

Latanya Sweeney attracts a lot of attention. It could be because of her deep affection for esoteric and cunning mathematics. Or maybe it is the black leather outfit she wears while riding her Honda VTX 1300 motorcycle around the sedate campus of Carnegie Mellon University, where she directs the Laboratory for International Data Privacy. Whatever the case, Sweeney suspects the attention helps to explain her fascination with protecting people's privacy. Because at the heart of her work lies a nagging question: Is it possible to maintain privacy, freedom and safety in today's security-centric, databased world where identities sit ripe for the plucking?

Several years ago Scott McNealy, chairman of Sun Microsystems, famously quipped, "Privacy is dead. Get over it." Sweeney couldn't disagree more. "Privacy is definitely not dead," she counters; those who believe it is "haven't actually thought the problem through, or they aren't willing to accept the solution."

Certainly privacy is under siege, and that, she says, is bad. Debates rage over the Patriot Act and data mining at the federal level, and states have a hodgepodge of reactive laws that swing between ensuring privacy and increasing security. Although identity theft began a slow decline in 2002, one recent study revealed that 8.4 million U.S. adults still suffered some form of identity fraud in 2006. "The problem grows as technologies explode," Sweeney says, and every problem requires a different solution, which is another way of saying that it is impossible to predict where new forms of privacy invasion will arise.

All this has kept Sweeney and her team busy the past six years wrestling some of today's thorniest confidentiality issues to the mat—identity theft, medical privacy and the rapid expansion of camera surveillance among them. Other academic labs tend to attack issues at a theoretical level; the 47-year-old Sweeney states that her group operates as a kind of digital detective agency staffed with a dedicated squad of programmers devising some seriously clever software. The researchers' approach is to technically fillet systems and then suggest ingenious but pragmatic solutions.

For example, Sweeney's Identity Angel program scours the Internet and quickly gathers thousands of identities by linking names in one database with addresses, ages and Social Security numbers scattered throughout others. Those four pieces of information are all anyone needs to snatch an identity and open a credit-card account. The lab routinely alerts vulnerable people so they can fix the problem.

Another program "anonymizes" identities. It was originally developed for the Department of Defense after the 9/11 attacks to help locate potential terrorists while still protecting the privacy of innocent citizens. The program prevents surveillance cameras from revealing an identity until authorities show they need the images to prosecute a crime. Unlike other software, the program does not pixelate or black out an individual's features but actually fabricates a new facial image from other faces in the database, making it impossible for humans or machines to identify.

The clever algorithms at the heart of Sweeney's lab go back to her days growing up in Nashville, when she would daydream about ways to create an artificially intelligent black box that she could talk to. "I spent hours fantasizing about that box," she recalls. Ten years later she parlayed her talent for mathematics and early fascination with artificial intelligence into scholarships that helped to pay her way to the Massachusetts Institute of Technology, a bastion in both fields. It would have seemed the perfect place to pursue her grade school dream of creating a smart machine. The problem was that Sweeney had just departed the polite world of a prim New England all-girls high school and was suddenly immersed in M.I.T.'s male-dominated geek culture, a transition that took her off guard. That, coupled with her experiences with a racially insensitive professor whom she never seemed able to please, led Sweeney to drop out and start up her own software consulting business.

After a decade in the business world, Sweeney returned to college, completing her undergraduate degree at Harvard University. She then earned her master's and doctorate in computer science at M.I.T., the first African-American woman to do so. "When I came back, I told them I didn't plan on taking any more crap," she laughs.

It was on her return to M.I.T. that Sweeney first fell into the orbits of privacy and security. She had won a fellowship at the National Library of Medicine, and to show her appreciation, she volunteered to help several Boston hospitals improve protection of their medical records, a concern that was surfacing as the Internet ballooned in the mid-1990s. Sweeney wrote a program called Scrub System that tapped her expertise in artificial intelligence to ingeniously search patient records, treatment notes and letters between physicians. Standard search-and-replace software had generally found 30 to 60 percent of personal, identifying information. Scrub System "understands" what constitutes a name, address or phone number and eliminates 99 to 100 percent of the revealing data.

The software won accolades from medical associations. "Her research was highly influential," says Betsy Humphreys, deputy director of the National Library of Medicine. "A lot of people didn't see how different life was [in the Internet age].... Latanya's work raised their awareness." With Scrub System, "I thought I had solved the privacy problem," Sweeney says sheepishly. But the truth was, "I really didn't understand a thing about privacy."

This realization hit home when one day she reviewed the medical history of a young woman. "At age two this girl was sexually molested, at age three she stabbed her sister with scissors, at four her parents got divorced, at five she set fire to her home," Sweeney recounts. Clearly, "removing the explicit identifiers wasn't what [privacy] was about." It was about the bread crumbs of information we leave behind in records strewn all over the Internet—in medical forms, credit applications, resumes and other

documents. Nothing specifically identified the girl in the report, but the scraps of information were unique, and Sweeney was pretty sure she could use them to reidentify her—and almost anyone else.

Programs such as Identity Angel have proved Sweeney correct, and she has spent plenty of time finding ways to invade privacy, sometimes getting the jump on the bad guys, sometimes not. She tells the story of a banker indicted in Maryland who cross-referenced information in publicly available hospital discharge records with his own client list to see if any of his clients had cancer. If they did, he called in their loans. In a project using data from the state of Illinois, Sweeney's lab found a way to reidentify patients with Huntington's disease even after all information about the patients had been deleted from their records. Huntington's is caused by the repetition of a short sequence of DNA. The more this sequence repeats, the earlier the age of onset. Sweeney's lab combined those data with hospital discharge records, which included patients' ages, to accurately link 90 percent of the Huntington's patients with DNA records on file. Abuses may be rare, Sweeney admits, but both cases show how ugly things can get if one database is used to leverage the information in another.

The real solution, Sweeney says, does not lie in her lab or in any other. Ultimately engineers and computer scientists will have to weave privacy protection into the design and usability of their new technologies, up front. If they do, "society can [then] decide how to turn those controls on and off," Sweeney remarks. Otherwise we might all need to ride a motorcycle to get a few private moments.

[Rights & Permissions](#)

CHIP WALTER is author of *Last Ape Standing: The Seven-Million-Year Story of How and Why We Survived* (Walker

[More by Chip Walter](#)



This article was originally published with the title "A Little Privacy, Please" in Scientific American Magazine Vol. 297 No. 1 (July 2007)

doi:10.1038/scientificamerican072007-44a7gf4cvFvJ5FzBb9rZhN

[View This Issue](#)

Expand Your World with Science

Learn and share the most exciting discoveries, innovations and ideas shaping our world today.

[Subscribe](#) [Sign up for our newsletters](#) [See the latest stories](#) [Read the latest issue](#) [Give a Gift Subscription](#)

Follow Us:

- [Return & Refund Policy](#)
- [About](#)
- [Press Room](#)
- [FAQs](#)

- [Contact Us](#)
- [International Editions](#)

- [Advertise](#)
- [Accessibility Statement](#)
- [Terms of Use](#)

- [Privacy Policy](#)
- [California Consumer Privacy Statement](#)
- [Use of cookies/Do not sell my data](#)

Scientific American is part of Springer Nature, which owns or has commercial relations with thousands of scientific publications (many of them can be found at www.springernature.com/us). Scientific American maintains a strict policy of editorial independence in reporting developments in science to our readers.

© 2024 SCIENTIFIC AMERICAN, A DIVISION OF SPRINGER NATURE AMERICA, INC.
ALL RIGHTS RESERVED.

