SENSOR NATION

> > > SOCIETY

It's alarming! It's no big deal! How your personal information is being collected and protected, used and misused

BY JEAN KUMAGAI & STEVEN CHERRY

SENSORS & SENSIBILITY

WHEN CANADIAN TOURIST BYUNGSOO SON picked up a rental car from a Payless office in San Francisco last November and set off with his wife and son on a 12-day tour of the California coast, Las Vegas, and the Grand Canyon, he had no idea how pricey that trip would be. Upon dropping off the car, he was floored when the expected US $260 charge turned out to be a whopping $3400, the result of a $1-a-mile fee that kicked in when Son crossed the California-Nevada border. Accompanying the bill was a detailed map of the family's route, made possible by the Global Positioning System tracking device installed in the car. Son had never bothered to read all of the fine print in his rental contract—who does, really?—which mentioned the out-of-state penalty and the possible presence of a tracking device.

Get used to it. One-fourth of rental cars in the United States now have GPS tracking installed, and over the last several years, at least two other companies have used the devices to fine errant drivers. If the car were stolen, or it broke down in a desert or a snowstorm, the trackers could be a lifesaver, the rental companies say. Some renters, if asked, might even appreciate a map of their trip as a souvenir. But having your every move tracked like a fugitive's? Most drivers, surely, would object.

Here's the problem with information: it spreads. "Once information exists, it's virtually impossible to limit its use," says David L. Sobel, general counsel of the Washington, D.C.–based Electronic Privacy and Information Center. "You have all this great data lying around, and sooner or later, somebody will say, 'What else can I do with it?' "

Over the last several years, new tracking and monitoring technologies coupled with new data-mining initiatives and a more permissive attitude toward surveillance have made it possible to deploy many creative, and intrusive, uses of our personal information. Life is undoubtedly made more convenient by key-chain tags that let you pay for gas right at the pump, wireless payment systems that let you drive through tolls without stopping, and fingerprint authentication systems that ensure you are who you say you are. Where there have been problems, they've tended to be more annoying than horribly invasive—an erroneous charge to your credit card or an unsolicited pitch for a new whitening toothpaste based on your past purchases.

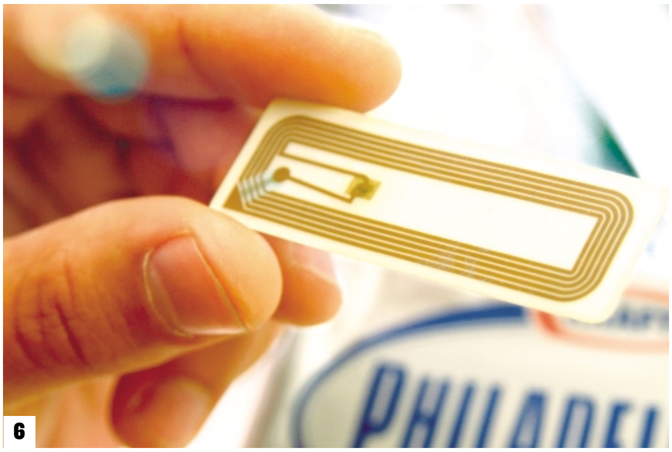
But as new technologies and uses of data are being added seemingly every day, the potential for greater abuse is growing, say Sobel and other privacy experts. Meanwhile, legal protections are lagging far behind. What Son's rental car company did might have been a little sleazy, but it was perfectly legal.

Already, you're giving away more information than you probably realize. At the office, wireless security cards track your comings and goings; your employer could be keeping tabs on your e-mail, phone calls, and maybe even your keystrokes. When you surf the Web, government agencies and businesses can see which sites you visit, if they care to look. Emergency initiatives like Enhanced 911 in the United States and Enhanced 112 in Europe can pinpoint your location through your cellphone. Use a credit card or a loyal-shopper card, and your every purchase is logged. If you're visiting the United States from abroad, you now surrender your digitized fingerprints and photo at the

JEFF GRUNEWALD

1

2



6

5

border. And nearly everywhere you go—from the bus stop to the parking lot to the ATM to the fitting room—surveillance cameras are watching you.

