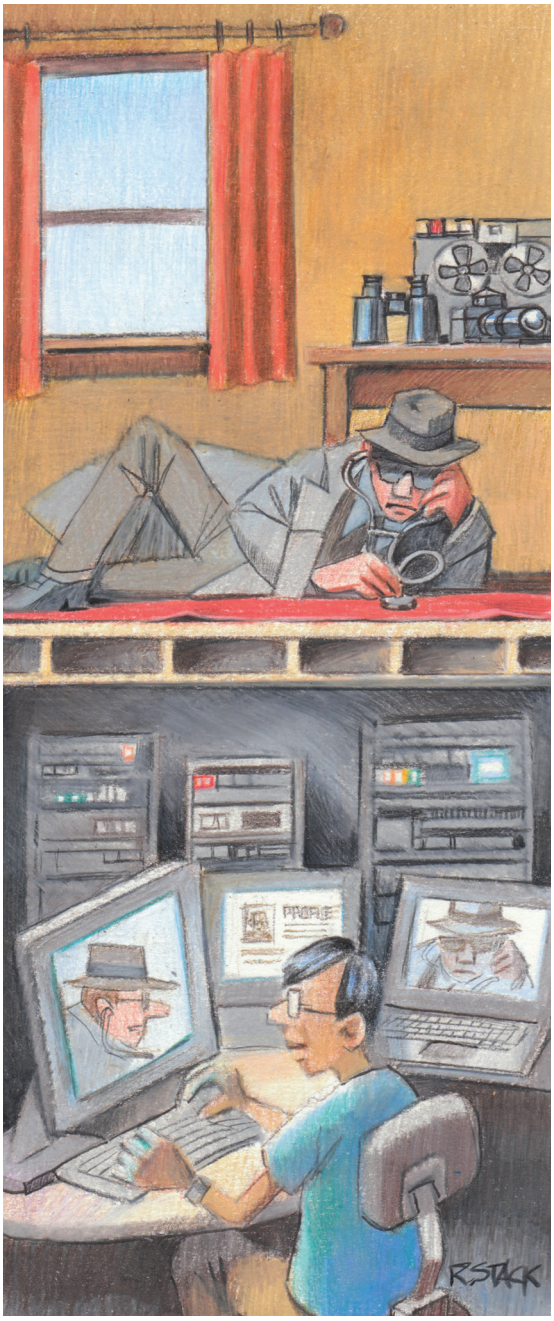
Editors: Angela Sasse, A.Sasse@cs.ucl.ac.uk | Alessandro Acquisti, acquisti@andrew.cmu.edu

The Eyes Have It:
Surveillance and How It Evolved

Shari Lawrence Pfleeger | Dartmouth College



Lately, many news media outlets have been laced with articles about surveillance. Amplified by the continuing revelations from Edward Snowden and others, we're becoming more concerned about who's watching us and why, and what happens to the information being collected. But is this a new phenomenon, or is it simply an extension of what has been going on for centuries? In this article, I explore the answers to these questions, as part of this special issue's goal of informing our readers about the back story of surveillance, so that we can design, build, and use technology responsibly and in context.

In the Beginning

Almost from the beginning of recorded history, data have been collected about us. For instance, in Numbers and Deuteronomy, the Bible describes how representatives of several tribes were sent out to examine the land of Canaan, to determine how well it would support habitation. They were asked to collect information about its current state: air quality, soil quality, and tree cover, and whether the land was fruitful. But they were also asked to describe the people—their number, stature (short or tall), and ability to work hard—and the way they lived (in tents or houses, within walls or without). This evaluation of people and their surroundings was intended to determine whether Canaan was indeed a "land of milk and honey."

India claims to have the oldest formal census, devised in the third century BCE. Its principles of government, called the Arthashastra, mandated collection of population, economic, and agricultural statistics as the basis for taxation. During India's early Moghul period (1500s), the government also kept careful records of land, production, population, famines, and more. This data collection was further extended during the rule of Akbar the Great (mid-1500s) to include industry, wealth, and other information (http://censusindia.gov.in/Data_Products/Library/Indian_perceptive_link/History_link/censushistory.htm).

A census taken in Roman times (circa 7 CE) determined which men were fit for military service. Later generations continued to see value in data collection about people and property. These types of surveillance, which were much like today's census-taking, provided data that were seen as a public good. They informed decision-making about productive crop-planting, land distribution, and the best sites for defensive structures and soldiers.

