaviors relevant to the baby food category through their online behavior. Instead of targeting primarily by demographics, you can target consumers who are showing or have shown an interest in your category. … Gator offers several vehicles to display your ad or promotional message. You decide when and how your message is displayed to consumers exhibiting a behavior in your category.[7]

Many individual sites aim to provide similar services to marketers, though on a more limited scale. Many collect names and email addresses and use an "opt out" approach to gather targets for email advertising by themselves or "affiliates" on topics that ostensibly relate to the site themes. Some sites link their online knowledge of individuals with data collected offline. Typically, the more prestigious sites sell that information only in aggregate to advertisers. So, for example, an online newspaper may offer to send an ad for a client to all its users who are male and own a home. Because the newspaper site serves the ad, the advertiser does not know the names of those who receive it—unless they click on the ad and respond with their names to an offer. Some well-known sites may also have deals with companies that serve ads on their sites and share the revenues. These firms place their own cookies into the computers of those who visit the websites and then track people's activities into the many other sites that affiliate with the ad-serving firms. Some of them may try to coax names and email addresses from consumers that click on their ads even if the site on which their ads appeared did not.

The idea that consumers' electronic actions are increasingly transparent has alarmed some. Critics of these sorts of activities come at them with a variety of concerns from a variety of viewpoints. Many emphasize the danger that some kinds of personal information may fall into the hands of companies or people who could take advantage of the consumer. In the wake of the anti-terror PATRIOT Act, critics also worry that various government agencies will expand the tracking and generalizing about consumers on the web that had until recently seemed to be the domain of business. They point out

[5] [http://www.gatorcorporation.com/advertise/qtr/page_2.html?mp14], accessed on May 29, 2003.
[6] [http://www.gatorcorporation.com/advertise/qtr/page_3.html?mp14], accessed on May 29, 2003.
[7] [http://www.gatorcorporation.com/advertise/qtr/page_4.html?mp14], accessed on May 29, 2003.

the profound damage that errors or names on suspect lists can cause individuals and families.

Others note that sites' application of email addresses in the service of marketing has helped the proliferation of unwanted email on the web, adding to a spam epidemic that has internet users and their service providers steaming. More sociologically-inclined analysts underscore that the invisible nature of much of the tracking and sorting can lead marketers to make generalizations about consumers that the consumers don't know and don't agree with. Inferences drawn from demographics and web-surfing habits can encourage discrimination in the kinds of editorial and advertising materials a site shows consumers. Such activities will become more intense as technologies to mine data, analyze data, and tailor based on the conclusions become more efficient and cost-effective. As they expand, the activities may well lead people to feel anxious not only that they are being tracked but that they are being treated differently—for example, given different discounts—than others because of who they are and what their "clickstream" says about them.

Law professor Jeffrey Rosen poses the humanistic critique bluntly. Paraphrasing the Czech writer Milan Kundera, he suggests that "by requiring citizens to live in glass houses without curtains, totalitarian societies deny their status as individuals." He goes on to note that spying on people without their knowledge is an indignity. It fails to treat its objects as fully deserving of respect, and treats them instead like animals in a zoo, deceiving them about the nature of their own surroundings."[8]

Those concerned about the secondary use and sharing of data about individuals point to the European Union's rather stringent prohibitions against using data in ways for which they were not originally gathered. In the U.S., no such broad rules apply, though in the late 1990s the Federal Trade Commission advanced a set of "Fair Information Practices" reflective of principles that had been advanced in the early 1980s by the Office for Economic Cooperation and Development. These would mandate certain levels of data security on websites, provide notice to potential users of sites about the way data will be collected and used, give the users choice about allowing that collection, and provide them with access to data that have been collected to find out what firms know and determine their accuracy. They, in turn, had been the basis for guiding the FTC's enforcement of a "Safe Harbor" agreement with the European Union, whereby U.S. companies wanting to use personally identifiable data about EU citizens in the U.S. had to recognize these practices in the EU though not in the U.S.[9]

As FTC Commissioner Orson Swindle recalled in late 2002, U.S. regulatory officials tended to encourage industry self-regulation rather than the legislative mandating of these practices. "Use of the Internet for marketing and attempts to address online privacy concerns were still in their infancy, and the Commission believed that the private sector

[8] Jeffrey Rosen, "The Eroded Self," *New York Times Magazine*, April 30, 2000.

9 See D. Brown, and J Blevins, "The safe-harbor agreement between the United States and Europe: a missed opportunity to balance the interests of e-commerce and privacy online?" *Journal of Broadcasting and Electronic Media* 46:4 (December 2002), p. 565.

would continue on its own toward better privacy practices than what federal regulation might require. More specifically, it seemed inappropriate in these formative years to prescribe regulations that would impose nontrivial costs without also achieving clear benefits."[10]

By 2000, however, three of the five members of the Commission believed that industry had made insufficient progress toward developing genuine, pragmatic privacy protections for consumers. They formally recommended that the Congress enact laws to codify the Fair Information Practice principles. Congress agreed with the naysayers, however, and no such law was passed. Instead, the Federal Trade Commission has used Section 5 of the Federal Trade Commission Act (which deals with unfair and deceptive practices) to prosecute websites that present fraudulent claims about information protection.[11]

An extreme example of the computer industry's riposte to such concerns about privacy came from Sun Microsystems chief executive Scott McNealy in February 1999 when someone pointed out that a new Sun product might allow people to track its users' movements. "You have zero privacy anyway," McNealy told a questioner. "Get over it."[12] The comment, which *The New York Times* used as its quotation of the day not long after he made it,[13] raised consternation within the business community as well as outside it.

The more typical corporate response to concerns about online consumer privacy has been to express agreement with the goal of protecting personal information while at the same time arguing that government intervention on consumers' behalf could be catastrophic to industry growth. A *New York Times* report in 2001 concluded that "Lawmakers . . . are bolstered in their efforts to slow the march of legislation by a flood of new studies and surveys sponsored by high-technology companies, questioning consumer attitudes about privacy and giving multibillion-dollar estimates of the costs of complying with such laws."[14] So, for example, a study in 2001 by Robert Hahn of the American Enterprise Institute, a conservative research center in Washington, concluded that complying with privacy legislation proposals would cost companies $30 billion. A spokesperson for the Association for Competitive Technology, which paid for the Hahn study, used the findings to argue that "the costs associated with regulation appear to be higher than the benefits achieved by regulation."[15]

---

10 Orson Swindle, "Perspectives on Privacy Law and Enforcement Activity in the United States," *Privacy & Information Law Report*, 3:4 (December, 2002).

11 Critics have argued that U.S. legislative venues for reinforcing consumer privacy rights in general are insufficient. The United States does not have a federal privacy law. Moreover, tort law does not protect the disclosure of personal data unless the data could be construed as libel or potentially embarrassing. The mere gathering of data is not actionable in courts unless the practice of gathering itself is arguably too intrusive. See Jessica Litman, "Information privacy/information property," Stanford Law Review, (2000) vol. 52, pp. 1283-1313.

12 Richard Morochove, "Sun Microsystems Lets Jini Out Of Bottle ," *Toronto Star*, February 4, 1999.

13 "Quotation of the Day," *New York Times*, March 3, 1999, Section A; Page 2; Column 6.

