You Thought Hackers Were Bad? Meet the Data Brokers

(5. kapitola knihy Marc Goodman: Future Crimes: Everything Is Connected, Everyone Is Vulnerable and What We Can Do About It First Edition Edition – vydalo nakladatelství Doubleday v únoru 2015)

Acxiom, Epsilon, Datalogix, RapLeaf, Reed Elsevier, BlueKai, Spokeo, and Flurry—most of us have never heard of these companies, but together they and others are responsible for a rapidly emerging data surveillance industry that is worth $156 billion a year. While citizens around the world reacted with shock at the size and scope of the NSA surveillance operations revealed by Edward Snowden, it's important to note that the $156 billion in annual revenue earned by the data broker industry is twice the size of the U.S. government's intelligence budget. The infrastructure, tools, and techniques employed by these firms rest almost entirely in the private sector, and yet the depth to which they can peer into any citizen's life would make any intelligence agency jealous with envy.

Data brokers get their information from our Internet service providers, credit card issuers, mobile phone companies, banks, credit bureaus, pharmacies, departments of motor vehicles, grocery stores, and increasingly our online activities. All the data we give away on a daily basis for free to our social networks—every Like, poke, and tweet—are tagged, geo-coded, and sorted for resale to advertisers and marketers. Even old-world retailers are realizing that they have a colossal secondary source of income—their customer data—that may be even more valuable than the actual product or service they are selling. As such, companies are rushing to profit from this brand-new revenue stream and transform their data infrastructure from a cost center into a profit center. Though credit bureaus such as Experian, TransUnion, and Equifax have been with us for decades, our increasingly digitally connected online lifestyle enables new firms to capture every drop of data about our lives previously unthinkable and impossible.

Just one company alone, the Acxiom Corporation of Little Rock, Arkansas, operates more than twenty-three thousand computer servers that "are collecting, collating and analyzing" more than 50 trillion unique data transactions every year. Ninety-six percent of American households are represented in its data banks, and Acxiom has amassed profiles on over 700 million consumers worldwide. Each profile contains more than fifteen hundred specific traits per individual, such as race, gender, phone number, type of car driven, education level, number of children, the square footage of his or her home, portfolio size, recent purchases, age, height, weight, marital status, politics, health issues, occupation, and right- or left-handedness, as well as pet ownership and breed.

The goal of Acxiom and other data brokers is to provide what is alternatively called "behavioral targeting," "predictive targeting," or "premium proprietary behavioral insights" on you and your life. In plain English, this means understanding you with extreme precision so that data brokers can sell the information they aggregate at the highest price to advertisers, marketers, and other companies for their decision-making purposes. For example, showing an ad for Pampers to a nineteen-year-old male college student might very well be a waste of an executive's marketing budget, but the same information presented to a thirty-two-year-old pregnant housewife might result in hundreds of dollars of sales. To maximize the value of the digital intelligence they collect, data brokers are forever segmenting us into increasingly specific groupings or profiles. Welcome to the world of dataveillance.

Acxiom sells these consumer profiles to twelve of the top fifteen credit card issuers, seven of the top ten retail banks, eight of the top ten telecommunications companies, and nine of the top ten insurers. To command the billions it charges its advertising customers every year, "Acxiom assigns you a 13-digit code and puts you into one of 70 'clusters' depending on your behavior and demographics." For example, people in cluster 38 "are most likely to be African American or Hispanic, working parents of teenage kids, and lower middle class and shop at discount stores." Someone in cluster 48 is likely to be "Caucasian, high school educated, rural, family oriented, and interested in hunting, fishing and watching NASCAR." These data are also sold to other third-party brokers who apply their own algorithms, further refining the data sets to create category lists of their own such as “Christian families,” “compulsive online gamblers,” “zero mobility,” and “Hispanic Pay Day Loan Responders.”

Those in the Christian family category might receive ads for Bibles and ChristianMingle.com, whereas gamblers and those deemed by an algorithm to have “zero mobility” would be targeted with ads for subprime lenders and debt-consolidation schemes. While being listed as a Christian family or as an urban Hispanic college-educated female might on the surface not appear too troubling, some data brokers have sold much more disturbing lists to advertisers and other parties unknown. For example, some brokers offer lists of seniors with dementia and people living with AIDS, while another firm, MEDbase200, has even auctioned off lists naming both victims of domestic violence and survivors of rape.

