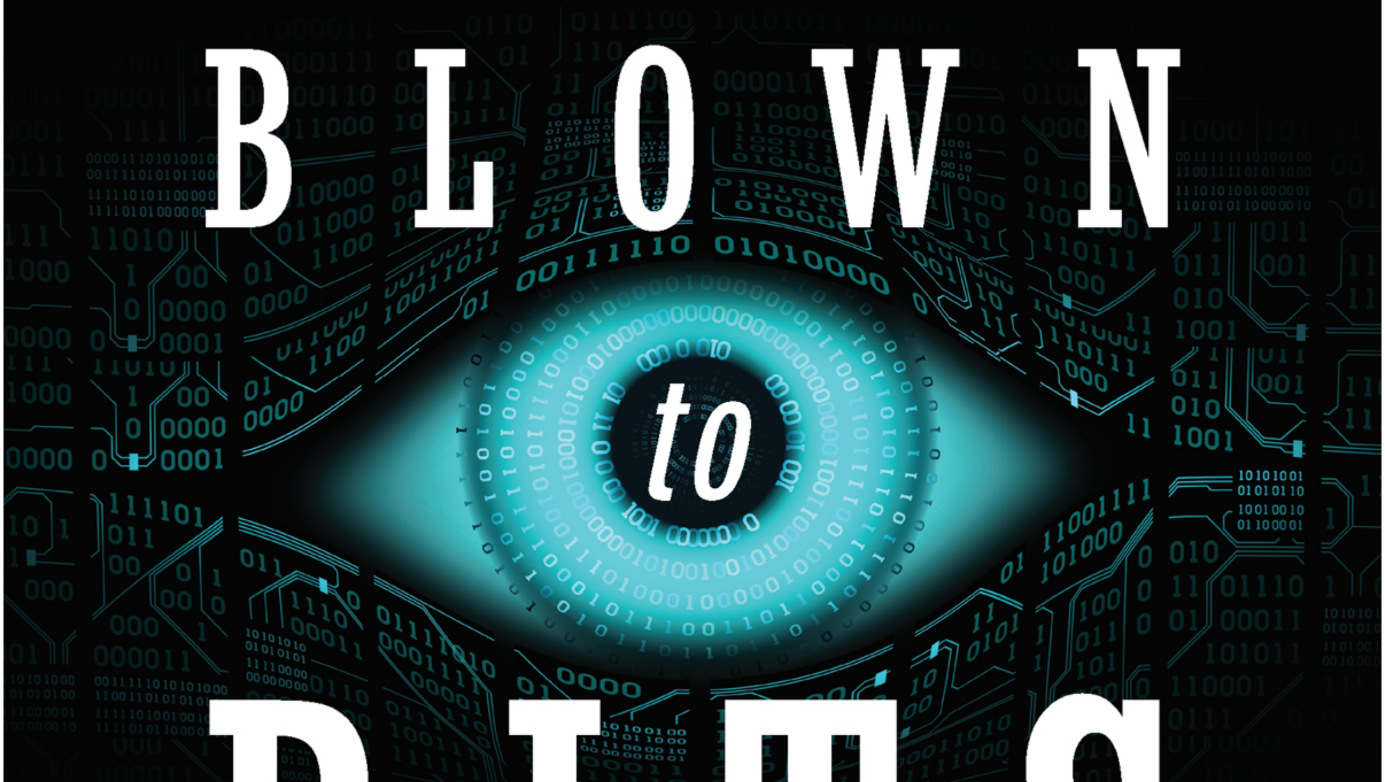
HAL ABELSON • KEN LEDEEN
HARRY LEWIS • WENDY SELTZER

BLOWN to BITS



[YOUR Life, Liberty, and Happiness After the Digital Explosion]

SECOND EDITION

Blown To Bits

Your Life, Liberty, and Happiness After the Digital Explosion

Second Edition

Hal Abelson
Ken Ledeen
Harry Lewis
Wendy Seltzer

Pearson

CHAPTER 2

Naked in the Sunlight

Privacy Lost, Privacy Abandoned

1984 Is Here, and We Like It

Fans attending Taylor Swift's packed Rose Bowl concert in the spring of 2018 saw her take the stage in a cloud of fog to sing hits from *Reputation*. As they entered or mingled between sets, some of those fans visited video kiosks to watch clips of the star's earlier performances and rehearsals, to get a behind-the-scenes glimpse of a favorite artist. What they didn't know was that the kiosk was watching them, too. The video booth was fitted with a camera that sent its visitors' images back to a "command post" in Nashville, where facial recognition software scanned them, reportedly looking for matches against a database of people who had stalked Swift in the past.[1] Were these images kept, or were they deleted securely? We don't know, just as we don't know how many other cameras capture us every day. Scanners like Swift's have been spotted at entrances to sports arenas, concert halls, and other entertainment venues. The public is often in the dark about their existence—and about policies related to how the images and other captured data are to be used, stored, or shared.

George Orwell's *1984* was published in 1948. Over subsequent years, the book became synonymous with a world of permanent surveillance, a society devoid of both privacy and freedom:

> ...there seemed to be no color in anything except the posters that were plastered everywhere. The black-mustachio'd face gazed down from every commanding corner. There was one on the house front immediately opposite. BIG BROTHER IS WATCHING YOU.[2]

The real 1984 came and went decades ago. Today, Big Brother's two-way telescreens would be amateurish toys today. Orwell's imagined London had cameras everywhere. His actual city now has at least half a million. Across the United Kingdom, there is one surveillance camera for every ten people.[3] The average Londoner is photographed hundreds of times a day by electronic eyes on the sides of buildings and on utility poles.

Yet there is much about the digital world that Orwell did not imagine. He did not anticipate that cameras are far from the most pervasive of today's tracking technologies. There are dozens of other kinds of data sources, and the data they produce is retained and analyzed. Cell phone companies know not only what numbers you call but where you have carried your phone. Credit card companies know not only where you spent your money but what you spent it on. Your friendly bank keeps electronic records of your transactions not only to keep your balance right but because it has to tell the government if you make huge withdrawals. When you go to a restaurant or a store, an app that has been quietly tracking your location asks you how you liked it, to feed your response into its recommendation-making engine.

The digital explosion has scattered the bits of our lives everywhere: records of the clothes we wear, the soaps we wash with, the streets we walk, and the cars we drive and where we drive them. And although Orwell's Big Brother had cameras, he didn't have search engines to piece the bits together, to find the needles in the haystacks. Wherever we go, we leave digital footprints, and computers of staggering capacity reconstruct our movements from those tracks. Computers reassemble the clues to form a comprehensive image of who we are, what we do, where we are doing it, and whom we are discussing it with.

Perhaps none of this would have surprised Orwell. Had he known about electronic miniaturization, he might have guessed that we would develop an astonishing array of tracking technologies. But there is something more fundamental that distinguishes the world of *1984* from the actual world of today. We have fallen in love with this always-on world. We accept our loss of privacy in exchange for efficiency, convenience, and small price discounts.

Attitudes have changed in the past decade. In a 2007 Pew/Internet Project report, 60% of Internet users were "not worried about how much information is available about them online," but by 2018, the ratio had flipped, and more than 60% "would like to do more to protect their privacy"; just 9% believe they have "a lot of control" over the information that is collected about them.[4] Although we're getting more worried about the loss of control over personal information, we're not sure there's much we can do about it.

In the world of bits, Big Brother has gotten both bigger and smaller. Technologically sophisticated nations like the United States and China have unprecedented ability to watch us, and they exercise that ability more often than we might like. Companies do, too. They've built new businesses around ubiquitous data collection, much of it geared toward marketing directly to us. Commercial data also forms a rich lode for government to mine, a public–private surveillance partnership.

We, too, are a part of the surveillance networks, keeping tabs on ourselves and one another. We invite apps to track our movements and smart assistants to listen in to our conversations. We record our changes of mood and chitchat with friends, and we snap photos of friends and strangers. About seven in ten adults have created profiles on social networking websites. Yet most are dissatisfied with the level of control they have over what happens to the data they post there.[5]

There are hints that the privacy tide may be changing, that we're not willing to trade privacy for the benefits of the digital world. Regulators are giving us new protections (although often not from government surveillance), and companies are now marketing privacy as a feature.

Bits Cubed: The Snowden Files

When a 29-year-old Edward Snowden met with journalists in the lobby of Hong Kong's Mira Hotel in June 2013, he told them to look for the guy with a Rubik's Cube.[6] They eventually did so and got a trove of classified documents and PowerPoint presentations describing massive U.S. government communications surveillance: a series of front-page stories for the journalists. Snowden, as a systems administrator for the National Security Agency (NSA), extracted gigabytes of material, copying it to micro-SD cards smaller than the stickers on his Rubik's Cube.