14 John Schwartz, "Government is Wary of Tackling Online Privacy," New York Times, September 6, 2002, Section C, page 1.

15 Schwartz, "Government is Wary of Tackling Online Privacy," page 1.

The *Times* report pointedly mentioned surveys "sponsored by high technology companies, questioning consumer attitudes about privacy." These studies argue consistently that although much of the public had certainly become concerned about online privacy, Americans are quite alert to the particulars of their information environment. They typically understand their information options, are aware of privacy policies, and are willing to negotiate privacy demands with companies who could offer them something in return.[16] Alan Westin's Privacy and American Business consultancy has been an important promulgator of this notion that Americans make cost-benefit analyses about whether to release their information online. Beginning 1995, his analyses of surveys conducted with the Harris research organization have promulgated a tri-partite division of the online public—*privacy unconcerned, privacy fundamentalists,* and *privacy pragmatists.*[17]

Looking back in 2003, Westin noted a sharp drop in the percentage of his *privacy unconcerned* group from 22% in 1999 to 8% two years later. A correspondingly higher percentage of Americans (56% in 2002 versus 34% in 1999) believed that most businesses did not "handle personal information they collect in a proper and confidential way." Nevertheless, Westin noted that the privacy pragmatists still formed by far the largest group of internet consumers, 58% in 2002. His description of their outlook reflects his position that most Americans take an informed cost-benefit tack in relation to their online information: "They examined the benefits to them or society of the data collection and use, wanted to know the privacy risks and how organizations proposed to control those, and then decided whether to trust the organization or seek legal oversight."[18]

This description of most Americans as aware of their online privacy options supported the line by internet industry players that an accurate privacy policy on every site is sufficient for allowing consumers to understand their information options in different sites. As a result of the Children's Online Privacy Protection Act (COPPA), the Federal Trade Commission mandated specific privacy practices and disclosures regarding children younger than 13 years. With respect to everyone else, however, the presence, form and content of privacy policies is optional, subject only to broad prescriptions for members of industry groups such as the Internet Advertising Bureau and the Direct Marketing Association. The result is a world of legalistically phrased privacy policies that typically start by assuring the consumer that the site cares about his or her privacy. The policies then run for many paragraphs; hedge with respect to many of their assurances; are ambiguous when it comes to the "affiliates" with whom they share information; don't necessarily report whether a site purchases data offline about its registered users; generally caution that the privacy policy can change at any time (sometimes telling consumers that the site will inform them when that happens); and

[16] On the development of this contention, see Oscar Gandy, "Public Opinion Surveys and the Formation of Public Policy," *Journal of Social Issues* 59:2 (2003) 283-299.
[17] A good summary is in Alan F. Westin, "Social and Political Dimensions of Privacy," *Journal of Social Issues* 59:2 (2003) 431-453.
[18] Westin, "Social and Political Dimensions of Privacy," pp. 445-446.

often note that by clicking on an ad link a consumer may be entering a world with a privacy policy totally different from the one they are reading.

Anecdotal conversations suggest that internet experts find privacy policies hard to read and difficult to understand.[19] A bold technological solution that has gained industry traction during the past few years is the Platform for Privacy Preferences (P3P). Its goal is to provide a web-wide computer-readable standard manner for websites to communicate their privacy policies automatically to people's computers. In that way visitors can know immediately when they get to a site whether they feel comfortable with its information policy.[20] A recent report by an AT&T Labs group found that while P3P's adoption by websites is growing, especially on the most popular sites, fewer than 10% of websites offer it.[21]

One reason that sites eschew P3P is that it requires them to transform their privacy policies into a number of straightforward answers to multiple choice questions. P3P consequently does not allow for the ambiguities, evasions and legal disclaimers that are hallmarks of such documents. Note, too, that the P3P approach does not have a facility for ensuring that websites answer the questions accurately or truthfully.

In the absence of a widespread technological solution, those concerned about the state of information privacy on the internet lobby for legislation[22] at the same time that they try to educate people about how to understand what goes on. There certainly are lots of places for people to learn what happens to their information online and how to keep it secure. The popular press continually beats a refrain about the dangers of the internet for information privacy, sometimes with links to online locations to learn more. Websites of organizations as varied as the Electronic Privacy Information Center (EPIC), Privacy.org (a joint project of EPIC and Privacy International), the Center for Democracy and Technology, Internet Education Foundation, AARP, Consumer's Union and the U.S. Federal Trade Commission have exhorted consumers (and citizens) to take specific steps to protect their privacy online.

---

[19] For an examination of privacy policies in children's websites, see Joseph Turow, Privacy Policies on Children's Websites: Do They Play By the Rules?" Philadelphia: Annenberg Public Policy Center, March 2002. [http://www.appcpenn.org/internet/family]

[20] P3P "user agents" are built into the Internet Explorer 6.0 and Netscape Navigator web browsers. An ingenious AT&T program called *Privacy Bird* is a P3P user agent that works with Internet Explorer 5.01 and higher. It displays a bird icon on the browser that changes color and shape to indicate whether or not a web site's P3P policy matches a user's privacy preferences. The beta-version software is free. See http://www.privacybird.com/.

[21] Lorrie Faith Cranor, Simon Byers, and David Kormann, "An Analysis of P3P Deployment on Commercial, Government and Children's Web Sites as of May 2003." Technical report prepared for the may 14, 2003 Federal Trade Commission Workshop on Technologies for Protecting Personal Information. [http://www.research.att.com/projects/p3p/]

[22] For a list of "privacy, speech, and cyber-liberties bills in the 108th Congress," see the Electronic Privacy Information Center's site: http://www.epic.org/privacy/bill_track.html

ConsumerPrivacy.org, for example, provides an online guide to help readers "take control of the way your information is used."[23] Sections include a *"how to" guide to privacy, top things you can do to protect your privacy, kids' privacy, frequently asked questions,* and a *privacy glossary*. The Internet Education Foundation has a similarly wide-ranging resource called GetNetWise that is supported by various corporations. AARP provides a guide called "Online Shopping: A Checklist for Safer Cybershopping." The Federal Trade Commission issues *FTC FACTS for Consumers* that deal with internet privacy with such titles as "Dialing Up to the Internet: How to Stay Safe Online" and "Safe at Any Speed: How to Stay Safe Online If You Use High-Speed Internet Access." And EPIC provides an online guide to "practical privacy tools" that help internet users with such activities as surfing anonymously, eliminating cookies, achieving email and file privacy, and deleting files so that they can never be read.[24]

A question unanswered through all the debates about information privacy and the web is whether consumers understand these approaches and how to implement them. Marketers argue that privacy notices are invaluable in helping to ease concerns over sharing information. They look with optimism to a study conducted in Spring 2001 for the Privacy Leadership Initiative (a coalition of CEOs and organizations dedicated to improving consumer privacy online). It found that consumers were increasingly paying attention to online privacy statements (82% in April 2001 vs. 73% in December 2000).[25]

- But does concern over privacy and increased "attention" to privacy policies mean that people really understand what is happening to their information on the web?
- Are writers such as Alan Westin correct to suggest that Americans make knowledgeable, pragmatic cost-benefit analyses when they disclose data about themselves online?

This study explores these and other key questions.

[23] "Protect Your Privacy Now—Welcome to ConsumerPrivacyGuide!" ConsumerPrivacyGuide.org [http://www.consumerprivacyguide.org/], accessed on May 28, 2003.
[24] Electronic Privacy Information Center, EPIC Online Guide to Practical Privacy Tools," [http://www.epic.org/privacy/tools.html], accessed May 28, 2003.
[25] Beth Mack, "Keep It To Yourself," *Marketing News*, November 25, 2002, p. 21..

THE STUDY AND THE POPULATION

We decided to focus on U.S. adults who have and use internet connections at home. Surveys indicate that they can be found in about half of U.S. homes.[26] Of course, many people go online both at home and elsewhere, especially work, and we included them in our sample. We did not include adults who use the web only outside the home—at work or in the library, for example. The reason is that using the web in the home raises issues of personal control over information that may not be true elsewhere. Information technology personnel at work may install firewalls and filters so that employees may feel that their information is protected from outside intruders in ways that people who go online at home do not. At the same time, office workers may worry primarily about their company's surveillance of their internet activities. Adults who go online exclusively from non-domestic locations may consequently hold different concerns about privacy, and have different ways to deal with them, than those who also go online at home. This is an important topic that ought to be explored in a separate study.

