Why We Lost Our Privacy, or Gave It Away

Information technology did not cause the end of privacy, any more than automotive technology caused teen sex. Technology creates opportunities and risks, and people, as individuals and as societies, decide how to live in the changed landscape of new possibilities. To understand why we have less privacy today than in the past, we must look not just at the gadgets. To be sure, we should be wary of spies and thieves, but we should also look at those who protect us and help us—and we should also take a good look in the mirror.

We are most conscious of our personal information winding up in the hands of strangers when we think about data loss or theft. Reports like the one about the British tax office have become fairly common. The theft of information about 45 million customers of TJX stores, described in Chapter 5, "Secret Bits," was even larger than the British catastrophe. In 2003, Scott Levine, owner of a mass email business named Snipermail, stole more than a billion personal information records from Acxiom. Millions of Americans are victimized by identity theft every year, at a total cost in the tens of billions of dollars annually. Many more of us harbor daily fears that just "a little bit" of our financial information has leaked out, and could be a personal time bomb if it falls into the wrong hands.

Why can't we just keep our personal information to ourselves? Why do so many other people have it in the first place, so that there is an opportunity for it to go astray, and an incentive for creative crooks to try to steal it?

We lose control of our personal information because of things we do to ourselves, and things others do to us. Of things we do to be ahead of the curve, and things we do because everyone else is doing them. Of things we do to save money, and things we do to save time. Of things we do to be safe from our enemies, and things we do because we feel invulnerable. Our loss of privacy is a problem, but there is no one answer to it, because there is no one reason why it is happening. It is a messy problem, and we first have to think about it one piece at a time.

We give away information about ourselves—voluntarily leave visible footprints of our daily lives—because we judge, perhaps without thinking about it very much, that the benefits outweigh the costs. To be sure, the benefits are many.

Saving Time

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 is also 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 from inside the car, is an RFID, powered with a battery so identifying information can be sent to the 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. All that 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—when, accurate to the second. It also shows where you got off and how far you went. Informing you of the mileage 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 as it was charging for 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 five 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 five 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!" These records can be subpoenaed, and have been, 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!

Saving Money

Sometimes it's money, not time, which motivates us to leave footprints. Such is the case with supermarket loyalty cards. If you do not want Safeway to keep track of the fact that you bought the 12-pack of Yodels despite your recent cholesterol results, you can make sure it doesn't know. You simply pay the "privacy tax"—the surcharge for customers not presenting a loyalty card. The purpose of loyalty cards is to enable merchants to track individual item purchases. (Item-level transactions are typically not tracked by credit card companies, which do not care if you bought Yodels instead of granola, so long as you pay the bill.) With loyalty cards, stores can capture details of cash transactions as well. They can process all the transaction data, and draw inferences about shoppers' habits. Then, if a lot of people who buy Yodels also buy Bison Brew Beer, the store's automated cash register can automatically spit out a discount coupon for Bison Brew as your Yodels are being bagged. A "discount" for you, and more sales for Safeway. Everybody wins. Don't they?

As grocery stores expand their web-based business, it is even easier for them to collect personal information about you. Reading the fine print when you sign up is a nuisance, but it is worth doing, so you understand what you are giving and what you are getting in return. Here are a few sentences of Safeway's privacy policy for customers who use its web site:

> Safeway may use personal information to provide you with newsletters, articles, product or service alerts, new product or service announcements, saving awards, event invitations, personally tailored coupons, program and promotional information and offers, and other information, which may be provided to Safeway by other companies. ... We may provide personal information to our partners and suppliers for customer support services and processing of personal information on behalf of Safeway. We may also share personal information with our affiliate companies, or in the course of an actual or potential sale, re-organization, consolidation, merger, or amalgamation of our business or businesses.

Dreary reading, but the language gives Safeway lots of leeway. Maybe you don't care about getting the junk mail. Not everyone thinks it is junk, and the

company does let you "opt out" of receiving it (although in general, few people bother to exercise opt-out rights). But Safeway has lots of "affiliates," and who knows how many companies with which it *might* be involved in a merger or sale of part of its business. Despite privacy concerns voiced by groups like C.A.S.P.I.A.N. (Consumers Against Supermarket Privacy Invasion and Numbering, `www.nocards.org`), most shoppers readily agree to have the data collected. The financial incentives are too hard to resist, and most consumers just don't worry about marketers knowing their purchases. But whenever purchases can be linked to your name, there is a record, somewhere in a huge database, of whether you use regular or super tampons, lubricated or unlubricated condoms, and whether you like regular beer or lite. You have authorized the company to share it, and even if you hadn't, the company could lose it accidentally, have it stolen, or have it subpoenaed.