In the Middle

Formal record-keeping about us, our property, and our habits continued through the Middle Ages. Shortly after William the Conqueror invaded England in 1066, he commissioned a survey of the counties south of the Rivers Tees

Copublished by the IEEE Computer and Reliability Societies

and Ribble—the border with Scotland at the time. The resulting Domesday Book, completed in August 1086, describes 13,418 settlements: towns, villages, and hamlets, organized by manor, to reflect who owned the land. The 2 million inhabitants were ruled by 200 elite members; the crown, church, and nobility owned three-quarters of the land. This detailed description of life and land painted a valuable picture not only for William but also for the generations that followed.

The nature of surveillance changed over time. In addition to overt stock-taking, nations began to scrutinize people and their activities in secret. The Romans used the Caesar cipher to hide their communications; the Arabs developed the principles of cryptanalysis, but their use of it waned with their civilization's decline. However, during the Renaissance, Western countries rediscovered cryptanalysis, and "the new nation-states used it to read the messages that foreign ambassadors in their capitals were sending to their home countries. By the 1700s, clandestine mail-opening and cipher-solving centers called black chambers existed in most of the monarchies of Europe."[1]

But other kinds of more direct surveillance began to grow, too. In 1785, while in Russia with his brother, Jeremy Bentham described his idea for a Panopticon: a model prison where a guard can see all prisoners but can't be seen himself. The prison design involved a circular building with the inmates housed around the periphery, facing the center, and a watchman in the middle, hidden inside a room. Bentham called it "a new mode of obtaining power of mind over mind, in a quantity hitherto without example."[2]

Bentham had hoped that Czarina Catherine the Great would have found such a prison appealing, but it was never adopted in Russia. Indeed, although the Panopticon was never built, the idea of an all-seeing, anonymous guard has persisted in the criminal justice and cybersecurity realms. It appeals not only because the inmates can be watched but also because they can't know for sure when and whether they're being watched—the watchman can't watch everywhere at once—which means that the inmates must always act as though they're under surveillance, even when they're not. Thus, good behavior is encouraged, even when surveillance isn't actually happening.

In the Recent Past

Despite US Secretary of State Henry Stimson's protest that "Gentlemen do not read each other's mail," surveillance of written communication continued in the 1800s. But so did direct surveillance, even in the home. As the Soviet Union gained power and became more industrial during the first half of the 20th century, millions of people left the country to work in the city, especially Moscow. The government began to provide housing for these workers, often placing them in apartments that were abandoned by rich or aristocratic families. "Most people in Moscow lived in communal apartments; seven or more families crammed together where there had been one, sharing one kitchen and one bathroom." Victor Shenderovich notes that, "Communal kitchens were not places where you would bring your friends. I think that was one of the ideas for creating a communal kitchen. There would be a watchful eye of society over every communal apartment. People would report on each other. You would never know who would be reporting" (www.npr.org/blogs/thesalt/2014/05/20/314054405/how-russias-shared-kitchens-helped-shape-soviet-politics).

And technological "listening" continues today. We read every day about government data capture, as special "black" programs such as Total Information Awareness, as normal government operations such as accumulation of tax records or land use information, and as economic espionage, where governments intrude on commercial data stores to find out about intellectual property or negotiation plans. For instance, "when the Clinton administration was locked in a high-stakes negotiation in the 1990s to reach an accord with Japan, it bugged the Japanese negotiator's limousine" (www.nytimes.com/2014/05/20/us/us-treads-fine-line-in-fighting-chinese-espionage.html). More recently, the US listened in on Petrobras to understand Brazil's likely future energy policy (http://nyti.ms/1j6nJVq).

In fact, some government agencies are directed to listen in. Britain's Government Communications Headquarters captures transmitted signals of various sorts, as do the US National Security Agency, France's Direction Générale de la Sécurité Extérieure, and the Australian Security Intelligence Organisation. Daniel Soar explains how the NSA now openly acknowledges that its primary job is signals intelligence (www.lrb.co.uk/v35/n20/daniel-soar/how-to-get-ahead-at-the-nsa): "'the production of foreign intelligence through the collection, processing and analysis of communications or other data, passed or accessible by radio, wire or other electromagnetic means'. 'Communications or other data' that is 'passed or accessible' by 'electromagnetic means': that's anything emitted or received by a phone, computer, fax, radio, guidance system or satellite, or data that travels along any kind of cable, whether dedicated to voice signals or internet payloads or banking transactions or supposedly secure diplomatic, government and

military communications. It's anything with a pulse."

