HAL ABELSON • KEN LEDEEN
HARRY LEWIS • WENDY SELTZER

BLOWN to BITS

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

SECOND EDITION

Blown To Bits

Your Life, Liberty, and Happiness After the Digital Explosion

Second Edition

Hal Abelson
Ken Ledeen
Harry Lewis
Wendy Seltzer

Pearson

CHAPTER 3

Who Owns Your Privacy?

The Commercialization of Personal Data

What Kind of Vegetable Are You?

It didn't raise eyebrows when Aleksandr Kogan offered "This is Your Digital Life" as a Facebook quiz app. Quiz apps are a staple of Facebook marketing, enticing users to participate and then harvesting marketing data. These apps—which are enticing, seductive, and highly effective—have spawned an entire subindustry of quiz-marketing tools and specialists.

About 270,000 Facebook users installed Kogan's app and took its personality test, in the process giving the app access to their contacts to invite them to follow suit. Kogan's ostensible motivation was academic research—studying how emojis are used to convey emotion. But what he did with all the data he collected was quite different. Through Kogan's app, the firm Cambridge Analytica harvested data on more than 50 million people. Cambridge Analytica used that information to help presidential candidate Donald Trump's campaign target audiences for digital advertising and fundraising, model voter turnout, identify markets to air television ads, and even plan Trump's travel. Cambridge Analytica asserted that its "psychographic profiles" helped to identify likely voters and the kinds of messages that would sway them to vote Trump.[1]

But how did a quarter million people downloading an app turn into data spillage from 50 million? Through the porous privacy model of Facebook apps. Each of the 270,000 users who installed the app was connected to an average of 200 friends. "This is Your Digital Life" based its assessment not so much on the quiz as on the history of pages "liked." The quiz was a pretext to

obtain access to users' likes and those of their contacts. Facebook permitted that data shoveling in 2015—although it says Kogan violated the program's terms by sharing profile data with Cambridge Analytica.

Your privacy is not your own. Even if you rejected "This is Your Digital Life," any of your friends—or the apps they installed—could have compromised *your* data. This has parallels in the non-digital world as well, of course. (Consider the old saying "Two people can keep a secret if one of them is dead.") But offline, you may have better intuitions about it. You know not to share a story with the gossipy neighbor until you're ready to be asked questions by strangers in the supermarket. Online, it took a long time for Facebook's privacy settings to gain simple audience controls, and not until after the Cambridge Analytica scandal did the social network stop allowing apps to traverse the social graph, slurping up the network of friend connections.

Leave Me Alone

More than a century ago, two lawyers raised the alarm about the impact technology and the media were having on personal privacy:

> Instantaneous photographs and newspaper enterprise have invaded the sacred precincts of private and domestic life; and numerous mechanical devices threaten to make good the prediction that "what is whispered in the closet shall be proclaimed from the house-tops."

This statement is from the seminal *Harvard Law Review* article on privacy published in 1890 by Boston attorney Samuel Warren and his law partner, Louis Brandeis, later to be a justice of the U.S. Supreme Court (where, as we saw, he dissented in defense of privacy in *Olmstead v. U.S.*).[2] Warren and Brandeis went on to say,

> Gossip is no longer the resource of the idle and of the vicious, but has become a trade, which is pursued with industry as well as effrontery. To satisfy a prurient taste the details of sexual relations are spread broadcast in the columns of the daily papers. To occupy the indolent, column upon column is filled with idle gossip, which can only be procured by intrusion upon the domestic circle.

New technologies made this garbage easy to produce, and then the supply created the demand. And those candid photographs and gossip columns were not merely tasteless; they were bad. Sounding like modern critics of mindless

reality TV, Warren and Brandeis raged that society was going to hell in a handbasket because of all that stuff that was being spread about:

> Even gossip apparently harmless, when widely and persistently circulated, is potent for evil. It both belittles and perverts. It belittles by inverting the relative importance of things, thus dwarfing the thoughts and aspirations of a people. When personal gossip attains the dignity of print, and crowds the space available for matters of real interest to the community, what wonder that the ignorant and thoughtless mistake its relative importance. Easy of comprehension, appealing to that weak side of human nature which is never wholly cast down by the misfortunes and frailties of our neighbors, no one can be surprised that it usurps the place of interest in brains capable of other things. Triviality destroys at once robustness of thought and delicacy of feeling. No enthusiasm can flourish, no generous impulse can survive under its blighting influence.

The problem Warren and Brandeis perceived was that it was hard to say just why such invasions of privacy should be unlawful. In individual cases, you could say something sensible, but the individual legal decisions were not part of a general regime. The courts had certainly applied legal sanctions for defamation—publishing malicious gossip that was false—but then what about malicious gossip that was true? Other courts had imposed penalties for publishing an individual's private letters—but on the basis of property law, just as though the individual's horse had been stolen rather than the words in his letters. That did not seem to be the right analogy either. No, they concluded, such rationales didn't get to the nub. When something private is published about you, something has been taken from you, you are a victim of theft—but the thing stolen from you is part of your identity as a person. In fact, privacy was a right, they said, a "general right of the individual to be let alone." That right had long been in the background of court decisions, but the new technologies had brought this matter to a head. In articulating this new right, Warren and Brandeis were, they asserted, grounding it in the principle of "inviolate personhood," the sanctity of individual identity.

Privacy and Freedom

The Warren–Brandeis articulation of privacy as a right to be left alone was influential, but it was never really complete. Throughout the twentieth century, there were simply too many good reasons for *not* leaving people alone, and there were too many ways in which people *preferred* not to be left alone. And in the United States, First Amendment rights stood in tension with privacy rights. As a general rule, the government cannot stop me from saying

anything truthful. In particular, it usually cannot stop me from saying what I lawfully discover about your private affairs. Yet the Warren–Brandeis definition worked well enough for a long time because, as Robert Fano put it, "The pace of technological progress was for a long time sufficiently slow as to enable society to learn pragmatically how to exploit new technology and prevent its abuse, with society maintaining its equilibrium most of the time."[3] By the late 1950s, the emerging electronic technologies, both computers and communication, had destroyed that balance. Society could no longer adjust pragmatically because surveillance technologies were developing too quickly.

