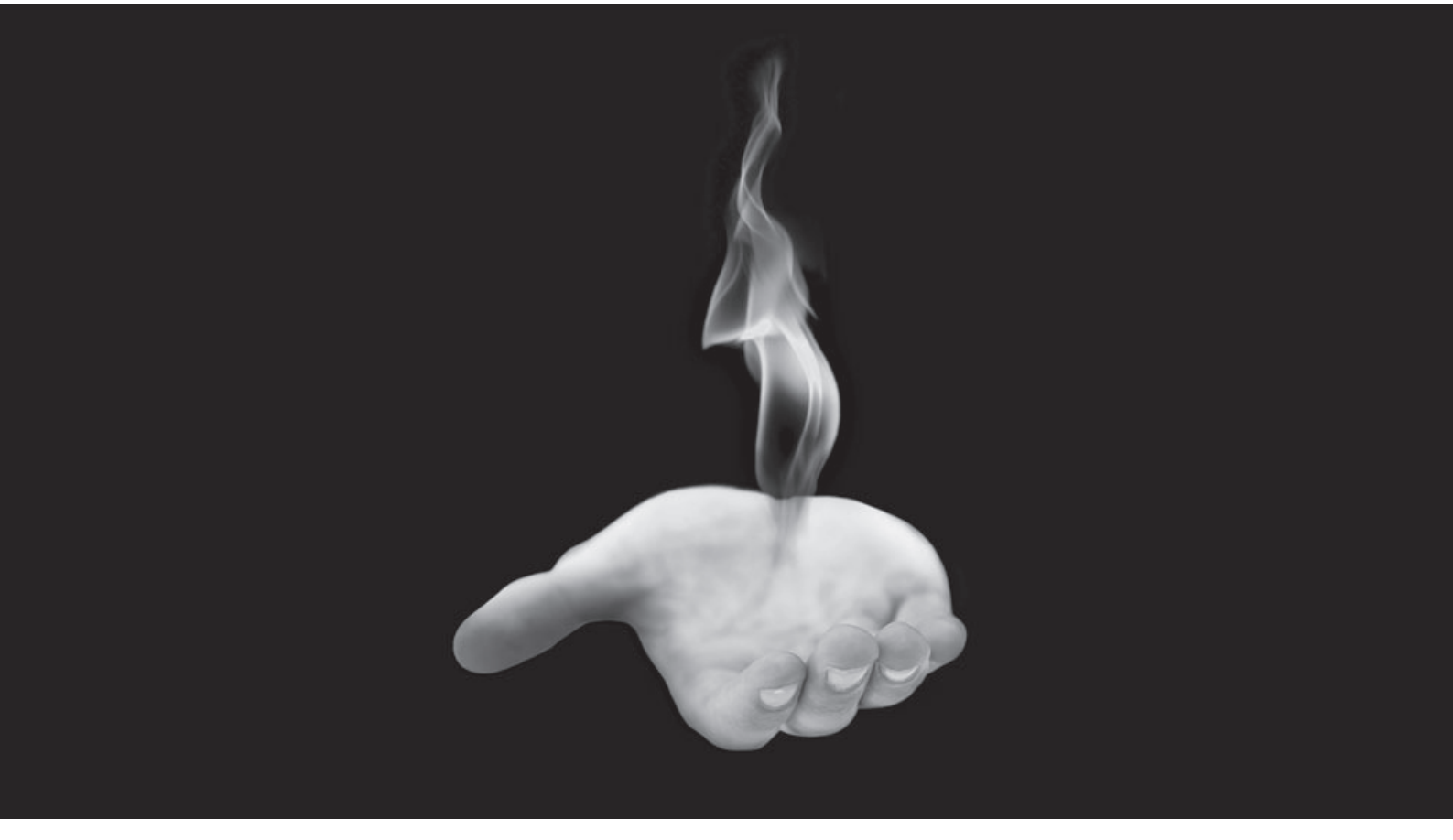
2

Privacy



2.1 Privacy Risks and Principles

2.1.1 What Is Privacy?

After the fall of the communist government in East Germany, people examined the files of Stasi, the secret police. They found that the government had used spies and informants to build detailed dossiers on the opinions and activities of roughly six million people—a third of the population. The informers were neighbors, co-workers, friends, and even family members of the people they reported on. The paper files filled an estimated 125 miles of shelf space.[1]

Before the digital age, surveillance cameras watched shoppers in banks and stores. And well into the era of computers and the Internet, pharmacies in Indiana disposed of hundreds of prescriptions, receipts, and order forms for medicines by tossing them into an open dumpster. Private investigators still search household garbage for medical and financial information, details of purchases, evidence of romantic affairs, and journalists' notes.

Computer technology is not necessary for the invasion of privacy. However, we discuss privacy at length in this book because the use of digital technology has made new threats possible and old threats more potent. Computer technologies—databases, digital cameras, the Web, smartphones, and global positioning system (GPS) devices, among others—have profoundly changed what people can know about us and how they can use that information. Understanding the risks and problems is a first step toward protecting privacy. For computer professionals, understanding the risks and problems is a step toward designing systems with built-in privacy protections and less risk.

There are three key aspects of privacy:

- Freedom from intrusion—being left alone
- Control of information about oneself
- Freedom from surveillance (from being followed, tracked, watched, and eavesdropped upon)

For the most part, in this book, we view privacy as a good thing. Critics of privacy argue that it gives cover to deception, hypocrisy, and wrongdoing. It allows fraud. It protects the guilty. Concern for privacy may be regarded with a suspicious "What do you have to hide?" The desire to keep things private does not mean we are doing anything wrong. We might wish to keep health, relationship, and family issues private. We might wish to keep religious beliefs and political views private from some of the people we interact with. Privacy of some kinds of information can be important to safety and security as well. Examples include travel plans, financial data, and for some people, simply a home address.

Privacy threats come in several categories:

- Intentional, institutional uses of personal information (in the government sector primarily for law enforcement and tax collection, and in the private sector primarily for marketing and decision making)
- Unauthorized use or release by "insiders," the people who maintain the information
- Theft of information
- Inadvertent leakage of information through negligence or carelessness
- Our own actions (sometimes intentional trade-offs and sometimes when we are unaware of the risks)

Privacy issues arise in many contexts. More topics with privacy implications appear in later chapters. We discuss spam, the intrusion of junk email and text messages, in Chapter 3. We address hacking and identity theft in Chapter 5. We discuss monitoring of workplace communications and other issues of privacy for employees in Chapter 6. Some privacy risks result from the fact that so much of the data stored about us is incorrect. Databases contain errors. Files are not updated. Records of different people with similar names or other similarities get comingled or confused. Chapter 8 discusses some of these problems. Privacy comes up again in Chapter 9, where we focus on the responsibilities of computer professionals.

It is clear that we cannot expect complete privacy. We usually do not accuse someone who initiates a conversation of invading our privacy. Many friends and slight acquaintances know what you look like, where you work, what kind of car you drive, and whether you are a nice person. They need not get your permission to observe and talk about you. Control of information about oneself means control of what is in other people's minds, phones, and data storage systems. It is necessarily limited by basic human rights, particularly freedom of speech. Nor can we expect to be totally free from surveillance. People see us and hear us when we move about in public (physically or on the Web).

If you live in a small town, you have little privacy; everyone knows everything about you. In a big city, you are more nearly anonymous. But if people know nothing about you, they might be taking a big risk if they rent you a place to live, hire you, lend you money, sell you automobile insurance, accept your credit card, and so on. We give up some privacy for the benefits of dealing with strangers. We can choose to give up more in exchange for other benefits such as convenience, personalized service, and easy communication with many friends. But sometimes, others make the choices for us.

I use many real incidents, businesses, products, and services as examples throughout this book. In most cases, I am not singling them out for special endorsement or criticism. They are just some of the many examples we can use to illustrate problems, issues, and possible solutions.

> *The man who is compelled to live every minute of his life among others and whose every need, thought, desire, fancy or gratification is subject to public scrutiny, has been deprived of his individuality and human dignity. [He] merges with the mass. . . . Such a being, although sentient, is fungible; he is not an individual.*
>
> —Edward J. Bloustein[2]

> *It's important to realize that privacy preserves not personal secrets, but a sense of safety within a circle of friends so that the individual can be more candid, more expressive, more* open *with "secrets."*
>
> —Robert Ellis Smith[3]

2.1.2 New Technology, New Risks

Computers, the Internet, and a whole array of digital devices—with their astounding increases in speed, storage space, and connectivity—make the collection, searching, analysis, storage, access, and distribution of huge amounts of information and images much easier, cheaper, and faster than ever before. These are great benefits. But when the information is about us, the same capabilities threaten our privacy.