And how much does the NSA capture? As much as it can. Barton Gellman and Ashkan Soltani analyzed documents released by Snowden that reveal the NSA's systems are "capable of recording '100 percent' of a foreign country's [such as the Bahamas] telephone calls, enabling the agency to rewind and review conversations as long as a month after they take place. … A senior manager for the program compares it to a time machine—one that can replay the voices from any call without requiring that a person be identified in advance for surveillance" (www.washingtonpost.com/world/national-security/nsa-surveillance-program-reaches-into-the-past-to-retrieve-replay-phone-calls/2014/03/18/226d2646-ade9-11e3-a49e-76adc9210f19_story.html).

Criminals have discovered the advantages of surveillance, too, especially as technology has made it easier. For example, "The Blackshades Remote Access Tool or RAT, was advertised on hacker forums for 40 bucks. A copy of RAT can be secretly installed if a user is enticed to click on a link. Preet Bharara, US attorney for the Southern District of New York … explained that the software can remotely monitor keystrokes and steal passwords. … It even gave users the ability to activate a computer's camera to spy on a person in a victim's own home" (www.npr.org/2014/05/20/314138852/malicious-software-probe-reveals-vast-criminal-network). Nicole Perlroth reports similar experiences: after having written about a case where Chinese hackers' software lurked in thermostats and office printers, she has covered up her computer's camera with masking tape and moved the TV out of her bedroom (http://bits.blogs.nytimes.com/2014/05/21/in-cybersecurity-sometimes-the-weakest-link-is-a-family-member/). These situations echo the eerie experiences and cautions related by Jeff Stein last year in this magazine,[3] as his computer began acting strangely once he published some national security–related articles.

Other well-intentioned data capture and analysis happen every day: when your doctor adds information to your medical record and when your fitness band tracks your exercise habits.

But technological surveillance isn't always government based or malicious. Google has announced that it might monitor devices such as refrigerators, car dashboards, thermostats, eyeglasses, and watches in order to provide targeted advertising (http://bits.blogs.nytimes.com/2014/05/21/google-plans-to-deliver-ads-through-your-thermostat-and-car/). And because Google has products and services on many other devices used throughout the day, the combined dataset will enable it to offer advertising on different devices at different times. Indeed, some technology analysts think Google is "the most committed to getting everything possible onto the internet, its mission being 'to organise the world's information and make it universally accessible and useful'" (www.lrb.co.uk/v33/n19/daniel-soar/it-knows).

Other well-intentioned data capture and analysis happen every day: when your doctor adds information to your medical record, when your fitness band tracks your exercise habits, and when your email application evaluates your deletion patterns to build a spam profile to shelter you from unwelcome messages are just a few examples. Similarly, surveillance cameras and sensors watch your behavior in stores and public spaces, the better to provide shopping convenience (in the first case) and public safety (in the second case). Indeed, Robert O'Harrow Jr. has written extensively about the degree to which data aggregators and brokers, such as Acxiom, amass information that is already in the public record.[4] Having it in one place, easily searchable, can be helpful to some, a bane to others.

How Much Data?

So how much data are we talking about? Facebook has 1.28 billion users (www.nytimes.com/2014/05/23/technology/facebook-offers-privacy-checkup-to-all-1-28-billion-users.html?from=homepage): "half of the Internet-connected population of the planet has an account. …There are 7 billion people on the planet, and nearly 7 billion mobile phones; 6 billion emails are sent every hour; 1.2 petabytes of data travel across the Internet every minute, the equivalent of two thousand years' worth of music playing continuously, the contents of 2.2 billion books." The magnitude of surveillance is equally intimidating: in 2006, the BBC reported that Britain was the most surveilled society, with 4.2 million closed-circuit television cameras, one for every 14 people (http://news.bbc.co.uk/2/hi/uk_news/6108496.stm).

But the nature of the data is just as important as the numbers. Even when this type of data collection was relatively new, a study of Britain showed that data was already being collected from

- closed-circuit cameras, with 300 appearances per person per day;
- automobile license plate cameras;
- store RFID tags;

- mobile phones;
- store loyalty cards;
- credit card transactions;
- transportation cards, such as London's Oyster Card;
- satellites;
- election records;
- health service records;
- personal video recorders;
- tapped telephones;
- hidden cameras and bugged technology;
- call-monitoring in the workplace;
- employee time monitoring, such as time clocks;
- Internet cookies, Web bugs, and beacons; and
- keystroke-capturing programs.