Among the biggest collectors and purveyors of your personal information are data aggregators like Acxiom Corp., in Little Rock, Ark., and ChoicePoint, in Alpharetta, Ga. It's their business to buy up information about ordinary citizens, correlate it with the billions of other records in their data warehouses, and then sell the information—to employers doing background checks, insurers and landlords doing credit checks, and, especially since 9/11, government agencies doing security checks.

For $20, you can see what others see: your ChoicePoint report listing your phone numbers, the current market value of the real estate you own, your car loans, any outstanding liens and judgments, and any pilot, maritime, radio, drug, and gun licenses you hold—plus the names, birth dates, and social security numbers of not just you, but your spouse, children, and parents, plus any friends with whom you've jointly filed legal documents.

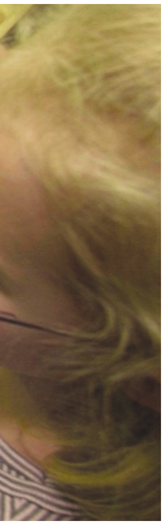
So what more is there to know? Plenty. While more traditional sources of information paint a picture of you in coarse strokes, newer and soon-to-emerge data-gathering technologies offer a much finer-grained image—where you are and what you're doing at any given time. These technologies include cheap and ubiquitous radio-frequency ID (RFID) tags, distributed and virtually invisible sensor networks, biometric scanners, and "smart" video surveillance. In the name of law enforcement, security, cost-saving, and convenience, commercial and government networks are digging ever deeper, gathering, sifting, and—increasingly—sharing data, uncovering what they hope will be pure, precise nuggets of information [see photo display, "Ways of Watching," for some current surveillance technologies]. Coupled with advances in networking, wireless communication, computation, and data mining, the result is that what we used to think of as deeply personal affairs are increasingly matters for public consumption.

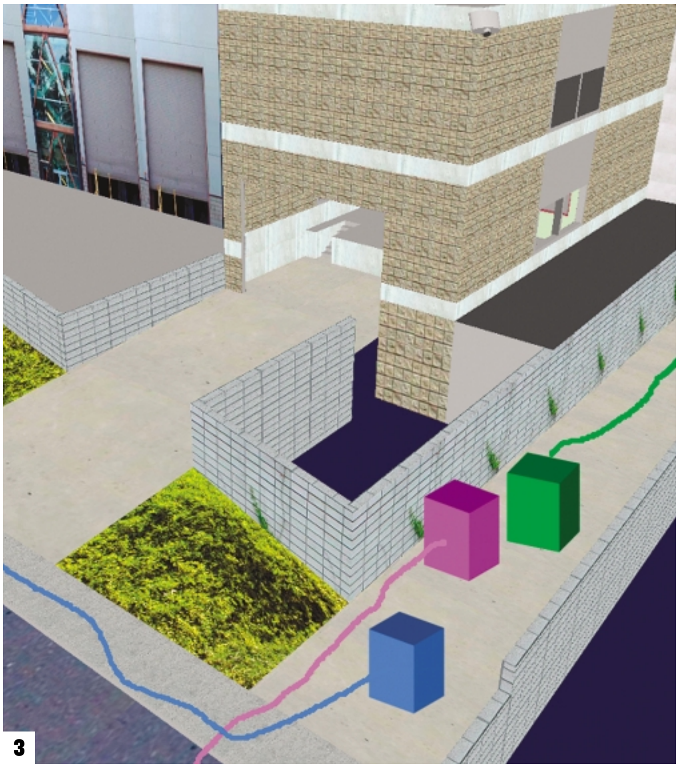
Against that crushing tide of data, a few researchers are conjuring up countermeasures. New database filters, for example, will let users search through sensitive information without uncovering personal data, while location-blocking algorithms and surveillance camera filters can obscure your exact position.