Today there are thousands (probably millions) of databases, both government and private, containing personal information about us. In the past, there was simply no record of some of this information, such as our specific purchases of groceries and books. Government documents like divorce and bankruptcy records have long been in public records, but accessing such information took a lot of time and effort. When we browsed in a library or store, no one knew what we read or looked at. It was not easy to link together our financial, work, and family records. Now, large companies that operate video, email, social network, and search services can combine information from a member's use of all of them to obtain a detailed picture of the person's interests, opinions, realtionships, habits, and activities. Even if we do not log in as members, software tracks our activity on the Web. In the past, conversations disappeared when people finished speaking, and only the sender and the recipient normally read personal communications. Now, when we communicate by texting, email, social networks, and so on, there is a record of our words that others can copy, forward, distribute widely, and read years later. Miniaturization of processors and sensors put tiny cameras in cellphones that millions of people carry everywhere. Cameras in some 3-D television sets warn children if they are sitting too close. What else might such cameras record, and who might see it? The wireless appliances we carry contain GPS and other location devices. They enable others to determine our location and track our movements. Patients refill prescriptions and check the results of medical tests on the Web. They correspond with doctors by email. We store our photos

and videos, do our taxes, and create and store documents and financial spreadsheets in a cloud of remote servers instead of on our own computer. Power and water providers might soon have metering and analysis systems sophisticated enough to deduce what appliances we are using, when we shower (and for how long), and when we sleep. Law enforcement agencies have very sophisticated tools for eavesdropping, surveillance, and collecting and analyzing data about people's activities, tools that can help reduce crime and increase security—or threaten privacy and liberty.

Combining powerful new tools and applications can have astonishing results. It is possible to snap a photo of someone on the street, match the photo to one on a social network, and use a trove of publicly accessible information to guess, with high probability of accuracy, the person's name, birth date, and most of his or her Social Security number. This does not require a supercomputer; it is done with a smartphone app. We see such systems in television shows and movies, but to most people they seem exaggerated or way off in the future.

All these gadgets, services, and activities have benefits, of course, but they expose us to new risks. The implications for privacy are profound.

> *Patient medical information is confidential. It should not be discussed in a public place.*
>
> —A sign, aimed at doctors and staff, in an elevator in a medical office building, a reminder to prevent low-tech privacy leaks.

Example: Search query data

After a person enters a phrase into a search engine, views some results, then goes on to another task, he or she expects that the phrase is gone—gone like a conversation with a friend or a few words spoken to a clerk in a store. After all, with millions of people doing searches each day for work, school, or personal uses, how could the search company store it all? And who would want all that trivial information anyway? That is what most people thought about search queries until two incidents demonstrated that it is indeed stored, it can be released, and it matters.

Search engines collect many terabytes of data daily. A terabyte is a trillion bytes. It would have been absurdly expensive to store that much data in the recent past, but no longer. Why do search engine companies store search queries? It is tempting to say "because they can." But there are many uses for the data. Suppose, for example, you search for "Milky Way." Whether you get lots of astronomy pages or information about the candy bar or a local restaurant can depend on your search history and other information about you. Search engine companies want to know how many pages of search results users actually look at, how many they click on, how they refine their search queries, and what spelling errors they commonly make. The companies analyze the data to improve

search services, to target advertising better, and to develop new services. The database of past queries also provides realistic input for testing and evaluating modifications in the algorithms search engines use to select and rank results. Search query data are valuable to many companies besides search engine companies. By analyzing search queries, companies draw conclusions about what kinds of products and features people are looking for. They modify their products to meet consumer preferences.

But who else gets to see this mass of data? And why should we care?

If your own Web searches have been on innocuous topics, and you do not care who sees your queries, consider a few topics people might search for and think about why they might want to keep them private: health and psychological problems, bankruptcy, uncontrolled gambling, right-wing conspiracies, left-wing conspiracies, alcoholism, anti-abortion information, pro-abortion information, erotica, illegal drugs. What are some possible consequences for a person doing extensive research on the Web for a suspense novel about terrorists who plan to blow up chemical factories?

In 2006, the federal government presented Google with a subpoena* for two months of user search queries and all the Web addresses† that Google indexes.‡ Google protested, bringing the issue to public attention. Although the subpoena did not ask for names of users, the idea of the government gaining access to the details of people's searches horrified privacy advocates and many people who use search engines. Google and privacy advocates opposed the precedent of government access to large masses of such data. A court reduced the scope of the subpoena, removing user queries.[4]

A few months later, release of a huge database of search queries at AOL showed that privacy violations occur even when the company does not associate the queries with people's names. Against company policy, an employee put the data on a website for search technology researchers. This data included more than 20 million search queries by more than 650,000 people from a three-month period. The data identified people by coded ID numbers, not by name. However, it was not difficult to deduce the identity of some people, especially those who searched on their own name or address. A process called *re-identification* identified others. Re-identification means identifying the individual from a set of anonymous data. Journalists and acquaintances identified people in small communities who searched on numerous specific topics, such as the cars they own, the sports teams they follow, their health problems, and their hobbies. Once identified, a person is linked to all his or her other searches. AOL quickly removed the data, but journalists,

---

* A subpoena is a court order for someone to give testimony or provide documents or other information for an investigation or a trial.

† We use the term Web address informally for identifiers, or addresses, or URLs of pages or documents on the Web (the string of characters one types in a Web browser).

‡ It wanted the data to respond to court challenges to the Child Online Protection Act (COPA), a law intended to protect children from online material "harmful to minors." (We discuss COPA in Section 3.2.2.)

researchers, and others had already copied it. Some made the whole data set available on the Web again.[5]*

Example: Smartphones

With so many clever, useful, and free smartphone apps available, who thinks twice about downloading them? Researchers and journalists took a close look at smartphone software and apps and found some surprises.

Some Android phones and iPhones send location data (essentially the location of nearby cell towers) to Google and Apple, respectively. Companies use the data to build location-based services that can be quite valuable for the public and for the companies. (Industry researchers estimate the market for location services to be in the billions of dollars.) The location data is supposed to be anonymous, but researchers found, in some cases, that it included the phone ID.

Roughly half the apps in one test sent the phone's ID number or location to other companies (in addition to the one that provided the app). Some sent age and gender information to advertising companies. The apps sent the data without the user's knowledge or consent. Various apps copy the user's contact list to remote servers. Android phones and iPhones allow apps to copy photos (and, for example, post them on the Internet) if the user permits the app to do certain other things that have nothing to do with photos. (Google said this capability dated from when photos were on removable memory cards and thus less vulnerable.[6] This is a reminder that designers must regularly review and update security design decisions.)

A major bank announced that its free mobile banking app inadvertently stored account numbers and security access codes in a hidden file on the user's phone. A phone maker found a flaw in its phones that allowed apps to access email addresses and texting data without the owner's permission. Some iPhones stored months of data, in a hidden file, about where the phone had been and when, even if the user had turned off location services. Data in such files are vulnerable to loss, hacking, and misuse. If you do not know the phone stores the information, you do not know to erase it. Given the complexity of smartphone software, it is possible that the companies honestly did not intend the phones to do these things.†

Why does it matter? Our contact lists and photos are ours; we should have control of them. Thieves can use our account information to rob us. Apps use features on phones that indicate the phone's location, the light level, movement of the phone, the presence of other phones nearby, and so on. Knowing where we have been over a period of time (combined with other information from a phone) can tell a lot about our activities and

* Members of AOL sued the company for releasing their search queries, claiming the release violated roughly 10 federal and state laws.

† The various companies provided software updates for these problems.

1. Files on hundreds of thousands of students, applicants, faculty, and/or alumni from the University of California, Harvard, Georgia Tech, Kent State, and several other universities, some with Social Security numbers and birth dates (stolen by hackers).
2. Names, birth dates, and possibly credit card numbers of 77 million people who play video games online using Sony's PlayStation (stolen by hackers). Another 24 million accounts were exposed when hackers broke into Sony Online Entertainment's PC-game service.
3. Records of roughly 40 million customers of TJX discount clothing stores (T.J. Maxx, Marshalls, and others), including credit and debit card numbers and some driver's license numbers (stolen by hackers).

More about the TJX incident: Section 5.2.5

4. Bank of America disks with account information (lost or stolen in transit).
5. Credit histories and other personal data for 163,000 people (purchased from a huge database company by a fraud ring posing as legitimate businesses).
6. Patient names, Social Security numbers, addresses, dates of birth, and medical billing information for perhaps 400,000 patients at a hospital (on a laptop stolen from a hospital employee's car).
7. More than 1000 Commerce Department laptops, some with personal data from Census questionnaires. (Thieves stole some from the cars of temporary Census employees; others, employees simply kept.)
8. Confidential contact information for more than one million job seekers (stolen from Monster.com by hackers using servers in Ukraine).

Figure 2.1 Lost or stolen personal information.[7]

interests, as well as with whom we associate (and whether the lights were on). As we mentioned in Section 1.2.1, it can also indicate where we are likely to be at a particular time in the future.

Some of the problems we described here will have been addressed by the time you read this; the point is that we are likely to see similar (but similarly unexpected) privacy risks and breaches in each new kind of gadget or capability.

Stolen and lost data

Criminals steal personal data by hacking into computer systems, by stealing computers and disks, by buying or requesting records under false pretenses, and by bribing employees of companies that store the data. Shady information brokers sell data (including cellphone records, credit reports, credit card statements, medical and work records, and location of relatives, as well as information about financial and investment accounts) that they obtain illegally or by questionable means. Criminals, lawyers, private investigators, spouses, ex-spouses, and law enforcement agents are among the buyers. A private investigator could have obtained some of this information in the past, but not nearly so easily, cheaply, and quickly.