Convenience of the Customer

The most obvious reason not to worry about giving information to a company is that you do business with them, and it is in your interest to see that they do their business with you better. You have no interest in whether they make more money from you, but you do have a strong interest in making it easier and faster for you to shop with them, and in cutting down the amount of stuff they may try to sell you that you would have no interest in buying. So your interests and theirs are, to a degree, aligned, not in opposition. Safeway's privacy policy states this explicitly: "Safeway Club Card information and other information may be used to help make Safeway's products, services, and programs more useful to its customers." Fair enough.

No company has been more progressive in trying to sell customers what they might want than the online store Amazon. Amazon suggests products to repeat customers, based on what they have bought before—or what they have simply looked at during previous visits to Amazon's web site. The algorithms are not perfect; Amazon's computers are drawing inferences from data, not being clairvoyant. But Amazon's guesses are pretty good, and recommending the wrong book every now and then is a very low-cost mistake. If Amazon does it too often, I might switch to Barnes and Noble, but there is no injury to me. So again: Why should anyone care that Amazon knows so much about me? On the surface, it seems benign. Of course, we don't want the credit card information to go astray, but who cares about knowing what books I have looked at online?

Our indifference is another marker of the fact that we are living in an exposed world, and that it feels very different to live here. In 1988, when a

How Sites Know Who You Are

1. **You tell them**. Log in to Gmail, Amazon, or eBay, and you are letting them know exactly who you are.
2. **They've left cookies on one of your previous visits**. A *cookie* is a small text file stored on your local hard drive that contains information that a particular web site wants to have available during your current session (like your shopping cart), or from one session to the next. Cookies give sites persistent information for tracking and personalization. Your browser has a command for showing cookies—you may be surprised how many web sites have left them!
3. **They have your IP address**. The web server has to know where you are so that it can ship its web pages to you. Your IP address is a number like 66.82.9.88 that locates your computer in the Internet (see the Appendix for details). That address may change from one day to the next. But in a residential setting, your Internet Service Provider (your *ISP*—typically your phone or cable company) knows who was assigned each IP address at any time. Those records are often subpoenaed in court cases.

If you are curious about who is using a particular IP address, you can check the American Registry of Internet Numbers (`www.arin.net`). Services such as `whatismyip.com`, `whatismyip.org`, and `ipchicken.com` also allow you to check your own IP address. And `www.whois.net` allows you to check who owns a domain name such as `harvard.com`—which turns out to be the Harvard Bookstore, a privately owned bookstore right across the street from the university. Unfortunately, that information won't reveal who is sending you spam, since spammers routinely forge the source of email they send you.

videotape rental store clerk turned over Robert Bork's movie rental records to a Washington, DC newspaper during Bork's Supreme Court confirmation hearings, Congress was so outraged that it quickly passed a tough privacy protection bill, The Video Privacy Protection Act. Videotape stores, if any still exist, can be fined simply for keeping rental records too long. Twenty years later, few seem to care much what Amazon does with its millions upon millions of detailed, fine-grained views into the brains of all its customers.

It's Just Fun to Be Exposed

Sometimes, there can be no explanation for our willing surrender of our privacy except that we take joy in the very act of exposing ourselves to public

view. Exhibitionism is not a new phenomenon. Its practice today, as in the past, tends to be in the province of the young and the drunk, and those wishing to pretend they are one or the other. That correlation is by no means perfect, however. A university president had to apologize when an image of her threatening a Hispanic male with a stick leaked out from her MySpace page, with a caption indicating that she had to "beat off the Mexicans because they were constantly flirting with my daughter."

And there is a continuum of outrageousness. The less wild of the party photo postings blend seamlessly with the more personal of the blogs, where the bloggers are chatting mostly about their personal feelings. Here there is not exuberance, but some simpler urge for human connectedness. That passion, too, is not new. What *is* new is that a photo or video or diary entry, once posted, is visible to the entire world, and that there is no taking it back. Bits don't fade and they don't yellow. Bits are forever. And we don't know how to live with that.

Bits don't fade and they don't yellow. Bits are forever. And we don't know how to live with that.

For example, a blog selected with no great design begins:

> This is the personal web site of Sarah McAuley. ... I think sharing my life with strangers is odd and narcissistic, which of course is why I'm addicted to it and have been doing it for several years now. Need more? You can read the "About Me" section, drop me an email, or you know, just read the drivel that I pour out on an almost-daily basis.

