CAN I GET SOME PRIVACY?

How much do Internet companies know about us, and what do they plan to do with the information? If only we knew.

BY **BRIAN EULE**

PHOTO ILLUSTRATION BY **JAMES PORTO**

Assuming you possess a cell phone and a computer and a credit card, the following scenario, or something like it, might sound familiar.

Your morning begins with coffee and a bagel and the morning paper, perhaps read on a laptop. You click on stories about Egyptian unrest, the firearms industry and *Downton Abbey*. Two other websites are open on your desktop. One of them shows your Facebook account. You notice that you've been "tagged" in a photo from last week's poker game, in a pose that suggests one too many beers. Meanwhile, a friend has sent you a link to an article in the *Onion* that zestfully parodies a well-known senator. You "like" it.

You head out for your daily commute. At the toll booth, a Fastrak device validates the code on your car and records the date and time of your arrival.

You stop for gas. You swipe your debit card. The pump asks for your ZIP code and you type it in. As the 20-gallon tank fills, you pull out your smartphone and do a quick search for a weekend flight to Chicago. Along with the flight schedules and airfares, an advertisement appears about a local concert at the same venue where you attended a performance last month.

In the first two hours of your day, computers have recorded that you are a likely watcher of PBS, you drink alcohol and you have a penchant for irreverent humor. They know you drive a large vehicle and probably have family in the Midwest. They know when you go to work and the route you take. It's 8 a.m. and you've already left a sizable virtual fingerprint.

Now add the dozens of other electronic transactions you make in a given day—every website you visit, every item you purchase online, all the searches you do, all the posts you make on social media sites—plus those of all your friends. Multiply that by hundreds of days of Internet activity. Throw in motor vehicle records, mortgage documents, credit scores, medical diagnoses. What does your profile look like now?

Data about all of us lives online, in "clouds," on our web browsers and in others' databases. Cell phones show our physical location and track the places we have been. Websites display the address and price of home purchases, along with the buyer and seller. Advertising agencies know the web pages we have visited and the text we have entered online. Increasingly, and with increasing sophistication, companies are collecting, analyzing and selling data about tens of millions of people. And most of those people have no idea when or how it's happening.

"I don't think that people understand all the information that's out there about them," says Jennifer Granick, director of civil liberties at Stanford Law School's Center for Internet and Society. "People might not think that you can put it all together, but they're wrong. It's increasingly easy to figure out who people are. There is a treasure trove of information out there that is available."

The interdisciplinary CIS is helping to expose the massive asymmetry between the average consumer's understanding and practices that might threaten their privacy. Its scholars, along with privacy advocates in the nonprofit sector, are pushing for more transparency and stricter industry standards in how data is collected and used.

Concern about privacy intrusions often originates from an innocuous-sounding source: cookies. So named because of the "crumbs" of information they collect, cookies are codes imbedded in a computer hard drive that track web activity. They are legal and in many ways beneficial. For example, cookies "remember" passwords so repeat users of a site don't have to type it in every time they return. They save user preferences and enable basic Internet conventions like a shopping cart that makes online buying easier and less time-consuming. But a third party, unbeknownst to the user, also can set cookies that follow that user from site to site, gathering information about him or her. The proliferation of this practice has spawned a new business category: data brokers. These companies harvest public records along with web activity of all kinds, then mash it up with algorithms designed to help clients target potential customers with advertisements. Although individual names aren't attached to this data, scholars say there is sufficient information to tease out a person's identity.

GRANICK

"Web browsing history is inextricably linked to personal information," wrote Jonathan Mayer, a Law School student and a PhD student in computer science, and Stanford computer science professor

AMANDA AVILA (2)

John C. Mitchell, in a paper last year for the Institute of Electrical and Electronics Engineers Symposium on Security and Privacy. "The pages a user visits can reveal her location, interests, purchases, employment status, sexual orientation, financial challenges, medical conditions, and more. Examining individual page loads is often adequate to draw many conclusions about a user; analyzing patterns of activity allows yet more inferences."