Such efforts are sorely needed. A new U.S. Department of Defense–sponsored report on the privacy implications of the government's data-mining activities warns of data-mining tools being "used by the government to scrutinize personally identifiable data concerning U.S. persons who have done nothing to warrant suspicion." It cautions that "they run the risk of becoming the 21st-century equivalent of general searches, which the authors of the Bill of Rights were so concerned to protect against."

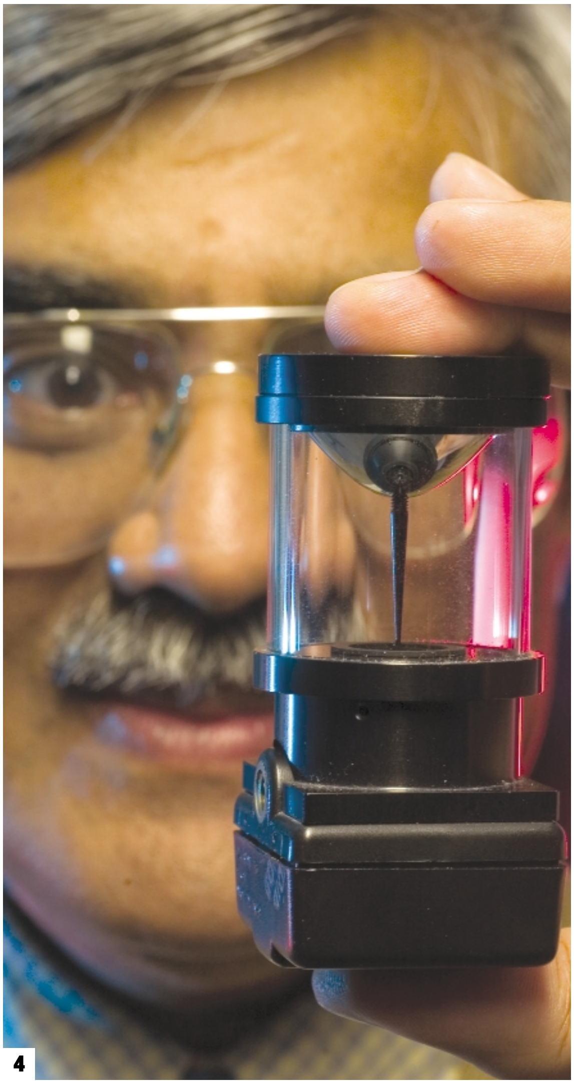
Absent legal and technological protections, the report concludes, current surveillance efforts threaten to chill the behavior of ordinary citizens, stifling not just "innocuous, everyday

TOP ROW: AL DIAZ/MIAMI HERALD/NEWSCOM; ERIK S. LESSER/GETTY; MOHAN TRIVEDI; JOEL ZWINK
BOTTOM ROW: WILLIAM THOMAS CAIN/GETTY; FUTURE STORE





3



4

WAYS OF WATCHING: Recent advances in sensors and computing let governments, companies, and other institutions track people as never before. [1] One in four rental cars in the United States has a Global Positioning System tracking device, which allows companies to track the vehicle's movements in real time. [2] Foreign visitors to the United States now have their fingerprints scanned and photos taken at the border. [3, 4] Computer vision expert Mohan Trivedi's surveillance-camera system automatically blocks people's identities, rendering them instead as colored cubes. Should a camera detect suspicious behavior—a person running amid a crowd of walkers, say—the system automatically switches and reveals the person's true image. [5] To make sure students are lunching in the cafeteria, some schools in Pennsylvania have installed fingerprint scanners. [6] Cheap, disposable radio-frequency ID tags affixed to goods allow them—and you—to be tracked wirelessly.

activities" but also religious expression, political dissent, and public discourse.