The Snowden revelations fueled a series of front-page stories in the *New York Times*, *Washington Post*, and *Guardian* in 2013.[7] They showed the NSA engaged in pervasive communications surveillance—not just of foreigners and suspected terrorists but of law-abiding American citizens. If you used Yahoo! Mail or Google Search or dozens of other popular services, you were swept up in the dragnet. While the U.S. Constitution and laws make a sharp distinction between U.S. citizens and "foreign persons" that limit the government's ability to spy on its citizens, the bits carried no such distinction, and citizens ended up in the same buckets.

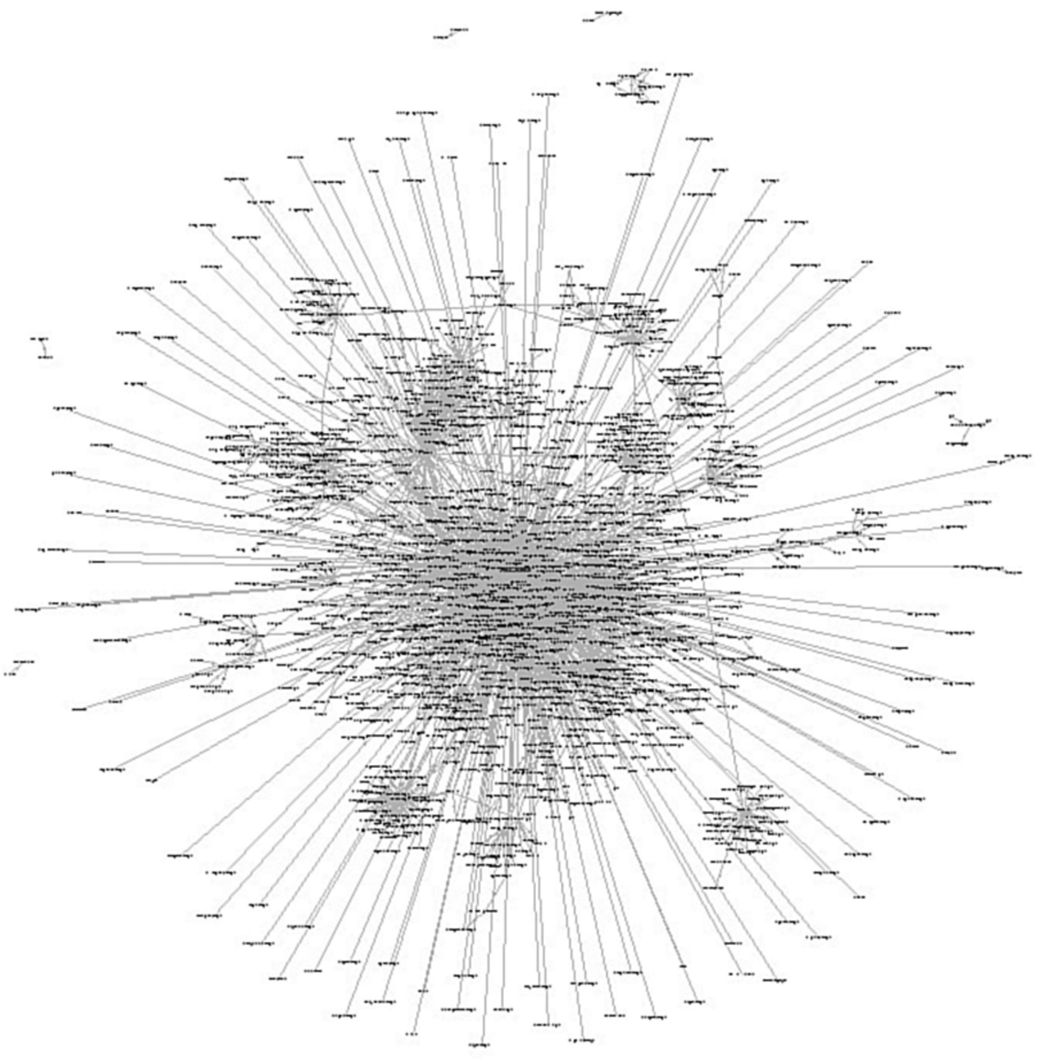
After the September 11, 2001, terrorist attacks, Congress passed new laws increasing spying powers. Notably, the USA PATRIOT Act authorized national security letters, which are secret demands for communications records; warrantless wiretaps of foreigners suspected of terrorist activity; and increasing ability to collect information on citizens any time obtaining foreign intelligence information is "a significant purpose" of the surveillance. Civil liberties groups expressed concern at the time that the act eliminated judicial checks and balances on surveillance,[8] but the act passed the Senate 98:1. Snowden's documents showed how far the NSA was pushing these new authorities.

The NSA exploited several properties of electronic communications. The popularity of centralized services for phone, email, search, and storage meant that taps at these corporate networks captured significant activity. The global nature of the Internet meant these taps could reach around the world from a few implant sites. A single request to Verizon for "business records" enabled the collection of millions of Americans' telephone call activity.[9] The Upstream program made full copies of everything carried along major domestic fiber-optic cable networks. Other top-secret warrantless data collection tools included XKEYSCORE[10] and EGOTISTICALGIRAFFE.[11]

U.S. officials defending the programs said they were only collecting metadata, not the contents of communications—the envelopes and addresses, not the letters inside. However, the web of contacts itself is tremendously informative. "We kill people based on metadata," said General Michael Hayden, former director of the NSA and the CIA.[12] William Binney, another ex-NSA whistleblower, left after the agency cut a program to conduct privacy-preserving searches.

The way we leave fingerprints and footprints is only part of what is new. We have always left a trail of information behind us—in our tax records, hotel reservations, and long-distance telephone bills. True, the footprints are far clearer and more complete today than ever before. But something else has changed: the harnessing of computing power to correlate data, to connect the dots, to put pieces together, and to create cohesive, detailed pictures from what would otherwise be meaningless fragments. The digital explosion does not just blow things apart. Like the explosion at the core of an atomic bomb, it blows things together as well. Gather up the details, connect the dots, and assemble the parts of the puzzle, and a clear picture will emerge.

Computers can sort through databases too massive and too boring to be examined with human eyes. They can assemble colorful pointillist paintings out of millions of tiny dots when any few dots would reveal nothing. When a federal court released half a million Enron emails obtained during the corruption trial, computer scientists quickly identified the subcommunities, and perhaps conspiracies, among Enron employees, using no data other than the pattern of who was emailing whom (see Figure 2.1). The same kinds of clustering algorithms work on patterns of telephone calls. You can learn a lot by knowing who is calling or emailing whom, even if you don't know what they are saying to each other—especially if you know the times of the communications and can correlate them with the times of other events.



Source: Enron, Jeffrey Heer, Figure 3 from http://jheer.org/enron/v1/

FIGURE 2.1 Diagram showing clusters of Enron emailers, indicating which employees carried on heavy correspondence with which others. The evident “blobs” may be the outlines of conspiratorial cliques.

The tale of Snowden and the NSA is two bits stories at once. Digital communication made it possible for the NSA to collect vast quantities of information, millions and millions of calls and emails, from just a few locations—something that would have been impossible if we were still communicating with regular phones and paper letters. And when Snowden took copies of everything, he could fit the equivalent of thousands of file cabinets of information into his pocket.

What can we do in the face of such government-directed surveillance? Snowden chose exposure, aiming for his disclosures to help support lawsuits against the programs and public pressure on lawmakers to rein in the NSA.

When he finished opening his Pandora's memory card of documents, he left us with a cause for hope: Math works. The NSA may have the world's best cryptographers and cryptanalysts, but the fundamental mathematics of encryption are still effective. The years since Snowden's disclosures have seen a dramatic increase in the use of encryption in basic Internet and web protocols[13] and in the applications that run on them.[14] End-to-end encryption enables us to reclaim some of the privacy that pervasive monitoring of unencrypted traffic unraveled.

"Reasonable Expectations of Privacy" Technology and the Fourth Amendment

Technological change has stood in tension with privacy before. When the Supreme Court first encountered the telephone wiretap in 1928, the president did not yet have a phone on his desk, although traffickers in illegal liquor (this was during Prohibition) had found the technology, and law enforcement wanted to listen in.[15] When the bootleggers challenged the tapping of their phone lines—alligator clips on physical wires outside homes and offices—the Court's majority put the telephone, which was high-tech at the time, in a frame they recognized, of physical intrusion and trespass. Without trespass, the Court held, there was no "search" or "seizure" and therefore no need for a warrant:

> The reasonable view is that one who installs in his house a telephone instrument with connecting wires intends to project his voice to those quite outside, that the wires beyond his house, and messages while passing over them, are not within the protection of the Fourth Amendment. Here those who intercepted the projected voices were not in the house of either party to the conversation.

Justice Brandeis, who did not agree, wrote in his dissent:

> Whenever a telephone line is tapped, the privacy of the persons at both ends of the line is invaded, and all conversations between them upon any subject, and although proper, confidential, and privileged, may be overheard. Moreover, the tapping of one man's telephone line involves the tapping of the telephone of every other person whom he may call, or who may call him."