No thank you, but be our guest. Or consider that there is a Facebook group just for women who want to upload pictures of themselves uncontrollably drunk. Or the Jennicam, through which Jennifer Kay Ringley opened her life to the world for seven years, setting a standard for exposure that many since have surpassed in explicitness, but few have approached in its endless ordinariness. We are still experimenting, both the voyeurs and viewed.

Because You Can't Live Any Other Way

Finally, we give up data about ourselves because we don't have the time, patience, or single-mindedness about privacy that would be required to live our daily lives in another way. In the U.S., the number of credit, debit, and bank cards is in the billions. Every time one is used, an electronic handshake records a few bits of information about who is using it, when, where, and for what. It is now virtually unheard of for people to make large purchases of

ordinary consumer goods with cash. Personal checks are going the way of cassette tape drives, rendered irrelevant by newer technologies. Even if you could pay cash for everything you buy, the tax authorities would have you in their databases anyway. There even have been proposals to put RFIDs in currency notes, so that the movement of cash could be tracked.

Only sects such as the Amish still live without electricity. It will soon be almost that unusual to live without Internet connectivity, with all the fingerprints it leaves of your daily searches and logins and downloads. Even the old dumb TV is rapidly disappearing in favor of digital communications. Digital TV will bring the advantages of video on demand—no more trips to rent movies or waits for them to arrive in the mail—at a price: Your television service provider will record what movies you have ordered. It will be so attractive to be able to watch what we want when we want to watch it, that we won't miss either the inconvenience or the anonymity of the days when all the TV stations washed your house with their airwaves. You couldn't pick the broadcast times, but at least no one knew which waves you were grabbing out of the air.

Little Brother Is Watching

So far, we have discussed losses of privacy due to things for which we could, in principle anyway, blame ourselves. None of us really needs a loyalty card, we should always read the fine print when we rent a car, and so on. We would all be better off saying "no" a little more often to these privacy-busters, but few of us would choose to live the life of constant vigilance that such resolute denial would entail. And even if we were willing to make those sacrifices, there are plenty of other privacy problems caused by things others do to us.

The snoopy neighbor is a classic American stock figure—the busybody who watches how many liquor bottles are in your trash, or tries to figure out whose Mercedes is regularly parked in your driveway, or always seems to know whose children were disorderly last Saturday night. But in Cyberspace, we are all neighbors. We can all check up on each other, without even opening the curtains a crack.

Public Documents Become VERY Public

Some of the snooping is simply what anyone could have done in the past by paying a visit to the Town Hall. Details that were always public—but inaccessible—are quite accessible now.

In 1975, Congress created the Federal Election Commission to administer the Federal Election Campaign Act. Since then, all political contributions have been public information. There is a difference, though, between "public" and "readily accessible." Making public data available on the Web shattered the veil of privacy that came from inaccessibility.

Want to know who gave money to Al Franken for Senate? Lorne Michaels from Saturday Night Live, Leonard Nimoy, Paul Newman, Craig Newmark (the "craig" of craigslist.com), and Ginnie W., who works with us and may not have wanted us to know her political leanings. Paul B., and Henry G., friends of ours, covered their bases by giving to both Obama and Clinton.

The point of the law was to make it easy to look up big donors. But since data is data, what about checking on your next-door neighbors? Ours definitely leaned toward Obama over Clinton, with no one in the Huckabee camp. Or your clients? One of ours gave heartily to Dennis Kucinich. Or your daughter's boyfriend? You can find out for yourself, at www.fec.gov or fundrace.huffingtonpost.com. We're not telling about our own.

Hosts of other facts are now available for armchair browsing—facts that in the past were nominally public but required a trip to the Registrar of Deeds. If you want to know what you neighbor paid for their house, or what it's worth today, many communities put all of their real estate tax rolls online. It was always public; now it's accessible. It was never wrong that people could get this information, but it feels very different now that people can browse through it from the privacy of their home.

If you are curious about someone, you can try to find him or her on Facebook, MySpace, or just using an ordinary search engine. A college would not peek at the stupid Facebook page of an applicant, would it? Absolutely not, says the Brown Dean of Admissions, "unless someone says there's something we should look at."