The result was a landmark study of privacy by the Association of the Bar of the City of New York, which culminated in the publication, in 1967, of a book by Alan Westin, titled *Privacy and Freedom*.[4] (Fano was reviewing Westin's book when he painted the picture of social disequilibrium caused by rapid technological change.) Westin proposed a crucial shift of focus.

Brandeis and Warren had seen a loss of privacy as a form of personal injury, which might be so severe as to cause "mental pain and distress, far greater than could be inflicted by mere bodily injury." Individuals had to take responsibility for protecting themselves. "Each man is responsible for his own acts and omissions only." But the law had to provide the weapons with which to resist invasions of privacy.

Westin recognized that the Brandeis–Warren formulation was too absolute, in the face of the speech rights of other individuals and society's legitimate data-gathering practices. Protection might come not from protective shields but from control over the uses to which personal information could be put. "Privacy," wrote Westin, "is the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others." Westin proposed:

> ...what is needed is a structured and rational weighing process, with definite criteria that public and private authorities can apply in comparing the claim for disclosure or surveillance through new devices with the claim to privacy. The following are suggested as the basic steps of such a process: measuring the seriousness of the need to conduct surveillance; deciding whether there are alternative methods to meet the need; deciding what degree of reliability will be required of the surveillance instrument; determining whether true consent to surveillance has been given; and measuring the capacity for limitation and control of the surveillance if it is allowed.[5]

So even if there were a legitimate reason why the government, or some other party, might know something about you, your right to privacy might limit what the knowing party could do with that information.

This more nuanced understanding of privacy emerged from the important social roles that privacy plays. Privacy is not, as Warren and Brandeis had it, the right to be isolated from society; privacy is a right that makes society work.

Fano mentioned three social roles of privacy. First, "the right to maintain the privacy of one's personality can be regarded as part of the right of self-preservation"—the right to keep your adolescent misjudgments and personal conflicts to yourself, as long as they are of no lasting significance to your ultimate position in society. Second, privacy is the way society allows deviations from prevailing social norms, given that no one set of social norms is universally and permanently satisfactory—and, indeed, given that social progress requires social experimentation. And third, privacy is essential to the development of independent thought; it enables some decoupling of the individual from society so that thoughts can be shared in limited circles and rehearsed before public exposure.

Philosopher Helen Nissenbaum similarly grounds privacy in social being, describing privacy as "contextual integrity."[6] Privacy depends on a match between data flows and the expectations and norms of the setting in which information was generated and shared. When Facebook invites you to friend your therapist or a fellow patient, that's a context violation. Online spaces offer the opportunity to multiply contexts: You can be one persona on your Instagram feed and another in the classroom. But online spaces also threaten context collapse, as Stacy Snyder found way back in the days of Myspace, when her photograph captioned "drunken pirate" on what she thought was a merely social post cost her a teaching degree.[7]

Privacy is the way society allows deviations from prevailing social norms, given that social progress requires social experimentation.

The explosive growth in digital technologies has radically altered our expectations about what will be private and shifted our thinking about what *should* be private. It has made privacy violations easier and potentially more numerous. Indeed, it is remarkable that we no longer blink at intrusions that a decade ago would have seemed shocking. Unlike with the story of secrecy, there was no single technological event that caused the change, no privacy-shattering breakthrough—only a steady advance on several technological fronts that ultimately passed a tipping point.

Sensor devices got cheaper, better, and smaller. Tiny cameras, GPS units, and microphones have gone from the stuff of spy museums to the banality of everyday carry. Once they became useful consumer goods, we seemingly worried less about their uses as surveillance devices. Instead of trying to come up with a unifying theory of privacy and its value, we find ourselves piecing together privacy from feelings of discomfort and regret amid the abundance. It's that much harder when we're the ones bringing spies into our own homes and those of our friends, when we trade privacy against conviviality and convenience.

Smile While We Snap!

Big Brother had his legions of cameras, and the City of London has theirs today. But for sheer photographic pervasiveness, nothing beats the cameras in the cell phones in the hands of everyday people. Flying out before the Fourth of July, Helen was asked to switch seats with another woman who wanted to be seated with her boyfriend. She took her seat a row up and struck up a conversation with her new seatmate, unaware that the row behind was filming them as romance. The pair she had helped were tweeting the flight, hashtagged #PlaneBae, and the story soon made the rounds of television morning shows. Innocent fun, it might seem, but not for Helen, who stated (through lawyers),

> Without my knowledge or consent, other passengers photographed me and recorded my conversation with a seatmate. They posted images and recordings to social media, and speculated unfairly about my private conduct.
>
> Since then, my personal information has been widely distributed online. Strangers publicly discussed my private life based on patently false information.
>
> I have been doxxed, shamed, insulted and harassed. Voyeurs have come looking for me online and in the real world.[8]

The massive dissemination of cheap cameras coupled with universal access to the Web enables a kind of vigilante justice—a ubiquitous Little-Brotherism, in which we can all be detectives, judges, and corrections officers. Bloggers can bring global attention to ordinary citizens.

For every lens aimed deliberately, there are also scores more watching unattended: public and private observation and surveillance. Main Street is lined with security cameras peeping from store windows and police surveillance cameras, some of which even offer public viewing. Leafy Lane may be watching, too, thanks to networks of Ring doorbells and vigilant neighbors in Nextdoor groups. Coupled with automated facial recognition, the wired streets could be building dossiers on us all.