Hacking: Section 5.2

Another risk is accidental (sometimes quite careless) loss. Businesses, government agencies, and other institutions lose computers, disks, memory cards, and laptops containing sensitive personal data (such as Social Security numbers and credit card numbers) on thousands or millions of people, exposing people to potential misuse of their information and lingering uncertainty. They inadvertently allow sensitive files to be public on the Web. Researchers found medical information, Social Security numbers, and other sensitive personal or confidential information about thousands of people in files on the Web that simply had the wrong access status.

The websites of some businesses, organizations, and government agencies that make account information available on the Web do not sufficiently authenticate the person accessing the information, allowing imposters access. Data thieves often get sensitive information by telephone by pretending to be the person whose records they seek. They provide some personal information about their target to make their request seem legitimate. That is one reason why it is important to be cautious even with data that is not particularly sensitive by itself.

More about authentication techniques: Section 5.3.2

Figure 2.1 shows a small sample of incidents of stolen or lost personal information (the Privacy Rights Clearinghouse lists thousands of such incidents on its website). In many incidents, the goal of thieves is to collect data for use in identity theft and fraud, crimes we discuss in detail in Chapter 5.

A summary of risks

The examples we described illustrate numerous points about personal data. We summarize here:

- Anything we do in cyberspace is recorded, at least briefly, and linked to our computer or phone, and possibly our name.
- With the huge amount of storage space available, companies, organizations, and governments save huge amounts of data that no one would have imagined saving in the recent past.
- People often are not aware of the collection of information about them and their activities.
- Software is extremely complex. Sometimes businesses, organizations, and website managers do not even know what the software they use collects and stores.[8]
- Leaks happen. The existence of the data presents a risk.
- A collection of many small items of information can give a fairly detailed picture of a person's life.
- Direct association with a person's name is not essential for compromising privacy. Re-identification has become much easier due to the quantity of personal information stored and the power of data search and analysis tools.

- If information is on a public website, people other than those for whom it was intended will find it. It is available to everyone.
- Once information goes on the Internet or into a database, it seems to last forever. People (and automated software) quickly make and distribute copies. It is almost impossible to remove released information from circulation.
- It is extremely likely that data collected for one purpose (such as making a phone call or responding to a search query) will find other uses (such as business planning, tracking, marketing, or criminal investigations).
- The government sometimes requests or demands sensitive personal data held by businesses and organizations.
- We often cannot directly protect information about ourselves. We depend on the businesses and organizations that manage it to protect it from thieves, accidental collection, leaks, and government prying.

2.1.3 Terminology and Principles for Managing Personal Data

We use the term *personal information* often in this chapter. In the context of privacy issues, it includes any information relating to, or traceable to, an individual person. The term does not apply solely to what we might think of as sensitive information, although it includes that. It also includes information associated with a particular person's "handle," user name, online nickname, identification number, email address, or phone number. Nor does it refer only to text. It extends to any information, including images, from which someone can identify a living individual.

Informed consent and invisible information gathering

The first principle for ethical treatment of personal information is *informed consent*. There is an extraordinary range to the amount of privacy different people want. Some blog about their divorce or illnesses. Some pour out details of their romantic relationships on television shows or to hundreds of social network friends. Others use cash to avoid leaving a record of their purchases, encrypt all their email,* and are angry when someone collects information about them. When a business or organization informs people about its data collection and use policies or about the data that a particular device or application collects, each person can decide, according to his or her own values, whether or not to interact with that business or organization or whether to use the device or application.

Invisible information gathering describes collection of personal information without the person's knowledge. The important ethical issue is that if someone is not aware of the collection and use, he or she has no opportunity to consent or withhold consent. We gave

* Encrypting data means putting it in a coded form so that others cannot read it.

several examples involving smartphones and their apps in the previous section. Here are examples from other contexts.

- A company offered a free program that changed a Web browser's cursor into a cartoon character. Millions of people installed the program but then later discovered that the program sent to the company a report of the websites its users visited, along with a customer identification number in the software.[9]
- "Event data recorders" in cars record driving speed, whether or not the driver is wearing a seatbelt, and other information.
- "History sniffers" are programs that collect information about a person's online activity based on the different colors a browser uses to display sites recently visited.
- Software called *spyware*, often downloaded from a website without the user's knowledge, surreptitiously collects information about a person's activity and data on his or her computer and then sends the information over the Internet to the person or company that planted the spyware. Spyware can track someone's Web surfing for an advertising company or collect passwords and credit card numbers typed by the user. (Some of these activities are illegal, of course.)


Sophisticated snooping technologies: Section 2.2.2

When our computers and phones communicate with websites, they must provide information about their configuration (e.g., the Web browser used). For a high percentage of computers, there is enough variation and detail in configurations to create a "fingerprint" for each computer. Some companies provide device fingerprinting software for combating fraud and intellectual property theft and for tracking people's online activity in order to target advertising. Both collection of configuration information and building of activity profiles are invisible. Financial firms that use device fingerprinting for security of customer accounts are likely to say so in a privacy policy. We are less likely to know when someone is using it to build marketing profiles.

Whether or not a particular example of data collection is invisible information gathering can depend on the level of public awareness. Some people know about event data recorders in cars; most do not.[10] Before the release of AOL user search data described in Section 2.1.2, collecting search query data was an example of invisible information gathering; for many people it still is. Many businesses and organizations have policy statements or customer agreements that inform customers, members, and subscribers of their policy on collecting and using personal data, but many people simply do not read them. And if they read them, they forget. Thus, there can be a significant privacy impact from the many automated systems that collect information in unobvious ways, even when people have been informed. However, there is an important distinction between situations where people are informed but not aware and situations where the information gathering is truly covert, such as in spyware and in some of the smartphone apps we described in Section 2.1.2.


A legal remedy for secret data collection: Section 5.2.6

Cookies

Cookies are files a website stores on a visitor's computer.[11] Within the cookie, the site stores and then uses information about the visitor's activity. For example, a retail site might store information about products we looked at and the contents of our virtual "shopping cart." On subsequent visits, the site retrieves information from the cookie. Cookies help companies provide personalized customer service and target advertising to the interests of each visitor. They can also track our activities on many sites and combine the information. At first, cookies were controversial because the very idea that websites were storing files on the user's computer without the user's knowledge startled and disturbed people. Today, more people are aware of cookies and use tools to prevent or delete them. In response, some companies that track online activity developed more sophisticated "supercookies" that recreate deleted cookies and are difficult to find and remove.

Secondary use, data mining, matching, and profiling

> *My most private thoughts, my personal tragedies, secrets about other people, are mere data of a transaction, like a grocery receipt.*
>
> —A woman whose psychologist's notes were read by an insurer.[12]

Secondary use is the use of personal information for a purpose other than the one for which the person supplied it. Examples include sale of consumer information to marketers or other businesses, use of information in various databases to deny someone a job or to tailor a political pitch, the Internal Revenue Service searching vehicle registration records for people who own expensive cars and boats (to find people with high incomes), use of text messages by police to prosecute someone for a crime, and the use of a supermarket's customer database to show alcohol purchases by a man who sued the store because he fell down.

Data mining means searching and analyzing masses of data to find patterns and develop new information or knowledge. The research using social network data and smartphone data that we described in Section 1.2.1 are examples. *Matching* means combining and comparing information from different databases, often using an identifier such as a person's Social Security number or their computer's Internet address to match records. *Profiling* means analyzing data to determine characteristics of people most likely to engage in certain behavior. Businesses use these techniques to find likely new customers. Government agencies use them to detect fraud, to enforce other laws, and to find terrorists. Data mining, computer matching, and profiling are, in most cases, examples of secondary use of personal information.

We will see examples of secondary use throughout this chapter. One of the controversial issues about personal information is the degree of control people should have over secondary uses of information about them. The variety of uses illustrated by the few examples we gave above suggests that quite different answers are appropriate for different users and different uses.

After informing people about what personal information an organization collects and what it does with that information, the next simplest and most desirable privacy policy is to give people some control over secondary uses. The two most common forms for providing such choice are *opt out* and *opt in*. Under an opt-out policy, one must check or click a box on a contract, membership form, or agreement or contact the organization to request that they not use one's information in a particular way. If the person does not take action, the presumption is that the organization may use the information. Under an opt-in policy, the collector of the information may not use it for secondary uses unless the person explicitly checks or clicks a box or signs a form permitting the use. (Be careful not to confuse the two. Under an opt-out policy, more people are likely to be "in," and under an opt-in policy, more people are likely to be "out," because the default presumption is the opposite of the policy name.) Opt-out options are now common. Responsible, consumer-friendly companies and organizations often set the default so that they do not share personal information and do not send marketing emails unless the person explicitly allows it—that is, they use the opt-in policy. Particularly in situations where disclosing personal information can have negative consequences and it is not obvious to a customer that the organization might disclose it, a default of nondisclosure without explicit permission (that is, an opt-in policy) is the responsible policy.