But he was in the minority; for decades, warrantless wiretapping was lawful.

The Court's ruling in *Olmstead v. United States* increased the vulnerability of telephonic communications to police snooping, but it also publicly exposed that lack of privacy. Criminals, judges, and the general public learned that their conversations were liable to be tapped. As the telephone itself became more widely used, the legal rule triggered responses. States passed wiretap acts to protect by statute what the Constitution would not, and in 1934, Congress included anti-interception prohibitions in The Communications Act, section 605.

When Charles Katz came before the Court in 1967 to challenge the wiretapping of his (illegal wagering) conversation from a public telephone booth, the times, technologies, and legal norms had all changed. The telephone was part of everyday life, for personal and intimate communications as well as businesses both lawful and unlawful. The public and the justices themselves had experience to color their views of the technology. Asked again "whether a public telephone booth is a constitutionally protected area," the Court said that was the wrong formulation: It wasn't place but context. Telephone calls now demanded greater protection, even when conducted from the relative publicity of a glass-walled "public" phone booth. Justice Harlan, concurring in the judgment throwing out Katz's wiretap, articulated the test that still defines the Fourth Amendment's privacy protection: a "twofold requirement, first that a person have exhibited an actual (subjective) expectation of privacy and, second, that the expectation be one that society is prepared to recognize as 'reasonable.'"[16]

Location, Location, Location

Buy a navigation-equipped car, and it will listen to precisely timed signals from satellites reporting their positions in space. The Global Positioning System (GPS) calculates locations based on the satellites' locations and the times their signals are received. The 24 satellites spinning 12,500 miles above the earth enable your car to locate itself within 25 feet, at a price so low that most new cars have it as a standard feature. What was once a military secret now comes free in every smartphone.

If you carry a GPS-enabled cell phone, your friends can find you if that is what you want. If your GPS-enabled rental car has a radio transmitter, you can be found whether you want to be or not. Car leasing companies are adding transponders, including auto-immobilizers, to enable remote repossession, without even sending a repo man to the site. Those who fall behind in their car payments may suddenly find themselves unable to get to or from work.

GPS enables you to determine your location anywhere on earth, and even a low-end cell phone serves as a rudimentary positioning system. If you are traveling in settled territory—anyplace where you can get cell coverage—you move from the range of one cell tower to the range others, pinging the towers as you go. Triangulation among these signals can be used to locate you. The location is less precise than that supplied by GPS—only within ten city blocks or so—but the fact that it is possible at all means that a pattern of your behavior can be built, or photos can be stamped with identifying information about where they were shot, as well as when and with what camera.

Timothy Carpenter was given away by the bits from his cell phone, and then he was given a second chance by the law. A string of robberies of Radio Shack and T-Mobile stores in the Detroit area led to the arrest of four men in April 2011.[17] One of them confessed and gave the FBI the cell phone numbers of his accomplices; he also let law enforcement collect recently called numbers from his phone. With that evidence, prosecutors obtained an order directing wireless phone carriers to disclose information and cell-site location histories on the called numbers. They concluded that a phone registered to Timothy Carpenter had been near four store locations at the times when the stores were robbed. Carpenter was taken to court, and at trial, several confederates testified that he had been the leader of the robbery operation. With the corroboration of the cell-site mapping data, he was sentenced to more than 100 years in prison.

Carpenter appealed his case to the U.S. Supreme Court, arguing that the use of cell-site location data amounted to a "search,"[18] which could be conducted only with a warrant based on probable cause—not the mere order prosecutors had used to obtain records from the wireless carriers.

In 2018, the Court agreed with Carpenter: Because "cell phone location information is detailed, encyclopedic, and effortlessly compiled," the equivalent of a long-term digital "tail," individuals should have a reasonable expectation of privacy that their location history will not be exposed without a search warrant. Just because we're all carrying detailed location trackers, and those devices locate us to third parties in order to function, doesn't mean law enforcement gets automatic access to our map history. As it had earlier in *Katz*, the Court said that new technological capabilities shouldn't upend the balance between law enforcement and the public. Bits might trace our every move, but police need the judicial oversight of a search warrant to see them.

Lawyers and technologists argue about the balance between their disciplines. Post-Snowden revelations, they wonder whether we can trust the government to check its own power, or if records supposed to be available only by application to a neutral magistrate for a search warrant will instead be handed over or gathered in bulk. In 2013, after Edward Snowden revealed the existence of a secret, and extensive, data collection program code-named "PRISM," the NSA argued that records aren't effectively "collected" until they are searched, even once they are gathered in data banks. But while you can encrypt your conversations, it's much harder to hide the metadata of your digital footprints. (Tor Project's onion routing, https://www.torproject.org, is the best option.) To protect the privacy of those activities that must be public to be effective or that depend on interactions with others we don't necessarily trust to keep our secrets, we need the force of law and social norms.

Black Boxes: Not Just for Airplanes Anymore

On April 12, 2007, John Corzine, governor of New Jersey, was heading back to the governor's mansion in Princeton to mediate a discussion between Don Imus, the controversial radio personality, and the Rutgers University women's basketball team.[19]

His driver, 34-year-old state trooper Robert Rasinski, headed north on the Garden State Parkway. He swerved to avoid another car and flipped the governor's Chevy Suburban. Governor Corzine, who had not fastened his seatbelt, broke 12 ribs, a femur, his collarbone, and his sternum. The details of exactly what happened were unclear. When questioned, Trooper Rasinski said he was not sure how fast they were going—but we *do* know. He was going 91 in a 65-mile-per-hour zone. There were no police with radar guns around; no human was tracking his speed. We know his exact speed at the moment of impact because his car, like 30 million other cars in America, had a black box—an event data recorder (EDR) that captured every detail about what was going on just before the crash. An EDR is an automotive "black box" like the ones recovered from airplane crashes.

EDRs started appearing in cars around 1995, and they now appear in almost all models. Your insurance company is probably entitled to its data if you have an accident. Yet most people do not realize that EDRs exist, unless they've gotten an offer from their insurance company to give up real-time data rather than pay higher premiums.

EDRs capture information about speed, braking time, turn signal status, seat belts: information needed for accident reconstruction, to establish responsibility, or to prove innocence. CSX Railroad was exonerated of all liability in the death of the occupants of a car when its EDR showed that the car was stopped on the train tracks when it was hit. Police generally obtain search warrants before downloading EDR data—but not always; in some cases, they do not have to. When Robert Christmann struck and killed a pedestrian on October 18, 2003, Trooper Robert Frost of the New York State Police downloaded data from the car at the accident scene. The EDR revealed that Christmann had been going 38 miles per hour in an area where the speed limit was 30. When the data was introduced at trial, Christmann claimed that the state had violated his Fourth Amendment rights because it had not asked his permission or obtained a search warrant before retrieving the data. That was not necessary, ruled a New York court. Taking bits from the car was not like taking something out of a house, and no search warrant was necessary.[20]

Bits mediate our daily lives. It is almost as hard to avoid leaving digital footprints as it is to avoid touching the ground when we walk, and even if we lived our lives without walking, we would unsuspectingly be leaving fingerprints.

It is almost as hard to avoid leaving digital footprints as it is to avoid touching the ground when we walk.

Searching Location

Download your location history from Google or Facebook and look at the picture it paints. Does anything there make you nervous (whether it should or not)? What would you have difficulty explaining? Have you ever changed the settings from their account defaults? Should you?

Saving Time: Electronic Tolling and License Plate Readers

For commuters who use toll roads or bridges, the risk–reward calculation is not even close. Time is money, and time spent waiting in a car also means anxiety and frustration. If there is an option to get a toll booth transponder, many commuters will get one, even if the device costs a few dollars up front. Cruising past the cars waiting to pay with dollar bills is not just a relief; it actually brings the driver a certain satisfied glow.

The transponder, which the driver attaches to the windshield inside the car, is an RFID-enabled device powered with a battery that sends information to a sensor several feet away as the driver whizzes past. The sensor can be mounted in a constricted travel lane, where a toll booth for a human toll taker might have been. Or it can be mounted on a boom above traffic so the driver doesn't even need to change lanes or slow down. And what is the possible harm? Of course, the state is recording the fact that the car has passed the sensor; that is how the proper account balance can be debited to pay the toll. When the balance gets too low, the driver's credit card may get billed automatically to replenish the balance, which only makes the system better—no fumbling for change or doing anything else to pay for your travels.

The monthly bill—for the Massachusetts Fast Lane, for example—shows where and when you got on the highway, accurate to the second. It also shows how far you traveled on the highway and where you got off. Fast Lane also informs you of the mileage, which is another useful service because Massachusetts drivers can get a refund on certain fuel taxes if the fuel was used on the state toll road. Of course, you do not need a PhD to figure out that the state also knows when you got off the road, to the second, and that with one subtraction and one division, its computers could figure out if you were speeding. Technically, in fact, it would be trivial for the state to print the appropriate speeding fine at the bottom of the statement and to bill your credit card for that amount at the same time it charges you for the tolls. That would be taking convenience a bit too far, and no state does it—yet.