Looking at images on the Web is now a leisure activity that anyone can do at any time, anywhere in the world. Using Google Street View, you can sit in a café in Tajikistan and identify a car that was parked in my driveway when Google's camera came by (perhaps months ago). From Seoul, you can see what's happening right now, updated every few seconds, in Piccadilly Circus or on the strip in Las Vegas. These views were always available to the public, but cameras plus the Web change the meaning of "public."

Some of the intrusions into our privacy come because of the unexpected, unseen side effects of things we do quite voluntarily. While the Fourth Amendment protects us from overreach of government surveillance, there is only patchwork legal consideration of private information gathering in the United States. Companies routinely gather and infer information about individuals and use it to customize product offerings and advertisements. As the saying goes, if you're not paying, you're the product.

Footprints and Fingerprints

As we do our daily business and lead our private lives, we leave footprints and fingerprints. We can see our footprints in mud on the floor and in the sand and snow outdoors. We would not be surprised that anyone who went to the trouble to match our shoes to our footprints could determine, or guess, where we had been. Fingerprints are different. It doesn't even occur to us that we are leaving them as we open doors and drink out of tumblers. Those who have guilty consciences may think about fingerprints and worry about where they are leaving them, but the rest of us don't.

THE UNWANTED GAZE

The Unwanted Gaze by Jeffrey Rosen (Vintage, 2000) details many ways in which the legal system has contributed to our loss of privacy.

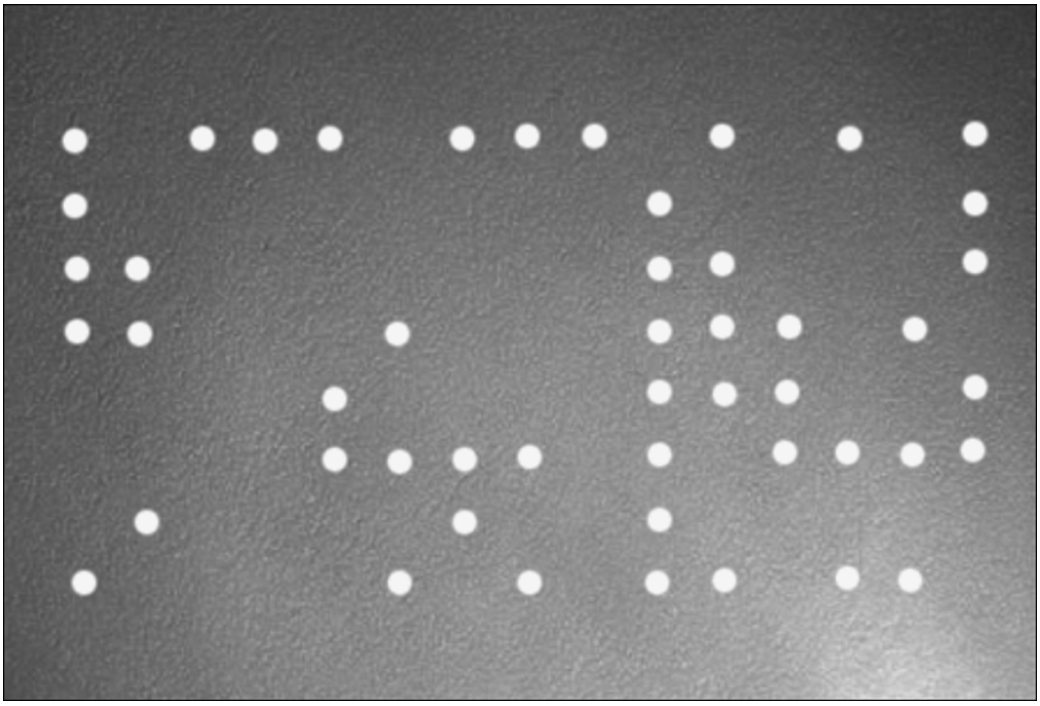
In the digital world, we all leave both electronic footprints and electronic fingerprints—data trails we leave intentionally, and data trails of which we are unaware or unconscious. The identifying data may be useful for forensic purposes. Because most of us don't consider ourselves criminals, however, we tend not to worry about that. What we don't think about is that the various small smudges we leave on the digital landscape may be useful to someone else—someone who wants to use the data we left behind to make money or to get something from us. It is therefore important to understand how and where we leave these digital footprints and fingerprints.

Tracing Paper

If I send an email or download a web page, it should come as no surprise that I've left some digital footprints. After all, the bits have to get to me, so some part of the system knows where I am. In the old days, if I wanted to be anonymous, I could write a note, but my handwriting might be recognizable, and I might leave fingerprints (the oily kind) on the paper. I might have typed, but Perry Mason regularly solved crimes by matching a typewritten note with the unique signature of the suspect's typewriter. More fingerprints.

So, today I would laser print the letter and wear gloves. But even that might not suffice to disguise me. Researchers at Purdue have developed techniques for matching laser-printed output to a particular printer.[9] They analyze printed sheets and detect unique characteristics of each manufacturer and each individual printer—fingerprints that can be used, like the smudges of old typewriter hammers, to match output with source. It may be unnecessary to put the microscope on individual letters to identify what printer produced a page.

The Electronic Frontier Foundation has demonstrated that many color printers nearly invisibly encode the printer serial number, date, and time on every page they print (see Figure 3.1). Therefore, when you print a report, you should not assume that no one can tell who printed it.



Source: Electronic Frontier Foundation, http://w2.eff.org/Privacy/printers/docucolor/

FIGURE 3.1 Fingerprint left by a Xerox DocuColor 12 color laser printer. The dots are very hard to see with the naked eye; the photograph was taken under blue light. The dot pattern encodes the date (2005-05-21), the time (12:50), and the serial number of the printer (21052857).