The depth and extent of the commercial data collection and surveillance economy were highlighted in early 2014 when a grieving father in Lindenhurst, Illinois, received a sales flyer in the mail from the retailer OfficeMax. Printed on the address label were the words “Mike Seay, Daughter Killed in Car Crash,” followed by the man’s home address. OfficeMax had indeed reached the right guy: Seay’s seventeen-year-old daughter had been killed in a car crash with her boyfriend the year prior. When Seay called OfficeMax to complain about the incident, the manager refused to believe him and dismissed the allegation as “impossible.” It wasn’t until a local NBC reporter in Chicago ran the story that OfficeMax acknowledged the error was “the result of a mailing list rented through a third-party provider.” Eventually, Seay received a phone call from a lower-level OfficeMax executive who apologized for the incident but refused Seay’s repeated requests to name the data broker responsible for the incident. Nor would the executive reveal whether the company held similar data on other prospective customers. Seay’s story is obviously troubling, especially because he was an infrequent OfficeMax shopper who had only on occasion purchased printer paper in the store.

The incident highlights some serious questions about the data broker industry. For example, what other deeply personal data does OfficeMax have on its customers? For the data broker that sold the information in the first place, what else might its massive data banks reveal about you and your family? Brother an alcoholic? Mother diagnosed schizophrenic? Thirteen-year-old daughter with an eating disorder? What regulations exist to limit what data brokers can do with this information, and what can you do if the information they hold on you is incorrect? There are hardly any regulations, as it turns out. It’s reminiscent of the plotline from Franz Kafka’s famous novel The Trial, in which a man is arrested without being informed why and only later learns that a mysterious court has a secret dossier on him, which he cannot access. Today’s modern data brokers, unlike credit-reporting agencies, are almost entirely unregulated by the government. There are no laws, such as the Fair Credit Reporting Act, that require them to safeguard a consumer’s privacy, correct any factual errors, or even reveal what information is contained within their systems on you and your family.

As a result of Seay’s experience and those of thousands more like him, Congress, led by Senator Jay Rockefeller of West Virginia, the Federal Trade Commission, and the Consumer Financial Protection Bureau, have begun to investigate the nature and scope of the multibillion-dollar data broker industry. Any meaningful regulatory changes will be vehemently opposed by the data brokers; there is just too much money to be made. Moreover, once the data is out there, it is virtually impossible to put the proverbial toothpaste back in the tube. In the interim, Acxiom and others continue their stockpiling of your information. In late 2013, Acxiom’s CEO, Scott Howe, proudly announced that his firm had collected nearly 1.1 billion third-party cookies and had identified and profiled the mobile devices of more than 200 million customers. “Our digital reach will soon approach nearly every Internet user in the US,” Howe affirmed.

By mining public databases and aggregating that knowledge with all the personal information people share, wittingly and unwittingly, about themselves, their friends, and their families on social media, companies such as Acxiom have been able to deploy the most comprehensive intelligence surveillance system that has ever existed into the lives of nearly every American alive today. This technological feat represents the “new normal” of our data surveillance society and is part of what the former vice president Al Gore dubbed the “stalker economy” while speaking at the 2013 South by Southwest interactive festival in Austin, Texas.

Gore is right. As should be obvious by now, surveillance is the business model of the Internet. You create “free” accounts on Web sites such as Snapchat, Facebook, Google, LinkedIn, Foursquare, and PatientsLikeMe and download free apps like Angry Birds, Candy Crush Saga, Words with Friends, and Fruit Ninja, and in return you, wittingly or not, agree to allow these companies to track all your moves, aggregate them, correlate them, and sell them to as many people as possible at the highest price, unencumbered by regulation, decency, or ethical limitation. Yet so few stop and ask who else has access to all these data detritus and how it might be used against us. Dataveillance is the “new black,” and its uses, capabilities, and powers are about to mushroom in ways few consumers, governments, or technologists might have imagined.

