6 Privacy

Ask people what they think of when they hear the phrase "breach of privacy," and you'll get some predictable responses. Common among these might be examples of police or military surveillance, closed circuit television cameras watching your every move, paparazzi snooping into private family moments with long camera lenses, or a government agency intercepting personal messaging services. These examples fit with the understanding most people have of their right to privacy. Most people would probably take the right to mean something like "the right to be let alone."[1] But there's more to it than that, and in this chapter we want to help you better understand the different dimensions of privacy, how new technologies affect these dimensions, why this matters, and how these issues relate to you as a citizen.

Dimensions of Privacy

It might come as a surprise that, given the cross-cultural importance of privacy, there is still no internationally settled definition of what it actually is. In 1948, members of the United Nations adopted the Universal Declaration of Human Rights, which set out the fundamental human rights they agreed must be universally protected, including the right to privacy in Article 12: "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honor and reputation. Everyone has the right to the protection of the law against such interference or attacks."[2]

Unfortunately, as fine a sentiment as this is, countries have been free to interpret this standard as they've seen fit. Europe has perhaps traveled farthest along the path to securing robust privacy rights for its citizens. Most jurisdictions haven't come nearly as far. In any case, trying to secure international

consensus on the meaning of the term isn't the only problem. Getting *anyone* to define the concept (let alone over 150 countries) is a tall order.

The privacy scholar Julie Inness once lamented that privacy is a concept in chaos. One of the entries in the index to her book on privacy reads, "Quagmire, privacy as deep."[3] But chaotic though it may be, Inness still felt it was salvageable—a concept whose many jagged fragments could be made to fit together.[4] This is because, even if privacy has a broad range of meanings, not all of which are mutually accommodating, all of them do seem to share a preoccupation with intimacy and dignity. For Inness, privacy boils down to our natural concern for maintaining a degree of autonomy with respect to matters we care about, such as intimacy in our relationships.

Years later, the legal scholar Daniel Solove would try his own hand at a definition, but unlike Inness, he chafed at essentialism. Taking his cue from the famous Austrian philosopher Ludwig Wittgenstein, Solove doubts we'll ever find the unique, irreducible essence of a concept like privacy. As far as he's concerned, we're better off abandoning this search altogether, redirecting our attention to the concrete consequences of privacy violations. Accordingly, he devised a taxonomy of privacy *harms*. In his scheme there are harms that arise from information collection (such as surveillance), harms that arise from information processing (such as aggregation, identification, or insecure processing), harms arising from information dissemination (such as unlawful disclosure or breach of confidence), and harms arising from invasion (such as physical intrusion).[5] A harms-based approach is certainly useful, focusing on what's likely to matter most to people when they think about their privacy. But Solove's particular approach also raises questions the moment we bring machine learning into the picture. If datasets don't relate to anyone in particular—they are usually anonymized aggregations of many individual profiles run together—and if they aren't likely to be used against any of the individuals whose data go into them, then what exactly is the harm? A predictive model can, of course, be weaponized against someone if the algorithm is stacked against them, for example, if by using a particular software package a person's gender or ethnicity makes it almost certain they'll be rejected for a job. But it isn't obvious that the harm here results from a breach of privacy per se. It certainly doesn't seem to be a breach of *that person's* privacy (at least not straightforwardly). And even if your personal information *is* being funneled back "against" you, what exactly is the "harm" if all you're having to contend with are Netflix and Amazon recommendations?

This isn't to say that focusing on harms is misguided, but it is a reminder to make sure any working definition takes our brave new world of big data seriously.

Here, we're going to be pragmatic. We're not going to pretend privacy means just one thing or results in one basic kind of harm (and nor did Solove, for that matter). Instead, there are at least four things privacy can mean—four *dimensions* of privacy, if you will—each of which results in a distinctive kind of harm:

1. **Bodily privacy** secures a person's *bodily integrity* against nonconsensual touching or similar interference;
2. **Territorial privacy** protects a person's *ambient space* from intrusion and surveillance;
3. **Communication privacy** protects a person's *means of communicating* against interception;
4. **Informational privacy** prevents *personal information* being collected, processed, or used against its owner's wishes (otherwise known as "data protection").[6]