What does happen right now, however, is that toll transponder records are introduced into divorce and child custody cases. You've never been within 5 miles of that lady's house? Really? Why have you gotten off the highway at the exit near it so many times? You say you can be the better custodial parent for your children, but the facts suggest otherwise. As one lawyer put it, "When a guy says, 'Oh, I'm home every day at 5, and I have dinner with my kids every single night,' you subpoena his E-ZPass and you find out he's crossing that bridge every night at 8:30. Oops!" Such records have been subpoenaed hundreds of times in family law cases. They have also been used in employment cases, to prove that the car of a worker who said he was working was actually far from the workplace.

But most of us aren't planning to cheat on our spouses or our bosses, so the loss of privacy seems like no loss at all—at least compared to the time saved. Of course, if we actually *were* cheating, we *would* be in a big hurry and might take some risks to save a few minutes!

Massachusetts toll roads eliminated toll takers in 2017. Drivers can save some money by equipping their cars with transponders, but if they don't have a transponder, never fear: "We will bill you," the state announces in billboards along the highway. There are no cash lanes now. Gantries equipped with both transponder antennas and automated license plate readers record every car or truck that passes through. To stay anonymous, you need to take the slow road.

The License Plate Tells More Than You Think

In June 2018, southern California mall operator Irvine Company was found to be collecting the license plate numbers of vehicles entering its parking areas. When a visitor, 14-year-old Zoe Wheatcroft, dug deeper into the company's "privacy policy," she found that Irvine was not only collecting license plate information but sharing it with law enforcement, in a database that might be accessed by agents from Immigration and Customs Enforcement (ICE).[21] When word got out, Irvine and Vigilant, the database company, responded that their policy was in fact narrower and more restrictive but gave customers no way to know that a shopping trip wouldn't put them in surveillance crosshairs.

Automated license plate recognition is a form of mass surveillance enabled by cheaper and more sophisticated cameras, software, and network capabilities. Automatic cameras capture images of license plates, convert the plate numbers into plaintext characters, and annotate the images with time, date, and GPS-derived location before transmitting and storing each instance. The data stream may be queried in real time, as in a search for a wanted criminal or stolen vehicle, or it may be retrieved later to give a picture of shoppers' demographics or a particular shopper's travel pattern.

Loose Fitbits Sink Ships?

The Strava fitness-mapping application offers a connection to users' GPS-enabled smartphones, watches, and Fitbit devices in order to enable athletes to track their runs, cycle routes, and other activities. Strava combined the data into a "heatmap" visualization, aggregating more than a billion activity logs into colored streaks across a map. While the Strava team highlighted a few recreation images on their blog—the Ironman triathlon swim off Hawaii, mountain biking in Whistler, British Columbia—a researcher, noting what appeared to be the outlines of military bases in Afghanistan, posted screenshots to Twitter and reminded people "turning off data sharing is an option."[22] Strava's CEO followed up with a blog post pointing to explanations of the privacy settings and promising to work with military and government officials "to address potentially sensitive data."[23]

Of course, one can say that soldiers in sensitive locations should turn off their location reporting—which means they need to know that their devices and applications have that setting and consider its consequences. But the Strava heatmap may be only the most visible and most easily changed of the places we leave these trails. Cell phones build location maps as they ping nearby towers; frequently accessed websites have logs of the IP addresses from which they are viewed (from which the site operator can map corresponding geolocation); and many mobile apps collect location information to target advertising. Individual data points may seem harmless, but points gathered over time and space can paint a detailed picture of travel patterns or home life—and even secret military strategy.

Big Brother, Abroad and in the United States

Big Brother really is watching today, and his job has gotten much easier, thanks to the digital explosion. In China, which has a long history of tracking individuals as a mechanism of social control, the millions of residents of Shenzhen are being issued identity cards, which record far more than the bearer's name and address. According to a report in the *New York Times*,[24] the cards document the individual's work history, educational background, religion, ethnicity, police record, medical insurance status, landlord's phone number, and reproductive history. Touted as a crime-fighting measure, the new technology—developed by an American company—will come in handy in dealing with cases of street protests and individual activities deemed suspicious by the authorities. The sort of record keeping that used to be the responsibility of local authorities is becoming automated and nationalized as the country prospers and its citizens become increasingly mobile. The technology makes it easier to know where everyone is, and the government is taking advantage of that opportunity. In Xinjiang, where the Uighur minority faces especially strict scrutiny, police have an app that can flag when someone has stopped using a smartphone or avoids the front door. Facial recognition is targeted at Uighurs, who are made to pass through checkpoints that Han (the ethnic majority elsewhere in China) are permitted to avoid. Chinese tracking is far more detailed and pervasive than Britain's system of ubiquitous surveillance cameras.

Identifying Citizens—Without ID Cards

In the age of global terrorism, democratic nations are resorting to digital surveillance to protect themselves, creating hotly contested conflicts with traditions of individual liberty. In the United States, the idea of a national identification card prompts a furious libertarian reaction from parties not usually outspoken in defense of individual freedom. Under the REAL ID Act of 2005, uniform federal standards were to be implemented for state-issued driver's licenses. Although it passed through Congress without debate, the law is opposed by at least 18 states. Resistance pushed back the implementation timetable multiple times. In 2018, 13 years later, only 37 states met the REAL ID rules. Finally, in 2019, states were told their final extension would expire, and only REAL ID–compliant documents would be accepted for federal identification by October 2020. Then COVID-19 hit, and the deadline was extended yet again. Yet even fully implemented, REAL ID would fall far short of the true national ID preferred by those charged with fighting crime and preventing terrorism.

As the national ID card debate continues in the United States, the FBI is making it irrelevant by exploiting emerging technologies. There would be no need for anyone to carry an ID card if the government had enough biometric data on Americans—that is, detailed records of their fingerprints, irises, voices, walking gaits, facial features, scars, and earlobe shapes. Gather a combination of measurements on individuals walking in public places, consult the databases, connect the dots, and—bingo!—their names pop up on the computer screen. No need for them to carry ID cards; the combination of biometric data would pin them down perfectly.

As the national ID card debate continues in the United States, the FBI is making it irrelevant by exploiting emerging technologies.

Well, only imperfectly at this point, but the technology is improving. And the data is already being gathered and deposited in the data vault of the FBI's Criminal Justice Information Services database in Clarksburg, West Virginia. The database already holds some 75 million sets of fingerprints, and the FBI processes 100,000 requests for matches every day. Any of 900,000 federal, state, and local law enforcement officers can send a set of prints and ask the FBI to identify it. If a match comes up, the individual's criminal history can be accessed in the database, too.

But fingerprint data is hard to gather; mostly it is obtained when people are arrested. The goal of the project is to get identifying information on nearly everyone—and to get it without bothering people too much. For example, a simple notice at airport security could advise travelers that, as they pass through airport security, a detailed "snapshot" will be taken as they enter the secure area. The traveler would then know what is happening and could have refused (and stayed home). As an electronic identification researcher puts it, "That's the key. You've chosen it. You have chosen to say, 'Yeah, I want this place to recognize me.'"[25] The project eliminates the issue of REAL ID controversies, as all the data being gathered is, in some sense at least, offered voluntarily.

San Francisco, California, the epicenter of the technology boom, moved in the other direction, banning law enforcement use of facial recognition technology.[26] The Board of Supervisors heard concerns that the technologies were biased, lacked transparency, and could be abused by government. At the same time, however, data-based identification flourishes in private hands. The company PatronScan asserts that its database of IDs swiped at bars contains more than 60 million IDs across 200 cities. PatronScan not only checks that bar-goers are of legal drinking age but maintains a blacklist of patrons flagged for "bad behavior."[27]

Friendly Cooperation Between Big Siblings

In fact, there are two Big Brothers, and they often work together. We are, by and large, glad they are watching—if we are aware of it at all. Only occasionally are we alarmed about their partnership.

The first Big Brother is Orwell's—the government. And the other Big Brother is the industry about which most of us know very little: the business of aggregating, consolidating, analyzing, and reporting on the billions of individual transactions, financial and otherwise, that take place electronically every day. Of course, the commercial data aggregation companies are not in the spying business; none of their data reaches them illicitly. But they do know a lot about us, and what they know can be extremely valuable, both to businesses and to the government.

The new threat to privacy is that computers can extract significant information from billions of apparently uninteresting pieces of data, in the way that mining technology has made it economically feasible to extract precious metals from low-grade ore. Computers can correlate databases on a massive level, linking government data sources together with private and commercial ones to create comprehensive digital dossiers on millions of people. With their massive data storage and processing power, they can make connections in the data, by using brute force rather than ingenuity. And the computers can discern even very faint traces in the data—traces that may help track payments to terrorists, set insurance rates, or simply help us make sure our new babysitter is not a sex offender.

And so we turn to the story of the government and the aggregators.