Analyzing You

Each of us now leaves a trail of digital exhaust throughout our day—an infinite stream of phone records, text messages, browser histories, GPS data, and e-mail that will live on forever. The analysis of this information allows companies to find prospective customers with much higher degrees of accuracy and at greater value than previously possible. For example, let’s say that you’re interested in taking a family vacation to Miami Beach. You search for flights on Kayak. Later you go into a store and buy a bathing suit using your credit card. The data retrieved from the purchase of the swimsuit combined with your browsing data reinforce the likelihood that you are interested in booking a hotel room in Miami. As a result of this behavioral analysis, you now have a quantifiable data value to the hotels in Miami, which can outbid each other for your business, in real time, by presenting advertising that reaches you with highly relevant messages and offers based on your intended behavior.

Google Now, which promises “just the right information at just the right time,” is another example of deep analysis applied to large data sets. The Google Now app provides consumers with wonderfully convenient information that helps them capture and leverage all of the data invisibly swirling around them. Once users agree to Google’s ToS, Google Now will show them when their friends are nearby, provide traffic alerts, determine the quickest travel routes home and to work, automatically furnish the morning’s weather report, and keep track of favorite sports teams and update their scores in real time. Google Now automagically tells you when your flight is delayed and when your gate has changed and offers flight alternatives when available. Because Google Now knows where all your appointments are located and monitors traffic jams along all your intended routes in real time, the app will alert you at your current location, advising you to leave early if you hope to make your next appointment on time. Using a technique known as geo-fencing, Google Now will analyze your to-do list and match it against your persistently tracked location in order to alert you as you drive past the grocery store that you need to buy milk. To enjoy this cornucopia of information and bounty of convenience, you just need to provide Google Now with access to your entire online digital footprint, including your Gmail in-box, Web searches, hotel bookings, flight plans, full contact lists, friends’ birthdays, restaurant reservations, and calendar appointments, as well as your physical location at all times via the GPS on your mobile phone. From this massive data set, Google (and others) can re-create what intelligence analysts call your pattern of life, knowing and mapping your physical location over time as well as what you are doing and with whom. How terribly convenient.

But what else might Google or any other company that had access to your pattern of life be able to determine? Let’s say, for example, your mobile phone was on the nightstand in the same home as your wife’s telephone six nights a week. From these data, it would be logical to conclude that the owners of the two cell phones lived together and were likely sleeping together. But what if one night a week your mobile phone was on a nightstand next to another woman’s mobile phone? What might that suggest to Google or others about your fidelity? An analysis of your locational data and that of the phones (and apps) around you is an excellent approximation of the strengths and bonds of your personal and professional networks. When your data exhaust patterns are studied over time, many more revelations about your life become possible. For example, researchers in the U.K.

studied the past whereabouts of mobile phone users and using basic data analysis techniques were able to determine to within twenty meters of accuracy where a mobile phone user would be at the same time twenty-four hours later, a very useful tool for both advertisers and stalkers. Your phone today knows not only where you've been but also where you are going.

Analysis of your social network and its members can also be highly revealing of your life, politics, and even sexual orientation, as demonstrated in a study carried out at MIT. In an analysis known as Gaydar, researchers studied the Facebook profiles of fifteen hundred students at the university, including those whose profile sexual orientation was either blank or listed as heterosexual. Based on prior research that showed gay men have more friends who are also gay (not surprising), the MIT investigators had a valuable data point to review the friend associations of their fifteen hundred students. As a result, researchers were able to predict with 78 percent accuracy whether or not a student was gay. At least ten individuals who had not previously identified as gay were flagged by the researchers' algorithm and confirmed via in-person interviews with the students. While these findings might not be troubling in liberal Cambridge, Massachusetts, they could prove problematic in the seventy-six countries where homosexuality remains illegal, such as Sudan, Iran, Yemen, Nigeria, and Saudi Arabia, where such an "offense" is punished by death. A study of fifty-eight thousand Facebook users published by the National Academy of Sciences demonstrated that by merely studying their Likes, one could determine intimate details and personality traits with surprising accuracy. The rigorous study carried out in conjunction with the University of Cambridge predicted whether users had a high or low IQ, were emotionally stable, or came from a broken home. The challenge with the data we are leaking is that, as has been shown numerous times, others can pick up our digital bread crumbs and interpret them without our knowledge in ways that can cause us harm.