This "divide and conquer" strategy makes clear how new forms of privacy breach can emerge in a particular dimension. For example, modern informational privacy (data protection) principles arose partly in response to the vast and rapidly expanding amounts of personal information that both governments and private corporations held about individuals. Concerns over the adequacy of existing national laws and international standards resulted in litigation and, ultimately, new laws imposing clear obligations on agencies holding personal information.* Such agencies must now be careful to ensure that individuals who give them personal information retain power over that information, including the power to access, correct, use, and (if necessary) delete it. New forms of communication privacy have emerged too. Immediately following the Snowden revelations, for example, legal scholars and civil libertarians debated whether there could be a right to privacy with respect to electronic communications.

*For example, in 1980 the German Constitutional Court ruled on the validity of population census data, holding that in the age of data processing, individual rights to self-determination required protection from the unlimited collection, use, and storage of personal information.

Although each of these privacy dimensions is important—and AI has implications for every one of them—in this chapter, we aren't going to discuss bodily, territorial, or communication privacy in much detail. Instead, our focus will be on informational privacy, as this is the form of privacy most obviously and directly threatened by the advent of big data. Besides, advances in AI and machine learning may mean that these four dimensions will converge increasingly in years to come. It's a safe bet that ever larger and more sophisticated datasets will be used to enhance both state and corporate powers of surveillance and intrusion. To take just one example, face recognition software is already a kind of surveillance technology, and it depends crucially on access to high quality training data.

Informational Privacy and AI

The overarching concern about AI and privacy arises from the ways in which predictive algorithms and other forms of machine learning can recognize patterns in our personal information, perhaps even ones that we ourselves cannot see. In doing so, these technologies can peer into parts of our private lives that no other human would be able to see. This is, again, the *overarching* concern. There are more specific concerns too, of course, the two most important of which we'll consider here. But it's worth keeping the big picture in view as we look at these more specific problems.

A fundamental principle of data protection law is that "data should be collected for named and specific purposes."[7] But purpose limitation strikes at the heart of the big data business model. As we've seen, developing the capabilities of machine learning techniques requires access to training data—lots and lots of training data (see chapter 1). Privacy advocates worry in particular about *how* personal information is collected, especially on the internet. It is rarely used for "named and specific purposes," and rarely given with informed consent.[8] Oh, sure, companies like Facebook *do* mention purposes, but they are often so general as to be meaningless; for example, to provide and improve their products and services—in other words, to implement Facebook!

Knowing whether users who visit websites to book flights will next visit websites to book accommodation and then other websites to book rental vehicles is valuable information. Hotels and rental car companies are willing to pay for this information because it can be used to push targeted

advertisements to travelers. Individuals are often peacefully unaware that the knowledge of what they do on the internet is valuable and that they are allowing others to collect and profit from the use of this information. And even when they know or suspect that their information will be sold to third parties, they won't know *exactly* how those third parties will use it, and indeed *who* those third parties are. The cynical interpretation is that people are being exploited. "If it's free on the internet, *you're* the product," as the saying goes.

The internet isn't the only data collection point of interest. Data can be gathered from what might seem to be the most unlikely places. Many washing machines are now part of the "internet of things," and come equipped with sensors that can generate information about wash times, wash cycles, and other matters that can be downloaded and used to predict maintenance, repairs, faults, and energy consumption.[9] Although such data collection may be beneficial from a maintenance and design point of view—knowing that most people use only two or three of the wash functions on a machine might lead to a more efficient design, for example—it might also be valuable to third parties. The local energy company (say) might be interested to encourage you not to use your washing machine during peak electricity usage hours. A manufacturer of laundry detergent might be interested in hawking related products, such as fabric softeners. Again, these third parties will be willing to pay for this information, and, no less than before, consumers will be frequently unaware of the kinds of data being collected about them, how much data are being collected, and what uses they're being put to.