Acxiom is the country's biggest customer data company. Its business is to aggregate transaction data from all those swipes of cards in card readers all over the world. This amounted to more than a hundred billion transactions in 2018.[28] The company uses its massive amounts of data about financial activity to support the credit card industry, banks, insurers, and other consumers of information about how people spend money. Unsurprisingly, after the War on Terror began, the Pentagon also got interested in Acxiom's data and the ways the company gathers and analyzes it. Tracking how money gets to terrorists might help find the terrorists and prevent some of their attacks.

ChoicePoint is the other major U.S. data aggregator. ChoicePoint has more than 100,000 clients, which call on it for help in screening employment candidates, for example, or determining whether individuals are good insurance risks.

Acxiom and ChoicePoint are different from older data analysis operations in the scale of their operations. Quantitative differences have qualitative effects, as we said in Chapter 1; what has changed is not the technology but rather the existence of rich data sources. Forty years ago, credit cards had no magnetic stripes. Charging a purchase was a mechanical operation; the raised numerals on the card made an impression through carbon paper so you could have a receipt, and the top copy went to the company that issued the card. Today, if you charge something using your CapitalOne card, the bits go instantly not only to CapitalOne but to Acxiom and other aggregators. The ability to search through huge commercial data sources—including not just credit card transaction data but phone call records, travel tickets, and banking transactions, for example—is another illustration that more of the same can create something new.

Privacy laws do exist, of course. For a bank, or a data aggregator, to post your financial data on its website would be illegal. But privacy is still developing as an area of the law, and it is connected to commercial and government interests in uncertain and surprising ways.

A critical development in privacy law was precipitated by the presidency of Richard Nixon. In what is generally agreed to be an egregious abuse of presidential power, Nixon used his authority as president to gather information on those who opposed him—in the words of his White House counsel at the time, to "use the available federal machinery to screw our political enemies." Among the tactics Nixon used was to have the Internal Revenue Service audit the tax returns of individuals on an "enemies list," which included members of Congress, journalists, and major contributors to Democratic causes. Outrageous as it was to use the IRS for this purpose, it was not illegal, so Congress moved to ban it in the future.

The Privacy Act of 1974 established broad guidelines for when and how the federal government can assemble dossiers on citizens it is not investigating for crimes. The government has to give public notice about what information it wants to collect and why, and it has to use what it collects only for those reasons.

The Privacy Act limits what the government can do to gather information about individuals and what it can do with records it holds. Specifically, it states, "No agency shall disclose any record which is contained in a system of records by any means of communication to any person, or to another agency, except pursuant to a written request by, or with the prior written consent of, the individual to whom the record pertains, unless...." If the government releases information inappropriately, even to another government agency, the affected citizen can sue for damages in civil court. The protections provided by the Privacy Act are sweeping—although not as sweeping as they may seem. Not every government office is in an "agency"; the courts are not, for example. The act requires agencies to give public notice of the uses to which they will put the information, but the notice can be buried in the Federal Register, where the public probably won't see it unless news media happen to report it. Then there is the "unless" clause, which includes significant exclusions. For example, the law does not apply to disclosures for statistical, archival, or historical purposes; civil or criminal law enforcement activities; Congressional investigations; or valid Freedom of Information Act requests.

Despite the act's exclusions, government practices changed significantly because of this law. Then, a quarter century later, came 9/11. *Law enforcement should have seen it all coming*, was the constant refrain as investigations revealed how many unconnected dots were in the hands of different government agencies. *It all could have been prevented if the investigative fiefdoms* *had been talking to each other. They should have been able to connect the dots.* But they could not—in part because the Privacy Act restricted interagency data transfers. A response was badly needed. The Department of Homeland Security was created to ease some of the interagency communication problems, but that government reorganization was only a start.

In January 2002, just a few months after the World Trade Center attack, the Defense Advanced Research Projects Agency (DARPA) established the Information Awareness Office (IAO) with a mission to:

> imagine, develop, apply, integrate, demonstrate, and transition information technologies, components and prototype, closed-loop, information systems that will counter asymmetric threats by achieving total information awareness useful for preemption; national security warning; and national security decision making. The most serious asymmetric threat facing the United States is terrorism, a threat characterized by collections of people loosely organized in shadowy networks that are difficult to identify and define. IAO plans to develop technology that will allow understanding of the intent of these networks, their plans, and potentially define opportunities for disrupting or eliminating the threats. To effectively and efficiently carry this out, we must promote sharing, collaborating, and reasoning to convert nebulous data to knowledge and actionable options.

Vice Admiral John Poindexter directed the effort that came to be known as "Total Information Awareness" (TIA). The growth of enormous private data repositories provided a convenient way to avoid many of the prohibitions of the Privacy Act. The Department of Defense can't get data from the Internal Revenue Service because of the 1974 Privacy Act. *But the government can buy the very same data it is barred from collecting from private data aggregators!* In a May 2002 email to Adm. Poindexter, Lt. Col Doug Dyer discussed negotiations with Acxiom:

> Acxiom's Jennifer Barrett is a lawyer and chief privacy officer. She's testified before Congress and offered to provide help. One of the key suggestions she made is that people will object to Big Brother, wide-coverage databases, but they don't object to use of relevant data for specific purposes that we can all agree on. Rather than getting all the data for any purpose, we should start with the goal, tracking terrorists to avoid attacks, and then identify the data needed (although we can't define all of this, we can say that our templates and models of terrorists are good places to start). Already, this guidance has shaped my thinking.

> Ultimately, the U.S. may need huge databases of commercial transactions that cover the world or certain areas outside the U.S. This information provides economic utility, and thus provides two reasons why foreign countries would be interested. Acxiom could build this mega-scale database.

The *New York Times* broke the story in October 2002. As Poindexter had explained in speeches, the government had to "break down the stovepipes" separating agencies and get more sophisticated about how to create a big picture out of a million details, no one of which might be meaningful in itself. The *Times* story set off a sequence of reactions from the Electronic Privacy Information Center and civil libertarians. Congress defunded the office in 2003—but that was not the end of the idea.

The key to TIA was data mining: looking for connections across disparate data repositories, finding patterns, or "signatures," that might identify terrorists or other undesirables. The General Accountability Office report on Data Mining (GAO-04-548) reported on a survey of 128 federal departments.[29] It described 199 separate data mining efforts, of which 122 used personal information.

Although IAO and TIA went away, Project ADVISE at the Department of Homeland Security continued with large-scale profiling system development. Eventually, Congress demanded that the privacy issues concerning this program be reviewed as well. In his June 2007 report (OIG-07-56), Richard Skinner, the DHS inspector general, stated that "program managers did not address privacy impacts before implementing three pilot initiatives," and a few weeks later, the project was shut down. But ADVISE was only one of a dozen data-mining projects going on in DHS at the time.

Similar privacy concerns led to the cancellation of the Pentagon's TALON database project. That project sought to compile a database of reports of suspected threats to defense facilities as part of a larger program of domestic counterintelligence.

Despite these privacy concerns, as Edward Snowden revealed, many surveillance and data mining programs simply carried on under the radar.

The government creates projects, the media and civil liberties groups raise serious privacy concerns, the projects are canceled, and new projects arise to take their place. The cycle seems to be endless. In spite of Americans' traditional suspicions about government surveillance of their private lives, the cycle seems to be almost an inevitable consequence of Americans' concerns about their security and the responsibility that government officials feel to use the best available technologies to protect the nation. Corporate databases often contain the best information on the people about whom the government is curious.

Data Collection, Data Breach

Storage is cheap, but security is difficult. One of the depressingly common events in the digital era is data breach. A customer database is exposed, and user accounts or credit cards are misused until the breach is rectified. Data breach notification laws in many states now provide some transparency, as well as incentive for companies storing data to clean up to avoid class action lawsuits.

Amid numerous breaches, Equifax and OPM stand out. Equifax, one of the big credit-reporting companies, stores records of credit card account payment histories. If you go to take out a car loan or a mortgage, the lender will check your credit score with Equifax. In September 2017, Equifax announced a data breach that exposed the personal information—names and dates of birth, Social Security numbers, physical addresses, and other personal information that could lead to identity theft and fraud—of 147 million people, more than half the adult population of the United States.[30] The Federal Trade Commission complaint alleged that Equifax failed to take basic network security measures, including failing to update database software when notified of an access control vulnerability. This occurred even as the company had a privacy policy promising it implemented "reasonable physical, technical and procedural safeguards" to protect consumer data. Equifax settled the FTC complaint with an agreement to pay at least $575 million and potentially up to $700 million. As part of the settlement, affected consumers were offered free credit monitoring services. Those trying to exclude themselves from future databases, however, were told "You cannot opt out of this data collection."[31]

As the human resources arm of the U.S. government, the Office of Personnel Management collects a great deal of sensitive information: identification, background checks, and fingerprints.[32] Over 21 million of these records were stolen when OPM's data stores were breached in 2014. When people's credit cards are stolen, they get new cards. When their Social Security numbers are taken, they can be enrolled in credit monitoring services. But you can't be issued a new set of fingerprints.

