Introduction

What Is Data Privacy and Why Is It Important?

In February 2021, social media giant Facebook entered into a settlement agreement in a class action lawsuit brought by over 1.5 million of its users for infringements of an Illinois state law prohibiting the use of photo tags and other biometric data without the permission of those users. The case was settled for $650 million to be distributed among any Facebook members who chose to participate in compensation, potentially providing at least $345 to each affected user. The law in question is the Illinois Biometric Information Privacy Act, which requires companies to get permission before harvesting and using people’s biometric data, including digital “faceprints” and fingerprints. Facebook has routinely used facial recognition technology to find “friends” of its users for targeted advertising and other purposes. Since the litigation started in 2015, it has modified its practices on photo tagging.

This case is one of the more high-profile examples of increasing concerns by consumers about digital privacy in recent years. Amid the global COVID-19 pandemic, allegations of election interference, racial injustices, and identity theft, personal privacy concerns have bubbled to the forefront of the public imagination. Congress seems to be taking seriously the need for a comprehensive national privacy law for the first time since the dawn of the digital age. However, enacting, implementing, and enforcing such a law is fraught with technological and political challenges. This book attempts to explain these challenges, the genesis of privacy worries in the digital age, and the situations in our daily lives that most threaten our personal privacy.

The aim is not to be alarmist, but simply to explain, in user-friendly terms, when governments, businesses, and others may be harvesting your personal information, how they use it, and what you can do to monitor and
protect your personal data. Each chapter relates to a different aspect of our lives where privacy concerns may arise: privacy at home, privacy at work, privacy on social media, privacy and the government, privacy at school, and so on. And each chapter concludes with some tips and tricks for monitoring and protecting your privacy in the relevant context. Throughout the text, particularly important and/or unfamiliar terms will be in bold at first use.

First, let’s take a brief look at the background to some of the central concepts. What is personal data privacy and why does it matter? Let’s start with a familiar scenario.

Have you ever gone to a store and admired, say, a designer pair of shoes that you think your best friend might like? You snap a photo on your smartphone and text it to her. When you get home, you turn on your computer and check Facebook, where an ad pops up for other items by the same designer. Facebook suggests you might want to connect with the friend you texted. When you curl up in bed, you log in to your favorite streaming service, which suggests you might like to watch a new documentary about the designer.

While this scenario seems pretty mild, and maybe quite useful (you may want to watch that documentary, after all, or share it with your friend), does it bother you that your online service providers seem to know so much about you—your hobbies, your favorite books and TV shows, your location, your friends and family, your professional connections?

What if the information is inaccurate?
What if it is embarrassing?
What if it could expose you, or your family, to harm?

You may remember the 2012 news report about a distraught man storming into a Target store to complain that his teen daughter had received personally addressed coupons for baby clothes and cribs. He accused the store manager of encouraging teen girls to get pregnant. He called back later to apologize for the outburst, admitting that there had been some activities in his house that he was not aware of. His daughter was indeed pregnant, and Target’s directed marketing had outed her.²

Personal data is useful for more than just targeted marketing. Information about you can affect employment decisions, political decisions, and healthcare. Think about situations like the allegations of Russia hacking the 2016 presidential election, the Equifax data breach in 2017, or the
government’s use of data from home DNA tests to identify criminal suspects. All these activities rely on technology that collects and aggregates personal information.

Election hackers can target messages to voters likely to support a particular candidate, but who might not otherwise bother to vote, in order to galvanize them into action. To do so, the hacker must have access to information about how those people were likely to vote in the first place. The Equifax data breach implicated massive quantities of personal information that could be used for crimes like identity theft: using someone’s credentials to engage in credit card fraud, tax fraud, or health fraud. The police can, and have, used aggregated databases of home-DNA-test results to identify criminal suspects, often without a warrant, which raises significant due process concerns. And then there’s the increased use of facial recognition technology (or FRT), which is discussed in some detail in chapter 11, and which formed the basis of the Facebook example at the beginning of this introduction.

These activities have one commonality: reliance on massive amounts of personal data that are now routinely collected and aggregated by governments and businesses around the world. Most of us are aware that some businesses, like Facebook, rely on aggregating personal data to support their entire business model, capitalizing on targeted advertising. We will talk more about how targeted marketing works in chapter 4.

Governments and businesses routinely harvest and use our personal information, sometimes harmfully and sometimes helpfully. Contact tracing for those who may have been exposed to COVID-19 is an obvious example where gathering and processing personal information may be immensely helpful to controlling the spread of the virus.