THE TENSION BETWEEN technology and privacy isn't new, of course. One hundred and fourteen years ago, Samuel D. Warren and Louis D. Brandeis argued in the *Harvard Law Review* that "numerous mechanical devices threaten to make good the prediction that 'what is whispered in the closet shall be proclaimed from the housetops.'" Among the instances they cited was a "somewhat notorious case" involving an actress who'd been photographed surreptitiously while she performed in tights.

The Warren-Brandeis article, "The Right to Privacy," acknowledged that who we are largely consists of what is known about us, and it urged that our personal information therefore deserves some measure of privacy protection. Warren and Brandeis were especially concerned about an out-of-control press; were they alive today, they would undoubtedly be equally worried about out-of-control databases. From what we do, to where we go, to what we look like, to whom we socialize with, to what drugs we take and what magazines we read, each scrap of data may be insignificant—but taken together, they reveal a great deal.

Some of the most personal new information about you comes from loyal-shopper cards. While these programs purport to save you money, the store gets much more in return: when merchandise bar codes are scanned at the checkout, the purchase data gets correlated with the personal information connected to your loyalty card.

Katherine Albrecht, a Nashua, N.H.–based former marketing executive who now heads Consumers Against Supermarket Privacy Invasion and Numbering (CASPIAN), notes stores retain that sales data for years, looking for patterns in your purchases and making inferences about you, and they sell it to companies like Acxiom and ChoicePoint as well. "Cash registers are no longer adding machines with cash drawers," she says. "They're high-speed, data-collecting computers with connections to the Internet. And shopper loyalty cards tie that data to your identity. The whole goal is to figure out everything you can learn about your customer. We're creating a retail zoo, where customers are the exhibits."

For that reason, CASPIAN and other privacy groups have taken a hard line on the deployment of RFIDs in stores, libraries, and currency. Technologically speaking, an RFID tag is pretty simple: it's a small microchip coupled to a tiny radio antenna. The tags come in various shapes and sizes, but the smallest—Hitachi Ltd.'s μ-chip—is a speck about the size of a grain of salt, just 0.4 millimeter on a side. The cheapest tags now cost about 5 cents apiece, and some manufacturers predict they can bring that down to about 1 cent within five years.

The typical RFID tag can store no more than 128 bits, much of that memory taken up by the Electronic Product Code, a numeric designation that identifies the manufacturer, product, and serial number. Each tag is unique—that's right, eventually, every pair of

Dockers khaki pants, every can of Colgate Shave Cream, every box of Trojan Ultra Ribbed condoms will have its very own ID number. With bar codes, by contrast, all boxes of Ultra Ribbed condoms share a single product code.

Most RFID tags don't have a power source. Instead, when the tag comes within range of an RFID reader's electromagnetic field, it wakes up and uses the reflected energy to communicate with the reader. RFID readers, unlike their bar-code counterparts, can scan multiple tags simultaneously and at a distance—from a few centimeters to 10 or more meters.

In a pioneering foray into RFID technology, Wal-Mart Inc., headquartered in Bentonville, Ark., announced last year that it would make its top 100 suppliers use RFID tags on all pallets and cartons of goods. Stores in the Dallas–Fort Worth area started using the technology to track warehouse inventory this past spring. The retailing giant figures RFIDs can save it over $8 billion a year, out of total sales of $244 billion, mainly by reducing labor costs, theft, and errors.

When a shipment arrives at a Wal-Mart store, readers installed at the warehouse loading docks automatically scan each pallet and case of goods, transmitting the RFID data to an inventory control computer, which matches the serial number with, say, the type and number of cans of shaving cream on the pallet.

During the initial phases, Wal-Mart is using its own internal computer network to keep track of serial numbers. Eventually, though, it will probably switch over to a global Web-based network shared by other RFID-savvy retailers, suppliers, and manufacturers. In that scheme, the inventory computer would identify the serial number by sending a query over the Internet to something called an Object Name Service. These databases, operated by VeriSign Inc., Mountain View, Calif., the same company that keeps track of Internet domain names, act like a reverse telephone directory: upon receiving a serial number, they produce an address, namely the Internet Protocol (IP) address of a server where detailed information about the tagged item is located.