The number of new data sources—and the proliferation and interconnection of old data sources—is part of the story of how the digital explosion shattered privacy. But the other part of the technology story is about how all that data is put together.

Exponential growth—in storage size, processing speed, and communication speed—have changed the same old thing into something new. Blundering, stupidity, curiosity, malice, and thievery are not new. The fact that sensitive data about everyone in a nation could fit on a laptop *is* new. The ability to search for a needle in the haystack of the Internet *is* new. Easily connecting "public" data sources that used to be stored in file drawers in Albuquerque and Atlanta but are now both electronically accessible from Algeria—*that* is new, too.

Training, laws, and software all can help. But the truth of the matter is that, as a society, we don't really know how to deal with these consequences of the digital explosion. The technology revolution is outstripping society's capacity to adjust to the changes in what can be taken for granted.

Sometimes even public information is revealing. In Massachusetts, the Group Insurance Commission (GIC) is responsible for purchasing health insurance for state employees. When the premiums it was paying jumped one year, the GIC asked for detailed information on every patient encounter. And for good reason: All kinds of health care costs had been growing at prodigious rates. In the public interest, the state had a responsibility to understand how it was spending taxpayer money. The GIC did not want to know patients' names; it did not want to track individuals, and it did not want people to *think* they were being tracked. Indeed, tracking the medical visits of individuals would have been illegal.

So, the GIC data had no names, no addresses, no Social Security numbers, no telephone numbers—nothing that would be a "unique identifier" enabling a mischievous junior staffer in the GIC office to see who exactly had a particular ailment or complaint. To use the official lingo, the data was "deidentified"—that is, stripped of identifying information. The data did include the gender, birth date, zip code, and similar facts about individuals making medical claims, along with some information about why they had sought medical attention. That information was gathered not to challenge any particular person but to learn about patterns; if the truckers in Worcester are having lots of back injuries, for example, maybe workers in that region need better training on how to lift heavy items. Most states do pretty much the same kind of analysis of deidentified data about state workers.

Now this was a valuable data set not just for the Insurance Commission but for others studying public health and the medical industry in Massachusetts. Academic researchers, for example, could use such a large inventory of medical data for epidemiological studies. Because it was all deidentified, there was no harm in letting others see it, the GIC figured. In fact, it was such good data that private industry—for example, businesses in the health management sector—might pay money for it. And so the GIC sold the data to businesses. The taxpayers might even benefit doubly from this decision: The data sale would provide a new revenue source to the state, and in the long run, a more informed health care industry might run more efficiently.

But how deidentified was the material—really?

Latanya Sweeney was at the time a researcher at MIT. (She went on to become a professor at Carnegie Mellon University and then Harvard University.) She wondered how hard it would be for those who had received the deidentified data to "reidentify" the records and learn the medical problems of a particular state employee—for example, the governor of the Commonwealth.

Governor Weld lived, at that time, in Cambridge, Massachusetts. Cambridge, like many other municipalities, makes its voter lists publicly available for a charge of $15—and free for candidates and political organizations. For a particular precinct, you can obtain the records for only $.75. Sweeney spent a few dollars and got the voter lists for Cambridge. Anyone else could have done the same.

According to the Cambridge voter registration list, there were only six people in Cambridge with Governor Weld's birth date, only three of those were men, and only one of those lived in Governor Weld's five-digit zip code. Sweeney could use that combination of factors—birth date, gender, and zip code—to recover the Governor's medical records; she could therefore also recover the records of members of his family, since the data was organized by employee. This type of reidentification is straightforward. In Cambridge, in fact, birth date alone was sufficient to identify more than 10% of the population. Nationally, gender, zip code, and date of birth are all it takes to identify 87% of the U.S. population uniquely.

The data set contained far more than gender, zip code, and birth date. In fact, any of the 58 individuals who received the data in 1997 could have identified any of the 135,000 people in the database. "There is no patient confidentiality," said Dr. Joseph Heyman, president of the Massachusetts Medical Society. "It's gone."[33]

It is easy to read a story like this and scream, "Heads should roll!" But it is actually quite hard to figure out *who, if anyone, made a mistake*. Certainly collecting the information was the right thing to do, given that health costs are a major expense for all businesses and institutions. The GIC made an honest effort to deidentify the data before releasing it. Arguably the GIC might not have released the data to other state agencies. Data is a valuable resource, and once someone has collected it, the government is entirely correct in wanting it used for the public good. Forbidding such sharing would be like saying that every department of government should acquire its heating oil independently. Some might object to selling the data to an outside business—but only in retrospect; had the data really been better deidentified, whoever made the decision to sell the data might well have been rewarded for helping to hold down the cost of government.

Perhaps the mistake was the ease with which voter lists can be obtained. However, it is a tradition deeply ingrained in our system of open elections that the public may know who is eligible to vote and, indeed, who has voted. And voter lists are only one source of public data about the U.S. population. How many

It is easy to read a story like this and scream, "Heads should roll!" But it is actually quite hard to figure out who, if anyone, made a mistake.

21-year-old male Native Hawaiians live in Middlesex County, Massachusetts? In the year 2000, there were four. Anyone can browse the U.S. Census data, and sometimes it can help fill in pieces of a personal picture: Just go to factfinder.census.gov.

The mistake was thinking that the GIC data was truly deidentified when it was not. But with so many data sources available, and so much computing power that could be put to work connecting the dots, it is very hard to know just how much information has to be discarded from a database to make it truly anonymous. Aggregating data into larger units certainly helps; releasing data by five-digit zip codes reveals less than releasing it by nine-digit zip codes. But the coarser the data, the less it reveals also of the valuable information for which it was made available.

The Internet of Things

We have already observed that even privacy-conscious people surrender their privacy in exchange for convenience and small cost savings. Nowhere is this principle more evident than in the networking of light switches, refrigerators, and doorbells known as the Internet of Things (IoT). And it turns out that privacy is not the only thing we sacrifice when we let the Internet grow into everything we touch (and everything we no longer need to touch). The security of everything in the network can be compromised.

On October 21, 2016, the U.S. East Coast woke up to a massive Internet outage. Many popular websites for work and play, such as Twitter, Netflix, GitHub, and Reddit, wouldn't load.[34] It turned out that critical Internet services were under attack by hordes of machines elsewhere on the Internet. These machines were sending so many simultaneous requests that nameservers, key components of the Internet's traffic management infrastructure, couldn't keep up with the load. Trying to respond to malicious requests left them unable to answer legitimate users. Without nameservers to give directions, requesting computers couldn't find the sites, and so Twitter was "down" for users even while the service itself was still functioning.

Investigating researchers and engineers found that the requests were coming from an army of "smart" home devices: Internet-connected baby monitors, light bulbs, and routers. The devices' owners hadn't intended this activity and were mostly unaware of it. The gadgets had been infected by malicious software and enlisted in the Mirai botnet. Together, a group of devices operating with weak computing power on home Internet connections

> Unfortunately, this kind of mass attack has become common enough to get a name and an acronym: distributed denial of service (DDoS).

combined into a force strong enough to disrupt global Internet services. The malicious software, or *malware*, exploited some common security flaws—default administrative passwords that hadn't been changed and unpatched and outdated software on devices exposed directly to the Internet—to infect one device and then copy itself to other devices (an infection pattern known as a *worm*). Once installed, the malware turned each device into a waiting "bot," listening for commands.

On October 21, the controller directed the cohort to send a rapid stream of requests for domain names, which caused a burst of traffic to publicly accessible nameservers, including those of major nameservice provider Dyn DNS. Dyn reported that, under the attack, it was getting 10 to 20 times the normal volume of requests, which it estimated came from 100,000 malicious or infected end devices.[35] These requests, along with retry efforts from real end users who couldn't get through, overwhelmed Dyn's defenses and left the company's servers unable to respond to legitimate lookups.

What's New Here? Scale, Control, Connectedness, and Interoperability

The Internet of Things promises to connect the physical world much as the Internet of bits connects computers and data. Sometimes that means putting general-purpose connected computers into previously "dumb" devices like refrigerators. A smart refrigerator might not only warn you when you are running out of milk but contact your grocery store and have it delivered to your home and charged to your credit card. Other times, being "smart" means opening an interface by which one can remotely read from and control sensors (devices that see, hear, or otherwise perceive their environment) and actuators (devices that do something, like shut off the dryer). Smart thermostats, for example, can be triggered by motion detectors to turn on the heat or air conditioning when someone is in the room. These connected things enable a vision for automated factories and supply chains, smart homes and cities, and self-driving car fleets.