Our survey was carried out by International Communication Research/ICR from January 30 to March 21, 20003. To get a rough comparison of changes in privacy concerns we repeated questions that we had asked of a nationally representative sample of parents in 2000. We added new questions that explored people's understanding of privacy policies on the internet, whether they know how to protect their online information, whether they take steps to do that, what institutions they believe will help them control their information online, and whether or not they agree that certain policy approaches would be effective in helping people to protect information about themselves on the web.

Telephone interviews, which averaged 20 minutes, were completed with a nationally representative sample of 1,200 adults age 18 and older who said responded "yes" when asked "do you use the internet at home?" We used a nationally representative RDD (random digit dial) sample to screen households for adults age 18 or older who use the internet at home. We were able to determine that 53.3% of households that we phoned had at least one household member who met our eligibility requirements. Among those households, the percentage of eligible individuals who completed an interview, or the cooperation rate, was a remarkable 66.4%. The data were weighted by age, education, and race to the 2001 consumer population survey (CPS), which asked adults ages 18 or older questions similar to that used in the internet privacy study to ascertain internet use at home.[27]

---

[26] The CPS Internet and Computers survey (September 2001, N=143,000) found adults who use the internet at home in 54.9% households. A Centris study is more recent (February 1-28, 2003, N=7342) but also a bit more conservative because it asked respondents if they personally accessed the internet at home in the past 30 days. It found an incidence of 41%. For this survey we asked "do you use the internet at home?"

[27] Our unweighted data was actually remarkably similar on these categories to the CPS as well as Centris and Pew Internet and American Life surveys from 2002. We used the CPS because of its huge number of respondents (143,000) and reputation as the gold standard for weighting. The margin of error for reported percentages based on the entire sample of 1,200 is plus or minus 2.86 percentage points at the 95% confidence level. The margin of error is higher for smaller subgroups within this sample.

Tables 1 and 2 provide an introductory snapshot of the population we interviewed and its internet use. As Table 1 indicates, men and women are about equal in number; 77% designate their race as white (blacks and Hispanics together make up 13% of the total); about half are under age 45; and about half are parents of children under aged 18. Most have had at least some higher education, and while a substantial percentage say their household brings in more than $75,000 annually, a firm claim about this population's income distribution is difficult because one fifth of the respondents did not want to reveal it.

Table 2 indicates that almost half the adult population (46%) who use the internet at home has been going online from home for fewer than five years. Currently, 62% say they use dial-up phone connections to go online, but 36% of these individuals report already being connected via cable or DSL broadband. 97% of our sample has gone online at home during the past month; 49% say they have also used it at work during that time.

Adults who go online from home also seem to enjoy the experience. As Table 2 notes 77% agreed or agreed strongly with the statement that "the more years I have the web, the more interesting it becomes." It is understandable, then, that this population also reports being quite active on the internet. 53% of the adults say they go online several times a day from home or outside home (for example, at work or the library). Fully 75% report going online from somewhere at least once a day, and 47% say they do it from home for an hour or more on a "typical" day.

The table also indicates that the great majority of adults who use the web at home rank themselves in the middle (intermediate or advanced) rather than lowest or highest range (beginner or expert) of abilities when it comes to navigating the internet. Only 14% consider themselves beginners and only 13% call themselves experts. 42% consider themselves intermediates and 30% say they are advanced. More years online, using the Internet daily, staying online an hour or more, or going online at work all increase the likelihood a respondent will increase in expertise " at navigating the web. So do higher income levels and being male.[28]

[28] The optimal scaling regression method was used to explore these relationships with the ordinal dependent variable. The eight variables explained 32% of the variance. Interestingly, age shows a curvilinear relationship of age impact self-reported internet skill. That is, young people report high expertise; it drops as people get older; but then it rises again. Perhaps reported expertise increases because time spent with the internet increases among less busy older adults. More research is needed here.

Table 1: Characteristics of U.S. Adults Who "Use the Internet at Home"

	US Adults, Home Internet*
	(*N*=1,200)
Sex	%
Male	49
Female	51
Age	
18-34	33
35-44	24
45-54	21
55-64	11
65+	08
No answer	03
Race	
White	77
Black	07
Hispanic	06
Other	07
No answer	04
Education	
Less than high school (HS) grad	07
High school/tech school graduate	32
Some college	22
College graduate or more	39
Family Income	
Less than $40,000	24
$40K but less than $50K	10
$50K but less than $75K	19
$75K but less than $100K	13
$100K or more	13
No answer	21
Parental Status	
Parent of child below age 18	56
Not parent of child below age 18	44

* When the numbers don't add up to 100% it is because of a rounding error.

Table 2: Internet activity, interest and self-ranked expertise of U.S. adults who "use the internet at home"

	(*N*=1,200)
Online connection	**%**
Dial-up telephone	62
Cable modem	23
DSL	13
Another method	01
Don't Know	01
Years online at home	
One or less	09
Two	09
Three or four	28
Five	13
Six	08
Seven or more	28
Don't know	04
Response to "The more years I have the web, the more interesting it becomes."	
Agree strongly	44
Somewhat agree	33
Somewhat disagree	13
Strongly disagree	08
Neither agree nor disagree	02
Frequency online from anywhere	
Several times per day	53
About once a day	22
A few times per week	19
About once a week	04
About once a month	02
Few times a year	01
Went online last month at home or work**	
At home	97
At work	49
Typical daily time online at home	
Less than 15 minutes	12
More than15 minutes, less than 1hour	39
Between 1 and 2 hours	29
More than 2 hours	18
No response	03
Self-ranked expertise in navigating the internet	
Beginner	14
Intermediate	42
Advanced	30
Expert	13

* When the numbers don't add up to 100% it is because of a rounding error.
** These numbers don't add up to 100% because going online at work and home are not mutually exclusive.

ENDURING CONCERNS ABOUT WEB PRIVACY

Comparing this study with one of parents in 2000 suggests enduring concerns about web privacy. When presented with the statement "I am nervous about websites having information about me," 76% of the beginners, 74% the intermediates and 70% of advanced users agreed. The self-designated *experts* were more likely than the others to dispute the statement, but even 57% of them agreed that they are nervous. Overall, our population confirmed what other studies have found: a clear majority of Americans express worry about their personal information on the web.

This survey went beyond a one-question expression of concern, however, to explore the attitudes and knowledge that adults who go online at home hold about what happens to their information on the internet. To begin with a rough sense of whether ideas on this topic have changed in the past few years, we included thirteen statements that we had used in a study of a more limited population in the year 2000--online parents (see Table 3). For each of the assertions, we asked our respondents how much they agreed or disagreed along a five-point continuum, from agree strongly to disagree strongly.

Table 3 allows comparison of the answers given by adults who either don't have kids or whose kids are younger than age 6 with parents with youngsters at home who fall into an age bracket (6 through 18) that make them likely to use the internet. The table also allows comparison of the current sample of parents of "internet age" children their counterparts in our 2000 study. What is most interesting is how close the percentages are, not just between parents and non-parents of internet age kids in 2003 but also between the parents of 2000 and those of today. Quite logically, the two areas of greatest difference between those with and without internet-age kids relate to a somewhat greater likelihood that the parents of those who could go online worry about what teens and "family members" might reveal to websites. Perhaps the most interesting difference between 2003 and 2000 is that a smaller percentage of people three years ago agreed that that they trust websites not to share information when they say they won't (37% vs. 50%). Parents, at least, appear to have gotten more rather than less trusting. In general, though, the responses across groups and time were strikingly parallel to one another.