There was a sensible rationale behind this technology. The government wanted to make sure that office printers could not be used to turn out sets of hundred-dollar bills. The technology that was intended to frustrate counterfeiters makes it possible to trace every page printed on color laser printers back to the source. Useful technologies often have unintended consequences.

Many people, for perfectly legal and valid reasons, would like to protect their anonymity. They might be whistleblowers or dissidents. Perhaps they are merely railing against injustice in their workplace. Will technologies that undermine anonymity in political discourse also stifle free expression? A measure of anonymity is essential in a healthy democracy—and in the United States, anonymity has been a weapon used to advance free speech since the time of the Revolution. We may regret a complete abandonment of anonymity in favor of communication technologies that leave fingerprints.

The problem is not just the existence of fingerprints but that no one told us that we are creating them.

The problem is not just the existence of fingerprints but that no one told us that we are creating them.

When NSA contractor Reality Winner leaked classified information to The Intercept, she might have thought that sending a paper copy would thwart attempts to trace the leaks.[10] The Intercept had shared the document with NSA to verify its authenticity, and Winner was arrested a few days later. Initial reports speculated that she was traced through printer microdots, but the truth appears to have been even more mundane: NSA logs showed that only six accounts, including Winner's, had accessed the document, and Winner had used a personal account to contact The Intercept shortly beforehand.[11]

Advertising

If you ride the T in Boston, you'll see lots of advertisements for college and graduate programs. They all have phone numbers and URLs, and many direct you places like college.edu/recruiting/redline. That web address isn't saying the college has a special program on the Red Line, but it does have a special *advertising* program there. The "redline" at the end of the URL lets the college know that you were referred there by its subway ad. It might use that to direct you to the particular programs advertised on the poster and to track the effectiveness of this ad campaign.

Ads on the Web use the referring page as just one of many signifiers; others are less visible than the URL decoration visible on the subway poster. When you follow a link to open a web page in your browser, that click kicks off a series of events that starts with an electronic request for the web page and a request for any cookies the site may have set previously. All but the simplest of pages will then trigger requests for more subresources: images, fonts, scripts to make the page dynamic. A commercial site may have dozens of advertisements and tracking pixels, or "web bugs"—invisible elements that make your computer call out to yet another source for the purpose of tracking your activity.

How Sites Know Who You Are (An Incomplete List)

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 website wants to have available during your current session (about your shopping cart, for example) or from one session to the next. Cookies give sites persistent information for tracking and personalization. Your browser has a command for showing cookies; if you use it, you may be surprised how many websites 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. That address may change from one day to the next. But in a residential setting, your Internet service provider (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.
4. **You look like someone they already recognize.** Users who log in to Facebook often share a lot of detail about their lives and networks: friends and family connections, favorite bands and restaurants, political leanings—and that's just things they deliberately connect or "like." Facebook also creates shadow audiences, matching people on whom they have little information with others they already know, who share these characteristics.
5. **They've fingerprinted your browser and linked it to profiles from previous visits.** Websites can access lots of seemingly innocent details about your browser (which type, version, graphics encoding, language, and much more). These tend to remain fairly static, and often will uniquely identify a particular browser instance. This technique is simple, and remarkably accurate and effective.

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, IP address information won't reveal who is sending you spam, since spammers routinely forge the source of email they send you. In addition,

between the time you request a web page and its ads are displayed in your browser, there's often a real-time auction, in which your eyeballs (or at least the ad spaces in the web page your browser is about to display) are sold to the highest bidder. Ad networks collect the information from tracking pixels and page context to determine what ads to offer and how much to bid to place them in these auctions.

Why are these shoes following me around? Maybe you saw them on Instagram, tagged them on Pinterest, or searched for a new pair of sneakers on your favorite retailer's website. Maybe you even put them into a shopping cart before deciding they weren't in your budget at this time. Now, you can't seem to escape the shoes: whether you're reading the news or Facebooking with friends, there are the shoes, stalking you from the ad banners, urging you to click "buy."

Known in the trade as "retargeting," these ads are some of the products of real-time bidding. The marketer who dropped a tracking cookie in your browser during an earlier browsing session or the shopping visit you cut short is using it to identify you as a shoe-interested shopper and bidding to show you those ads in the hopes of luring you back to purchase. If you clicked through any of the ads, the marketer would register a "conversion" and factor this data further into your profile for future ad opportunities.

Web browsing users haven't taken all of this sitting quietly. The *Economist* calls data "the new oil," and browsers who are unwilling to be seen as gushers download ad blockers. As of early 2020, all of the major web browsers have incorporated tracker-blocking features or announced plans to limit third-party cookies.

Arvind Narayanan and his team at Princeton University have set up a laboratory for web measurement[12] and discovered new techniques for browser tracking. Through web "crawls," they find tracking techniques used in the wild to identify users and reidentify those who think they've cleared all previous interactions. One of the paradoxes of privacy on the Web is that browsers can be fingerprinted by their unique features, including features the user might enable with the goal of securing greater privacy. That means turning on such protections can make the privacy-seeking user stand out. In such cases, privacy depends on the actions of many to provide a crowd in which the privacy-seeking browser can blend. Standardized processes and well-thought-out default settings are necessary to preserve the opportunities for privacy.

Target Knows You're Pregnant

In 2012, as Charles Duhigg reported in the *New York Times*,[13] a man walked into a Minneapolis-area Target store, furiously asking to speak with the manager: "My daughter got this in the mail!" he said. "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?"

The store manager apologized to the Minneapolis man for their apparent mistake, but he returned a few weeks later with an apology of his own: His daughter was, in fact, pregnant. The store's predictive models had recognized the young woman's pregnancy even before her father had. Target's models didn't have access to her private information. They had the power of analytical tools and readily available data.