As sensors, actuators, and chips get cheaper, they grow in number. They also propagate down the value chain. When chips were expensive, it may have made sense to put them in costly equipment like airplanes, but today they are in doorbells. Even low-end smartphones are now smart enough to be at the center of a home appliance network. Capabilities that once were purchased only by factories and run by experts are now available to the general public. Sometimes smart gadgets are capable of doing much more than the purchaser realizes because the functionality is simplified for marketing reasons. At the same time, safety, reliability, and other less marketable features are given short shrift—and the manufacturers try to excuse their misplaced priorities by saying that devices must be kept small and operate on low power.

A light bulb or thermostat is often "set and forget": Once the device is functioning, its owner thinks of it as an appliance rather than as a small computer in need of security monitoring and software updates. Moreover, budget sellers of the devices might view them as one-time sales with no follow-on support, and even users who want to update software may find themselves with no option to do so. Another alternative is a suite of centrally managed devices, but this option may be more costly in terms of both dollars and customer privacy. Not everyone wants to share their lighting and temperature preferences, much less the audio and video streams from their baby monitors, with a company storing that data somewhere unknown and using it for who knows what.

Many IoT devices are always on, awaiting the moment when their owners will throw the switch to light a room. That makes them attractive hosts for writers of malware—programs designed with evil intent. The cleverest malware doesn't interfere with the devices' normal function; rather, it lurks invisibly, waiting for the "attack" command.

Threats: One-to-One Versus One-to-Many

When one home has a smart refrigerator, its behavior is interesting and important to only a handful of people.[36] An attacker could spoil a gallon of milk and make a mess of the kitchen, drain a bank account by ordering caviar instead of milk (if the right limits aren't set), or use the machine to cause local damage, possibly shorting the whole house's circuitry. Multiply the devices, however, and they can be leveraged to do damage beyond their neighborhood. The first *D* in DDoS, is for *distributed*. Replicating an attack from thousands of distributed devices can have an overwhelming cumulative effect. Denial of service can take many forms: requests for service that look legitimate but that are sent at a high volume; requests that take an unusually long time to fulfill; or requests that are malformed in such a way as to disable or crash the server to which they are sent. For example, consider what happens when all of a town's high schoolers call the local pizza parlor at the same Friday lunch time to ask the price of a slice with the works. A customer genuinely calling to order a pie will probably give up after a few busy signals.

What makes the IoT an "Internet" is standard protocols that enable the devices to communicate with one another and their controllers, Internet protocols, or special-purpose Wi-Fi or Bluetooth-based communications standards. Standard interfaces enable users to address multiple devices together, such as to plug a new light bulb in to an existing setup or add a freezer to a smart refrigerator. Designers might have anticipated that a connected baby monitor could share updates with grandparents or enable caregivers to listen from the far reaches of their homes; connectivity could enable a refrigerator to consult a weather report and order ice cream when the temperature hits 80 degrees. Connectivity can enable devices to get smarter over time, with software updates and new possibilities for interaction. However, unguarded connectivity can leave openings for intrusions like the Mirai worm, and the common interfaces and underlying software enable malware writers to "break once, run anywhere."

In December 2017, more than a year after attacks took Dyn offline, three men pleaded guilty to charges of computer fraud and abuse, admitting to having written the software behind Mirai: Paras Jha, an undergraduate studying computer science at Rutgers University in New Jersey, and two friends or associates. According to their plea, they first targeted their attack against online gaming servers for the popular Minecraft game, where they were attempting to overwhelm the servers to gain an advantage. Later, Jha started a business selling computer-protection services and launched attacks against Rutgers while taunting the university that it should be buying DDoS protection. Jha and his associates didn't necessarily intend to disrupt Dyn or the Internet at large, but after they posted the software's source code online, others modified and redeployed the malware, pointing it at new targets.

Who's Responsible for IoT Security?

When the Tesla Model 3 electric car was first reviewed by *Consumer Reports*, it got poor marks for braking.[37] "The Tesla's stopping distance of 152 feet from 60 mph was far worse than any contemporary car we've tested," wrote the reviewer. A week after publication of the *Consumer Reports* review, however, the car manufacturer sent an over-the-air software update to cars across the country, including those that had already been sold. The car's braking distance improved by 19 feet, performance comparable to that of other compact cars, prompting *Consumer Reports* to upgrade its review.[38] Tesla told *Consumer Reports* that it had updated software controlling the Model 3's antilock braking system.

This wasn't the first time an over-the-air update changed vehicle performance. While Hurricane Irma was heading for Florida, Tesla overrode software-defined range limitations for cars in the storm's path, enabling owners to escape further.[39] Both cases illustrate the blurred line between software and hardware and murky outlines of product boundaries. Physical features of the car were changed by a remote software update, and a car's owner might not have even been aware of the change or given an opportunity to accept or reject it. Few owners would reject longer range (a feature that was a costly upgrade outside Irma's wake), but what if the more consistent braking came at the cost of other performance features? Some owners found themselves thinking their cars were slower after the upgrade. What if public safety came at the price of speed? Should owners get to refuse updates?

Driving poses an externality problem. It's not just Tesla drivers who take risks if their cars don't stop in time; their vehicles are more dangerous to everyone who shares the road with them. We impose safety standards and inspection requirements on automobiles to reduce such risks and make roads safer. We might similarly impose a duty to upgrade on software and hardware users. If your software-enabled product is causing risks to others, if a safer alternative is found, you could be required to update, even if doing so would cause you some inconvenience or cost. Yet it's not just obviously dangerous and expensive objects like cars that require this caution; some of the connected devices taken over by the Mirai botnet were cheap toys. Some of their vendors might no longer be in business. Would this requirement change the nature of ownership?

Bruce Schneier speaks of the Internet of Things as a "world-sized robot," with sensors and actuators spanning the globe.[40] As the capabilities of this robot to cause harm—and actual examples of harm—multiply, he predicts that demands for regulation and liability will follow. Unless those who are building the technologies also build safeguards, the political and regulatory responses are likely to be blunt and may include prohibitions on connecting or using devices or broad restrictions on their use. Worse, regulations that do not account for the architectures and incentives of connectedness may fail to protect us.

Smart Cities: Efficiency, Individual Choice, Privacy, and Systemic Risks

An older man in a New York apartment complained that he was virtually imprisoned in his own home after the landlord installed app-controlled "smart locks" at the building's lobby entrance. Tony Mysak, 93 and blind in one eye, was unable to use the smartphone app required to open the lock. Mary Beth McKenzie, Tony's wife, objected to giving a record of her entries to the building and to Latch, the lock provider. The Latch app's privacy policy (since changed) noted that the app collected GPS location information that Latch might use for marketing purposes, as well as providing a record of door accesses and photographs to building management. When she asked for a physical key, the landlord laughed and offered only a smartcard. McKenzie and Mysak and a group of tenants had to sue their landlord to win the right to use keys instead of apps.[41]

The tenants had several complaints about the digitization of their front door. For some, it was the change in usability, from a familiar physical key to a new application. For others, it was the privacy, the sense that their entries and exits—and perhaps even their travels—would be tracked and compiled, not by a human and fallible doorman but in an impersonal corporate database that wasn't even visible to them. The new affordances of this system, such as the ability to let a guest or super into the building without leaving a key under the doormat, weren't enough to compensate tenants for their loss of control.

Scale this "smart" building up a few orders of magnitude, and you get the "smart city," full of embedded and networked sensors. Traffic lights might coordinate with cars and buses for efficiency; power meters might communicate with the electrical grid in real time to smooth demand.

The city of Toronto planned a revitalization of its industrial waterfront "from the Internet up." The new Quayside would be built as a "smart city" in a partnership between the city and Google/Alphabet's Sidewalk Labs. But as they engaged in giddy futuristic speculation, planners were surprised by the opposition their announcement sparked. People complained about privacy, security, and loss of control. Who gets to see the data generated by the digital infrastructure; who gets to make decisions based on it? Sadly, we won't learn the answer. In May 2020, Toronto and Google scrapped the project. The decision was taken amid the Coronavirus pandemic—but privacy advocates claimed credit.

> "This is a major victory for the responsible citizens who fought to protect Canada's democracy, civil and digital rights," said one opponent of the project. "Toronto will go down in history as one of the more disturbing planned experiments in surveillance capitalism"[42]—referring to the title of a best-selling business book.

Interconnection brings new privacy and security concerns. Who can learn when you're out of town by watching for changes in power usage patterns? Who can learn when you have company or take a hot shower? By monitoring the power signatures of home devices, a watcher could even see when you start the morning coffee pot or turn on the evening news.

The flow of bits, storage capacity, and processing power needed for analysis all tend to heighten the power disadvantage of individuals against governments and corporations. Privacy serves as a way of taking back some control, a zone of autonomy. In Orwell's imagined London, only O'Brien and other members of the Inner Party could escape the gaze of the telescreen. For now, individuals can employ a mix of mathematical and legal protections to shut out the watching eyes of Big Brother—at least most of the time.

Endnotes

[1] Sopan Deb and Natasha Singer, “Taylor Swift Keeping An Eye Out For Stalkers,” *New York Times*, December 15, 2018, C6, https://www.nytimes.com/2018/12/13/arts/music/taylor-swift-facial-recognition.html

[2] George Orwell, *1984* (Signet Classic, 1977), p. 2.