It's true that many companies are open about the terms of the deal they have struck with users: "If you allow us to mine your likes, shares and posts—and to sell what we learn to the highest bidder—we'll grant you access to our platform."* But not all instances of data mining occur with explicit consent. And even in cases where they do, this "consent" isn't necessarily free and voluntary. If you need to "consent" to data collection to access an essential

*Interestingly, Facebook doesn't directly sell any data—they're not "data brokers." Instead, they monetize their data by offering a *targeted ad placement* service to advertisers. They engage in analytics to home in on the most suitable customer groups for specific products, and then allow advertisers to place ads with these specific groups. It's as if Facebook were to say to a travel agency, "I know all the middle-aged Christian train enthusiasts. I won't tell you who they are, but if you pay me, I'll pass your message on to them."

service, like an online banking or health app, in what sense are you meaningfully consenting to the surrender of your personal information? You don't have a choice—you *have* to consent in order to access the essential service. Besides, long, complex terms and conditions that no one has the time to read and that don't clearly set out "named and specific purposes" make a mockery of genuinely free and informed consent.

The second consent-related issue thrown up by big data concerns the use of *inferred* data. Inferred or "derived" data (such as the suburb you live in derived from your postcode) can be distinguished from *collected* data (data that you explicitly, knowingly, provide, such as your name) and *observed* data (data that you passively, or implicitly, provide, such as your handwriting, accent, or keystroke rate). Machine learning and big data literally *exist* to facilitate the drawing of inferences. When a machine learning tool flags a submitted tax return as potentially fraudulent, for example, it's not doing so through directly ascertainable information (fraud per se), but on the basis of significant correlations it has learned between directly ascertainable information (reported income, losses, etc.) and the phenomenon of interest (i.e., fraud). In a reliable dataset of previous tax return information, larger-than-usual losses reported in consecutive tax years—itself directly ascertainable information—might be found to correlate strongly with known instances of tax fraud. In the language of chapter 1, tax fraud would be the "predicted variable," the inferred characteristic (like mortality given age, or weight given height). The question for privacy law is whether consent must be obtained for the use of this inferred information. Do inferred data count in the same way as primary (collected and observed) data?

This isn't just an academic worry. Consider machine learning techniques that use nonsensitive information to predict very sensitive information about an individual. Privacy advocates have raised concerns about the use of apparently unrelated data, such as information about location, social media preferences, screen time on different apps or phone activity, to aggregate and predict highly sensitive information (such as sexuality or political beliefs). One study found that the emotional states of computer users could be determined from such apparently innocuous information as their keystroke rate.[10] In 2017, a Stanford University study claimed that an algorithm could successfully distinguish between gay and straight men 81 percent of the time, and between gay and straight women 71 percent of

the time. The study used deep neural networks to extract features from over thirty-five thousand photographs—facial features that many of us would probably regard as nonsensitive (our faces are, after all, the one part of us always open for the world to see). The authors concluded that "given that companies and governments are increasingly using computer vision algorithms to detect people's intimate traits, our findings expose a threat to the privacy and safety of gay men and women."[11] There has been considerable skepticism about these particular findings on methodological grounds,[12] but follow-up models correcting various aspects of the original still appear to have some ability to identify sexual orientation from photographs.[13]

Beyond these two paramount data protection issues that machine learning has brought to the fore, a range of other, somewhat less definitive issues have also arisen in recent years. One is the potential for reidentification from anonymized datasets. Anonymization is used frequently in academia to protect the privacy of experimental subjects. It's easy to forget that psychological and medical experiments don't just happen. Among a myriad of other precautions and ethical protocols, researchers have to recruit willing participants, a task made marginally easier by being able to guarantee subjects that their personal information won't end up in the wrong hands. But scholars like Paul Ohm have emphasized just how weak anonymization techniques have become. Using publicly available information, Ohm says that, despite efforts to anonymize them, it's possible to make highly accurate predictions about the identity of specific individuals in a dataset.[14] This does not augur well for scientific research. Reidentification can be very easy. For example, you might suffer from a rare condition or live in a rural area where the number of people fitting your characteristics is low. In other cases in which reidentification may be more difficult because your personal information is not so uniquely identifiable, reidentification is still surprisingly technically straightforward. One French study found that 75 percent of mobile phone users within a dataset could be reidentified based on an individual's use of just two smart phone apps. The reidentification rate increased to 90 perecent if four, rather than two, smart phone apps were used.[15]

There are also questions about how existing data protection standards might apply to personal information used for machine learning and big data analytics. These questions relate to how individuals can access their information, who controls the information for the purposes of liability for correction,

exactly *where* obligations for ensuring accuracy fall in the data chain (particularly when information is repurposed or passed to a third party), what obligations for data deletion should be imposed, and how long data can be retained before it should be deleted. Only some of these questions have received answers in some jurisdictions. And even then, the answers aren't always decisive.