Like many other stores with loyalty cards or user accounts, Target built statistical models of shopper behavior to predict hot products for inventory and pricing and to make recommendations. Target correlated shopper purchase history based on an internal guest ID and purchased external data to supplement its logs. From those records, the company's statistician could derive patterns, noticing, for instance, that women in the second trimester of pregnancy would often purchase unscented moisturizing lotions and supplements. After watching this pattern play out many times, the store could anticipate future purchases of baby clothes and diapers from the earlier unscented lotion—and advertise to the mother-to-be at a time when her shopping habits were in flux—responding to a signal she didn't even know she was sending.

How can we solve a privacy problem that results from many developments, no one of which is really a problem in itself?

You Pay for the Mic, 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 most people carry their own radio-equipped microphones with them all the time or invite Alexa, Siri, Cortana, or Google into their homes.

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 form 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!

Danielle, an Amazon Echo customer in Portland, Oregon, was alarmed by a call from one of her husband's colleagues, who said, "Unplug your Alexa devices right now; you're being hacked."[14] The gadget, which was supposed to record only when triggered by the wake word "Alexa," must have heard both that and a "send message" command in Danielle's conversation. Her chat about hardwood floors turned into a voice message to a business acquaintance. A freak occurrence, perhaps, but one that may be repeated as we invite tiny networked recorders into more corners of our lives. German authorities banned "My Friend Cayla,"[15] a talking doll, over concerns about its spying and data-collecting abilities. To engage in conversation with children, Cayla uploaded the sounds she heard over the Internet. German parents were told to destroy the "illegal espionage apparatus." Meanwhile, here in the United States, your smart TV may be watching your viewing habits to tailor advertising. Vizio's CTO told the Consumer Electronics Show that TVs would cost more if it weren't for this revenue stream.[16]

Venmo: It All Adds Up

Earlier we discussed the tracking that credit cards enable in credit reporting bureaus and data analysis firms. Newer payment technologies bring the reporting directly to you. Venmo lets you send someone money or split a bill by entering the person's phone number. It's so easy that as you send money to friends or roommates using the Venmo app, you might not notice that these payment transactions are public, including any memo you write along with the payment. A researcher who found the feed correlated just a few of the threads among millions of transactions into "Venmo stories":[17] a student's fast food habit, a cannabis vendor's sales, a budding relationship? You might not mind sharing your passion for elote (seasoned corn) but might feel differently about recreational marijuana purchases, even in states where those are legal. The researcher, Hang Do Thi Duc, anonymized the details but notes that the feed, which includes everything except dollar values, remained accessible to any visitor to Venmo's public API. (Every page of the site Duc developed, publicbydefault.fyi, encourages Venmo users to change their privacy settings from the default to make transactions private between sender and recipient.)

DNA: The Ultimate Digital Fingerprint

In April 2018, the state of California arraigned Joseph James DeAngelo on a series of decades-old murder and rape charges. The Golden State Killer had been a cold case until an investigator uploaded DNA from a crime scene to a

public genealogy website, GEDmatch. The investigator created a fake profile for the unknown person whose recovered DNA he uploaded. After GEDmatch compared this person's DNA against its existing database to identify partial genetic matches, it showed profiles of people who were likely distant relatives of the suspected killer. Those names led to family trees and to genealogy that could be traced further through census records, obituaries, gravesites, and commercial and law enforcement databases. After these searches put a name to their suspect, investigators confirmed their suspicions by tracking him down and obtaining another DNA sample, from skin cells he left on the car door when he parked in a Hobby Lobby parking lot. That DNA matched the original crime scene samples.[18]

DeAngelo had not posted to the ancestry site, but because a parent passes roughly half of his or her genes to a child (notwithstanding a few mutations along the way), much of DeAngelo's genetic record could be read or revealed by relatives. If your family members explore their genetic profiles and family trees on GEDmatch, they are also exposing information about traits you might share. Your privacy can be invaded through no actions of your own. While the Genetic Information Nondiscrimination Act prohibits employers or health insurers from discriminating based on DNA, the law doesn't restrict numerous other ways DNA can be used.

The Golden State Killer case started a boom in DNA forensic genealogy. By the end of 2018, more than a dozen violent criminals and perpetrators of sexual assault had been identified through GEDmatch. But the site also heard privacy alarm and changed its terms of service to prohibit law enforcement matching of DNA profiles unless users opted in for their own records.

Fair Information Practice Principles

An earlier information revolution, set in rooms full of disk drives that sprouted in government and corporate buildings in the 1960s, set off a round of soul searching about the operational significance of privacy rights. What, in practice, should those holding a big data bank think about when collecting the data, handling it, and giving it to others?

In 1973, the Department of Health, Education, and Welfare issued "Fair Information Practice Principles" (FIPP), as follows:

> **Openness.** There must be no personal data record-keeping systems whose very existence is secret.
>
> **Disclosure.** There must be a way for a person to find out what information about the person is in a record and how it is used.

Secondary use. There must be a way for a person to prevent information about the person that was obtained for one purpose from being used or made available for other purposes without the person's consent.

Correction. There must be a way for a person to correct or amend a record of identifiable information about the person.

Security. Any organization creating, maintaining, using, or disseminating records of identifiable personal data must assure the reliability of the data for its intended use and must take precautions to prevent misuses of the data.

These principles were proposed for U.S. medical data but were never adopted. Nevertheless, they have been the foundation for many corporate privacy policies. Variations on these principles were codified in international trade agreements by the Organisation for Economic Co-operation and Development (OECD) in 1980 and within the European Union (EU) in 1995. In the United States, echoes of these principles can be found in some state laws, but federal laws generally treat privacy on a case-by-case, or "sectorial," basis. The 1974 Privacy Act applies to interagency data transfers within the federal government but places no limitations on data handling in the private sector. The Fair Credit Reporting Act applies only to consumer credit data but does not apply to medical data. The Video Privacy Act applies only to videotape rentals but not to on-demand movie downloads, which did not exist when the act was passed. Finally, few federal or state laws apply to the huge data banks in the file cabinets and computer systems of cities and towns. American government is decentralized, and authority over government data is decentralized as well.