Beyond reflecting concerns about outsiders invading their privacy, the pattern of answers are a springboard to four themes that speak to the major questions posed earlier:

The great majority of adults who go online at home reject the general proposition that their information is a currency for commercial barter. Only 21% agree that they like to give information to websites in exchange for offers, and only 16% agree that they will give out information only if paid. The answers mirror responses by the parent sample in 2000. They contradict analysts who characterize most Americans as quite open

Table 3: Among Adults Who Go Online at Home, the Percentage Who "Agreed" or "Agreed Strongly" With the These Statements:

	Total (N=1,200)	Non-Parents* in 2003 (N=775)	Parents* in 2003 (N=425)	Parents* in 2000 (N=902)
I should have a legal right to know everything that a web site knows about me.	94	94	95	95
Teenagers should have to get their parent's consent before giving out information online.	92	92	93	95
I am nervous about websites having information about me.	70	68	73	72
I look to see if a web site has a privacy policy before answering any questions.	71	69	72	72
My concern about outsiders learning sensitive information about me and my family has increased since we've gone online.	67	67	68	61**
I am more concerned about giving away sensitive information online than about giving away sensitive information any other way.	68	66	68	64
When I go to a web site it collects information about me even if I don't register	59	58	59	57
I would worry more about what information a teenager would give away to a web site than a younger child under 13 would.	58	53++	66	59
I trust web sites not to share information with other companies or advertisers when they say they won't.	49	50	50	37**
Web site privacy policies are easy to understand	47	45++	53	45**
I sometimes worry that members of my family give information they shouldn't about our family to web sites.	28	25++	35	37
I like to give information to web sites because I get offers for products and services I personally like.	23	21	25	17**
I will give out information to a website only if I am paid or compensated in some way.	16	16	17	10**

*Parents with children six to eighteen years. "Non-parents" means adults who do not have children six to eighteen years. ** indicates that the difference between the two samples of parents is significant statistically at the .05 level using the chi square statistic. ++ indicates that the difference between the 2003 sample of parents and non-parents is significantly statistically at the .05 level using the chi square statistic.

to giving up their information if the price is right. Philosophically, if not always in practice,[29] adults who use the web at home do not see their personal information as a commodity to be traded for online offers.

- **Most adults who go online at home know that websites track their behavior, but two in five are ignorant about the most basic aspect of information collection on the internet.** 59% are aware of what cookies do; they know that when they go online sites collect information on them even if they don't register. The flip side of the finding is that 40% of U.S. adults who use the internet at home are not aware of this most basic way that companies track their actions when they go online. Yet 76% of them say that "they look to see if a website has a privacy policy before answering any questions." In addition, 69% say they "always" or "sometimes" give their real email address to a website when it asks for personal information. Because privacy policies almost always mention cookies, the answers suggest that even though these people say they "look to see if a website has a privacy policy," the great proportion of online adults who aren't aware of what cookies do either don't actually read the policies or don't understand them.
- **The attitude statements also reveal that beyond being nervous over their sense of being tracked, most Americans want help to control their information.** 95% agree that they should have a legal right to know everything a website knows about them. Moreover, contrary to the U.S. government policy that teens are adults online, 92% of our respondents overwhelmingly agreed that teenagers should have to get parents' consent before giving out information online.

Comparison with the sample of parents in 2000 suggests that these key ideas are stable and generalizable. The current wider survey of all adults who use the web at home asked additional questions that aimed to deepen our understanding of them. The answers allow us to marshal more data to support the themes and add to them. We start with a question that relates to the second theme: What do adults who use the internet at home know and don't know about the way information about them is used on the web?

[29] Our 2000 study of parents found that 29% of parents with online connections at home said they would give their names, addresses, and preferences to a site of their "favorite" store in return for "a great free gift" worth up to $100 and a promise not to share the information with other companies. 71% of the parents said they would not. A Forrester report concluded in 2002 that one-third to one-half of consumers are willing to give up such information as their TV viewing history and their online surfing in exchange for a $5 monthly discount on their cable or ISP bill. Jed Kolko with James McQuivey and Jennifer Gordon, "Privacy for Sale: Just Pennies Per Day," Forrester Research *Technographics Research Brief*, June 11, 2002. The key question the Forrester study raises involves whether the respondents understood the uses that could be made of their data. The issue will be taken up in the conclusion to this paper.

NOT UNDERSTANDING DATA FLOW

Despite strong concerns about government and corporate intrusions, American adults who use the internet at home don't understand the flow of their data online. Our survey reveals a disconnect between their concern about information about them online and their knowledge about what websites do with it. Though they possess basic knowledge about the websites' acquisition and use of information about individuals, adults with internet connections at home are ignorant, even naïve, about the way data about them flows between companies behind their screens.

First, some additional privacy concerns: Our current study aimed to assess opinions about government surveillance that have arisen since the 2000 survey because of the World Trade Center destruction and the consequent "war on terrorism." As Table 4 indicates, a bit more than half of the adult population that goes online from home believes that "government agencies" are collecting information about them without their knowledge or consent. The online adults see some utility of for government surveillance. Depending on how the statement is phrased, 66% or 45% believe that the government should have the wherewithal to track evildoers (and even potential evildoers) online.

Table 4: Among Adults Who Go Online at Home, the Percentage Who "Agreed" or "Agreed Strongly" With the Following Statements:

	Total (N=1,200)
	%
Because of the war on terrorism, the government needs to make it easier for law enforcement to track users' online activities without their knowledge or consent.	66
US government agencies are collecting information about me online without my knowledge or consent.	52
In the interest of national security, the federal government should have the technology to find out what anyone is doing on the Internet at all times.	45
When a web site has a privacy policy, I know that the site will not share my information with other websites or companies.	57

And yet, the online-from-home population did not take this to mean that they were giving anyone the OK to collect information about *their* domains. Elsewhere in the interview, we asked respondents in two separate questions how concerned they would be if they found that the "US government" and "marketers" were "collecting information about

your household members' online activities without your knowledge or consent." 83% said they would be concerned if the government did it; 92% said they would be concerned if the snoopers were marketers.[30]

Although large proportions of the online-at-home adults voiced concern about their loss of privacy on the internet, much smaller percentages seem to have had actually tangled with the issue personally. Fully 82% of those interviewed said they had never had an incident where they worried about something a family member told a website. It may be that the concerns they described in the interviews came from media or interpersonal discussions without first hand experience to make them real. This seeming lack of a direct connection to personal privacy issues may explain how in a population where high proportions of adults who say they know how to register on sites (88%), understand that sites can track them (59%), and know how to change the privacy settings on their browser (64%), 57% mistakenly agree that the mere presence of a privacy policy means that a website will not share their information with other websites or companies.

The ignorance about privacy policies is, however, only the tip an iceberg of confusion about what goes with personal information behind the computer screen. The reactions of most online-at-home adults to a common way websites handle visitors' information indicate that they do not grasp the way their identifiable and anonymous data is collected, interrelated and used.

We presented the people interviewed with a supposed change in the information policy of a website that they had previously said they "like most or visit regularly from home." The goal was to gauge the acceptability of a common version of the way sites track extract and share information to make money from advertising. Unfortunately, it is impossible to determine an "average" or "typical" approach to information by websites. One reason is that it is not clear how to determine an average or typical website. More important, a website's approach to its visitors' information is by no means fully described in its privacy policy, long and tortuously worded though it may be. No law requires websites to disclose all aspects of their relationship to their visitors' information. The advertising trade press and conversations with people in the business, for example, makes clear that more than a few sites purchase offline data about individuals to append to data gathered during registration. The sites rarely divulge such transactions in their privacy policies, however.

Coming up with the description of a rather common privacy policy involved combining the experience of reading hundreds of privacy policies with a wide reading of the trade press on privacy-policy issues. The goal was to reflect the complex ways in which websites intend to explore patterns of visitors' personal and clickstream data with an eye toward selling them to advertisers. Most of the transactions using visitors' data are offered to advertisers in aggregate—that is, anonymously lumping people with one or another characteristic together for ad-targeting purposes. Some sites, however, do offer

[30] 50% of the respondents said they would be "very concerned" and 33% said they would be "somewhat concerned" if the government tracked them. 68% said they would be "very" and 24% "somewhat" concerned if marketers tracked them.

personally identifiable information directly to advertisers and say so in their privacy policies. Many sites say they share personally identifiable information only with so-called "affiliates"—though they rarely name them. Many more sites make it clear that if visitors click on advertising links, names given there (in contest registration, for example) may be used in ways counter to the website's policies. Websites also point out that they may change their policy at any time, and not all promise to keep previously collected data under the old regime. We strove to create an approach to personal information that would embody these data transactions along with their typical uncertainties and ambiguities without being too long.