Commercial establishments monitor us, too, by watching us online and in shops but also by placing sensors in our clothing. Nick Bilton describes a T-shirt from OMsignal of Montreal that monitors heart rate, breathing, caloric consumption, and stress level; it connects to your phone and to the Internet, to warn you if levels change dramatically or exceed a healthy threshold (http://bits.blogs.nytimes.com/2014/05/25/my-t-shirt-told-me-to-take-a-chill-pill/?ref=health). Lyon, France's Cityzen Sciences also produces textiles with embedded sensors; its T-shirt won an award this year for the most innovative product.

If you combine this information with what you choose to store on different devices, and then make it easy to search, some companies end up knowing where you are, with whom you communicate and how, and the content of all your email and voicemail messages. OPower, a new American company (www.opower.com), monitors your energy usage to let you know how it compares with typical usage by your friends and neighbors. Add to that what can be obtained from the public domain (your birthday, wedding day, the price you paid for your house), from other commercial vendors (your reading and music purchases, search terms, outstanding loans), and even from your website (your pictures, favorite restaurants, best friends), and it isn't hard to predict your behavior or know where to find you most of the time.

The vast amount of data available has been monetized by many companies. These days, it isn't just credit rating agencies that gather some of it up to assess your credit risk. In 2012, the US Federal Trade Commission requested data descriptions from nine data brokers (www.ftc.gov/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014): Acxiom, CoreLogic, Datalogix, eBureau, ID Analytics, Intelius, PeekYou, Rapleaf, and Recorded Future. These companies gather data about each of us from publicly available sources, including commercial, government, and online. The FTC discovered that the data brokers augment their raw data with "derived data": information inferred by putting together bits and pieces to support assumptions about you. For example, they assume that if you apply for a boating license, you have an interest in boating. Likewise, if you buy two Fiats, you might be loyal to that brand of car.

So how much data do these brokers use? One data broker's database has information on 1.4 billion consumer transactions and over 700 billion aggregated data elements; another data broker's database covers one trillion dollars in consumer transactions; and yet another data broker adds three billion new records each month to its databases. Most important, data brokers hold a vast array of information on individual consumers. For example, one of the nine data brokers has 3,000 data segments for nearly every US consumer.

Where Do We Go from Here?

James Gleick points out that, "We need to be aware that this kind of perfect, prosthetic memory that the Internet has created for us is a burden as much as it is a useful tool" (http://nyti.ms/1jBluz5). So what can we do about it? What choices do we have?

One choice is to simply opt out. But is opting out an option? Once upon a time you could do without email, without texting, without a cell phone, and without a smartphone, but such choices have become even harder today. As a small, but telling, example, consider that IEEE wants its magazines to have a presence on Facebook. But the person who creates that presence must also have a Facebook presence whether she wants one or not—and that holds true even for a committed privacy researcher. As Evgeny Morozov has observed (http://nyti.ms/1k5M02s), "While each of us can still choose not to be on Facebook, have a credit history or build a presence online, can we really afford not to do any of those things today? It was acceptable not to have a cellphone when most people didn't have them; today, when almost everybody does and when our phone habits can even be used to assess whether we qualify for a loan, such acts of refusal border on the impossible."

In other words, elements of tracking and monitoring pervade almost every aspect of our lives, so it's sometimes impossible to leave the surveillance behind. Even if you choose not to have a sensor in your car key that conveniently starts the engine as you approach the car, you might not be able to choose a bank account with no online presence (the data are stored so as to be available online, even when you have no intention of accessing your account that way) or a job where your phones are not monitored. Because

elements of tracking and monitoring pervade almost every aspect of our lives, it's sometimes impossible to leave the surveillance behind.

Advocates of data monitoring and prediction explain that the data are used to make improvements based on past behavior. But Morozov points out that predictive models will tell us what will happen, but they can't tell us why.

Moreover, data and predictive models can demonstrate only likely relationships. Understanding true cause and effect requires underlying theories that are then tested, developed, and revised. Auguste Comte, one of the fathers of sociology, pointed out that theories are essential to real science: "If it is true that every theory must be based upon observed facts, it is equally true that facts cannot be observed without the guidance of some theories. Without such guidance, our facts would be desultory and fruitless; we could not retain them: for the most part we could not even perceive them."[5]

Our choices are also affected by the decreasing transparency of surveillance. In the past, we could sometimes tell that we were being watched or followed; that awareness could influence our behaviors, including our decisions about what to do, where to go, and with whom. But now we're less and less able to know about the degree to which our actions are being scrutinized.