The United States is not lacking in privacy laws. But privacy has been legislated inconsistently and confusingly and in terms dependent on technological contingencies. There is no national consensus on what should be protected and how protections should be enforced. Without a more deeply informed collective judgment on the benefits and costs of privacy, the current legislative hodgepodge may well get worse in the United States.

The discrepancy between American and European data privacy standards threatened U.S. involvement in international trade because an EU directive would prohibit data transfers to nations, such as the United States, that do not meet the European "adequacy" standard for privacy protection. In 2000 the European Commission created a "safe harbor" for American businesses with multinational operations, but the European Court of Justice declared it inadequate to protect the rights of European data subjects. In 2016, the FTC developed an alternative, Privacy Shield, with a salient enforcement difference: "While joining the Privacy Shield Framework will be voluntary, once an eligible company makes the public commitment to comply with the Framework's requirements, the commitment will become enforceable under U.S. law."[19]

In 2020, The Court of Justice of the European Union (CJEU) ruled that even Privacy Shield was inadequate, because European citizens' data in the United States would be subject to U.S. government surveillance.[20]

Privacy as a Basic Right

Browse the Web on a visit to Europe, and you may notice a profusion of pop-ups and banners. Every site, it seems, wants you to consent to the use of cookies and the "processing of your data," assertedly to improve your browsing experience. While European law takes a stronger view of personal privacy as a fundamental right, European advertisers are just as eager to gather personal data as those in the United States. These banners are the means of asking for "consent to data processing," as the E-Privacy Directive required.

In 2018, the General Data Protection Regulation (GDPR) established specific individual rights in personal data and obliged businesses to give individuals ("data subjects") the ability to control the use of that data. Those who collect or process personal data must be able to justify the privacy intrusion based on consent or another "legitimate purpose"; for example, an email provider needs the email addresses of your contacts in order to send emails to their destination, but it doesn't need their home addresses. Individuals even have the right to withdraw consent, demanding that providers erase the data collected about them. Because the GDPR asserts extraterritorial reach, applying to European citizens wherever they are physically located, many providers outside Europe have also adopted cookie-consent requests and adapted their data handling to be able to respond to data deletion requests.

Despite the paper promise of European law, as of 2020, enforcement has been limited. Only one major fine has been issued, against Google for 50 million euros (roughly $54 million), or about one-tenth of what Google generates in a single day's ad sales. Without investigation of the hundreds of complaints raised by citizens to their national data protection authorities, it is difficult to say whether Europeans have more privacy online or just more pop-ups to click through.

It is, unfortunately, too easy to debate whether the European omnibus approach is more principled than the U.S. piecemeal approach, when the real question is whether either approach accomplishes what we want it to achieve. The Privacy Act of 1974 assured us that obscure statements would be buried deep in the Federal Register, providing the required official notice about massive governmental data collection plans; it was better than nothing but provided "openness" only in a narrow and technical sense. Most large corporations doing business with the public have privacy notices, and virtually no one reads them. Only 0.3% of Yahoo! users read its privacy notice in 2002, for example. In the midst of massive negative publicity that year, when Yahoo! changed its privacy policy to allow advertising messages, the number of users

who accessed the privacy policy rose only to 1%. None of the many U.S. privacy laws prevented the warrantless wiretapping program instituted by the Bush administration, nor the cooperation with it by major U.S. telecommunications companies.

Indeed, cooperation between the federal government and private industry seems more essential than ever before for gathering information about drug trafficking and international terrorism—because of yet another technological development. Twenty years ago, most long-distance telephone calls spent at least part of their time in the air, traveling by radio waves between microwave antenna towers or between the ground and a communication satellite. Government eavesdroppers could simply listen in. Now many phone calls travel through fiber-optic cables instead, and the government is tapping this privately owned infrastructure.

High privacy standards have a cost. They can limit the public usefulness of data. Public alarm about the release of personal medical information has led to major legislative remedies. The Health Insurance Portability and Accountability Act (HIPAA) was intended both to encourage the use of electronic data interchange for health information and to impose severe penalties for the disclosure of "protected health information," a very broad category including not just medical histories but, for example, medical payments. The bill mandates the removal of anything that could be used to reconnect medical records to their source. HIPAA is fraught with problems in an environment of ubiquitous data and powerful computing. Connecting the dots by assembling disparate data sources makes it

Ever Read Those "I Agree" Documents?

Companies can do almost anything they want with your information, as long as you agree. It seems hard to argue with this principle, but the deck can be stacked against the consumer who is "agreeing" to the company's terms. Sears Holding Corporation (SHC), the parent of Sears, Roebuck and Kmart, gave consumers an opportunity to join "My Sears Holding Community," which the company describes as "something new, something different…a dynamic and highly interactive online community…where your voice is heard and your opinion matters." When you went online to sign up, the terms appeared in a window on the screen.

The scroll box held only 10 lines of text, and the agreement was 54 boxfuls long. Deep in the terms was a detail: You were allowing Sears to install software on your PC that "monitors all of the Internet behavior that occurs on the computer…, including…filling a shopping basket, completing an application form, or checking your…personal financial or health information." So your computer might send your credit history and AIDS test results to SHC, and you had said that was fine!

extremely difficult to achieve the level of anonymity that HIPAA sought to guarantee. But help is available, for a price, from a whole industry of HIPAA compliance advisors. If you search for HIPAA online, you will likely see advertisements for services that will help you protect your data and also keep you out of jail.