Fair information principles

Privacy advocates have developed various sets of principles for protection of personal data. They are often called Fair Information Principles or Fair Information Practices.[13] Figure 2.2 presents such a list of principles. Informed consent and restrictions on secondary uses show up in the first and third principles. You will rarely see the last point in Figure 2.2 included among Fair Information Principles, but I consider it an important one. Some companies and organizations turn over personal data to law enforcement agents and government agencies when requested. Some do so only if presented with a subpoena or other court order. Some challenge subpoenas; some do not. Some inform their customers or members when they give personal data to the government; some do not. The entity that holds the data decides how far to go in protecting the privacy of its members or customers. The individual whose data the entity might release is rarely aware of the government request. Thus, the entities that hold the data have a responsibility to those people. Planning ahead for various possible scenarios, developing a policy, and announcing it (and following it) are all part of responsible management of other people's personal data.

1. Inform people when you collect information about them, what you collect, and how you use it.
2. Collect only the data needed.
3. Offer a way for people to opt out from mailing lists, advertising, and other secondary uses. Offer a way for people to opt out from features and services that expose personal information.
4. Keep data only as long as needed.
5. Maintain accuracy of data. Where appropriate and reasonable, provide a way for people to access and correct data stored about them.
6. Protect security of data (from theft and from accidental leaks). Provide stronger protection for sensitive data.
7. Develop policies for responding to law enforcement requests for data.

Figure 2.2 Privacy principles for personal information.

Many businesses and organizations have adopted some version of Fair Information Practices. Laws in the United States, Canada, and European countries (among others) require them in many situations. These principles are reasonable ethical guidelines. However, there is wide variation in interpretation of the principles. For example, businesses and privacy advocates disagree about what information businesses "need" and for how long.

It can be difficult to apply the fair information principles to some new technologies and applications. They do not fully address privacy issues that have arisen with the increase of cameras in public places (such as police camera systems and Google's Street View), the enormous amount of personal information people share in social networks, and the ubiquity and power of smartphones. For example, when someone puts personal information in a tweet to thousands of people, how do we determine the purpose for which he or she supplied the information? Can any recipient use the information in any way? How widely distributed must information be before it is public in the sense that anyone can see or use it? Even when people have agreed to share information, consequences of new ways of sharing or new categories of information can be unexpected and problematic. For example, in Section 2.3.2 we discuss default settings for features in social networks that have significant consequences.

Employers search employee social media: Section 6.3.1

2.2 The Fourth Amendment, Expectation of Privacy, and Surveillance Technologies

In George Orwell's dystopian novel *1984*, Big Brother (the government) could watch everyone via "telescreens" in all homes and public places. There was little crime and little

political dissent—and no love and no freedom. Today, the government does not have to watch every move we make, because so many of our activities leave data trails in databases available to government agencies.* When Big Brother wants to take a direct look at us and our activities, he uses sophisticated new surveillance tools. In this section, we consider the impact of these tools on privacy and look into their compatibility with constitutional and legal protections from government intrusions.

2.2.1 The Fourth Amendment

> *The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.*
>
> —Fourth Amendment, U.S. Constitution

The U.S. Constitution protects a right to privacy from government intrusion, most explicitly in the Fourth Amendment. The U.S. Supreme Court has interpreted other parts of the Bill of Rights to provide a constitutional right to privacy from government in other areas as well. England has a similar tradition, as expressed in William Pitt's colorful statement in 1763:

> The poorest man may in his cottage bid defiance to all the force of the Crown. It may be frail; its roof may shake; the wind may blow through it; the storms may enter; the rain may enter—but the King of England cannot enter . . . .[14]

Here, we look at how databases, surveillance technology, and popular consumer gadgets threaten this right. Although the discussion in this section is in the context of the U.S. Fourth Amendment and U.S. Supreme Court rulings, the new technological risks of intrusion by governments are similar in other countries.

The Fourth Amendment sets limits on the government's rights to search our homes and businesses and to seize documents and other personal effects. It requires that the government have probable cause for the search and seizure. That is, there must be good evidence to support the specific search. Two key problems arise from new technologies. First, much of our personal information is no longer safe in our homes or the individual offices of our doctors and financial advisors. We carry a huge amount of personal information on smartphones and laptops. Much personal information is in huge databases outside of our control. Many laws allow law enforcement agencies to get information from nongovernment databases without a court order. Federal privacy rules allow law enforcement agencies to access medical records without court orders. The USA PATRIOT

* The use of myriad personal-data systems to investigate or monitor people is sometimes called *dataveillance*, short for "data surveillance."

Act (passed after the terrorist attacks in 2001) eased government access to many kinds of personal information, including library and financial records, without a court order. The second factor weakening Fourth Amendment protections is that new technologies allow the government to search our homes without entering them, to search our persons from a distance without our knowledge, and to extract all the data on a cellphone (including deleted data and password protected data) in less than two minutes at a traffic stop.

As we consider all the personal information available to government agencies now, we can reflect on the worries of Supreme Court Justice William O. Douglas about the potential abuse from government access to only the records of someone's checking account. In 1974, he said:

> In a sense a person is defined by the checks he writes. By examining them agents get to know his doctors, lawyers, creditors, political allies, social connections, religious affiliation, educational interests, the papers and magazines he reads, and so on ad infinitum. These are all tied in to one's social security number, and now that we have the data banks, these other items will enrich that storehouse and make it possible for a bureaucrat—by pushing one button—to get in an instant the names of the 190 million Americans who are subversives or potential and likely candidates.[15]

Today's readers should not miss the irony of the last sentence: 190 million was almost the entire population of the United States at the time.

With each new data storage or search technology, law enforcement agencies and civil libertarians argue the question of whether the Fourth Amendment applies. In the next few sections, we discuss such technologies and some principles the Supreme Court has established.

> *When the American Republic was founded, the framers established a libertarian equilibrium among the competing values of privacy, disclosure, and surveillance. This balance was based on technological realities of eighteenth-century life. Since torture and inquisition were the only known means of penetrating the mind, all such measures by government were forbidden by law. Physical entry and eavesdropping were the only means of penetrating private homes and meeting rooms; the framers therefore made eavesdropping by private persons a crime and allowed government to enter private premises only for reasonable searches, under strict warrant controls. Since registration procedures and police dossiers were the means used to control the free movement of "controversial" persons, this European police practice was precluded by American governmental practice and the realities of mobile frontier life.*
>
> —Alan F. Westin, *Privacy and Freedom*[16]

2.2.2 New Technologies, Supreme Court Decisions, and Expectation of Privacy

> *The principles laid down in this opinion . . . apply to all invasions on the part of government and its employees of the sanctity of a man's home and the privacies of life. It is not the breaking of his doors, and the rummaging in his drawers, that constitutes the essence of the offense; but it is the invasion of his indefeasible right of personal security, personal liberty and private property.*
>
> —Justice Joseph Bradley, *Boyd v. United States*, 1886.

"Noninvasive but deeply revealing" searches

The title above is from Julian Sanchez's description of a variety of search and detection technologies.[17] Many sound like science fiction; they are not. These technologies can search our homes and vehicles but do not require police to physically enter or open them. They can search our bodies beneath our clothes from a distance without our knowledge. What restrictions should we place on their use? When should we permit government agencies to use them without a search warrant?

Noninvasive but deeply revealing search tools (some in use and some in development) include particle sniffers that detect many specific drugs and explosives, imaging systems that detect guns under clothing from a distance, devices that analyze the molecular composition of truck cargo without opening the truck, thermal-imaging devices (to find heat lamps for growing marijuana, for example), and devices that locate a person by locating his or her cellphone. These devices have obvious valuable security and law enforcement applications, but the technologies can be used for random searches, without search warrants or probable cause, on unsuspecting people. As Sanchez points out, we live "in a nation whose reams of regulations make almost everyone guilty of some violation at some point."[18] Before the government begins using these tools on, say, ordinary people bringing medications home from Canada, making their own beer, or keeping a banned sweetener or saturated fat in their home (or whatever might be illegal in the future), it is critical for privacy protection that we have clear guidelines for their use—and, in particular, clarification of when such use constitutes a search requiring a search warrant.

Supreme Court decisions and expectation of privacy

Several Supreme Court cases have addressed the impact of earlier technology on Fourth Amendment protection. In *Olmstead v. United States*,[19] in 1928, the government had used wiretaps on telephone lines without a court order. The Supreme Court allowed the wiretaps. It interpreted the Fourth Amendment to apply only to physical intrusion and only to the search or seizure of material things, not conversations. Justice Louis Brandeis

dissented, arguing that the authors of the Fourth Amendment did all they could to protect liberty and privacy—including privacy of conversations—from intrusions by government based on the technology available at the time. He believed that the court should interpret the Fourth Amendment as requiring a court order even when new technologies give the government access to our personal papers and conversations without entering our homes. In *Katz v. United States*, in 1967, the Supreme Court reversed its position and ruled that the Fourth Amendment does apply to conversations and that it applies in public places in some situations. In this case, law enforcement agents had attached an electronic listening and recording device on the outside of a telephone booth to record a suspect's conversation. The court said that the Fourth Amendment "protects people, not places," and that what a person "seeks to preserve as private, even in an area accessible to the public, may be constitutionally protected." To intrude in places where a reasonable person has a reasonable expectation of privacy, government agents need a court order.