We read the result to five web experts from academia, business, government and social advocacy groups who agreed that what we would be presenting was a common version of a site's approach to information. Accordingly, we integrated the hypothetical scenario into the questionnaire. After several questions asking them about the type of website, whether or not they registered to get in, whether or not they pay a subscription to use it, and if so, how much, we posed the situation this way.

> SUPPOSE THE WEB SITE THAT YOU LIKE MOST AND USE REGULARLY SAYS THAT IN ORDER FOR IT TO CONTINUE OPERATING IT MUST CHARGE USERS $6 A MONTH.[31] IF YOU PAY, THE SITE WILL SHOW YOU ADS BUT IT WILL NOT USE PERSONAL INFORMATION ABOUT YOU TO MAKE MONEY FROM OUTSIDE ADVERTISERS. OR YOU CAN GET THE SITE FOR FREE IN EXCHANGE FOR ALLOWING THE WEB SITE TO USE PERSONAL INFORMATION ABOUT YOU TO MAKE MONEY FROM ADVERTISERS. IT WILL LEARN ABOUT YOU BY GETTING YOUR NAME AND MAIN EMAIL ADDRESS, BY BUYING PERSONAL INFORMATION ABOUT YOU, AND BY TRACKING WHAT YOU LOOK AT ON THE SITE. THE SITE WILL NOT DIRECTLY TELL ADVERTISERS MOST OF THE INFORMATION IT LEARNS, THOUGH IT MAY TELL ADVERTISERS YOUR EMAIL ADDRESS. IT WILL SEND ADS TO YOU FOR ITS ADVERTISERS BASED ON THE INFORMATION IT LEARNS. FOR EXAMPLE, IF YOU CLICK ON FOOTBALL LINKS, IT MAY CONCLUDE THAT YOU LIKE SPORTS, BELONG TO A PARTICULAR AGE GROUP, AND PROBABLY DRINK BEER. THE SITE WILL SEND YOU ADS ON THE SITE, THROUGH EMAIL AND MAYBE THROUGH POSTAL MAIL, BASED ON THE INFORMATION IT LEARNS.
>
> SO, IF THE SITE YOU LIKE MOST AND USE REGULARLY SAYS IT MUST CHARGE YOU OR USE YOUR INFORMATION TO MAKE MONEY FROM ADVERTISERS, WHAT WOULD YOU DO? WOULD YOU
>
> 1 AGREE TO PAY TO USE THE SITE SO THAT THE SITE CANNOT USE YOUR PERSONAL INFORMATION TO MAKE MONEY FROM ADVERTISERS?
> 2 AGREE TO GET THE SITE FOR FREE IN EXCHANGE FOR ALLOWING THE SITE TO USE YOUR PERSONAL INFORMATION TO MAKE MONEY FROM ADVERTISERS?

[31] If the respondent was already paying, we changed this amount to the number he/she had previously given plus a sliding extra number of dollars based on the existing payment; it typically came to $2 extra. 11% of the respondents told us they were paying to use their valued site. Monthly payments ranged from $2 to $100; the average monthly payment reported was $21.

3 LOOK FOR A SUBSTITUTE WEB SITE THAT DOES NOT CHARGE? OR
4 GIVE UP LOOKING FOR THAT TYPE OF CONTENT ON THE WEB?

[IF THE RESPONDENT CHOSE #3, WE THEN EXTENDED THE SCENARIO TO FORCE A CHOICE, AS FOLLOWS:]

SUPPOSE YOU CANNOT FIND A SUBSTITUTE WEB SITE THAT DOES NOT CHARGE, WHAT WOULD YOU DO THEN? WOULD YOU--

1 AGREE TO PAY TO USE THE SITE SO THAT THE SITE CANNOT USE YOUR PERSONAL INFORMATION TO MAKE MONEY FROM ADVERTISERS?
2 AGREE TO GET THE SITE FOR FREE IN EXCHANGE FOR ALLOWING THE SITE TO USE YOUR PERSONAL INFORMATION TO MAKE MONEY FROM ADVERTISERS?
3 GIVE UP LOOKING FOR THAT TYPE OF CONTENT ON THE WEB?

Table 5 presents the initial answers from the respondents who could think of websites that they "like most or visit regularly from home."[32] Note that only 10% agreed to continue getting the site for free in return for agreeing to this common version of the way sites handle personal information from advertising. Oddly, 21% said straight out they would give up looking for that type of content on the web when presented with such a choice. Perhaps they were angry that a site would give them this sort of choice. 18% said they would rather pay to use the site than agree to give up their information, while almost half—48%—suggested that they would try to retain their information and money by looking for a substitute site.

Table 5: If the site … says it must charge you or use your information …, what would you do?"*

	Total (N=919)
	%
Agree to get site for free and give up information	10
Agree to pay to use the site	18
Look for substitute site that doesn't charge	48
Give up looking for that content on the web	21
Don't know / refused	03
Total	100

* See text for explanation.

When the second question blocked this way out, only a small percentage of those stymied decided to use the marketing deal for free access to the valued site. Table 6 presents the

[32] Approximately 12% (140) of the 1200 people in the same could not think of such a site, so they were not asked the questions. In addition, an error caused another 142 people in our sample were not to get the questions. (The error did not systematically bias the kinds of people who received the hypothetical scenario.) Overall, then, 918 respondents answered this set of questions.

final decisions of all the respondents—the people who did and those who did not first say they would look for a substitute site. The central finding is that 85% of our sample did not accept an approach to privacy that is common on today's internet. Moreover, while 27% said they would pay for the site, a bit more than half—54%—contended that when presented with this website approach to their information they would rather give up looking for that type of content on the web than either pay or accept the information policy.

Table 6: Final decisions of all respondents regarding scenario*

	Total (N=919)
	%
Agree to get site for free and give up information	15
Agree to pay to use the site	27
Give up looking for that content on the web	54
Don't know / refused	04
Total	100

* See text for explanation.

The massive rejection of what is actually a common version of the way sites track, extract, and share information to make money from advertising suggests that adults who go online at home overwhelmingly do not understand the flow, manipulation and exchange of their data invisibly during and after they go online. Other findings indicate that a substantial subset of the people who refused to barter their information is especially ignorant about information activities on the web. Among the 85% who did not accept the marketing deal, about half (53%) had earlier said they gave or would be "very" or "somewhat" likely to give the valued site their real name and email address. Yet those bits of information are what a site needs to begin creating a stream of data about them—the very flow (personally identifiable or not) that they refused to allow in response to the scenario. Moreover, 63% of the people who said they had given up these data had also agreed that the mere presence of a website privacy policy means that it won't share data with other firms. Bringing these two results together suggests that least one of every three of our respondents who refused to barter their information either do not understand or do not think through basic data-collection activities on the internet.[33]

[33] As it turns out, the 15% of our sample who accepted the marketing deal did understand privacy policies and data collection any better than the others. 67% believed that when a web has a privacy policy if will not share knowledge (not a statistically significant difference from those who rejected the deal), though 58% indicated an awareness of cookies (not a statistically significant difference with the others). 39% both knew of cookies and misunderstood the presence of privacy policies—also not different from the other group. What makes these people stand from the 85% is not their knowledge; they too seem ignorant and confused. It is, rather, their seeming willingness to give up data whether or not they know what is happening to that information: 80% of this group (compared to 53% of the other) had earlier indicated they had or would likely give their real name and email address to the site.