But I've Got Nothing to Hide

In December 2009, when CNBC's Maria Bartiromo asked Google's own CEO, Eric Schmidt, about privacy concerns resulting from Google's increasing tracking of consumers, Schmidt famously replied, "If you have something that you don't want anybody to know, maybe you shouldn't be doing it in the first place." Schmidt, and others, dismiss privacy concerns by saying that if you haven't done anything wrong, you should not be afraid of people (corporations, governments, or your neighbors) knowing what you are doing.

This sentiment has been echoed by Facebook's CEO, Mark Zuckerberg, who has argued that "privacy is no longer the social norm." While privacy may no longer be the norm—at least for the general public—in his own life, Mr. Zuckerberg seems to treasure privacy quite a bit. In late 2013, it was revealed that the Facebook CEO spent $30 million to buy the four homes surrounding his own property in order to ensure his privacy would remain free from intrusion or disturbance. Facebook's chief operating officer, Sheryl Sandberg, too has suggested that your assertion of any privacy rights is in contrast with "true authenticity." Sandberg notes that "expressing authentic identity will become even more pervasive in the coming years … And yes, this shift to authenticity will take getting used to and it will elicit cries of lost privacy." Convenient for Schmidt, Zuckerberg, and Sandberg that these "naturally occurring shifts" in social norms are tied to their personal and professional bottom lines, which directly benefit from monetizing you and the mountains of information you are leaking to the fullest extent possible, as a result of their highly one-sided ToS.

But "I have nothing to hide" is absolutely the wrong way to think about our new dataveillance society. It is a false dichotomy of choice: either we accept total surveillance, or we are criminals worthy of suspicion. If proponents of the "nothing to hide" argument meant what they said, then they would logically not object to our filming them having sex with their spouses, publishing their tax returns online, and projecting video of their toilet use on the Jumbotron of a crowded stadium, right? After all, they have nothing to hide. The fact is that each of us has private special moments in our lives, made exceptional by limiting with whom we share such intimacies.

For those who believe the fallacy of nothing to hide, perhaps a lesson in something to fear might be appropriate, for all of us have details in our lives we would rather not share. For example, Google Voice, Skype, your mobile phone carrier, and any number of government agencies have records of anyone who has ever phoned an abortion clinic, a suicide hotline, or a local chapter of Alcoholics Anonymous. Data aggregators know who has searched for “slutty cheerleaders,” “Viagra,” or “Prozac” across any of their electronic devices. While all these behaviors may be perfectly legal, no doubt they have repercussions in our society should the information come to light.

Given that Google and Facebook alone have hundreds of petabytes of data on their users stored in perpetuity, perhaps it is more worthwhile to question not what any of us may have to hide today but what we might wish to keep private in the future—and if Facebook existed in 1950, how might history judge an off-color joke today? What future crime might you be convicted of without ever knowing you were in fact violating the law? Did you drive across the border to New Jersey or Delaware to save on taxes when buying back-to-school clothes for the kids? Your cell-phone and credit card receipts document your tax evasion. That photograph on Twitpic of the family dinner showing your twenty-year-old son drinking a glass of wine—evidence of alcohol furnished to a minor. As the computer security researcher Moxie Marlinspike pointed out, there are “27,000 pages of federal statutes” in the United States and another “10,000 administrative regulations. You probably do have something to hide, you just don’t know it yet.”

Privacy Risks and Other Unpleasant Surprises

As Wired’s Mat Honan and the grieving father Mike Seay discovered, our personal data can end up in the hands of those who we assuredly would prefer did not have access to such information. The combination of our social data commingled with public databases, cookies, beacons, and locations can lead to a series of unintended and even harmful consequences. Put another way, your data are increasingly promiscuous. They flow from one system to the next, from database to database, obscured and distributed across cloud-based networks around the world, shared, processed, and sold. But as we have learned from the real world, promiscuity can often lead to social diseases and other unintended consequences.