Consider facial recognition software. Joseph Atick, one of the field's pioneers, is also responsible for some of the rise in the now-booming biometrics business. Atick is proud of how his technology has prevented fraud and saved lives, but now that it's a $7.2 billion business sector, he's worried about how biometrics are being used: "What were those companies' policies for retaining and reusing consumers' facial data? Could they identify individuals without their explicit consent? Were they running face-matching queries for government agencies on the side?" (www.nytimes.com/2014/05/18/technology/never-forgetting-a-face.html). As ordinary citizens walk in the streets, they're being tracked—and decisions are being made about them.

Indeed, the 2014 FTC report describes how data brokers are mining data to provide marketing products, risk reduction products, and people search products. The marketing products tailor advertising campaigns to our suspected preferences. Charles Duhigg reports how Target offered baby product advertising to women it suspected were pregnant—even before the women knew (www.nytimes.com/2012/02/19/magazine/shopping-habits.html?_r=1&pagewanted=all). Risk reduction products attempt to detect fraud and verify identity, so that someone applying for a credit card or loan matches his or her claimed identity. And people search products offer publicly available information about consumers. The report points out that, "If a consumer is denied the ability to conclude a transaction based on an error in a risk mitigation product, the consumer can be harmed without knowing why. In such cases, the consumer is not only denied the immediate benefit, but also cannot take steps to prevent the problem from recurring."

Both Atick and the FTC acknowledge that many of these technologies add to our convenience and quality of life. Although Atick sees convenience in these kinds of uses, he describes an application to make his case for caution. NameTag, available for use with Google Glass, matches a name, occupation, and Facebook profile information to the face you're looking at. "We are basically allowing our fellow citizens to surveil us," Atick said. "Unlike fingerprinting or other biometric techniques, face recognition can be used at a distance, without people's awareness; it could then link their faces and identities to the many pictures they have put online. … Is faceprinting as innocuous as photography, an activity that people may freely perform? Or is a faceprint a unique indicator, like a fingerprint or a DNA sequence, that should require a person's active consent before it can be collected, matched, shared or sold? Dr. Atick is firmly in the second camp" (www.nytimes.com/2014/05/18/technology/never-forgetting-a-face.html).

In fact, several camps have formed around issues like this, with many technologists sparring verbally with civil libertarians. Atick has taken the middle ground, advocating actions that will preserve the ability to remain anonymous in public. For example, he suggests that public notices be placed wherever face recognition is being performed; that a consumer be asked for permission whenever a face print is given a unique identifier; and that face prints be used only for the reasons for which the consumer has given permission—much like the Fair Information Principles. Atick hopes that such steps will keep products and people from covertly linking real-world citizens with their online personas.

The FTC makes similar recommendations, suggesting not only some data broker best practices (such as privacy by design) but also that the US Congress update existing laws and implement new ones that would: "(1) allow consumers to access their own information; (2) allow consumers to suppress the use of this information; (3) disclose to consumers the data brokers' sources of information, so that, if possible, consumers can correct their information at the source; and (4) disclose any limitations of the opt-out option, such as the fact that close matches of an individual's name may continue to appear in search results."

The bottom line is that surveillance continues to intrude into more and more of life's crevices. Techniques abound for identifying, collecting, and storing the physical and digital crumbs that we leave as we lead our lives. Informed debate about the role of surveillance—and policies related to it—must acknowledge that many of these situations aren't new. Rather, their consequences have been amplified by technology, which in turn has amplified the consequences' reach. It is against this backdrop that we've prepared this special issue, focusing not just on the technology used for surveillance but also on how it might affect the very social fabric from which our communities are woven. ■

References

1. D. Kahn, *Codes: Secrets of the New Cryptology*, Macmillan, 1983.
2. J. Bentham, *The Panopticon Writings*, Verso, 1995.
3. J. Stein, "The End of National Security Reporting?," *IEEE Security & Privacy*, July/Aug. 2013, pp. 64–68.
4. R. O'Harrow Jr., *No Place to Hide*, Free Press, 2006.
5. A. Comte, *The Positive Philosophy of Auguste Comte Freely Translated and Condensed by Harriet Martineau*, (original in French 1855) AMS Press, 1974.

Shari Lawrence Pfleeger is editor in chief of *IEEE Security & Privacy* magazine. Contact her at pfleeger@dartmouth.edu.

cn *Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.*

Have an idea for a future article? Email editors Angela Sasse (A.Sasse@cs.ucl.ac.uk) and Alessandro Acquisti (acquisti@andrew.cmu.edu).