Tracking is an assault of liberty

by Nicholas Carr

In a 1963 Supreme Court opinion, Chief Justice Earl Warren observed that "the fantastic advances in the field of electronic communication constitute a great danger to the privacy of the individual." The advances have only accelerated since then, along with the dangers. Today, as companies strive to personalize the services and advertisements they provide over the Internet, the surreptitious collection of personal information is rampant. The very idea of privacy is under threat.

Most of us view personalization and privacy as desirable things, and we understand that enjoying more of one means giving up some of the other. To have goods, services and promotions tailored to our personal circumstances and desires, we need to divulge information about ourselves to corporations, governments or other outsiders.

This tradeoff has always been part of our lives as consumers and citizens. But now, thanks to the Net, we're losing our ability to understand and control those tradeoffs—to choose, consciously and with awareness of the consequences, what information about ourselves we disclose and what we don't. Incredibly detailed data about our lives are being harvested from online databases without our awareness, much less our approval.

Even though the Internet is a very social place, we tend to access it in seclusion. We often assume that we're anonymous as we go about our business online. As a result, we treat the Net not just as a shopping mall and a library but as a personal diary and, sometimes, a confessional. Through the sites we visit and the searches we make, we disclose details not only about our jobs, hobbies, families, politics and health, but also about our secrets, fantasies, even our peccadilloes.

But our sense of anonymity is largely an illusion. Pretty much everything we do online, down to individual keystrokes and clicks, is recorded, stored in cookies and corporate databases, and connected to our identities, either explicitly through our user names, credit-card numbers and the IP addresses assigned to our computers, or implicitly through our searching, surfing and purchasing histories.

A few years ago, the computer consultant Tom Owad published the results of an experiment that provided a chilling lesson in just how easy it is to extract sensitive personal data from the Net. Mr. Owad wrote a simple piece of software that allowed him to download public wish lists that Amazon.com customers post to catalog products that they plan to purchase or would like to receive as gifts. These lists usually include the name of the list's owner and his or her city and state.

Using a couple of standard-issue PCs, Mr. Owad was able to download over 250,000 wish lists over the course of a day. He then searched the data for controversial or politically sensitive books and authors, from Kurt Vonnegut's "Slaughterhouse-Five" to the Koran. He then used Yahoo People Search to identify addresses and phone numbers for many of the list owners.

Mr. Owad ended up with maps of the United States showing the locations of people interested in particular books and ideas, including George Orwell's "1984." He could just as easily have published a map showing the residences of people interested in books about treating depression or adopting a child. "It used to be," Mr. Owad concluded, "you had to get a warrant to monitor a person or a group of people. Today, it is increasingly easy to monitor ideas. And then track them back to people."

What Mr. Owad did by hand can increasingly be performed automatically, with data-mining software that draws from many sites and databases. One of the essential characteristics of the Net is the interconnection of diverse stores of information. The "openness" of databases is what gives the system much of its power and usefulness. But it also makes it easy to discover hidden relationships among far-flung bits of data.

In 2006, a team of scholars from the University of Minnesota described how easy it is for data-mining software to create detailed personal profiles of individuals—even when they post information anonymously. The software is based on a simple principle: People tend to leave lots of little pieces of information about themselves and their opinions in many different places on the Web. By identifying correspondences among the data, sophisticated algorithms can identify individuals with extraordinary precision. And it's not a big leap from there to discovering the people's names. The researchers noted that most Americans can be identified by name and address using only their ZIP Code, birthday and gender—three pieces of information that people often divulge when they register at a website.

The more deeply the Net is woven into our work lives and leisure activities, the more exposed we become. Over the last few years, as social-networking services have grown in popularity, people have come to entrust ever more intimate details about their lives to sites like Facebook and Twitter. The incorporation of GPS transmitters into cellphones and the rise of location-tracking services like Foursquare provide powerful tools for assembling moment-by-moment records of people's movements. As reading shifts from printed pages onto networked devices like the Kindle and the Nook, it becomes possible for companies to more closely monitor people's reading habits—even when they're not surfing the Web.

"You have zero privacy," Scott McNealy remarked back in 1999, when he was chief executive of Sun Microsystems. "Get over it." Other Silicon Valley CEOs have expressed similar sentiments in just the last few months. While Internet companies may be complacent about the erosion of personal privacy—they, after all, profit from the trend—the rest of us should be wary. There are real dangers.

First and most obvious is the possibility that our personal data will fall into the wrong hands. Powerful data-mining tools are available not only to legitimate corporations and researchers, but also to crooks, con men and creeps. As more data about us is collected and shared online, the threats from unsanctioned interceptions of the data grow. Criminal syndicates can use purloined information about our identities to commit financial fraud, and stalkers can use locational data to track our whereabouts.

The first line of defense is, of course, common sense. We need to take personal responsibility for the information we share whenever we log on. But no amount of caution will protect us from the dispersal of information collected without our knowledge. If we're not aware of what data about us are available online, and how they're being used and exchanged, it can be difficult to guard against abuses.

A second danger is the possibility that personal information may be used to influence our behavior and even our thoughts in ways that are invisible to us. Personalization's evil twin is manipulation. As mathematicians and marketers refine data-mining algorithms, they gain more precise ways to predict people's behavior as well as how they'll react when they're presented with online ads and other digital stimuli. Just this past week, Google CEO Eric Schmidt acknowledged that by tracking a person's messages and movements, an algorithm can accurately predict where that person will go next.

As marketing pitches and product offerings become more tightly tied to our past patterns of behavior, they become more powerful as triggers of future behavior. Already, advertisers are able to infer extremely personal details about people by monitoring their Web-browsing habits. They can then use that knowledge to create ad campaigns customized to particular individuals. A man who visits a site about obesity, for instance, may soon see a lot of promotional messages related to weight-loss treatments. A woman who does research about anxiety may be bombarded with pharmaceutical ads. The line between personalization and manipulation is a fuzzy one, but one thing is certain: We can never know if the line has been crossed if we're unaware of what companies know about us.

Safeguarding privacy online isn't particularly hard. It requires that software makers and site operators assume that people want to keep their information private. Privacy settings should be on by default and easy to modify. And when companies track our behavior or use personal details to tailor messages, they should provide an easy way for us to see what they're doing.

The greatest danger posed by the continuing erosion of personal privacy is that it may lead us as a society to devalue the concept of privacy, to see it as outdated and unimportant. We may begin to see privacy merely as a barrier to efficient shopping and socializing. That would be a tragedy. As the computer security expert Bruce Schneier has observed, privacy is not just a screen we hide behind when we do something naughty or embarrassing; privacy is "intrinsic to the concept of liberty." When we feel that we're always being watched, we begin to lose our sense of self-reliance and free will and, along with it, our individuality. "We become children," writes Mr. Schneier, "fettered under watchful eyes."

Privacy is not only essential to life and liberty; it's essential to the pursuit of happiness, in the broadest and deepest sense. We human beings are not just social creatures; we're also private creatures. What we don't share is as important as what we do share. The way that we choose to define the boundary between our public self and our private self will vary greatly from person to person, which is exactly why it's so important to be ever vigilant in defending everyone's right to set that boundary as he or she sees fit.