Privacy experts generally don't object to using RFID tags to streamline warehouse operations. And Wal-Mart has said it plans to use RFIDs only in its warehouses, although certain items, like TV sets and computer printers, will still have their tags when they hit the retail floor. But most industry observers are convinced that the tags will eventually replace bar codes on individual goods—retailers like Wal-Mart won't be able to wring out every last efficiency until they do. And once the tags are on individual items, stores will inevitably link you to what you buy, creating databases of everything you've purchased from them. The information will be simply too valuable to toss out: software will scan the databases, looking for patterns in your purchases; make inferences about you and other shoppers; and possibly send other merchants or service providers your way.

IN A NOD TO PRIVACY, each RFID tag contains the seeds of its own destruction: a 24-bit "destroy" code that, if triggered by a reader, will render the tag unreadable. But disabling the tags would preclude many of the useful applications that manufacturers are developing: smart washing machines that read tags in clothing and automatically adjust their cycles or networked medicine cabinets that know when your prescriptions need refilling.

In the rush to make our lives more convenient, though, we shouldn't ignore the possible unintended consequences, argues Albrecht. Without any regulation, for example, law enforcement could use RFIDs to monitor people's behavior. Police now routinely videotape public protests; in the future, they'll be able to walk around with RFID readers and collect the serial numbers from people's clothing and other tagged items they're carrying. Matching those serial numbers with retailers' records would yield a list of protesters' names, addresses, and so on. Or police could just look for the serial numbers themselves, at an airport security checkpoint, say. "That tube of strawberry Chapstick was at the World Bank protest! Pull that passenger aside!" Though that level of surveillance may be way down the road, says Albrecht, its implications are unsettling.

NOR WILL RFID TAGS BE THE ONLY WAY to surreptitiously identify you. Soon there'll be another: through Internet Protocol addresses. Right now, those numbers mainly identify intelligent devices like computers and PDAs, and the device may not use the same Internet address today as the one it used yesterday.

But Internet engineers are now rolling out a newer version of addressing called IPv6. This scheme uses addresses that are 128 bits long, instead of the current 32. Through the miracle of binary arithmetic, that yields 3 x 10^{38} addresses—enough to assign each sensor, widget, and appliance on the planet its very own permanent IP address, thus creating what IPv6's proponents have termed an "Internet of things." With every streetlight, parking meter, and video camera potentially broadcasting information about itself and everything it interacts with, you'll know much more about everything around you.