Although the Supreme Court's decision in *Katz v. United States* strengthened Fourth Amendment protection in some ways, there is significant risk in relying on reasonable "expectation of privacy" to define the areas where law enforcement agents need a court order. Consider the two technologies in the box nearby. One tracks private actions in public view; the other tracks people in private places.

As well-informed people come to understand the capabilities of modern surveillance tools, we might no longer expect privacy from government, in a practical sense. Does that mean we should not have it? The Supreme Court recognized this problem in *Smith v. Maryland*, in which it noted that, if law enforcement reduces actual expectation of privacy by actions "alien to well-recognized Fourth Amendment freedoms," this should *not* reduce our Fourth Amendment protection. However, the Court has interpreted "expectation of privacy" in a very restrictive way. For example, it ruled that if we share information with businesses such as our bank, then we have no reasonable expectation of privacy for that information (*United States v. Miller*, 1976). Law enforcement agents do not need a court order to get the information. This interpretation seems odd. We do expect privacy of the financial information we supply a bank or other financial institution. We expect confidentiality in many kinds of information we share with a few, sometimes carefully selected, others. We share our Web activity with ISPs, websites, and search engine companies merely by typing and clicking. We share many kinds of personal information at specific websites where we expect it to be private. Is it safe from warrantless search?

In *Kyllo v. United States* (2001), the Supreme Court ruled that police could not use thermal-imaging devices to search a home from the outside without a search warrant. The Court stated that where "government uses a device that is not in general public use, to explore details of the home that would previously have been unknowable without physical intrusion, the surveillance is a 'search.'" This reasoning suggests that when a technology becomes more widely used, the government may use it for surveillance without a warrant.

Tracking cars and cellphones

Law enforcement agents track thousands of people's locations each year. Sometime they have a court order to do so, and some times they do not. Do they need one? We describe two key cases as examples.

In 2012, the Supreme Court decided *U.S. v. Jones*, its first major case of digital technology surveillance. Does the Fourth Amendment prohibit police from secretly attaching a GPS tracking device to a person's vehicle without a search warrant? The police said no; they could have observed the suspect's car as it moved about on public streets. They argued the GPS device is a labor-saving device. The Court disagreed. There are two arguments in favor of Fourth Amendment protection in this case. First, a vehicle is one of a person's "effects" that the Fourth Amendment explicitly protects. Second, tracking a person's location for a month, 24 hours a day, as in this case, goes beyond someone observing the car pass by in public; it violates a person's expectation of privacy. The Court agreed (unanimously) with the first argument. Police need a search warrant to attach a surveillance device to a private vehicle. The justices recognized that expectation of privacy would be a key issue in tracking cases where directly attaching a device is not necessary, but the majority chose to leave a decision about that to future cases.*[20]

The police had one argument against expectation of privacy in *U.S. v. Jones*: the vehicle drove around in public view. Suppose a person is at home, at a friend's or lover's home, inside a church or a health facility, or in any private space. Law enforcement agencies use a device to locate a person by locating his or her cellphone, even when the person is not actively using the phone.† Police do not need to enter private premises or physically attach anything to a person's property. Thus, expectation of privacy is a key issue here. Law enforcement agenies argue that cellphone tracking (which they have used more than 1000 times, according to a *Wall Street Journal* investigation) does not require a search warrant because a person who uses a cellphone service has no expectation of privacy about the location data the phone transmits to cell towers. This view might surprise most cellphone owners. The Supreme Court has not yet heard a case about this technology.

* Four justices wrote an opinion that the tracking also violated expectation of privacy.

† The device pretends to be a cell tower. Agents drive around with it and get the target phone to connect to it in several locations. They then triangulate on the phone from the data the device collects.

This standard may allow time for markets, public awareness, and technologies to develop to provide privacy protection against the new technology. Is it a reasonable standard—a reasonable adaptation of law to new technology? Or should the court have permitted the search? Or should the government have to satisfy the requirements of the Fourth Amendment for every search of a home where a warrant would have been necessary before the technology existed?

> *Our use of these new technologies doesn't signal that we're less interested in privacy. The idea of the government monitoring our whereabouts, our habits, our acquaintances, and our interests still creeps us out. We often just don't know it's going on until it's too late.*
>
> —Judge Alex Kozinski[21]

2.2.3 Search and Seizure of Computers and Phones

> *Privacy in group association may . . . be indispensable to preservation of freedom of association, particularly where a group espouses dissident beliefs.*
>
> —The Supreme Court, ruling against the state of Alabama's attempt to get the membership list of the National Association for the Advancement of Colored People (NAACP) in the 1950s[22]

The NAACP's membership list was not on a computer in the 1950s. It undoubtedly is now. We consider several issues about how the Fourth Amendment applies to searches of computers, phones, and other electronic devices. How far does a search warrant extend when searching a computer? When is a search warrant needed?

The Fourth Amendment requires that search warrants be specific about the object of the search or seizure. Courts traditionally take the view that if an officer with a warrant sees evidence of another crime in plain view, the officer may seize it and prosecutors may use it. But the amount of information or evidence that might be in plain view in a house or office is small compared to what is on a computer. A computer at a business will have information about a large number of people. Membership lists, business records, medical records, and myriad other things can be on the same computer that law enforcement agents may search with a search warrant for specific, limited items. Access by law enforcement agents to all the data on a computer or device can be a serious threat to privacy, liberty, and freedom of speech.

How should we interpret "plain view" for a search of computer or smartphone files? A broad interpretation—for example, "all unencrypted files"—invites abuse. Agents could get a warrant for a small crime for which they have supporting evidence, and then go on fishing expeditions for other information. This thwarts the Fourth Amendment's requirement that a warrant be specific. In one case, while searching a man's computer with a search warrant for evidence of drug crimes, an officer saw file names suggesting illegal content not related to the warrant. He opened files and found child pornography. An appeals court said the names of files might be considered to be in plain view, but the contents of the files were not.[23] Although the crime in this case is a very unpleasant one, the principle protects us from abuses by the police.

In an investigation of the use of performance-enhancing drugs by professional baseball players, law enforcement agents obtained a search warrant for computer files of laboratory records on drug tests for 10 specific players. The lab files they seized contained records on hundreds of baseball players, hockey players, and ordinary people who are not athletes. The agents found that more than 100 baseball players tested positive for steroid use. This case received much attention in the news when the names of prominent players who allegedly tested positive leaked to the news media. A federal appeals court ruled that the information on all but the original 10 players was beyond the scope of the search warrant and the government was wrong to seize it.[24]

Suppose law enforcement agents have a search warrant for a computer but find that the files are encrypted. Must the owner supply the encryption key? The Fifth Amendment to the U.S. Constitution specifies that a person cannot be forced to testify against himself. However, courts sometimes allow the government to require a person to provide keys or combinations to a safe. Rulings in federal courts have been inconsistent about whether such a requirement can apply to encryption keys. (In many cases, law enforcement agents decrypt the files by other means.)

More about encryption: Section 2.5.1

> *What happened to the Fourth Amendment?*
> *Was it repealed somehow?*
>
> —A judge, commenting on the seizure of lab records for drug tests[25]

Phones and laptops

A mobile phone might contain contacts, numbers for calls made and received, email, text messages, documents, personal calendars, photos, a history of Web browsing, and a record of where the phone has been. For many people, the phone is a traveling office, containing proprietary and confidential information. A lawyer's phone might contain information about clients and cases—legally protected from access by police.

Police may search an arrested person (without a search warrant) and examine personal property on the person (in pockets, for example) or within his or her reach. Is a search warrant required before the police can search the contents of the person's cellphone? *Should* a search warrant be required?

This seems like a classic "no-brainer." The vast collection of information on a cellphone is the kind of information the Fourth Amendment is intended to protect. A judge who ruled against a cellphone search said the justifications for permitting police to search an arrested person were to find and take weapons and to prevent the person from hiding or destroying evidence. Once the police have custody of a phone, it is safe from destruction and police must wait until they have a search warrant before retrieving information from the phone. The Ohio Supreme Court ruled that searching an arrested person's phone

without a search warrant is unconstitutional:* people have an expectation of privacy for the contents of their phones.[26]

But the California Supreme Court ruled otherwise. It said that search of the contents of a cellphone was permitted because the phone was personal property found on the arrested person. Police have searched cellphones taken from arrested people in dozens of cases without warrants. Eventually, a case raising this issue will be heard by the U.S. Supreme Court. The result will have profound implications for privacy. In the meantime, lawyers suggest leaving a cellphone out of reach while driving.