The converging results point to a confusion about the nature of information gathering on the web. Although web users seem to be responding to public discussions of cookies as repositories of specific data about them—and while that in itself (rather than bad personal experience) seems to make them concerned—they do not understand that this collection of individual bits of information relates to a larger set of activities that involve the tracking, mining, and sharing of data. When they learn about it—as when we read them the scenario—they refuse to accept it as legitimate.

We found additional evidence that a substantial majority the online-at-home adults does not understand—and would reject—the complex ways websites and marketers extract and interrelate data about them. Those findings came as the result of a second scenario we created for the 440 people who said that they would go to a substitute site for favored content rather than pay or give up information. We told them to suppose that they agreed to let the substitute site track their movements and link them to other information about them. We then asked what their reaction would be if the focus of the information tracked would be their fashion preferences, political interests, health or medical history, gender, and financial information. Would they agree to pay so as not to be tracked, allow tracking and get the site for free, or give up looking for that content on the web?

As other studies have found, we noted variations in people's sensitivities to different topics when it comes to privacy. For both financial information and health or medical history, 84% of the respondents said they would give up looking for favorite content on the web than pay for the site or allow that information to be tracked and shared by marketers. When it came to political preferences, 75% said that if those were tracked they would give up looking for their favorite content on the web. With gender and fashion preferences, a smaller percentage contended they would abandon favorite content on the web. Even there, though, substantially more than half of the respondents (63% and 67%, respectively) say they would leave the web rather than pay or be tracked was high.

When one considers that people often give out their gender, fashion preferences, and even political preferences to websites and pollsters, these numbers appear bizarrely high. That is particularly the case considering that an average of 61% of those who said they would give up looking for content earlier said that they had or would likely share their real name and email address with the site. The pattern of answers suggests that their concern went beyond the nature of the information that would be released about them. Rather, it reflected worries about—perhaps even indignation over—what they learned regarding the website's tracking, manipulation, and sharing of data about them.

NOT TAKING STEPS TO LEARN

Not only do adults who use the web at home tend to be confused about data-collection activities, they tend not to take steps to learn about ways to control their information online. When asked how often they searched for "instructions on how to protect information about yourself on the web?" 64% answered never, while 25% said "a few times; 5% said "only once" and 6% said "many times." In answer to another question, 40% of adults who use the internet at home also told us that they know "almost nothing" about how to stop websites from collecting information about them.

We turned to the 60% of the population who said that they know more than "almost nothing"—that is, those who indicated at least some understanding about controlling their online information. We asked them whether they feel they have applied what they do know in ways that are sufficient. Only 5% agreed that they had carried out "everything that needs to be done" to stop websites from "collecting personal information" without their "knowledge or consent." The majority of people who have at least some knowledge about privacy control said they have done "some but not enough" to stop information collection. 20% said they have carried out either very little or nothing of what needs to be done.

Table 7 presents specifics about what all our respondents said they have actually ever carried out in relation to controlling their information. Fully 65% said that the have erased unwanted cookies at least once. This finding is consistent with our earlier realization that a clear majority of the sample is aware that cookies are a key component of information retrieval. The percentage applied other privacy tools drops steeply from there, however. 43% said that they have used filters to block unwanted email, 23% said they have used software that looks for spyware, and an even smaller percentage said they have used anonymizers—"software that hides your computer's identity from websites that they visit."

To gauge how experienced individuals are with the range of these practices, we gave them scores based on the number they reported performed. Four points went to people who said they have carried out all of these activities, three to those who have done three of them, and so on. We found that fully 25% had not carried out any of these information-controlling activities (we called them *highly inexperienced*). 31% had carried out one task (*inexperienced*). 25% were in the middle with two of the four (*neither experienced nor inexperienced*), only 11% fell into the *experienced* slot, and an even smaller 8% claimed to be *highly experienced*—having at least some skill at carrying out four of the four information-controlling activities.

Table 7: Have you ever--

	Yes %	No %	Don't Know %	Total % **
Erased all or some of the unwanted cookies on your computer?* (N=1200)	65	33	2	101
Used filters to block unwanted email? (N=1200)	43	57	1	101
Used software that looks for spyware on your computer.* (N=1200)	23	76	2	101
Used software that hides your computer's identity from web sites that you visit. (N=1200)	17	81	2	100

* If respondent asked what cookies are, the interviewer said, "Files internet firms place in your computer to track your movements on the web. If respondent asked what spyware is, the interviewer said, "Software that records every keystroke made on a computer."
** Total percentages exceed 100 because of rounding error.

One might expect that the amount people say they know or do to control their information would relate to the way they rank their ability to navigate the internet. And, in fact, a much higher proportion of those rated as highly experienced or experienced compared to everyone else (27% versus 8%) said that they know "a lot" about stopping web sites from collecting their personal information without consent. Similarly, 40% of the experienced categories compared to 20% said they know "some" about the subject. The same tendencies applied when we asked the people who said they knew more than "almost nothing" about how to control their information. People who were ranked *highly experienced* or *experienced* were far more likely than the others to say they carry out "everything that needs to be done" or "some but not enough" as opposed to very little or nothing.

For those who want to encourage more citizens to control their information online, an obvious path is to cultivate internet users who are experienced with privacy-protecting technologies. At present only 19% of adults who go online from home fall into either the *highly experienced* or *experienced* categories. The rest—from *neither experienced nor inexperienced* through *highly inexperienced*—are both much less knowledgeable and much less active about controlling their online data.

Unfortunately, we could not find out what characteristics or activities foretell whether or not a person will be more or less experienced in this regard. We used a statistical technique called optimal scaling regression. It helped us explore whether a variety of background characteristics that we expected would encourage concern with online privacy would, in fact, predict a higher score on privacy-tool experience. In addition to demographic characteristics such as age, income, race, education, and gender, and region of the country, we were interested in whether having a child aged six to seventeen who uses the internet leads someone to learn more privacy tools. We also thought that incidence of internet use and self-reported ability to navigate the web might pay important roles in leading a person to be privacy-tool experienced.[34]

[34] In our model, *incidence of internet use* involved three variables—years on the internet (prior to 1997 to present—2003), use/non-use of the internet at home during the past month, daily vs. weekly use of the

It turned out that among all the variables, only the time spent online (specifically, weekly versus daily and spending more than one hour on the internet) could be seen to impact involvement with privacy tools. Our statistical technique indicated, however, that even these variables predicted only 7% of the factors that drive experience with them. Overall, our model accounted for just 11% of the variance and so explains little about why certain individuals learn a number of ways to control their information online and others do not.

internet, and spending minutes vs. hours online. Linear relationships were test for age and income. Curvilinear relationship was also tested for age.

AGREEING WITH STRAIGHTFORWARD SOLUTIONS

Possibly because of their ignorance of what happens to their information online and how to control it, adults who use the internet at home agree widely and strongly when presented with solutions that let them know straightforwardly what is going on.

They strongly support regulations that force more disclosure from online entities. We have already seen in Table 3 that 95% of adults who use the internet at home agreed or agreed strongly that they should have the legal right to know everything websites know about them. 92% agreed or agreed strongly that teens should be required to get their parent's consent before giving out information online. The table does not reflect the intensity of those answers: 86% percent agreed *strongly* with the first proposition and 76% agreed strongly with the second. 80% also agreed strongly and an additional 14% simply "agreed" with the statement, not presented in Table 2, that "websites should be required to ask my permission before sending ads to me."