In an incident not too dissimilar from the OfficeMax debacle, a Minneapolis man learned his daughter was pregnant, not from her, but from his local Target store. The discovery was made when Target began sending the fifteen-year-old girl coupons for items that did not meet her father’s approval. Armed with the coupons and a letter addressed to his daughter, the father furiously marched into Target and began berating the store manager. “My daughter got this in the mail! … She’s still in high school, and you’re sending her coupons for baby clothes and cribs? Are you trying to encourage her to get pregnant?” A few days later, the man phoned the store to apologize, noting, “There have been some activities in my house I haven’t been completely aware of. She’s due in August. I owe you an apology.” But how in the world did Target know the girl was pregnant? Through its pregnancy prediction algorithm of course, which aggregated a customer’s entire purchase history with demographic statistics purchased from data brokers. Target reasoned that if it could find those women before their second trimester of pregnancy and hook them as customers, it would receive the lion’s share of their purchases, not just for baby wipes, cribs, and diapers, but for toys and clothes as the infants aged through adolescence. After an in-depth study by Target’s statisticians, Target noticed that women in the baby registry were “buying larger quantities of unscented lotion at the beginning of their second trimester in addition to vitamin supplements such as calcium, magnesium and zinc.” In total, Target was able to identify twenty-five products that, when analyzed together, allowed it to assign each shopper a “pregnancy prediction score.” When this model ran against the millions of women in Target’s customer databases, thousands and thousands of pregnant women were identified before any other companies had made the connection. Target and the company’s marketers were ecstatic with this discovery. Less enthralled was the father of that fifteen-year-old girl in Minneapolis who would learn of his forthcoming grandchild via a corporate coupon mailer. Given Target’s 2013 hacking, in which the financial data of 110 million of its customers were compromised, what guarantees do consumers have that the vast additional troves of highly personal data in Target’s vaults also won’t be stolen? Can customers trust Target or any other large retailer with the volumes of data it collects, stores, and analyzes? Likely not, and therein lies the problem.

The risks to our personal data come not just from hackers but, as more and more people are finding out, from big-data analytics as well. Previously, many of the data aggregated were held in limbo, as our collection abilities well surpassed our ability to make sense of all that had been collected. That is now changing, and the data we leak on social media sites like Facebook are showing up in unexpected ways. One such person affected was Bobbi Duncan, a twenty-two-year-old lesbian student at the University of Texas at Austin. She came from a strict Christian family and worked hard to keep her sexual orientation from her parents. As she began to understand herself better, she joined a number of student groups on campus, including the Queer Chorus, as a means of meeting other gay and lesbian students at her school. When she joined the organization, the president of the Queer Chorus welcomed Bobbi by adding her to the group's Facebook discussion page, which he was able to do without her permission (there is no setting in Facebook to prevent a third party from adding you to his or her group). When he did so, Facebook sent an automatic system notification to Bobbi's entire list of friends—including her father—notifying them that she had joined the Queer Chorus. Two days after receiving the notification, Bobbi's father wrote a reply on his Facebook page: "To all you queers. Go back to your holes and wait for GOD. Hell awaits you perverts. Good luck singing there." Facebook outed a closeted lesbian and caused her parents to disown her. In response to the irreparable harm she suffered, Bobbi was unequivocal in her stance: "I blame Facebook … It shouldn't be somebody else's choice what people see of me."

When you are the product of Internet and social media companies, the challenge you face is that data you provide in one context can be used in unexpected ways in another, with notable consequences. Such is the case with the highly popular "free" dating site OkCupid. Users seeking dates are asked to fill out questionnaires on the site, and most presume, wrongly, that the data they provide remain exclusively within the OkCupid system, used solely for the purposes of finding a suitable match for a date. Yeah, right! To allegedly get the best matches, OkCupid asks users a bevy of deeply personal questions about their number of past sexual partners, whether they support abortion rights, whether they own a firearm, if they would sleep with somebody on a first date, if they smoke cigarettes, and if they drink alcohol frequently or use illegal drugs (including which drugs and how often). At least that's what users see on their screens when completing their profiles …