However, those of us concerned about how much data is gathered about us on a daily basis may feel confused and powerless against faceless database operators cataloguing the finer details of our lives. The law is not much comfort, especially in the United States, which has no comprehensive federal privacy-protection law. Congress continues to debate the need for such a law, particularly in the wake of the Mueller Report with its focus on manipulation of data collected through social media, like Facebook, to influence election outcomes, hitting at the heart of our democracy. Some states, like
California, have moved forward in the absence of federal law to create their own comprehensive privacy laws, but these laws are limited to situations affecting residents of the state in question.

By contrast, most European countries have historically protected privacy as a basic human right. The European Convention on Human Rights (ECHR), a treaty in force throughout Europe—or at least in those countries that signed it—accepts privacy as a fundamental human right that is to be protected in all walks of life, with respect to all types of information, including but not limited to social, political, health, and financial. European Union member countries have also traditionally taken a strong stance on the protection of personal privacy, initially under the 1995 Data Protection Directive, and more recently under the General Data Protection Regulation (GDPR), which came into force in May 2018. We will consider the GDPR in some detail in chapter 8.

This book focuses on the American position on individual privacy, with some reference to other countries where comparisons are useful to aid our understanding of what is, or isn’t, possible in terms of privacy protection, as well as in regard to privacy issues that may arise across national borders. An obvious example is international travel: if one country’s law prohibits airlines from collecting certain data about passengers (like race, religion, political affiliation, or health information) and another country’s law requires it, how can that be resolved as a legal or practical matter?

Further, this book focuses on the individual level—what you can do to help monitor and protect your privacy. The idea is to empower readers to understand their rights and responsibilities in relation to protecting their personal information—and that of others—and to emphasize that sometimes you have more control than you might expect. Before we get into the meat of the discussion, we should also consider why, historically, privacy has been such a difficult issue for American lawmakers to embrace. Obviously, a big part of the explanation has to do with political priorities, but our history is very different from that of the European Union, both constitutionally and politically.

Privacy Protection: The Legal Background
Unlike the situation in many other countries, the U.S. Constitution does not include any clear right to privacy. Additionally, our powerful First Amendment protections of free speech—that is, speech free from
government interference—have been regarded as limiting laws that restrict what we can say about each other. Many other countries protect both speech and privacy as basic constitutional rights, so the courts and lawmakers in those countries are constitutionally required to balance the two. However, in the United States we have typically prioritized speech over privacy—because of the lack of a clear privacy right in our Constitution. This means that there is less protection here against the collection and use of personal information than in other countries.

In the absence of a clear constitutional right to privacy, lawmakers in the United States have developed the concept in a piecemeal way. The first major recognition of a general American need for a law on privacy dates back to the late nineteenth century, with the publication of a foundational law-review article by Samuel Warren and Louis Brandeis that attempted to define the concept. Warren and Brandeis famously described privacy as “the right to be let alone.”

That definition seems simplistic now. But it’s interesting that in the many, many, many years since that article was published, American law has never really come to grips with what privacy means, which is one of the reasons we struggle so much to protect our privacy in the digital age. It was not until the 1950s and 1960s that there were any other serious attempts to create legal privacy rights, and those came in the form of disharmonized state privacy torts.

Torts are laws that impose obligations on members of society not to harm others. They include negligence, trespass, and defamation. We will look at the privacy torts in more detail in chapter 2. In parallel to the tort law discussions are criminal laws developed to protect suspects against unreasonable searches and seizures, based on the due process provisions in the Constitution. These laws are the focus of much of chapter 11, including the extent to which due process protects our privacy in the face of mass government surveillance initiatives.

In the latter part of the twentieth century, some federal laws came into play to protect privacy in specific contexts, including privacy in the workplace, privacy in healthcare settings, and financial and consumer privacy. We will discuss those laws in later chapters. However, the history of American privacy law suggests a “reactive” approach: courts and Congress have taken
steps to address specific privacy problems when they arise, rather than tackling privacy in a comprehensive way.

TECHNOLOGY AND PRIVACY
While the law has struggled to keep pace with individual privacy, technology has charged ahead. The ability of computer networks to collect, organize, disseminate, and aggregate all kinds of information, including a lot of personal information, has, in recent years, been referred to under the label **Big Data**. Bernard Marr, in his *Beginner’s Guide to Big Data*, points out that it is not just human interactions that enable the collection and collation of huge amounts of data about individuals; computers interact with each other, too. Marr puts it like this: “We generate data whenever we go online, when we carry our GPS-equipped smartphones, when we communicate with our friends through social media or chat applications, and when we shop. You could say we leave digital footprints with everything we do that involves a digital transaction, which is almost everything.” This explains the scenario described earlier—how our connected devices link information about a designer pair of shoes to other aspects of our lives.