The respondents also agree that government regulations would be effective if they gave people leverage with online entities to control information about themselves. That sentiment came through in a series of questions toward the end of the interview. As the next-to-last questions before requesting basic demographic information, we asked about three potential policies in the following way:[35]

> COMPANIES SOMETIMES COMBINE ALL OF THE PERSONAL INFORMATION THEY COLLECT ABOUT YOU FROM YOUR ONLINE ACTIVITIES AT DIFFERENT SITES INTO A PROFILE OF YOU WITHOUT YOUR KNOWLEDGE OR CONSENT. PLEASE TELL ME IF YOU THINK ***A LAW THAT REQUIRES WEBSITE PRIVACY POLICIES TO HAVE UNDERSTANDABLE RULES AND THE SAME FORMAT*** WOULD BE VERY EFFECTIVE, SOMEWHAT EFFECTIVE, NOT VERY EFFECTIVE, OR NOT AT ALL EFFECTIVE WAY TO REGULATE THESE ACTIVITIES.
>
> [AFTER THE ANSWER:] HOW ABOUT ***A LAW THAT REQUIRES COMPANIES THAT COLLECT PERSONAL INFORMATION ONLINE TO HELP PAY FOR COURSES THAT TEACH INTERNET USERS HOW TO PROTECT THEIR PRIVACY ONLINE***?
>
> [AFTER READING THE CHOICES AND GETTING THE ANSWER:] HOW ABOUT ***A LAW THAT GIVES YOU THE RIGHT TO CONTROL HOW WEBSITES USE AND SHARE THE INFORMATION ABOUT YOU***? [READ CHOICES AND GET ANSWER.]

As Table 8 indicates, broad support emerged for all three policies. There is an important difference, however, in the response to the third policy in relation to the first two.

[35] The policies in italics were actually rotated so that different respondents received them in a different order. The actual last question before soliciting the demographic information was "when the current generation of teenagers in America reaches adult hood, do you think it will be much more, a little more, a little less or much less concerned about protecting information collected online than adults today?"

Compared to a law that would help them learn how to control their privacy, substantially more of those interviewed believed that legislation requiring easy-to-understand rules and the right to control information would be "very effective." Although people do not dismiss the possibility that formal learning about privacy tools can help society deal with information control, they seem to believe that government and corporate action that helps them learn straightforwardly what is going on is preferable.

Table 8: Among adults who go online at home, the percentage responses to the policies' probable effectiveness

	How Effective?*				
	Very %	**Somewhat %**	**Neither Effective nor Ineffective* %**	**Not Very %**	**Not at All %**
A law that requires website policies to have easy to understand rules and the same format. (N=1200)	40	46	0.5	8	4
A law that gives you the right to control how websites use and share the information they collect about you. (N=1200)	41	43	0.5	10	5
A law that requires companies that collective personal information online to help pay for courses that teach internet users how to protect their privacy online. (N=1200)	28	46	0.5	15	10

* Those small numbers who said "don't know" (2% and less) are not included. The people who said "neither effective nor ineffective" volunteered that answer.

CONFLICTED ABOUT WHETHER INSTITUTIONS WILL HELP

Yet online-at-home adults feel conflicted about whether the government or key corporate institutions will help them with their information privacy or take it away. We learned that by comparing two related sets of answers in our interviews. Each set asked about the same six institutions—the respondent's internet service provider (ISP), banks or credit card companies, major advertisers, Microsoft[36], privacy protection software, and "the government." We asked the person interviewed to "think about your ability during the next five years to control personal information online." In the first question set, the respondent was asked for every institution to note on a "on a scale of 1 to 5, with 5 being most important and 1 being least important, how important a role" that institution "will play in helping or teaching you to protect your information online." In the second set, for every institution the respondent was asked to note on a "on a scale of 1 to 5, with 5 being most likely and 1 being least likely, how likely will" that institution "be to release or share information about you by accident or on purpose without your knowledge or consent."

Table 9 lays out the average (mean) answers on the scale of 1 to 5 that each institution received for each question. In the interviews, numbers 1 and 2 indicated low levels of importance on the set of questions about the institution's role in protecting information. The numbers also indicated low levels of likelihood on the set of questions about the institution's likelihood to disclose information. 4 and 5 indicated high levels of importance or likelihood. We interpreted a response of 3 to mean neither high nor low.

As Table 9 indicates, adults who go online at home tend to consider major advertisers the least important of the six institutions to help them protect their information and the most likely to disclose it without consent. The adults also tend to see makers of privacy protection software as the most important of the six institutions to help them protect their information and the least likely to disclose it without consent.

The findings about advertisers and makers of privacy protection software are not really surprising. Concern about spam, the popular press' focus on marketers' use of cookies on the web, and a long history of distrust of advertisers in U.S. society make it logical that people would consider them least helpful in protecting information and most likely to disclose it. Similarly, constant injunctions in the press about the importance of virus protection software have given that part of the internet industry a favorable image that may well have rubbed off on "privacy protection software makers." It should be noted—and the means suggest—that these sentiments were by no means unanimous. Only 45% of the respondents indicated through a 1 or 2 that advertisers would be unimportant to helping protect their privacy. 32% thought they would be important (a 4 or 5), while 21% believed neither. And, while 64% did agree that advertisers would likely share their information, 17% said it was unlikely and 18% said neither. Roughly the same

[36] Though it is only one company, Microsoft's fundamental influence on the digital world led us to include it here even though our other examples were groups of organizations.

numbers—but reversed for the two questions—apply to the privacy-software manufacturers.

Table 9: How important will institutions be for helping protect your information? How likely will institutions be to release your information?

	Mean Response on Protect	Mean Response on Release	Difference Between Means	Effect Size
Major advertisers (N=1175*1185)	2.78	3.79	-1.01	-.88
Microsoft (N=1165*1156)	3.45	3.20	.25	.10
The government (1179*1171)	3.53	3.26	.27	.24
Banks/credit card companies (N=1189*1181)	3.75	3.32	.43	.34
Internet service providers (N=1189*1183)	3.68	3.19	.49	.47
Makers of privacy protection software (N=1177*1165)	3.86	2.97	.89	1.18

On "protect": 5 is "most important." On "release": 5 is most likely. See text. The means in every pair are statistically significant using the paired-samples t test. Standard deviations going down the first column of means are 1.471, 1.331, 1.382, 1.390, 1.247, and 1.164. Standard deviations going down the second column of means are 1.371, 1.284, 1.411, 1.413, 1.283, and 1.350. The different N for each variable and column reflects that "don't know" and "refused" were not calculated in the means.

Lack of homogeneity in these answers also applies to the other institutions in Table 9. What is particularly noteworthy about Microsoft, the government, banks/credit card companies, and internet service providers, however, is that all their means in the table exceed 3 (that is, they fall in the "important" and "likely" range) on both the first and second of questions. Moreover, the differences in these means, while statistically significant, are small—less than .5. Their *effects size*, a widely accepted measure of the extent to which these differences between means really make a difference, range from relatively small (for Microsoft and the government) to small-to-moderate (banks/credit card companies and internet service providers).[37]

Taken together, these findings indicate two related points: First, respondents tend to rank the institutions as somewhat more important for protecting their information as for having the likelihood to disclose it. But two, the effect sizes reflect that the proportions of respondents who believe the institutions are important for helping them protect their information are not that different from the proportions who believe that they will likely disclose their information without people's knowledge or consent. An example with percentages might make the point a bit clearer: While 51% of the respondents said that the government would be important to helping protect privacy, 44% said that the government would likely disclose information about them.

An obvious question then arises: What proportion of respondents believes both? That is, how many suspect an institution that actively helps them pursue their privacy concerns also surreptitiously discloses their information? By contrast, how many respondents trust

[37] The effects size was calculated by dividing each mean in the pair by its standard deviation (to standardize it) and then subtracting the resulting two numbers.

an institution to actively help them pursue their privacy concerns without then disclosing their information? And more: How many do not trust the institution to help them, are caught in a conflict about the institution's information protecting and disclosing activities, or for some reason have not formed a strong opinion on the relationship between the institution and their privacy?

To answer, we created a new variable that merged the answers to the two sets of questions on each institution. If a respondent answered that an institution would be important in helping to protect information online and then said it would be unlikely to disclose information, we considered that the person *trusts* the institution to actively help with information privacy. If a respondent answered that the institution were unlikely to help in protecting information but then said it would be likely to disclose information, we considered that the person *does not trust* the institution to actively help with information privacy. If the person indicated that the institution was "unimportant" with helping to protecting information *and* "unlikely" to release it—or "neither"—we considered the respondent felt *neither trusting nor untrusting* toward the institution when it came to information privacy. Finally, if the respondent indicated that the institution would be important in helping to protect online information but then also indicated that the same institution would likely disclose personal information, we considered that person *conflicted*.