At the same time as HIPAA and other privacy laws have safeguarded our personal information, they are making medical research costly and sometimes impossible to conduct. It is likely that classic studies such as the Framingham Heart Study, on which much public policy about heart disease was founded, could not be repeated in today's environment of strengthened privacy rules. Dr. Roberta Ness, president of the American College of Epidemiology, reported that "there is a perception that HIPAA may even be having a negative effect on public health surveillance practices."[21]

The five FIPP principles, and the spirit of transparency and personal control that lay behind them, have doubtless led to better privacy practices. But they have been overwhelmed by the digital explosion, along with the insecurity of the world and all the social and cultural changes that have occurred in daily life. Fred H. Cate, a privacy scholar at Indiana University, characterizes the FIPP principles as almost a complete bust:

> Modern privacy law is often expensive, bureaucratic, burdensome, and offers surprisingly little protection for privacy. It has substituted individual control of information, which it in fact rarely achieves, for privacy protection. In a world rapidly becoming more global through information technologies, multinational commerce, and rapid travel, data protection laws have grown more fractured and protectionist. Those laws have become unmoored from their principled basis, and the principles on which they are based have become so varied and procedural, that our continued intonation of the FIPP mantra no longer obscures the fact that this emperor indeed has few if any clothes left.[22]

Only sects such as the Amish still live without electricity. It is 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 "over-the-air" TV is rapidly disappearing in favor of digital communications.[23]

Digital TV brings the advantages of video on demand, but with a steep privacy cost. Your television service provider records everything you watch, and when. It is so attractive to be able to watch what we want when we want to watch it that we don'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.

Privacy as a Right to Control Information

Privacy is complex and under attack from our peers, our own devices, and governments and corporate marketers. The bits are everywhere; there is simply no locking them down, and no one really wants to do that anymore. The meaning of privacy has changed, and we do not have a good way of describing it. It is not the right to be left alone, because not even the most extreme measures will disconnect our digital selves from the rest of the world. It is not the right to keep our private information to ourselves because the billions of atomic factoids don't lend themselves to being simply and uniquely classified as either private or public.

Which would we prefer: the new world with digital fingerprints everywhere and the constant awareness that we are being tracked, or the old world with few digital footprints and a stronger sense of security from prying eyes? And what is the point of even asking the question when the world cannot be restored to its old information lockdown?

The bits are everywhere; there is simply no locking them down, and no one really wants to do that anymore.

In a world that has moved beyond the old notion of privacy as a wall around the individual, we could instead regulate those who would inappropriately *use* information about us. If I post a YouTube video of myself dancing in the nude, I should expect to suffer some personal consequences. Ultimately, as Warren and Brandeis said, individuals have to take responsibility for their actions. But society has drawn lines in the past around which facts are relevant to certain decisions and which are not. Perhaps, the border of privacy having become so porous, the border of relevancy could be stronger. As Daniel Weitzner explains:

> New privacy laws should emphasize usage restrictions to guard against unfair discrimination based on personal information, even if it's publicly available. For instance, a prospective employer might be able to find a video of a job applicant entering an AIDS clinic or a mosque. Although the individual might have already made such facts public, new privacy protections would preclude the employer from making a hiring decision based on that information and attach real penalties for such abuse.[24]

There can still be principles of accountability for the *misuse* of information. Some ongoing research is outlining a possible new web technology to help ensure that information is used appropriately when it is known. Perhaps automated classification and reasoning tools, developed to help connect the dots in networked information systems, can be retargeted to limit inappropriate

use of networked information. A continuing border war is likely to be waged, however, along an existing free speech front: the line separating my right to tell the truth about you from your right not to have that information used against you. In the realm of privacy, the digital explosion has left matters deeply unsettled.

Paul Ohm posits a "database of ruin":

> Almost every person in the developed world can be linked to at least one fact in a computer database that an adversary could use for blackmail, discrimination, harassment, or financial or identity theft.[25]

We must, through a combination of law, technology, and norms of behavior, find ways to avoid a mutually assured privacy destruction.

A few beacons of hope come from state lawmakers, most notably in California, and a growing culture of privacy among engineers. Some corporate privacy notices are still boilerplate, but others give the impression that privacy is a product feature, designed to add value for users and respond to their needs.

Always On

In *1984,* the pervasive, intrusive technology could be turned off:

> As O'Brien passed the telescreen a thought seemed to strike him. He stopped, turned aside and pressed a switch on the wall. There was a sharp snap. The voice had stopped.
>
> Julia uttered a tiny sound, a sort of squeak of surprise. Even in the midst of his panic, Winston was too much taken aback to be able to hold his tongue.
>
> "You can turn it off!" he said.
>
> "Yes," said O'Brien, "we can turn it off. We have that privilege....Yes, everything is turned off. We are alone."

Sometimes we can still turn it off today—and should. But mostly we don't want to. We don't want to be alone; we want to be connected. We find it convenient to leave it on, to leave our footprints and fingerprints everywhere, so we will be recognized when we come back. We don't want to have to keep retyping our name and address when we return to a website. We like it when the restaurant remembers our name, perhaps because our phone number showed up on caller ID and is linked to our record in their database. We appreciate buying grapes for $1.95/lb instead of $3.49, just by letting the

store know that we bought them. We may want to leave it on for ourselves because we know it is on for criminals. Being watched reminds us that they are watched as well. Being watched also means we are being watched over.

And perhaps we don't care that so much is known about us because that is the way human society used to be: In kinship groups and small settlements, knowing everything about everyone else was a matter of survival. Having it on all the time may resonate with inborn preferences we acquired millennia ago, before urban life made anonymity possible. Still, today, privacy means something very different in a small rural town than it does on the Upper East Side of Manhattan.

We cannot know what the cost will be of having it on all the time. Just as troubling as the threat of authoritarian measures to restrict personal liberty is the threat of voluntary conformity. As Fano astutely observed, privacy allows limited social experimentation—the deviations from social norms that are much riskier to the individual in the glare of public exposure, but which can be, and often have been in the past, the leading edges of progressive social changes. With it always on, we may prefer not to try anything unconventional and stagnate socially by collective inaction.