What they don't see is the fifty or so companies with whom OkCupid shares this information, including ad firms, data brokers, and marketers. To understand the extent of the data leakage, Ashkan Soltani, a digital privacy specialist who used to work at the Federal Trade Commission, created a dummy account on OkCupid. Using several free privacy browser plug-in tools, including Collusion and mitmproxy, Soltani was able to observe that the answers provided by OkCupid's users were parsed and forwarded to dozens of data brokers in real time. When Soltani completed his test OkCupid profile and clicked that he frequently used drugs, he was able to observe a cookie file that shared his purported drug usage with a data broker known as Lotame. You think you're just filling out a confidential profile for a "free" online dating service; in reality, you've been had and are instead detailing information that you would never otherwise share with any marketing company or data broker. It's a huge ruse: dating is just the "cover story" for massive data extraction. In an ensuing investigation into Soltani's research by NPR, both OkCupid and Lotame declined to comment on the matter. Such is the state of affairs in the world's unregulated data broker industry. Who else might be willing to pay for OkCupid's archive of your drug use and sexual history? An insurance company, prospective employers, or perhaps the government after that DUI incident you had last June?

Even when you have "nothing to hide," your continually tracked social network graph and location can come back to bite you and even affect your financial status. A handful of tech start-ups have begun to use the quality of the friends in your social network to determine whether or not you are a good credit risk. One such company, Lenddo, determines if you are friends with somebody who is behind on paying back her loans and how often you interact with that person. As a result, your creditworthiness can drop because of whom you've friended on Facebook. If your pals on Google+ and Pinterest are deadbeats, chances are you may be too (according to the big-data gods). Facebook may become the next FICO credit-scoring agency as financial data aggregators take full advantage of your social data feeds to rate your financial stability. So as your mom used to warn you, choose your friends wisely.

The fact is that we are all contributing to our own digital pollution. Just as in the twentieth century people thought nothing of pouring industrial waste into a river or tossing garbage onto the street, so too do we fail to comprehend the long-term consequences of our digital actions today. The current state of affairs stems from our fundamental misunderstanding of the bargain we have made for so-called free online services.

Opening Pandora’s Virtual Box

People share their most intimate thoughts and secrets online as if they were having a private conversation with a trusted friend. If only the legal system agreed. In the United States, social networks are considered to be public spaces, not private ones, and any information shared there is covered under the so-called third-party doctrine, which in plain English means that users have no reasonable expectation of privacy in the data their service providers (cell-phone companies, ISPs, cable companies, and Web sites) collect on them.

This noted exception to the Fourth Amendment’s prohibition on unreasonable search and seizure means that any data you post online in any format (regardless of your privacy settings) or any data that are collected by the third parties with whom you have an agreed-upon business relationship are not considered private. Nor does it meet the constitutional standard of “private papers” but rather forms part of the business records of the institution in possession of the data. Shocking though this may be, it is the current state of jurisprudence in the United States, with noted and profound impact on all citizens both online and off. As a result, your data leak to places you would never want them to, and you cannot claw them back, no matter how hard you try.

Accordingly, the word “Facebook” appeared in a full one-third of divorce filings in 2011. All of this provides excellent fodder for the 81 percent of divorce attorneys who admit searching social media sites for evidence that can be used against their clients’ spouses. For instance, all the data shared on Facebook and Twitter and all the cell-phone call records and GPS locational data that neatly recorded whose cell phone was next to whose and when become fair game in the battle royal that can be divorce proceedings. The pictures innocently taken of you at all those parties over the years, blurry-eyed with drink in hand, now become evidence of unfit parenting, a nugget of gold for opposing counsel during cross-examination. That profile you created on OkCupid indicating you were single (which was shared via your browser’s cookies with fifty marketing companies)—perfectly admissible when your wife brings it up during your divorce hearing. When a husband complains that his wife is an inattentive and an unfit mother, he has new powerful evidence to support his claims in the form of subpoenaed records documenting the hundreds of hours she logged on FarmVille and in World of Warcraft, times coinciding with all of her children’s soccer and baseball games that she missed. But the data we’re leaking affect us not only during divorce but in our jobs as well.