Table 10: Trust / distrust that institution will help protect information online and not release it without knowledge or consent.

	Distrust %	Neither %	Trust %	Conflicted %
Major advertisers (N=1198)	40	34	4	23
Microsoft (N=1189)	15	50	12	23
The government (N=1191)	17	43	13	26
Banks/credit card companies (N=1198)	16	35	18	31
Internet service providers (N=1196)	16	35	18	31
Makers of privacy protection software (N=1188)	8	45	25	23

The different N for each variable reflects when respondents said "don't know" or "refused" on both "protect" and "release." See text.

Table 10 presents the results of this analysis for all six institutions. It shows that with the exception of major advertisers, straight trust or distrust is not the mode when it comes to information privacy. Between one-third and half of the respondents simply sit on the fence, not believing that they can trust or distrust an institution when it comes to privacy. Between one-third and one quarter of the rest are conflicted about how these key institutions of the digital world relate to their privacy. They seem to feel that while institutions will help them with control their information online, those same institutions (or other parts of them) will also take that information privacy away.

CONCLUDING REMARKS

The findings in this report must be dispiriting for those who believe in giving citizens the wherewithal to control their information on the internet. We found that despite their strong concerns about online privacy, most adults who use the internet at home misunderstand the purpose of a privacy policy. Just as important, our findings indicate that despite fairly wide awareness that websites collect information about them, adults who use the internet at home are fundamentally unaware of data flow: how organizations glean bits of knowledge about individuals online, interconnect those bits, link them to other sources of information, and share them with other organizations.

This ignorance of data flow stands at the heart of the imbalance of power that currently exists when it comes to controlling personal information online. In many ways, it is the ability to mine and manipulate data about individuals that makes interactive digital media such as the internet so attractive to marketers and governments. The activity is in relative infancy, but it is likely to grow enormously in presence and profits during the coming decades. Marketers and media firms, for example, see increased sophistication in real-time transactional databases as critical to the success of audience targeting, content-tailoring, and customer relationship management activities of the twenty-first century.[38]

When consumers are unaware of the data flows that take place behind their screens, they cannot really engage in the kinds of informed cost-benefit analyses that writers such as Alan Westin suggest take place when consumers "pragmatically" give up information about themselves. What consumers can't evaluate are the costs involved when marketers or governments hitch seemingly trivial information the consumers have allowed them to track, such TV viewing habits or fashion interests, to other knowledge in order to create powerful profiles about them. Correct or not, the profiles can impact people's lives in ways they can't control for lack of knowledge. Online and offline media might change content depending on what the media firms and their advertisers "know" about them. The consumers might receive different ads and different discounts than they had in the past. Government agencies might pay more or less attention to them than to others.

This study found that when adults who use the internet at home are brought face-to-face with a common approach to collecting, interconnecting and using their online information, they overwhelmingly reject it. It is also important to note, however, that these people don't go out of their way to learn what is going on with their online information. 64% say they have never searched for instructions on how to "protect information" about themselves on the web. Large percentages of online-at-home adults have little, if any, experience with basic internet privacy tools.

Why haven't these people tried to understand what happens to their information online and what to do about it? One reason may simply be that they have many other things to

[38] See Joseph Turow, "Marketing Trust and Surveillance in the New Media World," presented at *The New Politics of Surveillance and Visibility* conference, University of British Columbia, May 23-25, 2003.

do—56% are parents of a child under age 18, for example. Our survey also suggests a more basic, though related, reason: so far, they personally haven't suffered from it.

Recall that 82% of those interviewed said they had never had an incident where they worried about something a family member told a website. Recall, too, our finding that 77% of the respondents said that the more years they have the web, the more interesting it becomes. Add to those findings both a misperception that all privacy policies provide at least some security and the fact that data flows take place invisibly, behind the screen, while a person is engaged with what is on it. In this context, it is not at all difficult to understand why adults who say they are concerned about the collection of information online without their permission nevertheless know and do little about it.

Based on these findings, one wonders whether it is realistic to believe that most American consumers can be educated successfully about ways to protect their online information. The ignorance we found comes at a time when news and entertainment media constantly din people about online dangers. Moreover, there are currently many places online and off for people to learn about privacy protection tools. It may be that it will take a data-gleaning disaster—with publicity matching that of Enron's meltdown—to energize people to learn how to control their information. An alternative view is that technologies to extract and manipulate information about audiences for digital interactive media are becoming ever-more complex. Competitors vie with each other for the best approaches while trying to get around privacy-enhancing technologies. Perhaps it may be too much to expect ordinary people to keep up. It seems clear that, at the very least, that people need active help in protecting their information.

From that standpoint, it is particularly disconcerting that we found that such a small percentage of adults who use the internet at home trust key internet-related institutions to actively aid them protect their information while not also disclosing it without their consent. The largest percentage claims no strong stance on the subject—they neither trust nor distrust—while the second-largest proportion believes that institutions talk differently from different sides of their mouths: one side helps protect personal information while the other accidentally or purposefully releases personal information to outsiders without permission.

Adults who use the internet at home, then, know that they do not have the knowledge to control their information and are not sure whether major entities who have that knowledge will act in consumers' best interests. It therefore makes sense that when offered policy choices our respondents overwhelmingly agree with solutions that let them know straightforwardly what is going on. They strongly support regulations that force more disclosure from online entities. They also strongly agree on the effectiveness of government regulations that give people leverage with online entities to control information about themselves.

Bringing together this study's findings suggests that three policy initiatives are needed to address citizens' desire to control their information in direct, straightforward ways:

- First, federal legislation ought to require all websites to integrate the P3P protocols into their privacy policies. That will provide a web-wide computer-readable standard for websites to communicate their privacy policies automatically to people's computers. Visitors can know immediately when they get to a site whether they feel comfortable with its information policy. An added advantage of mandating P3P is that the propositional logic that makes it work will force companies to be straightforward in presenting their positions about using data. It will greatly reduce ambiguities and obfuscations about whether and where personal information is taken.
- Second, federal legislation ought to mandate data-flow disclosure for any entity that represents an organization online. The law would work this way: When an internet user begins an online encounter with a website or commercial email, that site or email should prominently notify the person of an immediately accessible place that will straightforwardly present (1) exactly what information the organization collected about that specific individual during their last encounter, if there was one; (2) whether and how that information was linked to other information; (3) specifically what other organizations, if any, received the information; and (4) what the entity expects will happen to the specific individual's data during this new (or first) encounter. Some organizations may then choose to allow the individuals to negotiate which of forthcoming data-extraction, manipulation and sharing activities they will or won't allow for that visit.
- Third, the government should assign auditing organizations to verify through random tests that both forms of disclosure are correct—and to reveal the results at the start of each encounter. The organizations that collect the data should bear the expense of the audits. Inaccuracies should be considered deceptive practices by the Federal Trade Commission.

The three proposals follow the widely recognized Federal Trade Commission goals of providing users with access, notice, choice, and security over their information. Companies will undoubtedly protest that these activities might scare people from allowing them to track information and raise the cost of maintaining databases about people online. One response is that people, not the companies, own their personal information. Another response is that perhaps consumers' new analyses of the situation will lead them to conclude that such sharing is not often in their benefit. If that happens, it might lead companies that want to retain customers to change their information tracking-and-sharing approaches.

The issues raised here about citizen understanding of privacy policies and data flow are already reaching beyond the web to the larger digital interactive world of personal video recorders (such as TiVo), cell phones, and personal digital assistants. At a time when technologies to extract and manipulate consumer information are becoming ever-more complex, citizens' ability to control their personal information must be both more straightforward and yet more wide-ranging than previously contemplated.