AI, Privacy, and the Consumer

So how, in practice, will all this affect your privacy and the privacy of your family, friends, and others in your community? What should your expectations of privacy be? And given how little knowledge or power you might feel you have, will your expectations of privacy matter anyway?

First up, let's consider the targeted ads we mentioned earlier. These ads are focused on almost every aspect of your online life, from that weekend away you just had to those new shoes you'd like to buy. They are a key part of online sales and product and service promotion, and used by a vast array of companies worldwide. Why does this particular application of machine learning create consumer privacy risks? There are consent-related objections to it, as we saw. But is that all?

One way in which targeted ads pose consumer privacy risks is in their potential for discrimination. The fact is that these practices aren't solely aimed at affecting your selection of a product or service from a range of choices. The techniques used to promote these can also affect *whether you are offered certain choices in the first place.*

Take housing. In early 2019, the US government filed a lawsuit against Facebook, alleging that its targeted advertising algorithms violated the Fair Housing Act by discriminating against some people, using data mining practices to restrict which users were able to view its housing-related ads.[16] The Fair Housing Act makes discrimination in housing and housing-related services illegal. For example, it's illegal to advertise that housing is available to people of a certain race, color, national origin, religion, sex, or marital status. Ben Carson, Secretary of the Department for Housing and Urban Development, put it simply: "Facebook is discriminating based on who people are and where they live. Using a computer to limit a person's housing choice can be just as discriminatory as slamming a door in someone's face."[17] This effect can be amplified if those offering targeted ads have

significant market dominance. Facebook is estimated to control about 20 percent of online advertising in the United States.*

Public lawsuits are helpful because they enable us to monitor at least some of what is happening and to have judicial oversight of practices that affect consumers. But commentators point out that recent lawsuits against Facebook by the National Fair Housing Alliance, American Civil Liberties Union, and other civil society groups have been settled out of court, and, in some of those cases, the terms of the settlements remain private. This makes it harder to figure out the specific privacy-protective measures, if any, that such litigation is forcing companies to take.[18]

Now you might think that this whole business of online discriminatory advertising isn't itself *squarely* a privacy issue. It's really a cocktail of anti-discrimination, fair trading, and human rights issues jumbled together. Still, let's not forget that discriminatory advertising is *discriminating* (in the sense of *discerning*), based on technology that predicts the kind of person you *are*—technology, indeed, that infers things about you that you may not want others to know. In some countries, being gay is illegal and punishable by death. "Gaydar" software that "outs" you at an airport security checkpoint wouldn't merely be inconvenient; it could be life-threatening.

Let's consider another way in which machine learning and natural language processing tools in particular are using information in ways that affect your privacy as a consumer: training data that are used to develop consumer behavior prediction tools. This is clearly a privacy issue. Our behaviors, intentions, and innermost proclivities are being predicted, and possibly even manipulated (see chapter 7). As we've discussed, those developing machine learning tools rely on existing datasets to train and test them. These datasets range in size, quality, and diversity, and are used in many different ways depending on the type of AI being developed. So where does the data come from? You might be surprised to learn that you have probably already given your information to a training dataset. For example, if you've ever made a call to your insurance, phone, or electricity company,

*Similar action is taking place in other countries. In 2018, for instance, Privacy International filed complaints with UK, French, and Irish data protection authorities against seven companies, complaining of their use of personal data for targeted advertising and exploitative purposes.

you've probably contributed to the collection of this kind of training data. You might recall an irritating automatically-generated voice message, telling you that your call "may be recorded for quality and training purposes." Imagine hundreds and thousands, perhaps millions, of those calls being made available to train natural language processing tools that recognize all kinds of useful things: differences between male and female voices, when someone is angry, when they are upset, the typical questions customers ask, and the typical complaints they make.

Thinking of canceling your insurance policy and switching to another provider? Based on your behavior, machine learning tools can anticipate this. Customer "churn prediction" or attrition rates are prominent performance indicators in many companies, and being able to predict and reduce likely churn can provide a significant business advantage. The models can be trained on the behavioral data of thousands of previous customers who have switched accounts and those who have stayed. Using this information to predict likely customer churn, a list of "at risk" customers can be generated and sent to an accounts manager for review and action.