New participatory websites create even bigger opportunities for information-sharing. If you are about to go on a blind date, there are special sites just for that. Take a look at www.dontdatehimgirl.com, a social networking site with a self-explanatory focus. When we checked, this warning about one man had just been posted, along with his name and photograph: "Compulsive womanizer, liar, internet cheater; pathological liar who can't be trusted as a friend much less a boyfriend. Total creep! Twisted and sick—needs mental help. Keep your daughter away from this guy!" Of course, such information may be worth exactly what we paid for it. There is a similar site, www.platewire.com, for reports about bad drivers. If you are not dating or driving, perhaps you'd like to check out a neighborhood before you move in, or just register a public warning about the obnoxious revelers who live next door to you. If so, www.rottenneighbor.com is the site for you. When we

typed in the zip code in which one of us lives, a nice Google map appeared with a house near ours marked in red. When we clicked on it, we got this report on our neighbor:

> you're a pretty blonde, slim and gorgeous. hey, i'd come on to you if i weren't gay. you probably have the world handed to you like most pretty women. is that why you think that you are too good to pick up after your dog? you know that you are breaking the law as well as being disrespectful of your neighbors. well, i hope that you step in your own dogs poop on your way to work, or on your way to dinner. i hope that the smell of your self importance follows you all day.

For a little money, you can get a lot more information. In January 2006, John Aravosis, creator of `Americablog.com`, purchased the detailed cell phone records of General Wesley Clark. For $89.95, he received a listing of all of Clark's calls for a three-day period. There are dozens of online sources for this kind of information. You might think you'd have to be in the police or the FBI to find out who people are calling on their cell phones, but there are handy services that promise to provide anyone with that kind of information for a modest fee. The *Chicago Sun Times* decided to put those claims to a test, so it paid $110 to `locatecell.com` and asked for a month's worth of cell phone records of one Frank Main, who happened to be one of its own reporters. The *Sun Times* did it all with a few keystrokes—provided the telephone number, the dates, and a credit card number. The request went in on Friday of a long weekend, and on Tuesday morning, a list came back in an email. The list included 78 telephone numbers the reporter had called—sources in law enforcement, people he was writing stories about, and editors in the newspaper. It was a great service for law enforcement—except that criminals can use it too, to find out whom the detectives are calling. These incidents stimulated passage of the Telephone Records and Privacy Act of 2006, but in early 2008, links on `locatecell.com` were still offering to help "find cell phone records in seconds," and more.

If cell phone records are not enough information, consider doing a proper background check. For $175, you can sign up as an "employer" with ChoicePoint and gain access to reporting services including criminal records, credit history, motor vehicle records, educational verification, employment verification, Interpol, sexual offender registries, and warrants searchers—they are all there to be ordered, with *a la carte* pricing. Before we moved from paper to bits, this information was publicly available, but largely inaccessible. Now, all it takes is an Internet connection and a credit card. This is one

PERSONAL COMPUTER MONITORING SOFTWARE

PC Pandora (www.pcpandora.com) enables you to "know everything they do on your PC," such as "using secret email accounts, chatting with unknown friends, accessing secret dating profiles or even your private records." Using it, you can "find out about secret email accounts, chat partners, dating site memberships, and more."

Actual Spy (www.actualspy.com) is a "keylogger which allows you to find out what other users do on your computer in your absence. It is designed for the hidden computer monitoring and the monitoring of the computer activity. Keylogger Actual Spy is capable of catching all keystrokes, capturing the screen, logging the programs being run and closed, monitoring the clipboard contents."

of the most important privacy transformations. Information that was previously available only to professionals with specialized access or a legion of local workers is now available to everyone.

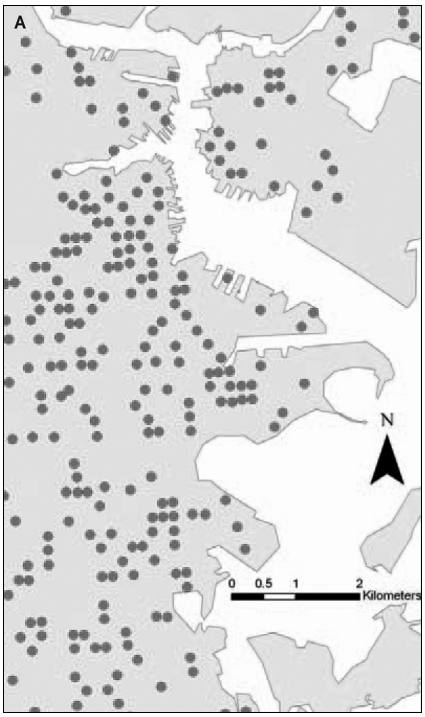
Then there is real spying. Beverly O'Brien suspected her husband was having an affair. If not a physical one, at a minimum she thought he was engaging in inappropriate behavior online. So, she installed some monitoring software. Not hard to do on the family computer, these packages are promoted as "parental control software"—a tool to monitor your child's activities, along with such other uses as employee monitoring, law enforcement, and to "catch a cheating spouse." Beverly installed the software, and discovered that her hapless hubby, Kevin, was chatting away while playing Yahoo! Dominoes. She was an instant spy, a domestic wire-tapper. The marketing materials for her software neglected to tell her that installing spyware that intercepts communications traffic was a direct violation of Florida's Security of Communications Act, and the trial court refused to admit any of the evidence in their divorce proceeding. The legal system worked, but that didn't change the fact that spying has become a relatively commonplace activity, the domain of spouses and employers, jilted lovers, and business competitors.