A survey conducted by Microsoft on the matter of online reputation found that 70 percent of human resource professionals had rejected a job candidate based on information they had uncovered during an online search. Worse, some employers are now demanding the social media passwords of job applicants and even current employees. Want to work for the Norman, Oklahoma, Police Department, the Maryland Department of Public Safety and Correctional Services, the city of Bozeman, Montana, or the Virginia State Police? Applicants in all of these jurisdictions were required to turn over their Facebook and other social media passwords as part of their so-called routine background checks. This includes providing prospective employers access to all your messages, photographs, and timelines, private and public, on Facebook, Google, Yahoo!, YouTube, and Instagram. While some states, including California, have barred such practices against employees, there is no federal law banning such practices, and it remains legal in 80 percent of American states, and so the data leak.

Increasingly, more and more teachers and school districts are demanding this information from students as well, without warrant of course. That’s what happened to a twelve-year-old Minnesota middle school girl who was accused of posting “inappropriate comments” on her Facebook account.

The student at the Minnewaska Area Middle School had posted that she “hated” a particular school official who was “always mean to her.” The girl was summoned to the principal’s office, where administrators, a school counselor, and a deputy sheriff were waiting for her and demanded that she divulge her Facebook password so that they could review all of her postings. Yes, of course a lawsuit is pending, but the growing number of egregious cases shows your children are leaking data that can come back and bite them as well.

Even college athletes at schools like the University of North Carolina and the University of Oklahoma are being required to provide their passwords on social media sites to their coaches as a condition of playing sports at the schools. Some college athletes have also been compelled to install monitoring software on their personal computers and phones from companies such as UDiligence, which tracks in real time the students’ activities to ensure “that collegiate athletic departments protect against damaging posts made by student-athletes.”

Governments are also getting in on the action. A survey by the International Association of Chiefs of Police of more than five hundred law enforcement organizations revealed that 86.1 percent of police departments now routinely include social media searches as part of their criminal investigations. The IRS too began training its investigators on how to use social networks to investigate taxpayers back in 2009, and Homeland Security’s Citizenship and Immigration Service instructed its agents in 2010 to use social media sites to “observe the daily life of petitioners and beneficiaries suspected of fraud.”

Federal agents can readily access your social data through a variety of means, by serving subpoenas, national security letters, and other administrative orders on your service providers, who under the third-party doctrine exception to the Fourth Amendment needn’t even notify you of the request. For example, AT&T revealed that in 2013 it received more than 300,000 requests for data relating to both civil and criminal cases. The demands for information came from state, federal, and local authorities and included nearly “248,000 subpoenas, nearly 37,000 court orders and more than 16,000 search warrants.” In 2009, Sprint disclosed that it had even created a law-enforcement-only portal that gave police the ability to “ping” (without warrant) any one of Sprint’s mobile phones in order to geo-locate users in real time—a feature that was used more than eight million times by police in a one-year period.

What data of yours the government doesn’t subpoena, it just buys. The NSA and other government agencies didn’t build their global eavesdropping and data-siphoning network from scratch; they purchased or otherwise obtained a complete copy of what the corporate world was already collecting. It makes perfect sense: Why build what they can just buy? ChoicePoint, now owned by Reed Elsevier, maintains seventeen billion records on businesses and individuals that it resells to its 100,000 clients, including to 7,000 federal, state, and local law enforcement agencies. Revelations from Edward Snowden alleged the Central Intelligence Agency pays AT&T $10 million a year for its call data and suggested Verizon too supplies data to the U.S. government. Commercial data brokers have lost no time in offering their paid subscription services to government agents, serving up the streams of information you have freely provided across your social networks.

A brilliant parody on the comedic Onion News Network lampooned the current state of affairs in a fake evening news report:

Congress today reauthorized funding for Facebook, the massive online surveillance program run by the CIA. According to reports, Facebook has replaced almost every other CIA information-gathering program since it was launched in 2004. [A mock CIA official noted,] “After years of secretly monitoring the public, we were astounded so many people would willingly publicize where they live, their religious and political views, an alphabetized list of all their friends, personal e-mail addresses, phone numbers, hundreds of photos of themselves, and even status updates about what they were doing moment to moment. It is truly a dream come true for the CIA. Much of the credit belongs to CIA agent Mark Zuckerberg, who runs the day-to-day Facebook operation for the agency.”