AT AN EXTREME, piecing together information that exists about each of us can be used for identity theft. But that's rare in comparison to more typical concerns regarding the lack of control over who sees what personal information, how they use it and what decisions they base on it. Aleecia M. McDonald, director of privacy at the CIS, notes that banks might charge a higher mortgage rate for a customer whose friends on Facebook had negative credit events. Or, web merchants might adjust the price of products based on a customer's ZIP code. Much of the concern, McDonald notes, resides in the uncertainty over how all of the information will eventually be employed.

It's not just the things they disclose that people find troubling; "it's also this data leakage about what they do online and what they're interested in, their intellectual history and then also their friends," McDonald says. "They don't know where the data is going, they don't know how it's used, and they don't know what happens 10, 20, 40, 50 years from now."

Inferences based on what a user does online and who their friends are can be misleading. Car insurance companies already vary premiums based on demographics, but what if a user's Internet searches also informed a risk assessment? Taken out of context, most of us have conducted searches that might look suspicious if revealed in raw form. Employers are allowed to ask a job applicant to log in and show them their Facebook page during an interview. What if they also could see your search history? Might a college reject an applicant based on additional information that now lives online?

Earlier this year, Facebook announced a feature it called "graph search" which allowed users to search for others who have "liked" various topics or checked in at specific locations. Privacy advocates howled. Here was information people might have voluntarily shared, but did not expect to be catalogued. Information once known only to close friends might now more easily be found by strangers—and paired with other information. The Electronic Frontier Foundation, a nonprofit that champions consumers' digital rights, used the example of a graph-search-enabled query for "People who work at Apple, Inc. who like Samsung Mobile," information that, if shared, might put those employees in an awkward position. For its part, Facebook is encouraging all users to revisit their privacy settings, which locks down some of what others could find via graph search.

MCDONALD

Google logs massive amounts of information about its users and, "regularly receives requests from governments and courts around the world to hand over user data," according to the company's transparency reports. In the second half of 2012, Google received requests for information on more than 33,000 users' accounts and complied with 66 percent of those.

An investigation by the *Wall Street Journal* in 2010 found that, "the nation's 50 top websites on average installed 64 pieces of tracking technology onto the computers of visitors, usually with no warning." Twelve of them, it noted, installed more than 100.

Privacy concerns may vary by age. McDonald speculates that younger generations might be most vigilant about protecting their privacy from their parents. The middle generation might be most concerned with what employers or health care providers might learn about them. Regardless of age, much of the issue centers around control, or lack of it.

"The question, on some level, is 'Whose data is it?'" McDonald says.

And the problem isn't confined to for-profit companies. Last October, Mayer noticed an article in the *New York Times* about the use of third-party trackers by the Obama and Romney campaigns. Both campaigns claimed they had safeguards in place to protect users' anonymity. Mayer didn't buy it. "This seemed pretty implausible to me," he says. "It was frustrating, at this level of politics, that they were making this claim."

So he fired up an open source platform he had created, called FourthParty, that measures dynamic web content—sites whose offerings vary based on different information provided by the user or the program—and monitors interactions with web applications. Mayer had to give himself a screen name, so he went with "Leland Stanford." Then he entered some information and tried to see what ended up in the page codes that got passed along.

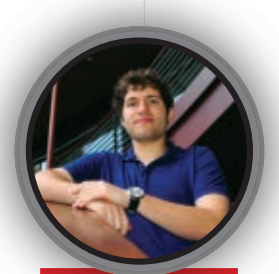
'(Consumers) don't know where the data is going, they don't know how it's used, and they don't know what happens 10, 20, 40, 50 years from now.'

Within a day, Mayer had confirmed his hunch. On both campaign sites, personal information—in some instances a user's name, in others an address or ZIP code—was included in the page web address that was given to the third-party trackers.