Marr continues by noting that part of the data-aggregation equation involves our devices generating their own collections of data and analyzing the data automatically—based on their original programming, of course:

On top of this, the amount of machine-generated data is rapidly growing too. Data is generated and shared when our “smart” home devices communicate with each other or with their home servers. Industrial machinery in plants and factories around the world is increasingly equipped with sensors that gather and transmit data. Soon, self-driving cars will take to the streets, beaming a real-time, four-dimensional maps [sic] of their surroundings back home from wherever they go.

Sounds like science fiction, doesn’t it? But the future is happening now. Self-driving cars are already in use in many cities, including Pittsburgh, where Lyft and Uber self-driving vehicles collect data about where customers are going and when. Big Brother really is watching us now, although it is not only one Big Brother (the government), but many, including companies, research labs, and nonprofit entities. Without necessarily realizing it, we often consent to data uses when we click “I agree” to access online services.
We are agreeing to contracts that include terms about what is done with our personal information. The fact that we don’t read the terms doesn’t mean they are not legally enforceable.

Big Data comprises more than names, addresses, email addresses, and telephone numbers. Today, Big Data—and the personal information from which it is aggregated—includes photographs, video, and location and voice data. It is also important to appreciate that Big Data, in and of itself, is neither good nor bad. Like most advances in technology, it has socially beneficial as well as potentially harmful uses, depending on who is using it and for what purposes. Remember that COVID-19 example? Knowing who has been exposed to the virus and being able to trace clusters and outbreaks is an essential part of addressing national and global health concerns.

The concept of Big Data is based on the idea that not only can an entity aggregate a lot of information, but that information can be used to predict social behaviors and gain new understandings of those behaviors. The more data you have, the more you can identify patterns of behavior and make predictions about the future. This could be as simple as figuring out which customers might be predisposed to buy which products, to enable more cost-effective marketing; or it could be as complex as ascertaining data about the universe to help NASA, and private companies like SpaceX, plan future space exploration missions.

As Marr points out, Big Data can serve many socially useful purposes, including research and development—for example, to help cure diseases like cancer by using demographic patient and genetic data; to prevent crime by analyzing patterns of where and when crimes are more likely to occur and deploying more resources to those areas; and to feed the hungry by aggregating agricultural data about crop yields at particular times in specific places.

Of course, on the flip side, Big Data may impinge on an individual’s privacy and the security of personal information, or manipulated to justify undesirable practices like racial or gender discrimination or election hacking. The harms to an individual from misuse of, or insecurity over, personal data can range from a general sense of unease—who knows what about me?—to actual harm like identity theft or damage to one’s credit score.

Many real-world harms could be prevented by prohibiting the manipulation of so much personal data. For example, a lot of health, housing, education,
and employment discrimination can be traced to the increased availability of applicants’ personal information on social networks. Cyberstalking and cyberharassment—which statistically tend to target those groups who have less power (e.g., women, children, people of color, members of the LGBTQIA+ community)—have also led to tangible physical and emotional harms, including tragic events like suicides. We will discuss some of these real-world tragedies in chapters 4 and 6.

Researchers have suggested that when negative consequences occur because of undesirable uses of personal information, the best approach is for lawmakers to address the resulting harms (e.g., discrimination, physical attacks, or emotional abuse) rather than regulating the information itself.⁹ In

BIG DATA, MACHINE LEARNING, AND ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI) and machine learning are often mentioned in conjunction with Big Data. This is because Big Data relies on self-learning computer programs (AIs) that can analyze data more efficiently, and at much greater speeds, than human analysts can manage.

Machine learning happens when computers are programmed to recognize certain patterns and automatically improve upon those patterns through experience. Machine learning is the basis of AI. Recently, the term AI has become more of a marketing ploy, when most references to AI really indicate machine learning. For example, the self-driving cars being developed by Tesla and Ford collect vast amounts of video data. They are programmed to recognize things like people, road signs, bicycles, and other vehicles. They are also programmed to identify the patterns these objects make and what to do when there is an anomaly in the pattern. The voluminous data points, when brought together, allow the program to make calculated predictions and safely navigate a vehicle through traffic that, to a human eye, seems random. Other uses of machine leaning include predictions about population growth, healthcare needs, environmental phenomena, land use, and space exploration.
other words, the focus should be on dealing with the resulting damage rather than regulating the source of information that led to the damage.

This is an interesting idea, although discrimination and other harms, short of physical violence, are often notoriously difficult to prove, particularly in cases where those discriminating in areas like healthcare, employment, housing, and so on can shield the reasons for their decisions behind more reasonable-sounding explanations—for example, it wasn’t that the applicant was African American, she simply wasn’t the best-qualified person for the job.