[3] Silkie Carlo, “Britain Has More Surveillance Cameras per Person Than Any Country Except China. That’s a Massive Risk to Our Free Society,” *Time*, May 17, 2019, https://news.yahoo.com/britain-more-surveillance-cameras-per-151641361.html.

[4] Lee Rainie, “Americans’ Complicated Feelings About Social Media in an Era of Privacy Concerns,” Pew Research Center, March 27, 2018, https://www.pewresearch.org/fact-tank/2018/03/27/americans-complicated-feelings-about-social-media-in-an-era-of-privacy-concerns/.

[5] Lee Raine, Americans’ complicated feelings about social media in an era of privacy concerns, Pew Research Center, March 27, 2018, https://www.pewresearch.org/fact-tank/2018/03/27/americans-complicated-feelings-about-social-media-in-an-era-of-privacy-concerns/.

[6] Edward Snowden, *Permanent Record* (Metropolitan Books, 2019).

[7] https://www.washingtonpost.com/investigations/us-intelligence-mining-data-from-nine-us-internet-companies-in-broad-secret-program/2013/06/06/3a0c0da8-cebf-11e2-8845-d970ccb04497_story.html.

[8] Kevin Bankston, “EFF Analysis of the Provisions of the USA PATRIOT Act,” Electronic Frontier Foundation, October 27, 2003, https://www.eff.org/deeplinks/2003/10/eff-analysis-provisions-usa-patriot-act.

[9] Glenn Greenwald, “NSA Collecting Phone Records of Millions of Verizon Customers Daily,” *The Guardian*, June 6, 2013, https://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order.

[10] Micah Lee et al., “A Look at the Inner Workings of NSA’s XKEYSCORE,” *The Intercept*, July 2, 2015, https://theintercept.com/2015/07/02/look-under-hood-xkeyscore/.

[11] Tom Bowman, “Why Does the NSA Keep an EGOTISTICALGIRAFFE? It’s Top Secret,” NPR, November 10, 2003, https://www.npr.org/2013/11/10/244240199/why-does-the-nsa-keep-an-egotisticalgiraffeits-top-secret.

[12] David Cole, “We Kill People Based on Metadata,” *The New York Review of Books*, May 10, 2014, https://www.nybooks.com/daily/2014/05/10/we-kill-people-based-metadata/.

[13] Stephen Farrell and Hannes Tschofenig, “Pervasive Monitoring Is an Attack,” Internet Engineering Task Force, RFC 7258, May 2014, https://tools.ietf.org/html/rfc7258.

[14] “HTTPS Encryption on the Web," Google Transparency Report, accessed May 18, 2020, https://transparencyreport.google.com/https/overview;%20https:=.

15 *Olmstead v. United States*, 277 U.S. 438 (1928), https://supreme.justia.com/cases/federal/us/277/438/.

16 *Katz v. United States*, 389 U.S. 347 (1967), https://supreme.justia.com/cases/federal/us/389/347/.

17 Eric Sandy, "Supreme Court Case Has Roots in Radio Shack Robberies in Michigan and Ohio," *Detroit Metro Times*, November 28, 2017, https://www.metrotimes.com/news-hits/archives/2017/11/28/supreme-court-case-has-roots-in-radio-shack-robberies-in-michigan-and-ohio.

18 *Carpenter v. United States*, 138 S. Ct. 2206 (2018), https://www.oyez.org/cases/2017/16-402.

19 Ellen Messmer, "Black Hat: Researcher Claims Hack of Processor Used to Secure Xbox 360, Other Products," *Network World,* February 2, 2010, https://www.networkworld.com/article/2243700/black-hat--researcher-claims-hack-of-processor-used-to-secure-xbox-360--other-products.html.

20 *People v. Christmann*, 861 N.W.2d 18 (2015), https://caselaw.findlaw.com/ny-justice-court/1143124.html.

21 Taylor Hatmaker, "California Malls Are Sharing License Plate Tracking Data with an ICE-Linked Database," TechCrunch, July 10, 2018, https://social.techcrunch.com/2018/07/10/alpr-license-plate-recognition-ice-irvine-company/.

22 Thomas Brewster, "Why Strava's Fitness Tracking Should Really Worry You," *Forbes*, January 29, 2018, https://www.forbes.com/sites/thomasbrewster/2018/01/29/strava-fitness-data-location-privacys-care/#46e488aa55c3.

23 James Quarles, "A Letter to the Strava Community," Strava, January 29, 2018 https://blog.strava.com/press/a-letter-to-the-strava-community/.

24 Chris Buckley and Paul Mozur, "How China Uses High-Tech Surveillance to Subdue Minorities," *The New York Times*, May 22, 2019, https://www.nytimes.com/2019/05/22/world/asia/china-surveillance-xinjiang. html.

25 Ellen Nakashima, "FBI Prepares Vast Database of Biometrics," *The Washington Post*, December 22, 2007, http://www.washingtonpost.com/wp-dyn/content/article/2007/12/21/AR2007122102544.html.

26 Kate Conger et al., "San Francisco Bans Facial Recognition Technology," *The New York Times*, May 14, 2019, https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html.

27 Susie Cagle, "This ID Scanner Company Is Collecting Sensitive Data on Millions of Bargoers," Medium, May 29, 2019, https://onezero.medium.com/id-at-the-door-meet-the-security-company-building-an-international-database-of-banned-bar-patrons-7c6d4b236fc3.

28 "The 2019 Federal Reserve Payments Study," Board of Governors of the Federal Reserve System, January 6, 2020, https://www.federalreserve.gov/paymentsystems/2019-December-The-Federal-Reserve-Payments-Study.htm.

29 GAO, U.S. Government Accountability Office, https://www.gao.gov/products/GAO-04-548.

30 "Equifax to Pay $575 Million as Part of Settlement with FTC, CFPB, and States Related to 2017 Data Breach," Federal Trade Commission, July 22, 2019, https://www.ftc.gov/news-events/press-releases/2019/07/equifax-pay-575-million-part-settlement-ftc-cfpb-states-related.

31 "Equifax Data Breach Settlement," Federal Trade Commission, July 11, 2019, https://www.ftc.gov/enforcement/cases-proceedings/refunds/equifax-data-breach-settlement.

32 OPM.GOV, Cybersecurity Resource Center, https://www.opm.gov/cybersecurity/cybersecurity-incidents/

33 Michael Lasalandra, "Panel told releases of med records hurt privacy," *Boston Herald*, March 20, 1997.

34 Manos Antonakakis et al., "Understanding the Mirai Botnet," *Proceedings of 26th USENIX Security Symposium*, April 16, 2017; "Mirai IoT Botnet Co-Authors Plead Guilty," Krebs on Security, December 13, 2017, https://krebsonsecurity.com/2017/12/mirai-iot-botnet-co-authors-plead-guilty/.

35 Scott Hilton, "Dyn Analysis Summary of Friday October 21 Attack," Oracle, 2016. http://dyn.com/blog/dyn-analysis-summary-of-friday-october-21-attack/.

36 C. J. Hughes, "The Latest in Apartment Technology: Fridge Cams and Robotic Valets," *The New York Times*, December 15, 2017, https://www.nytimes.com/2017/12/15/realestate/apartment-technology-fridge-cams-robotic-valets.html.

37 Patrick Olsen, "Tesla Model 3 Falls Short of a CR Recommendation," *Consumer Reports*, May 30, 2018, https://www.consumerreports.org/hybrids-evs/tesla-model-3-review-falls-short-of-consumer-reports-recommendation/.

38 Patrick Olsen, "Tesla Model 3 Gets CR Recommendation After Braking Update," *Consumer Reports*, May 30, 2018, https://www.consumerreports.org/car-safety/tesla-model-3-gets-cr-recommendation-after-braking-update/.

39 Andrew Liptak, "Tesla Extended the Range of Some Florida Vehicles for Drivers to Escape Hurricane Irma," *The Verge*, September 10, 2017, https://www.theverge.com/2017/9/10/16283330/tesla-hurricane-irma-update-florida-extend-range-model-s-x-60-60d.

40 Bruce Schneier, *Click Here to Kill Everybody: Security and Survival in a Hyper-Connected World* (WW Norton & Company, 2018).

41 Alfred Ng, "Tenants Win as Settlement Orders Landlords Give Physical Keys over Smart Locks," CNET, May 7, 2019, https://www.cnet.com/news/tenants-win-rights-to-physical-keys-over-smart-locks-from-landlords/.

42 Rob Gillies, Google Affiliate scraps plan for Toronto Smart City Project, *US News and World Report*, May 7, 2020, https://www.usnews.com/news/business/articles/2020-05-07/google-affiliate-scraps-plan-for-toronto-smart-city-project and Shoshana Zuboff, *The Age of Surveillance Capitalism,* PublicAffairs, 2019.