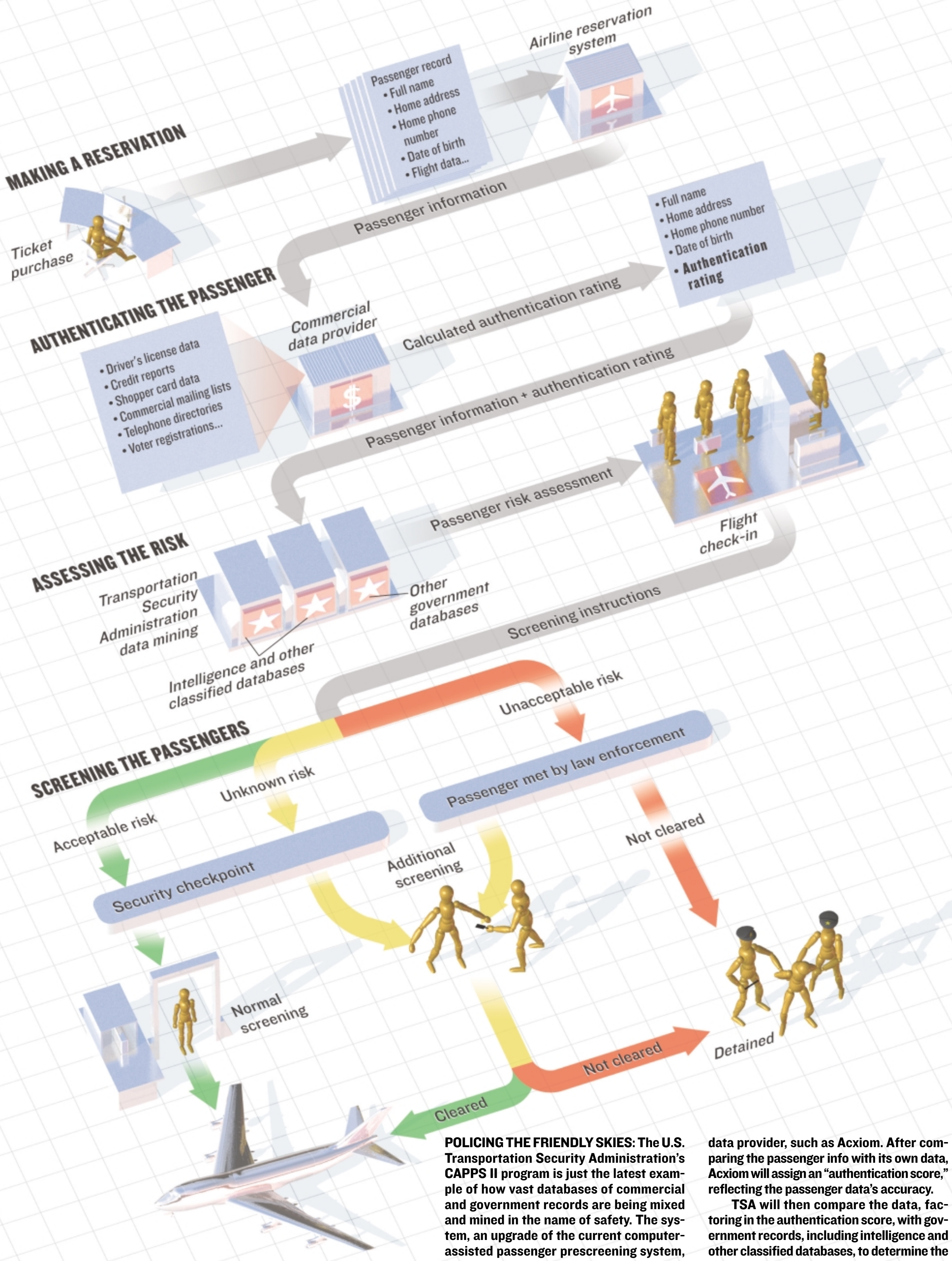
Of course, your environment will know a lot more about you as well. Indeed, every time your car or cellphone connects to the Internet, you'll reveal what you're doing and where you are. A Borders bookstore might send you a text message with a discount coupon as you pass by. Less benignly, your boss at work or your spouse at home will be able to watch in real time as you run errands around town, just as Payless tracked Byungsoo Son across the Nevada desert. And it's not too hard to imagine your IPv6 addresses winding up in your ChoicePoint profile, right alongside your phone numbers.

Though ChoicePoint mainly sells its data to other commercial entities, since 9/11 it has found an eager client in the U.S. government. As the recent Defense Department report makes clear, a wide variety of U.S. agencies would like to apply the same customer profiling and data-mining techniques perfected by companies like Wal-Mart and Amazon.com to pursue terrorists and other criminals.

The most notorious program was former Admiral John Poindexter's Total Information Awareness, officially cancelled in 2003. But many other data-mining projects are ongoing, the report noted, and all pose significant privacy risks. Among the projects cited were the Treasury Department's Financial Crimes Enforcement Network, aimed at catching money laundering; the MATRIX (Multistate Anti-Terrorism Information Exchange) system being used by several states and the Department of Homeland Security to link law enforcement records with other government and private-sector databases; and the U.S. Transportation Security Administration's revamped and expanded Computer-Assisted Passenger Prescreening System.