Also, a focus on only particular real-world harms doesn’t deal with the larger, underlying issue that makes many people uncomfortable: not knowing who holds what information about them—information that may even be inaccurate, and yet may impact their lives in tangible ways. Many dystopian novels and movies play on such fears. And those with this kind of power over information are not just in government. Today, we may be equally concerned about businesses that are collecting volumes of our personal information, and about what they may do with that information, for better or worse. George Orwell’s Big Brother has multiplied into a family of Big Brothers and Sisters monitoring our daily lives. Think about the runaway success, a few years ago, of Dave Eggers’s *The Circle*—a best-selling novel, adapted into a film starring Tom Hanks, in which a fictional company that resembles Facebook encourages its employees and customers to share all their information all the time, regardless of the consequences.

Despite many deep-seated real-world and fictional concerns about unbridled uses of our personal information, a lot of researchers and writers have suggested that we should, in fact, all be more transparent about our personal information. As dramatized in *The Circle*, one suggestion is that if we could all know everything about one another, this would level the playing field in many ways. The reasons for decisions on housing, employment, and other matters would be made clear and transparent. We would ultimately lose our fears about privacy and secrecy, because no one would have any significant expectations of privacy at all, in a world of transparency enforced for everyone.¹⁰

A significant problem with this approach comes down to aggregation of data. There is simply too much data for any human to process effectively at a
IS BIG BROTHER TRACKING ME?

Radio frequency identification (RFID) is a technology that uses electromagnetic fields to track digital tags attached to physical objects and devices. Unlike bar codes (those black and white codes affixed to products in the grocery store), RFID tags can be monitored at a distance. You do not need a digital reader in the proximity of the device to locate and “read” its information. RFID microchips are implanted in livestock and pets to help find them if they are lost. They are implanted into digital devices, notably automobiles, to find them if they become stolen or lost, or simply to track them for work or other purposes. Theme parks even use RFID tags to track where you are sitting in a restaurant in order to deliver food you preordered.

In terms of personal privacy concerns, it is, of course, possible to implant an RFID tag under a person’s skin or in the clothes they wear or items they carry. Concerns with RFID have led to various agreements internationally about limiting uses of the technology in relation to tracking and release of personal information. RFID tags mainly track product information. However, the collection of product information in relation to a particular person has the potential to create personal data profiles without the knowledge of the individual.

Similar concerns arise with respect to the cell site location information that is triangulated from cell towers, which enables cell phone service providers like AT&T, Verizon, and Sprint to locate a specific cell phone. In 2018, the Supreme Court held in essence that it was a violation of the Fourth Amendment’s constitutional guarantee against unreasonable searches and seizures for the government to search a suspect’s cell phone location records without a warrant, at least in the particular circumstances that arose in the case. This was a landmark decision that overturned previous law on governmental use of cell phone location records. We will look more at government intrusions into personal life in chapter 11.
given time. Having access to voluminous amounts of data does not necessarily help anyone make better decisions, at least not without machine learning as an aid. Additionally, most of us simply do not want everyone else in the world to have access to all information about us all the time. Our actions may be embarrassing or damning, or maybe we just want to be left alone. That is a big problem with Big Data: we never know who has access to what information about us, or how and when it might be used to harm, embarrass, or simply annoy us.

The bottom line is that digital technology enables more collection, aggregation, and use of personal information than ever before. Again, this is neither good nor bad, in and of itself—data can be used for both beneficial and harmful purposes. We may like the beneficial aspects of Big Data and fear the potential harms. On one hand, we may be uncomfortable with the thought of large amounts of our personal information being collected by entities outside our control. On the other hand, we may be okay with our information being collected if it is anonymized—although, with modern technology, it is frighteningly easy to de-anonymize data and identify the subject of specific information.

So how do we proceed from here?

Each chapter of this book will target a particular area of daily life, highlighting when and how personal information may be collected and used and making suggestions about how best to monitor and protect that information. In our digital reality, it is simply not realistic to attempt to claw back absolute privacy rights to personal information. Privacy can never be an absolute right in any event. Even in countries that protect privacy as a constitutional or human right, that right has to be balanced against rights and concerns like national security, public health, and free speech. We will look at where those balances play out currently in the United States, and how the situation in this country differs from that in other countries (and why).

Whether you choose to read from cover to cover or dip in and out of chapters that interest you, the pages that follow will offer some explanations, and maybe even some comfort, about your data and how it is used, along with many useful tips and tricks. Chapter 1 sets the scene by explaining, in simple terms, the issue of who (if anyone) “owns” our data and in what contexts. The answers are not as obvious as you might think!