As hilarious and spot-on as the faux news report was, the leakage of our personal information to both shadowy data brokers and government agencies is no joking matter. The cost of the surveillance economy, owing to great advances in technology, is dropping exponentially. Gone is the need for vast teams of special agents to follow you around, tailing you on foot and in vehicles as you traverse a city. Instead, one study has estimated that by using proxy surveillance technologies such as mobile phones, online activity, social data, GPS information, and financial transactions, the government now spends just “$574 per taxpayer, a paltry 6.5 cents an hour,” to track each and every American.

Upon learning the true extent of the NSA’s domestic and international spying prowess, the former head of the East German Stasi Wolfgang Schmidt admitted publicly that such a system “would have been a dream come true.” Schmidt noted that during his reign as head of the much-feared secret police service of the former German Democratic Republic, the Stasi could tap only forty telephones nationwide at a time, but clearly now technology had made it possible to monitor all calls and Internet data at all times. He cautioned, “It is the height of naivete to think that once collected this information won’t be used … This is the nature of secret government organizations. The only way to protect the people’s privacy is not to allow the government to collect their information in the first place.”

Knowledge Is Power, Code Is King, and Orwell Was Right

In George Orwell’s dystopian novel 1984, he depicted an omnipotent government surveillance state controlled by a privileged few elite who persecuted independent thinking as “thought crimes.” Though Orwell clearly would have foreseen the NSA debacle, it’s less clear he might have predicted Acxiom, Facebook, and Google. To that point, in those cases it wasn’t Big Brother government that “did something to us,” but rather we who did something to ourselves. We allowed ourselves to become monetized and productized on the cheap, giving away billions of dollars of our personal data to new classes of elite who saw an opportunity and seized it. We accepted all their one-sided ToS without ever reading them, and they maximized their profits, unencumbered by regulation or oversight. To be sure, we got some pretty cool products out of the deal, and Angry Birds is really fun. But now that we’ve given all these data away, we find ourselves at the mercy of powerful data behemoths with near-government-level power who do as they please with our information and our lives.

In his 1999 book Code and Other Laws of Cyberspace, the Harvard Law School professor Lawrence Lessig insightfully demonstrated that the instructions encoded in any software program, app, or platform shape and constrain the Internet, just as laws and regulations do. Thus, when Facebook or Google unilaterally changes its terms of service to allow your news feeds to become public or your photographs to be used in advertisements against your will, it is as if a new “law” has been passed. Code, is in effect, law.

Perhaps then the only way to opt out of such a system would be to close one’s account or never create one in the first place? Unfortunately, both approaches are problematic and increasingly impossible. A New York Times article previously noted that Facebook keeps all your data even after you’ve closed your account. Even if you chose not to participate in an online social network, your friends would continue to tag you in pictures, the GPS in your car would still track your location, and Target would track all of your purchases.

The unprecedented volumes of data about ourselves that we have entrusted to private companies are up for grabs, and once the genie is out of the bottle, there’s no putting it back in. The troika of opportunity created by our online data exhaust, ridiculous terms of service, and little or no regulation means that modern data brokers can surveil us with better-than-government-grade surveillance capabilities, capturing our every thought, photograph, and location and subjecting them to big-data analytics. As Mat Honan, Bilal Ahmed, Mike Seay, Bobbi Duncan, Leigh Van Bryan, and Emily Bunting all learned firsthand, there are social costs and risks associated with our continued data leakage. But privacy implications are just one of the great threats resulting from the exponential growth in data.

Hackers are hard at work stealing all of the social data you have dutifully reported on yourself and are successfully breaking into the computers of data brokers and Internet giants responsible for storing it all. As Sony, Target, and even the Department of Defense have learned, data stored in insecure information systems are data waiting to be taken. As such, all data gathered will eventually leak, with potent implications for our personal and professional lives and even for our safety and security.

The problem with our being the product as opposed to the customer of massive data brokers is that we are not in control of our data and thus not in control of our destiny. The continued aggregation of this information, unregulated and insecure, sits as a ticking time bomb, with our every thought and deed available for the picking by a new and emerging class of bad actors whose intents are far worse than selling us discounted diapers and adjusting our insurance rates. International organized crime groups, rogue governments, and even terrorists are rapidly establishing their own data brokerages and bolstering their analytic capabilities in order to take full advantage of the single largest bonanza that has ever come their way, with frightening implications for us all.