Mayer didn't think it was an intentional privacy breach, but he felt the parties should have known better than to claim they could keep the data anonymous.

Facebook presents a particular dilemma. The site is extraordinarily popular in part because it fosters connections by inviting people to share information. But its reach and aggressiveness in collecting user data are troubling, says Mayer. His research indicates roughly half of web browsers are logged into Facebook while users are visiting other pages. Each time those users visit a page that also has a Facebook icon, the information is sent back to Facebook. Even if the user doesn't click on that icon.

In the absence of strong controls, what are consumers to do to protect themselves? One strategy: Pay for privacy. Start-ups such as Reputation.com will scrub personal information from online databases for a fee. But while some people are willing to pay, critics say consumers need better options. "Having to pay a fee in order to engage in a retrospective effort to claw back personal information doesn't seem to us the right way to go about this," David Vladeck, then director of the Bureau of Consumer Protection at the Federal Trade Commission, said at a congressional hearing in 2010.



MAYER

Deleting cookies from one's computer is only a half measure. There are still other fingerprints left behind, Mayer says. Which version of which web browser they use, which Windows updates they have, which plugins they installed, the order of the updates they downloaded, and so on, all create a unique trail of sites visited. "Consumers by and large have no idea what's going on," he asserts.

Scholars at CIS are actively working to strengthen individuals' remedies. Each Wednesday, members of an international World Wide Web working group on tracking protection dial in to a conference call. Their mission is to "improve user privacy and user control by defining mechanisms for expressing user preferences around Web tracking and for blocking or allowing Web tracking elements." Representatives from academia and industry, including people from Microsoft, Apple, Facebook, Google and Mozilla, try to agree on a set of recommendations for the field. McDonald and Mayer both participate.

NARAYANAN

Much of the discussion stems from a relatively simple idea that Mayer and Arvind Narayanan, a former postdoc at Stanford, now an affiliate scholar at the CIS and professor at Princeton, helped demonstrate.

Around 2007, in response to increased tracking on the web, privacy advocates explored a Do Not Track program that would provide website users a means of blocking trackers. It would work much like the Do Not Call registry adopted to protect consumers from intrusive telephone marketers. It seemed more sensible to work from the user end, rather than having each company offer an opt-out, but many in the industry thought it was impossible to do.

Mayer and Narayanan began writing on the subject, describing on a blog how it would work: A header in an HTTP field, the building block of the web, would signal the computer not to collect information, thus enabling users to opt out of tracking of all kinds. They tried to show companies ways they could respond to protect their businesses. It is "a simple technology that is completely compatible with the existing web," they wrote. "We believe regulation is necessary to verify and enforce compliance with a user's choice to opt out of tracking." In a "Do Not Track Cookbook," which they posted online, Mayer and Narayanan proposed limiting identifiers to each website to prevent tracking from one place to another.

A 2010 FTC report recommended implementing a Do Not Track mechanism; several web browsers have adopted its use, but compliance is voluntary and its effectiveness has been limited.

UNLIKE SOME COUNTRIES

that have codified a comprehensive right to privacy, Jennifer Granick notes, the United States has no universal privacy law. Instead, it relies on a patchwork of regulations and the Fourth Amendment, which states: *"The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized."*

But the Fourth Amendment applies only to intrusions from the government. And most federal privacy statutes apply only to specific sectors, such as health care, education or communications and therefore fail to adequately protect personal data on the Internet. The oddest origin of such a statute relates to video rental records and stems from the days of Robert Bork's Supreme Court confirmation hearings.

In 1987, Michael Dolan, then a reporter for the *Washington City Paper,* an alternative weekly in Washington, D.C., walked into a local video store he knew Bork and his wife frequented and requested a list of the couple's video rentals. The subsequent article he wrote, describing Bork based on 146 videos he had presumably watched, did little to define the man, other than revealing a yen for Alfred Hitchcock

COURTESY ARVIND NARAYANAN (TOP); ELLICK CHAN

and Cary Grant. But it caused a stir among the nation's legislators, who were suddenly concerned about their own privacy. Within a year, Congress passed the Video Privacy Protection Act to prohibit "wrongful disclosure of video tape rental or sale records" without a customer's consent. The Act recently returned to the floor of Congress, with an amendment that makes it easier for companies like Netflix to have consumers share their online video viewing as a means of delivering suggestions that fit their tastes.