That might be well and good—perhaps you've been missing out on a better deal and really appreciate that call from the insurance company to see if you're still happy with their product. But when these kinds of datasets are used to create profiles of different types of people based on preselected categories of information, such as age, sex, medical history, location, family status, and so on, mispredictions will be rife. What if the profile generated for you is so wildly different from your actual situation that you miss out on some options completely?* What if, like Virginia Eubanks, your family's health insurance account is red-flagged by an algorithm for suspicious behavior because when your domestic partner was assaulted you (very reasonably) made a health insurance claim for domestic care services around the same time you switched jobs and took out a new policy? We know that algorithms can infer lots of uncanny things about you. This creepiness factor is one thing when the inferences are correct but quite another when the inferences are wrong.[19]

*You can check your Facebook profile to see what categories you've been placed in. Go to Settings > Privacy Shortcuts > More Settings > Ads > Your Information > Review and Manage Your Categories. You'll see some true stuff, and probably some weird stuff.

Companies are also combining dynamic price differentiation with data collection to enable algorithmic setting of real time service prices. For example, a report by Salesforce and Deloitte in 2017 found that although uptake of algorithms by major brand businesses was still low—with just over a third of such businesses adopting AI—among those that have taken up algorithmic tools, 40 percent were using them to tailor prices.[20]

Why is this important? Well, you might think you're roaming freely online, privately seeking out the information you want when and how you want it. In fact, your online life is increasingly being curated, filtered, and narrowed. In the process, not only is your sphere of privacy reducing, but the breadth of your participation in public life is also reducing, simply because your freedom to receive information of *any* kind is being restricted (see chapter 7). This doesn't just affect the opportunities and choices available to you online, either. It's increasingly affecting your offline existence, too, including at work (see chapter 9). One privacy implication is the detrimental effect of being constantly monitored while working. This is another domain where, as we mentioned earlier, different dimensions of privacy (here the territorial and informational dimensions) are converging. Your personal information can be used to help employers infer the movements and habits of "workers like you." The difference between being "watched" by a camera and being "known" by an algorithm is becoming less important.

AI, Privacy, and the Voter

In the first part of 2017, a news story broke about the UK's referendum on leaving the EU—a story that would change the landscape of political campaigning in the United Kingdom and around the world. Media reports emerged in the United Kingdom that Cambridge Analytica and related companies (which we'll just collectively call "Cambridge Analytica" for ease of reference) had assisted the Leave campaign by providing data services that supported micro-targeting of voters. Eighteen months later, in November 2018, the UK Information Commissioner, Elizabeth Denham, reported on her investigation of the story during which she had engaged 40 investigators who identified 172 organizations and 71 witnesses, issued 31 notices demanding information, executed two warrants, instigated one criminal prosecution and seized materials including 85 pieces of equipment, 22

documents and 700 terabytes of data (the equivalent of more than 52 billion pages of evidence).[21] The investigation uncovered how political campaigns use individuals' personal information to target potential voters with political messages and advertisements. It revealed the complex system of data sharing between data brokers, political parties and campaigns, and social media platforms. The commissioner concluded: "We may never know whether individuals were unknowingly influenced to vote a certain way in either the UK EU referendum or in the US election campaigns. But we do know that their personal privacy rights have been compromised by a number of players and that the digital electoral eco-system needs reform."[22]

Again a key concern was the use of data collected for one purpose for a completely different purpose without consent and in violation of Facebook policies. The commissioner found that Cambridge Analytica worked with university researcher and developer, Dr. Aleksandr Kogan, to establish a company (GSR) that contracted with Cambridge Analytica to develop a new app, *thisisyourdigitallife*. Users who logged into Facebook and authorized the app made their data—*and* those of their Facebook friends—available to GSR and Cambridge Analytica. (To prevent this sharing, the friends would have had to uncheck a field in their Facebook profile that was on by default, something hardly any users ever did.) The new app was able to access about 320,000 Facebook users who took detailed personality tests while logged into their Facebook accounts. In doing so, the app was able to collect the users' public profile (including birth date, current city, photos in which the user was tagged, pages they had liked, timeline and newsfeed posts, lists of friends, email addresses, and Facebook messages). Facebook estimated that the total number of users of the app, including affected Facebook friends, was approximately 87 million.[23]

Should We Give Up on Privacy?