Idle Curiosity

There is another form of Little Brother-ism, where amateurs can sit at a computer connected to the Internet and just look for something interesting—not about their neighbors or husbands, but about anyone at all. With so much data out there, anyone can discover interesting personal facts, with the

investment of a little time and a little imagination. To take a different kind of example, imagine having your family's medical history re-identified from a paper in an online medical journal.

Figure 2.4 shows a map of the incidence of a disease, let's say syphilis, in a part of Boston. The "syphilis epidemic" in this illustration is actually a simulation. The data was just made up, but maps exactly like this have been common in journals for decades. Because the area depicted is more than 10 square kilometers, there is no way to figure out which house corresponds to a dot, only which neighborhood.



Source: John S. Brownstein, Christopher A. Cassa, Kenneth D. Mandl, No place to hide—reverse identification of patients from published maps, *New England Journal of Medicine*, 355:16, October 19, 2007, 1741-1742.

FIGURE 2.4 Map of part of Boston as from a publication in a medical journal, showing where a disease has occurred. (Simulated data.)

At least that was true in the days when journals were only print documents. Now journals are available online, and authors have to submit their

figures as high-resolution JPEGs. Figure 2.5 shows what happens if you download the published journal article from the journal's web site, blow up a small part of the image, and superimpose it on an easily available map of the corresponding city blocks. For each of the seven disease locations, there is only a single house to which it could correspond. Anyone could figure out where the people with syphilis live.

Source: John S. Brownstein, Christopher A. Cassa, Kenneth D. Mandl, No place to hide—reverse identification of patients from published maps, *New England Journal of Medicine*, 355:16, October 19, 2007, 1741-1742.

FIGURE **2.5** Enlargement of Figure 2.4 superimposed on a housing map of a few blocks of the city, showing that individual households can be identified to online readers, who have access to the high-resolution version of the epidemiology map.

This is a re-identification problem, like the one Latanya Sweeney noted when she showed how to get Governor Weld's medical records. There are things that can be done to solve this one. Perhaps the journal should not use such high-resolution images (although that could cause a loss of crispness, or even visibility—one of the nice things about online journals is that the visually impaired can magnify them, to produce crisp images at a very large scale). Perhaps the data should be "jittered" or "blurred" so what appears on the screen for illustrative purposes is intentionally incorrect in its fine details. There are always specific policy responses to specific re-identification scenarios.

Every scenario is a little different, however, and it is often hard to articulate sensible principles to describe what should be fixed.

In 2001, four MIT students attempted to re-identify Chicago homicide victims for a course project. They had extremely limited resources: no proprietary databases such as the companies that check credit ratings possess, no access to government data, and very limited computing power. Yet they were able to identify nearly 8,000 individuals from a target set of 11,000.

The source of the data was a free download from the Illinois Criminal Justice Authority. The primary reference data source was also free. The Social Security Administration provides a comprehensive death index including name, birth date, Social Security Number, zip code of last residence, date of death, and more. Rather than paying the nominal fee for the data (after all, they were students), these researchers used one of the popular genealogy web sites, RootsWeb.com, as a free source for the Social Security Death Index (SSDI) data. They might also have used municipal birth and death records, which are also publicly available.

The SSDI did not include gender, which was important to completing an accurate match. But more public records came to the rescue. They found a database published by the census bureau that enabled them to infer gender from first names—most people named "Robert" are male, and most named "Susan" are female. That, and some clever data manipulation, was all it took. It is far from clear that it was wrong for any particular part of these data sets to be publicly available, but the combination revealed more than was intended.