For the most part, it is too late, realistically, ever to turn it off. We may once have had the privilege of turning it off, but we have that privilege no more. We have to solve our privacy problems another way.

The digital explosion is shattering old assumptions about who knows what. Bits move quickly, cheaply, and in multiple perfect copies. Information that used to be public in principle—for example, records in a courthouse, the price you paid for your house, or stories in a small-town newspaper—is now available to everyone in the world. Information that used to be private and available to almost no one—medical records and personal snapshots, for example—can become equally widespread through carelessness or malice. The norms and business practices and laws of society have not caught up to the change.

Endnotes

1 Matthew Rosenberg, Nicholas Confessore, and Carole Cadwalladr, "Firm That Assisted Trump Exploited Data of Millions," *New York Times*, March 18, 2018: A1, https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html.

2 Samuel A. Warren and Louis D. Brandeis, "The Right to Privacy," *Harvard Law Review* 4, no. 5 (December 15, 1890), https://groups.csail.mit.edu/mac/classes/6.805/articles/privacy/Privacy_brand_warr2.html.

3 Robert Fano, "Review of Alan Westin's *Privacy and Freedom*," *Scientific American* (May 1968): 148–152.

4 Alan F. Westin, *Privacy and Freedom* (Atheneum, 1967).

5 Ibid.

6 Helen Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (Stanford Law Books, 2009).

7 "Judge Sides with University Against Student-Teacher with 'Drunken Pirate' Photo," *The Chronicle of Higher Education*, December 4, 2008, https://www.chronicle.com/article/Judge-Sides-With-University/42066.

8 Taylor Lorenz, "Unidentified Plane-Bae Woman's Statement Confirms the Worst," *The Atlantic*, July 13, 2018, https://www.theatlantic.com/technology/archive/2018/07/unidentified-plane-bae-womansstatement-confirms-the-worst/565139/.

9 Emil Venere, "Printer Forensics to Aid Homeland Security, Tracing Counterfeiters," Purdue University, October 12, 2004, https://www.purdue.edu/uns/html4ever/2004/041011.Delp.forensics.html.

10 Michael M. Grynbaum and John Koblin, "Journalists Fear Effects of Arrest," *New York Times*, June 7, 2017: A19, https://www.nytimes.com/2017/06/06/business/media/intercept-reality-winner-russia-trump-leak.html.

11 Jake Swearingen, "Did the Intercept Betray Its NSA Source?", New York Magazine, June 6, 2017. https://nymag.com/intelligencer/2017/06/intercept-nsa-leaker-reality-winner.html.

12 "Web Privacy—Arvind Narayanan," accessed May 18, 2020, https://www.cs.princeton.edu/~arvindn/web-privacy/.

13 Charles Duhugg, "How Companies Learn Your Secrets," *The New York Times*, February 16, 2012, https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html.

14 "Amazon Alexa Can Accidentally Record and Share Your Conversations," *Vanity Fair*, May 24, 2018, https://www.vanityfair.com/news/2018/05/yes-amazons-alexa-can-secretly-record-and-share-conversations.

15 Katie Collins, "That Smart Doll Could be a Spy. Parents, Smash!", *CNET*, February 17, 2018, https://www.cnet.com/news/parents-told-to-destroy-connected-dolls-over-hacking-fears/.

16 Ben Gilbert, "There's a simple reason your new smart TV was so affordable: It's collecting and selling your data, and serving you ads," *Business Insider*, April 5, 2019, https://www.businessinsider.com/smart-tv-data-collection-advertising-2019-1.

17 Hang Do Thi Duc, Public By Default, Venmo Stories of 2017, HYPERLINK \h https://publicbydefault.fyi/.

18 Avi Selk, "The ingenious and 'dystopian' DNA technique police used to hunt the 'Golden State Killer' suspect," *Washington Post*, April 28, 2018, https://www.washingtonpost.com/news/true-crime/wp/2018/04/27/golden-state-killer-dna-website-gedmatch-was-used-to-identify-joseph-deangelo-as-suspect-police-say/.

19 "Fact Sheet: Overview of the EU–U.S. Privacy Shield Framework for Interested Participants," U.S. Department of Commerce, July 12, 2016, https://2014-2017.commerce.gov/sites/commerce.gov/files/media/files/2016/fact_sheet-_eu-us_privacy_shield_7-16_sc_cmts.pdf.

20 Court of Justice of the European Union, Press Release No 91/20, Luxembourg, July 16, 2020, Judgment in Case C-311/18: Data Protection Commissioner v Facebook Ireland and Maximillian Schrems. https://curia.europa.eu/jcms/upload/docs/application/pdf/2020-07/cp200091en.pdf.

21 "Privacy Rule Slows Scientific Discovery and Adds Cost To Research, Scientists Say," University of Pittsburgh Schools of the Health Sciences. https://www.sciencedaily.com/releases/2007/11/071113165648.htm. Also: Roberta B. Ness, MD, MPH, "Influence of the HIPAA Privacy Rule on Health Research," JAMA. 2007;298(18):2164-2170. doi:10.1001/jama.298.18.2164

22 Fred H. Cate, "The failure of Fair Information Practice Principles," in Jane K. Winn, ed., *Consumer Protection in the Age of the "Information Economy"* (Ashgate, 2006).

23 https://www.nielsen.com/us/en/insights/report/2019/nielsen-local-watch-report-the-evolving-ota-home/

24 Daniel. J. Weitzner, "Beyond Secrecy: New Privacy Protection Strategies for Open Information Spaces," in *IEEE Internet Computing* 11, no. 5 (September–October 2007): 96–95, https://dl.acm.org/doi/10.1109/MIC.2007.101

25 Paul Ohm, "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization," SSRN Scholarly Paper (Social Science Research Network, August 13, 2009), https://papers.ssrn.com/abstract=1450006.