With all of these concerns it is fair to ask, is privacy dead? Contrary to popular belief, demand for privacy will likely increase rather than decrease in the age of AI. Many surveys show that consumers value their online privacy. In the United States in 2015, Consumer Reports found that 88 percent of people regard it as important that no one is listening to or watching them. Echoing this finding, a Pew research study found that the majority of Americans believe it is important, or very important, to be able to

maintain their privacy in everyday life. In relation to online life, 93 percent of adults said that being in control of who can get information about them is important, and 95 percent believed that what is collected about them is important. However, the same survey found that nearly two thirds of people were not confident that their online activities would be kept secure by online advertisers, social media sites, search engine providers, or online video sites.[24]

Global civil society organizations such as Privacy International and the Electronic Frontier Foundation, as well as consumer rights groups, have long called on users to take more control of the collection and use of their personal information online. They have also advocated for simpler privacy protections. The fact is there *are* ways to limit the effects of online advertising, but surprisingly few implement them. Sometimes it's because companies make it hard. If you don't consent to your phone call being recorded, you might not be able to access a service. In other cases, even if you do take care with your online activities—for example, by staying away from certain platforms or not posting information about your political beliefs, health status, or social activities—algorithms can still predict what you might do next. Before the "friend's permission" feature in Facebook was disabled, your personal information may have been stored in your friends' social media accounts and therefore vulnerable if your friends chose to allow other apps to access their contacts. Another significant barrier is the lack of simple tools to help consumers.

Can We Protect Privacy in the Age of AI?

There are now much blurrier lines between personal and nonpersonal information, new forms of data exploitation, new forms of data collection, and new ways in which personal information is being used to create profiles and predictions about individuals through machine learning tools. The result is that the various dimensions of privacy we've discussed are both expanding and constricting in exciting, complex, and confusing ways. In light of all of this, you might well ask, can we better protect individual privacy? Taking a little more responsibility for our privacy settings is one step, but what else can be done?

Calls have been made to strengthen privacy in technology development. Ann Couvakian coined the phrase "privacy by design" to help conceptualize

the request to technology developers for privacy-enhancing technologies.[25] The EU's General Data Protection Regulation (GDPR) charted new territory in limiting the use of automated processing in certain circumstances and requiring individuals to be provided with information as to its existence, the logic involved and the significance and proposed consequences of the processing for the individual concerned. The GDPR also confers rights of correction and erasure.

Remember that machine learning's predictive accuracy depends on the future looking like the past (see chapter 1). If the data on which an algorithm has been trained and tested remain the same, the predictions will also remain the same, and that's fine. But data are rarely static. People change. They develop new skills, end relationships, form new ones, change jobs, find new interests. And some people will always fall through the cracks of what is considered typical for someone of a certain age, gender, sexuality, ethnicity, and so on. The result is that inferences can be based on data that are obsolete or in some other way "dirty." But given how pervasive and persistent inferential analytics is set to become, a key step would be to develop the law here even further—beyond the modest protections offered by the GDPR. What can we allow to be reasonably inferred, in what sorts of situations and under what kinds of controls (such as rights of access, correction, and challenge, which the GDPR already confers to some degree)? Recently, a right to reasonable inferences has been proposed to reorient data protection law toward the *outputs* of personal information as distinct from its collection and use (the traditional focus of data protection law).[26] The proposal would require data controllers to show why particular data are relevant to drawing certain "high risk" inferences, why the inferences themselves need to be drawn at all, and whether both the data and methods used to draw these inferences are statistically reliable.

Legal systems also need to come clean on the status of inferred data—should it attract the same protections as the primary (collected and observed) personal information on which it is based? It's been suggested that inferred data *can* be defined as personal information if the content, purpose, or result of the data processing relates to an identifiable individual.[27]

We've touched on issues of freedom and voluntariness in this chapter. In our next chapter, we'll probe more deeply into the ways that both legal and illegal uses of your personal information potentially compromise your freedom as a human agent.