Also known as CAPPS II, the new passenger screening system is to replace an existing one that uses secretive but ineffectual "no-fly" lists: a test at a U.S. airport this past January revealed that a person named "Osama bin Laden" could scamper right onto his flight, no questions asked. CAPPS II is designed to categorize prospective passengers into three groups: those deemed "acceptable to fly," those who present an "unknown" risk, and those who are "unacceptable to fly." [See illustration, "Policing the Friendly Skies."]

These lists will emerge as follows: several days before a flight, the reservation records for every passenger are sent to Acxiom or some other commercial data aggregator. The data, including name, address, birth date, and phone number, are checked against Acxiom's records. Depending on the number of discrepancies, Acxiom assigns each passenger an authentication score. The TSA then checks the reservation data against U.S. government data-



POLICING THE FRIENDLY SKIES: The U.S. Transportation Security Administration's CAPPS II program is just the latest example of how vast databases of commercial and government records are being mixed and mined in the name of safety. The system, an upgrade of the current computer-assisted passenger prescreening system, will vet preflight passengers.

To do that, it will extract information from an airline's reservation system, including name, home address, phone number, and date of birth. A few days before the flight, that information will get shipped to a commercial data provider, such as Acxiom. After comparing the passenger info with its own data, Acxiom will assign an "authentication score," reflecting the passenger data's accuracy.

TSA will then compare the data, factoring in the authentication score, with government records, including intelligence and other classified databases, to determine the passenger's risk status. Passengers deemed "acceptable to fly" will undergo just the standard X-ray screening at the airport; those considered "unknown" will get additional screening; and "unacceptable" passengers will, of course, not be allowed to fly.

BRYAN CHRISTIE

bases, factoring in the authentication score, to determine the passenger's risk status.

Although the TSA had planned to launch CAPPS II later this year, the program is far behind schedule, in part because protests over privacy violations have kept developers from getting realistic databases with which to test their software. In the meantime, the agency will test a voluntary screening system, known as Registered Traveler, this summer, though when this issue went to press, little was known about how it would work. When contacted by *IEEE Spectrum*, the TSA refused to discuss which databases it would mine, what mechanisms would be used to correct erroneous information, or even the names of the contractors researching and testing the system.

Whether voluntary or not, such systems bother privacy activists. "A system does all this data mining of disparate information and then spits out a name," says Sobel, of the Electronic Privacy and Information Center. "Does this person then bear a secret government-imposed tag, 'Possible Terrorist'? Does he have an opportunity to know about it and challenge it?"

That's not an idle concern. Prior to the much-contested 2000 presidential election, the state of Florida used a list of names purchased from a company called DBT Online (since acquired by ChoicePoint) to "cleanse" convicted felons from its voter registry. The list was so spotty that thousands of legitimate voters were dropped from the rolls; some were guilty only of misdemeanors, like public drunkenness, while others were simply victims of mistaken identity, including one county's own election supervisor.

A recent General Accounting Office report on CAPPS II worried about similar problems and noted that TSA currently doesn't require commercial data providers to fix errors. Passengers may not even be allowed to know who the data providers are. And, of course, classified government databases will be off-limits.

EVEN AS THE NEW SENSOR-LADEN WORLD strives to make our lives more transparent, a small cadre of researchers is striking back, creating technologies that enhance and protect privacy.

Some of this work strives to construct databases that don't compromise people's identities, a huge concern for antiterrorism investigations, such as the CAPPS II program. Teresa Lunt, a researcher at California's Palo Alto Research Center, is designing a "privacy appliance." Attached to a database, it filters data flowing in and out, acting like the network firewalls that block computer viruses and hacker intrusions.

Suppose, for example, the answer to a database query includes an auto repair record for a white Volvo station wagon that's registered in ZIP code 10001. Lunt's filter would first check other public databases, such as motor vehicle registrations, to see how many similar Volvos there are. If there were only one—thus revealing the owner's identity—the query response would be blocked until a court order or some other authorization was produced. The appliance would also create log files and audit trails, so that if anyone tried to interfere with the appliance, the intruder could be traced.