Customs and border officials search luggage when U.S. citizens return from another country and when foreigners enter the United States. Border officials search, and sometimes seize, laptops and phones of journalists, businesspeople, and other travelers. Is searching a laptop equivalent to searching luggage? Or, because of the amount and kind of personal information they contain, does searching them at the border require reasonable suspicion of a crime? A federal appeals court ruled that customs agents do not need reasonable suspicion of a crime to search laptops, phones, and other electronic devices. Lawsuits and debate on the issue are ongoing.[27]

2.2.4 Video Surveillance and Face Recognition

We are used to security cameras in banks and convenience stores. They help in investigations of crimes. Prisons use video surveillance systems for security. Gambling casinos use them to watch for known cheaters. Video surveillance systems monitor traffic and catch drivers who run red lights. In these cases, people are generally aware of the surveillance. After the 2001 terrorist attacks, the police in Washington, D.C., installed cameras that zoom in on individuals a half mile away.

Cameras alone raise some privacy issues. When combined with face recognition systems, they raise even more. Here are some applications of cameras and face recognition and some relevant privacy and civil liberties issues.

In the first large-scale, public application of face recognition, police in Tampa, Florida, scanned the faces of all 100,000 fans and employees who entered the 2001 Super Bowl (causing some reporters to dub it Snooper Bowl). The system searched computer files of criminals for matches, giving results within seconds. People were not told that their faces were scanned. Tampa installed a similar system in a neighborhood of popular restaurants and nightclubs. Police in a control room zoomed in on individual faces and checked for matches in their database of suspects.[28] In two years of use, the system did not recognize anyone that the police wanted, but it did occasionally identify innocent people as wanted felons.

* The court allowed for exceptions in certain kinds of emergencies.

The ACLU compared the use of the face recognition system at the Super Bowl to a computerized police lineup to which innocent people were subject without their knowledge or consent. Face recognition systems had a poor accuracy rate in the early 2000s,[29] but the technology improved, along with the availability of photos to match against (tagged photos in social networks, for example). A police officer can now snap a photo of a person on the street and run a cellphone app for face recognition. (Another app scans a person's iris and collects fingerprints.)

Some cities have increased their camera surveillance programs, while others gave up their systems because they did not significantly reduce crime. (Some favor better lighting and more police patrols—low tech and less invasive of privacy.) Toronto city officials refused to let police take over their traffic cameras to monitor a protest march and identify its organizers. In a controversial statement, the Privacy Commissioner of Canada argued that the country's Privacy Act required a "demonstrable need for each piece of personal information collected" to carry out government programs and therefore recording activities of large numbers of the general public was not a permissible means of crime prevention.[30]

England was the first country to set up a large number of cameras in public places to deter crime. There are millions of surveillance cameras in Britain. A study by a British university found a number of abuses by operators of surveillance cameras, including collecting salacious footage, such as people having sex in a car, and showing it to colleagues. Defense lawyers complain that prosecutors sometimes destroy footage that might clear a suspect.[31] Enforcing a curfew for young people is one of the uses of public cameras in England. This application suggests the kind of monitoring and control of special populations the cameras make easy. Will police use face recognition systems to track political dissidents, journalists, political opponents of powerful people—the kinds of people targeted for illegal or questionable surveillance in the past? In 2005, the British government released a report saying Britain's closed-circuit TV systems were of little use in fighting crime. The only successful use of the cameras was in parking lots where they helped reduce vehicle crime.[32] Later that year, photos taken by surveillance cameras helped identify terrorists who set off bombs in the London subway. After rioters burned and looted neighborhoods in England in 2011, police used recordings from street cameras and face recognition systems to identify rioters. It is rare for all the facts or strong arguments to support only one side of an issue. What trade-offs between privacy and identifying criminals and terrorists are we willing to make?

The California Department of Transportation photographed the license plates on cars driving in a particular area. Then it contacted the car owners for a survey about traffic in the area. Hundreds of drivers complained. These people objected vehemently to what they considered unacceptable surveillance by a government agency even when the agency photographed only their license plates, not their faces—for a survey, not a police action. Many ordinary people do not like being tracked and photographed without their knowledge.

Clearly, some applications of cameras and face recognition systems are reasonable, beneficial uses of the technology for security and crime prevention. But there is a clear need for limits, controls, and guidelines. How should we distinguish appropriate from inappropriate uses? Should international events such as the Olympics, which are sometimes terrorist targets, use such systems? Should we restrict technologies such as face recognition systems to catching terrorists and suspects in serious crimes, or should we allow them in public places to screen for people with unpaid parking tickets? Do people have the right to know when and where cameras are in use? In the United States, police must have a reason for requiring a person to be fingerprinted. Should similar standards apply to their use of face recognition and iris scanning? If we consider these issues early enough, we can design privacy-protecting features into the technology, establish well-thought-out policies for their use, and pass appropriate privacy-protecting legislation before, as the Supreme Court of Canada worries in the quote below, "privacy is annihilated."

> *To permit unrestricted video surveillance by agents of the state would seriously diminish the degree of privacy we can reasonably expect to enjoy in a free society. . . . We must always be alert to the fact that modern methods of electronic surveillance have the potential, if uncontrolled, to annihilate privacy.*
>
> —Supreme Court of Canada.[33]

> *This is a public meeting!*
>
> —Reporter Pete Tucker, upon his arrest for taking a photo with his cellphone at an open meeting of a U.S. government agency. Newsman Jim Epstein was then arrested for recording the arrest of Tucker on his own phone.[34]

2.3 The Business and Social Sectors

2.3.1 Marketing and Personalization

> *Acxiom provides complete and accurate pictures of customers and prospects, powering all marketing and relationship efforts.*
>
> —Acxiom website[35]

Marketing is an essential task for most businesses and organizations. It is one of the biggest uses of personal information—by businesses, political parties, nonprofit organizations, and advocacy groups. Marketing includes finding new customers, members, or voters

BEEP BEEP BEEP...
WELCOME TO PIZZA SHMIZZA, MR. OPUS! ...OF 996 MELBA LN.!
COOL! YOU KNOW ALL ABOUT ME!
RIGHT HERE ON MY SCREEN, SIR.
I LOVE OUR NEW INFORMATION SOCIETY!
I NEED ONE PEPPERONI, PLEASE.
NOT WITH THAT COLON, YOU DON'T.
YOUR SCREEN SHOWS MY COLONOSCOPY?
DUDE. I'VE SEEN CLEANER PIPES ON A HARLEY.
I'LL HAVE A GARDEN TOFU PIZZA THEN.
YOU CAN'T AFFORD IT. YOU MADE SOME BIG PURCHASES IN JUNE.
I DID NOT!!
"ONE DELUXE HUG-ME MOM INFLATABLE DOLL: $249.99"
7-10-05
HUG TIME
BERKELEYBREATHED.COM
Washington Post Writers Group
STOP WAVING YOUR LITTLE BATH SCRUBBIE.
YOU'RE WATCHING ME!!
AM NOT.
ARE TOO!!
NOT.
MAYBE I'M JUST PARANOID.
SOMEBODY NEEDS A BIKINI WAX!
LISTEN! MY PRIVACY IS--
© Breathed

Data mining and clever marketing[36]

Customers of the British retailing firm Tesco permit the company to collect information on their buying habits in exchange for discounts. The company identifies young adult males who buy diapers and sends them coupons for beer—assuming that, with a new baby, they have less time to go to a pub.

Target beats that. Target's data miners analyzed purchases of women who signed up for baby registries. They discovered that pregnant women tend to increase their purchases of a group of 25 products. So if a woman starts buying more of several of those products (e.g., unscented lotions and mineral supplements), Target starts sending coupons and ads for pregnancy and baby products. It can even time them for stages of the pregnancy.

To compete with Wal-Mart, Tesco aimed to identify those customers who were most price conscious and hence most likely to be attracted to Wal-Mart's low prices. By analyzing purchase data, the company determined which customers regularly buy the cheapest version of products that are available at more than one price level. Then they determined what products those customers buy most often, and they set prices on those products below Wal-Mart's.

Are these examples of desirable competition or scary intrusiveness and manipulation of consumers?

and encouraging old ones to continue. It includes advertising one's products, services, or cause. It includes how to price products and when and to whom to offer discounts.

Through most of the 20th century, businesses sent out catalogs and advertisements based on a few criteria (age, gender, and neighborhood, for example). Computers and the increased storage capacity of the 1980s and 1990s began a revolution in targeted marketing. Now, businesses store and analyze terabytes of data, including consumer purchases, financial information, online activity, opinions, preferences, government records, and any other useful information to determine who might be a new customer and what new products and services an old customer might buy. They analyze thousands of criteria to target ads both online and offline. Online retailers make recommendations to you based on your prior purchases and on those of other people with similar buying patterns. Websites greet us by name and present us with options based on prior activity at that site.