The law in general is still catching up to the technology. In early February, the California Supreme Court ruled that Apple could legally require some personal information as a means of validating users and preventing fraud. However, the majority opinion suggested that new laws might be necessary to adequately protect consumer privacy.

Narayanan tries to make a clear distinction between privacy research and privacy advocacy. He believes in an individual's choice, and thus transparency and consumer awareness are important. He also is quick to point out that technology advancements can improve privacy options. At the start of the privacy class he teaches each year, he shares an example.

The novel *Fifty Shades of Grey* might have been stigmatized by its graphic sexual content, Narayanan tells his students, but because it first was released as an e-book, people were able to read it on tablets or e-readers without other people knowing. Then, when the book became popular enough that there was no stigma attached, it was published in print.

"The narrative of technology killing privacy is, at best, dramatically overstated," Narayanan says. "For every example of technology hurting privacy, there's one of technology helping privacy." Another example: Self-checkout kiosks used in some large retailers and grocery stores that allow shoppers to make purchases without a store clerk knowing what they've bought.

These examples present an interesting paradox: While reading *Fifty Shades of Grey* on a Kindle feels more private, there is still an electronic record of the purchase. Compare that to buying it at a bookstore, with cash. A clerk might know you like steamy novels but that's where the "record" of your purchase ends. As technology is adopted more widely, old ways are made obsolete or, in some cases, disappear altogether. But that limits our ability to avoid the technology, and the attendant privacy concerns, if we chose to do so.

Solving the privacy conundrum would be easier if the solution didn't also encroach on the ability of companies to prosper, and to deliver new and interesting methods of entertainment, social engagement and commerce that consumers happily embrace. The same technological developments that raise privacy questions also add convenience to many ordinary tasks. They enable instantaneous communication. Social media sites work because of the participation of all of our friends, sharing photos and updates that we enjoy receiving. What's the answer?

Control and transparency were major themes of a 2012 government report titled "A Consumer Privacy Bill of Rights" that aimed to establish "a baseline of clear protections for consumers and greater certainty for companies." The report stated that "Consumers have a right to exercise control over what personal data companies collect from them and how they use it" as well as a right "to easily understandable and accessible information about privacy and security practices."

The report recognized and attempted to account for the benefits of data collection and to find ways of protecting privacy without thwarting innovation. But it warned that if companies don't adopt measures themselves, further regulatory scrutiny is likely. Those warnings are coming true. Last July Congress began an inquiry into data mining practices. In October, a similar probe was launched into nine data brokers.

The Electronic Frontier Foundation expects several pieces of legislation to go before Congress over the next year, including amendments to existing bills that would mandate a warrant for obtaining private electronic communications such as old emails. Minnesota Sen. Al Franken recently introduced The Location Protection Privacy Act of 2012 that would potentially prevent smartphone apps from tracking a cell phone's location and sending it to a third party without consent. Another major player is the Electronic Privacy Information Center, whose president and executive director Marc Rotenberg, JD '87, has testified before Congress on many issues related to consumer privacy.

"I think the next couple of years will be formative for the next decade after," CIS's McDonald says. But forecasts about how business interests and privacy concerns ultimately will be reconciled are cloudy at best. And the proverbial slippery slope is getting more treacherous all the time.

"I would expect that targeting advertising is just the beginning of what could be done with this data," McDonald says. She worries "that we will look back later on and go, 'remember when it was so simple? It was only advertising.'" ■

'For every example of technology hurting privacy, there's one of technology helping privacy.'

BRIAN EULE, '01, *is a frequent contributor to* STANFORD.