Ironically, Lunt's research had been funded by the Total Information Awareness project. When an irate U.S. Congress yanked the bulk of TIA's funding, Lunt's program was canned as well.

Latanya Sweeney, an assistant professor of computer science at Carnegie Mellon University in Pittsburgh, takes a somewhat different tack. Her privacy-enhancing software alters the results of database queries so they don't identify individuals. For example, it might reveal only the first three digits of a person's ZIP code, or give only a birth year instead of the exact date. Like Lunt's, Sweeney's federal funding for this year was tossed out with the TIA bathwater. But she continues to refine her software and has helped form a venture, DatAnon LLC, in Pittsburgh, to commercialize it.

Other researchers are also trying to keep overly personal information from being disseminated. With video surveillance cameras, only criminal activity is of interest, and yet the cameras pick up all kinds of activity. So Mohan Trivedi of the University of California, San Diego, has developed a surveillance system that blocks out images of people and other objects; dedicated processors on the cameras represent them instead as colored cubes. If a camera detects suspicious activity—a person running down the street when everyone else is walking, or two cars crashing into each other on a highway—it will switch and reveal the true image. The cameras are also arranged in an array, so that the system knows what it's tracking as objects or people move from one camera to the next—a skill not shared by the vast majority of the world's 31 million surveillance cameras.

Pinpointing your location is another thing sensors are good at. So-called smart building sensor networks, for example, can monitor people's locations in offices, even though that's not their intended function. Marco Gruteser, a doctoral student in computer science at the University of Colorado, Boulder, has developed anonymizing algorithms that dynamically sense how many people are present in a given region—what he calls user density—and then adjust the precision of the data so that it's still useful but not able to identify individuals.

It may be a losing battle, though. About a year and a half ago, Marc Langheinrich, a researcher at the Institute for Pervasive Computing at ETH Zurich, visited a handful of European labs involved in sensor-based computing, to ask the designers about their systems' privacy implications. "Most said either 'It's not my business, it's the lawmakers' ' or 'It's not my business, because it's not my field.' Others said that if they thought about privacy, it would get in the way of building their designs," he recalls. The result, he says, is that privacy protection becomes an afterthought.

Or maybe the designers' attitudes just reflect the changing world they live in. It's obvious that over time our expectations about privacy shift, sometimes by a little, sometimes by a lot. Unlike the shy actress cited by Warren and Brandeis, today's starlet wouldn't think twice about baring her legs for the cameras. In a world of ever more probing sensors and databases, we're growing accustomed to being watched. And as the technology grows cheaper and more powerful, it may give all of us, not just the powerful, the means for watching, too. Whether we ultimately end up with a privacy-free society, as envisioned in the next article, "We Like to Watch," remains to be seen.

What's already clear is that such a transition will not be smooth. "This is the danger of having too big a carrot in front of us," says Langheinrich. "These sensors are promising that we'll be super secure, super efficient, and have a super life. But nothing is foolproof. At the end of the day, we'll still have to defend against all this spotty data and all the potentials for abuse." ■

TO PROBE FURTHER

Jeffrey Rosen's *The Naked Crowd: Reclaiming Security and Freedom in an Anxious Age* (Random House, New York, 2004) provides a valuable backgrounder on privacy issues.

The report by the Technology and Privacy Advisory Committee of the U.S. Department of Defense, "Safeguarding Privacy in the Fight Against Terrorism," Washington, D.C., 2004, is available at http://www.cdt.org/headlines/20040517a.shtm.

See also the U.S. General Accounting Office's "Computer-Assisted Passenger Prescreening System Faces Significant Implementation Challenges," Washington, D.C., February 2004 (http://www.gao.gov/cgi-bin/getrpt?GAO-04-385).