To many, the idea that merchants collect, store, and sell data on their purchasing habits is disturbing. These activities impinge upon a key aspect of privacy: control of information about oneself. Privacy advocates and some consumers object to advertising based on consumer purchase histories and online activity. Marketers argue that finely targeted marketing is useful to the consumer and that it reduces overhead and, ultimately, the cost of products. L.L. Bean, a big mail-order business, says it sends out fewer catalogs as it does a better job of targeting customers. A Web ad company said users clicked on 16% of ads displayed based on the user's activity profile—many more than the 1% typical for untargeted Web ads. Another firm says that 20–50% of people used the personalized coupons it provided on screen or by email, compared with the 1–5% redemption rate for

newspaper inserts. The companies say targeting ads via personal consumer information reduces the number of ads overall that people will see and provides ads that people are more likely to want.[37] Many people like the personalization of ads and recommendations. Targeting is so popular with some people that Google advertised that its Gmail displays no *untargeted* banner ads.

Some kinds of less obvious personalization trouble people more (when they learn of them). The displays, ads, prices, and discounts you see when shopping online might be different from those others see. Some such targeting is quite reasonable: A clothing site does not display winter parkas on its home page for a shopper from Florida. Some sites offer discounts to first-time visitors. Some display any of hundreds of variations of a page depending on time of day, gender, location, and dozens of other attributes of a person's session. (Some sites guess a visitor's gender based on clicking behavior.[38]) If a person hesitates over a product, a site might offer something extra, perhaps free shipping. Is this collection and use of behavioral information an example of inappropriate invisible information gathering? When we shop in stores, sales clerks can see our gender and our approximate age. They can form other conclusions about us from our clothing, conversation, and behavior. Good salespeople in expensive specialty stores, car dealerships, flea markets, and third-world street markets make judgments about how much a potential customer will pay. They modify their price or offer extras accordingly. Is the complex software that personalizes shopping online merely making up for the loss of information that would be available to sellers if we were shopping in person? Are some people uneasy mainly because they did not realize that their behavior affected what appears on their screen? Are people uneasy because they did not realize that websites can determine (and store) so much about them when they thought they were browsing anonymously? Is the uneasiness something we will get over as we understand the technology better? Or are there privacy threats lurking in these practices?

Companies can use face recognition systems in video game consoles and televisions to target ads to the individual person who is playing a game or watching TV. What risks to privacy does this entail? Is it unethical to include such features? Will most people come to like the customization? Do they understand that if they see ads targeted to their interests, someone somewhere is storing information about them?

Our examples so far have been commercial situations. The Democratic and Republican parties use extensive databases on tens of millions of people to profile those who might vote for their candidates. The parties determine what issues to emphasize (and which to omit) in personalized campaign pitches. The databases include hundreds of details such as job, hobbies, type of car, and union membership.[39] One party might send a campaign flyer to a conservative union member that emphasizes its labor policy but does not mention, say, abortion, while another party might do the opposite.

The issue is informed consent

Technological and social changes make people uncomfortable, but that does not mean the changes are unethical. Some privacy advocates want to ban all advertising targeted by

online behavior. It should be clear that targeted or personalized marketing is not, in itself, unethical. Most of the legitimate concern has to do with how marketers get the data they use. In some cases there is consent, in some there is not, and in many the complexity of the situation makes consent unclear.

Collection of consumer data for marketing without informing people or obtaining their consent was widespread, essentially standard practice, until roughly the late 1990s. Sometimes, small print informed consumers, but often they did not see it, did not understand the implications, or ignored it. Gradually, public awareness and pressure for improvement increased, and data collection and distribution policies improved. Now websites, businesses, and organizations commonly provide explicit, multi-page statements about what information they collect and how they use the information. They provide opt-out and opt-in options. (Federal laws and regulations require specific privacy protections for financial and medical information.[40]) There are still many companies that get it wrong, whether out of lack of concern for people's privacy or by misjudging what people want. There is also a vast world of data collection over which we have little or no direct control. When someone consents to a company's use of his or her consumer information, the person probably has no idea how extensive the company is and how far the data could travel. Firms such as Acxiom (quoted at the beginning of this section), a large international database and direct-marketing company, collect personal data from a huge number of online and offline sources. Such companies that maintain huge consumer databases buy (or merge with) others, combining data to build more detailed databases and dossiers. They sell data and consumer profiles to businesses for marketing and "customer management." Most people do not know such firms exist.

Extensive and hidden tracking of online activity led to calls for a "Do Not Track" button in browsers. The exact meaning and effects of such buttons are yet to be determined. The idea is that users would have one clear place to indicate that they do not want their Web activity tracked and stored. Many advertisers, providers of popular Web browsers, and large Internet companies agreed to implement and comply with some version of Do Not Track.

Awareness varies among consumers, and many do not read privacy policies. Is it the user's responsibility to be aware of the data collection and tracking policies of a site he or she visits? Does a person's decision to interact with a business or website constitute implicit consent to its posted data collection, marketing, and tracking policies? How clear, obvious, and specific must an information-use policy be? How often should a site that runs (or allows third parties to run) tracking software remind users? Some people who allow extensive tracking and information collection might later regret specific decisions they made. Whose responsibility is it to protect them? Can we protect them without eliminating options for the people who use them sensibly? Potentially negative future consequences of choices we make now (such as not getting enough exercise) are common in life. We can educate consumers and encourage responsible choices. (At the end of the chapter, we list nonprofit organizations that help do this.) Respect for people's autonomy

means letting them make their own choices. Designing systems ethically and responsibly means including ways to inform and remind users of unobvious data collection, of changes in policies or features, and of risks.

Paying for consumer information

When businesses first began building extensive consumer databases, some privacy advocates argued that they should pay consumers for use of their information. In many circumstances, they did (and do) pay us indirectly. For example, when we fill out a contest entry form, we trade data for the opportunity to win prizes. Many businesses give discounts to shoppers who use cards that enable tracking of their purchases. Many offer to trade free products and services for permission to send advertising messages or to collect information. Some privacy advocates criticize such programs. Lauren Weinstein, founder of Privacy Forum, argues that among less affluent people the attraction of free services may be especially strong, and it "coerces" them into giving up their privacy.[41] People do not understand all the potential uses of their information and the long-term consequences of the agreements. On the other hand, such programs offer an opportunity for people with little money to trade something else of value (information) for goods and services they desire. Free-PC started the trend, in 1999, with its offer of 10,000 free PCs in exchange for providing personal information and watching advertising messages. Hundreds of thousands of people swamped the company with applications in the first day.

In any case, these early programs are dwarfed by the development of social networking, free video sites, and a huge number of other websites that provide information and services for free. People understand that advertising funds them. Gmail targets ads to individual users by analyzing the user's email messages. Some privacy advocates were horrified: it reads people's email! In exchange for permission to do so, Gmail provides free email and other services. Millions of people signed up. The success of these businesses and services shows that many people do not object to retailers using their purchase history or email and do not consider the intrusion of online ads to be extremely bothersome, nor their Web surfing to be particularly sensitive. Do they understand the potential consequences?

2.3.2 Our Social and Personal Activity

> *Broadcast Yourself.*
>
> —Slogan on YouTube's home page[42]

Social networks—what we do

There are two aspects of social networks to consider: our own responsibility for what we share (how we risk our privacy and that of our friends) and the responsibilities of the companies that host our information.

Many young people post opinions, gossip, and pictures that their friends enjoy. Their posts might cause trouble if parents, potential employers, law enforcement agents, or various others see them. An 18-year-old who posts sexy photos of herself in bathing suits is thinking about her friends viewing them, not potential stalkers or rapists. People who try to clean up their online personas before starting a job search find that it is hard to eliminate embarrassing material. Some social network apps ask for personal information—such as religion, political views, and sexual orientation—about one's friends as well as oneself. Do people think about how the information might be used and whether their friends would like it disclosed?

Why was it for so long standard practice to stop mail and newspaper delivery when going away on a trip? This one detail about location ("away from home") was important to protect from potential burglars. Yet, now, a great many people post their location (and that of their friends) to social networks.

Social networkers, with hundreds or thousands of network friends they never met, probably do not give enough thought to the implications of the personal information they make available. When someone initially chooses privacy settings, will that person later remember who is getting real-time reports on his or her status and activities?

Government agencies and businesses do many things wrong, but individuals also do not always exercise appropriate thought and care for their own privacy, future, and safety.

> *Polls show that people care about privacy.*
> *Why don't they act that way?*
>
> —Ian Kerr

Social networks—what they do

We use Facebook for our examples here because it has so many features and so many members, and because it has made instructive mistakes. The principles apply to other social media and other websites.

Facebook regularly introduces new services, new ways to share with friends and stay up-to-date on their activities. Several times, Facebook seriously misjudged how members would react and made poor choices. Some of the examples we describe quickly generated storms of criticism from tens of thousands to hundreds of thousands of members as well as from privacy advocates.

News feeds send recent changes in a member's personal information, friends list, and activities to that member's friends.[44] Facebook said it did not change any privacy settings when it introduced the feeds. It sends the information only to people the members had already approved and who could already see it if they looked. Within a day or two, hundreds of thousands of Facebook members protested vehemently. Why? The ease of accessing information can sometimes be more important than the fact that it is available somewhere. Many people do not check on their hundreds of friends regularly. The feeds, however, spread information to everyone instantly. Here is just one kind of instance where

it makes a difference: In the physical world, we might share information about the end of a relationship, a serious illness, or a family problem with a few, chosen, close friends. Gradually, as we adjust to the new situation, others might learn of the event. The feeds remove the emotionally protective delay.