The more re-identification problems we see, and the more *ad hoc* solutions we develop, the more we develop a deep-set fear that our problems may never end. These problems arise because there is a great deal of public data, no one piece of which is problematic, but which creates privacy violations in combination. It is the opposite of what we know about salt—that the component elements, sodium and chlorine, are both toxic, but the compound itself is safe. Here we have toxic compounds arising from the clever combination of harmless components. What can possibly be done about *that?*

Big Brother, Abroad and in the U.S.

Big Brother really is watching today, and his job has gotten much easier because of 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*, the cards will 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 case of street protests or any individual activity 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. Chinese tracking is far more detailed and pervasive than Britain's ubiquitous surveillance cameras.

You Pay for the Mike, We'll Just Listen In

Planting tiny microphones where they might pick up conversations of underworld figures used to be risky work for federal authorities. There are much safer alternatives now that many people carry their own radio-equipped microphones with them all the time.

Many cell phones can be reprogrammed remotely so that the microphone is always on and the phone is transmitting, even if you think you have powered it off. The FBI used this technique in 2004 to listen to John Tomero's conversations with other members of his organized crime family. A federal court ruled that this "roving bug," installed after due authorization, constituted a legal from of wiretapping. Tomero could have prevented it by removing the battery, and now some nervous business executives routinely do exactly that.

The microphone in a General Motors car equipped with the OnStar system can also be activated remotely, a feature that can save lives when OnStar operators contact the driver after receiving a crash signal. OnStar warns, "OnStar will cooperate with official court orders regarding criminal investigations from law enforcement and other agencies," and indeed, the FBI has used this method to eavesdrop on conversations held inside cars. In one case, a federal court ruled against this way of collecting evidence—but not on privacy grounds. The roving bug disabled the normal operation of OnStar, and the court simply thought that the FBI had interfered with the vehicle owner's contractual right to chat with the OnStar operators!

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 causes a furious libertarian reaction from parties not usually outspoken in defense of individual freedom. Under the REAL ID act of 2005, uniform federal standards are being implemented for state-issued drivers' licenses. Although it passed through Congress without debate, the law is opposed by at least 18 states. Resistance pushed back the implementation timetable first to 2009, and then, in early 2008, to 2011. 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 U.S., 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 the shape of their earlobes. 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 U.S., 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 55 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 is there 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.'" No REAL ID controversies, goes the theory; all the data being gathered would, in some sense at least, be offered voluntarily.

Friendly Cooperation Between Big Siblings

In fact, there are two Big Brothers, who often work together. And 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 governmental data sources together with private and commercial ones, creating comprehensive digital dossiers on millions of people. With their massive data storage and processing power, they can make connections in the data, like the clever connections the MIT students made with the Chicago homicide data, but 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 our insurance rates, or simply help us be sure that 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—in 2004, this amounted to more than a billion transactions a day. The company uses its massive 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 they gather and analyze 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, simply because of 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. Thirty 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, while 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 or 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 web site would be illegal. Yet 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 congressmen, 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 it 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.

In spite of its 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 inter-agency 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 they can both buy it 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. Yet 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 their survey of 128 federal departments. They 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 twelve 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.

The Transportation Security Administration (TSA) is responsible for airline passenger screening. One proposed system, CAPPS II, which was ultimately terminated over privacy concerns, sought to bring together disparate data sources to determine whether a particular individual might pose a transportation threat. Color-coded assessment tags would determine whether you could board quickly, be subject to further screening, or denied access to air travel.

The government creates projects, the media and civil liberties groups raise serious privacy concerns, the projects are cancelled, and new ones 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.

Technology Change and Lifestyle Change

New technologies enable new kinds of social interactions. There were no suburban shopping malls before private automobiles became cheap and widely used. Thirty years ago, many people getting off an airplane reached for cigarettes; today, they reach for cell phones. As Heraclitus is reported to have said 2,500 years ago, "all is flux"—everything keeps changing. The reach-for-your-cell phone gesture may not last much longer, since airlines are starting to provide onboard cell phone coverage.

The more people use a new technology, the more useful it becomes. (This is called a "network effect"; see Chapter 4, "Needles in the Haystack.") When one of us got the email address `lewis@harvard` as a second-year graduate student, it was a vainglorious joke; all the people he knew who had email addresses were students in the same office with him. Email culture could not develop until a lot of people had email, but there wasn't much point in having email if no one else did.

Technology changes and social changes reinforce each other. Another way of looking at the technological reasons for our privacy loss is to recognize that the social institutions enabled by the technology are now more important than the practical uses for which the technology was originally conceived. Once a lifestyle change catches on, we don't even think about what it depends on.