When Facebook began telling members about purchases their friends made, problems ranged from spoiling surprise gifts to embarrassing and worrisome disclosures. Should Facebook introduce such features turned "on" for everyone? Or should the company announce them and let members opt in with a click? When Facebook introduced a face recognition tool to help members tag friends in photos, the default was that the tool was on for all members. There was a way to opt-out, but many users were not aware of the new feature, so they did not know to opt out. Facebook's Places feature lets users tag friends who are at their location (whether or not the friend is actually there). What should the default settings be?

Angry members are not good for business. These incidents demonstrate the importance, from both an ethical perspective and a business perspective, of giving careful thought to the implications and risks of new features and the selection of default settings. Changes that might seem small and subtle can have big impacts on people's perceptions of privacy, on risk, and on feelings of comfort. People might be happy if a few friends tag them in a few photos, but they might be very uneasy if an automated system tags every photo they appear in. Quantity can make a difference in perceived quality (in particular, in one's feeling of control of information about oneself). In complex environments, such as social networks with their many features and members, an opt-in policy is preferable—that is, a policy where members must explicitly turn the feature on, or else it remains off. In complex environments, it is also valuable to have a range of options. For example, for a tagging feature (for location or photos), options can include informing the person and allowing removal of the tag, requesting permission for each tag before it goes live, and allowing a member to completely opt out of being tagged. (Facebook modified Places to include a range of levels of protection.)

According to the Federal Trade Commission (FTC), Facebook violated its stated policies in several instances: by giving users' IDs to advertisers along with data on user activity, by allowing third-party apps full access to member personal data, and by failing to delete some of a member's data when the member deleted the account. Such actions, in violation of a company's own statements about its practices, are deceptive; they thwart informed decisions and agreements. We might dislike, denounce, debate, and disagree about the ethics of some data practices. Deceptive practices are more clearly unethical (or unethical at a stronger level) than mistakenly or carelessly making poor choices about defaults.

Responsibility of free services

We should appreciate the astounding amount of free service available to us from social network companies—as well as search engines, communication systems such as Twitter, websites full of expert information, and so on. We can choose to use them or not. At

the same time, the businesses that run these free services have a responsibility to their users. If you invite your neighbors to use your car anytime they wish without asking, you have an ethical responsibility not to leave the keys in the car when the brakes are not working. It does not matter that you do not charge a fee. Companies may not, ethically, offer attractive services and then cause a significant risk of harm, especially when the risk is hidden or unexpected.

Life in the clouds

Soon after a woman started writing a personal blog, she discovered that someone she had not seen in years read it. This horrified her. Perhaps she thought only people to whom she gave the Web address read the blog. She did not realize that it showed up high in search results for her name.[45] Another woman liked the feature on a social network site that told her which members read her profile. She was surprised and upset to find that people whose profiles she read knew that she read them. After Facebook suggested that two women might want to be friends, one of them discovered that they were both married to the same man.

The first incident reminds us that some people do not know or understand enough about how the Web works to make good decisions about what to put there.* The second indicates that some people do not think carefully about it. It also illustrates a very common phenomenon: people often want a lot of information about others, but they do not want others to have access to the same kinds of information about themselves. The bigamist did not realize that Facebook would notice his two wives had something in common.

Some people include their birth date in online profiles or in résumés they post on job-hunting sites. Genealogy sites are very popular. People create family trees with complete profiles of family members, including birth dates and mother's maiden name. Medical and financial institutions used this same information (birth dates and mother's maiden name) to verify a customer's identity. We can change a disclosed password; we cannot change our birth date or mother's maiden name.

The Web is public. Most people are decent and harmless, but many are evil and dangerous. Pedophiles have websites that link to sites of Cub Scouts, Brownies (the young version of Girl Scouts), junior high school soccer teams, and so on—sites with pictures of children and sometimes names and other personal information. That is scary. It does not mean that such organizations should not put pictures on their websites. It suggests, however, that they consider whether to include children's names, whether to require registration for use of the site, and so on.

Years ago, when many homes had answering machines connected to telephones, some people, instead, used answering services. Messages left for them resided on recording machines at the service's business site. I recall my surprise that people were comfortable having their personal messages on machines outside their control. How quaint and old-

* In an unusual example of initiative, the woman studied the techniques used to rank search results and modified her blog so that it no longer showed up prominently in searches for her name.

fashioned that concern seems now. Our cellphone and email messages routinely reside on computers outside our home or office. Text messages are retrievable months later. After many incidents of exposure of embarrassing messages, we still see individuals, politicians, lawyers, celebrities, and businesspeople writing sensitive, rude, or compromising things in email, text, and tweets with the apparent belief that no one but the intended recipient will ever see them.

Millions of Americans prepare their tax returns online. Do they think about where their income and expenditure data are going? How long the data remain online? How well secured the data are? Small businesses store all their accounting information online (in the "cloud") on sites that provide accounting services and access from anywhere. Do the business owners check the security of the sites? Several medical websites provide an easy place for people to store their medical records. Various companies offer services where people store all their data (email, photos, calendars, files) on the company's servers, instead of on their own PC or laptop. You can store an inventory of your valuable property on the Web (for free) to help with insurance claims after a fire or tornado. The companies supplying this service might all be honest, but the data, if leaked or hacked, is a shopping list for thieves.

There are big advantages to all these services. They are convenient. We do not have to manage our own system. We do not have to do backups. We can get to our files from anywhere with Internet access. We can more easily share files and collaborate with others on projects. There are disadvantages too. We cannot access our files when the network is down or if there is a technical problem at the company that stores them. But the more serious risks are to privacy and security. We lose control. Outside our home, our files are at risk of loss, theft, misuse by employees, accidental exposure, seizure by government agencies, uses by the service provider described in an agreement or privacy policy we did not read, uses we ignored when signing up for the service, and later uses that no one anticipated. We might not care who else sees our vacation photos. We might decide the convenience of filling out tax forms online or storing our medical records online outweighs the risks. The point is to be aware and to make the decision consciously. For computer professionals, awareness of the risks should encourage care and responsibility in developing secure systems to protect the sensitive information people store online.

2.3.3 Location Tracking

Global positioning systems (GPS), cellphones, radio frequency identification (RFID) tags,* and other technologies and devices enable a variety of location-based applications—that is, computer and communications services that depend on knowing exactly where a person or object is at a particular time. Since the introduction of the iPhone, there has been an explosion in such applications. The applications are extraordinarily diverse

* RFID tags are small devices that contain an electronic chip and an antenna. The chip stores identification data (and possibly other data) and controls operation of the tag. The antenna transmits and receives radio signals for communicating with devices that read the tag.

and have significant benefits. However, they add detailed information about our current location and our past movements to the pool of information that computer systems store about us, with all the potential threats to privacy.

To analyze risks, we should always consider unintended, as well as intended, uses. Recall from Section 2.2.2 that law enforcement agencies locate people by locating their phone. Details of the technology are secret and the device is probably expensive. But that is temporary. Eventually there will be an app for that. So imagine that anyone can enter a person's ID number (perhaps a phone number) on their own mobile device and ask where that person is now. Or perhaps a device could sweep a particular location and detect identifying devices of the people there—or identify them by face recognition. Who might a person *not* want to get this information? Thieves. A violent spouse or ex-spouse. A divorce lawyer. An annoying or nosy neighbor. A stalker. Co-workers or business associates. Anyone else who might object to your religion, politics, or sexual behavior. The government. (Oh, we see that our new teacher is at a meeting of Alcoholics Anonymous. Who is in that medical marijuana store or gun store right now?) Extensive records of where we were provide more details to the ever-growing profiles and dossiers businesses and governments build about us. With fast search, matching, and analysis tools, they can add more detail about who we spend time with and what we are doing. In Chapter 1, we mentioned that researchers learn about social organization and the spread of disease (among other things) by studying huge amounts of cellphone data. Such statistical data can be extremely valuable to us all, but a cellphone identifies a person, and, thus, the tracking information (if associated with the phone's number or ID) is personal information and raises the usual issues of consent, potential secondary uses, risks of misuse, and so on. Care must be taken to ensure that such data are protected. If accessed surreptitiously, stolen, disclosed accidentally, or acquired by government agencies, records of our location and movements pose threats to privacy, safety, and liberty. Privacy and industry organizations are developing guidelines for use of location-tracking applications to implement principles in Figure 2.2 and protect against some of the risks.[46]


Tracking employees at work: Section 6.3.2

Studying the behavior of customers in a store or other facility is a big potential application of location tracking. For example, a supermarket or an amusement park might want to analyze customer traffic patterns within the facility to plan a better layout, to determine how much time people spend inside, or to analyze waiting times. The privacy implications and risks of monitoring people's movements vary from little to great depending on how the tracking system does its work. Suppose, for example, an amusement park such as Disneyland wants to study visitor traffic patterns, detect crowds and long lines, and so on. It can do so with a location-emitting ticket that people get when they enter and discard when they leave. It need have no information connected to the person or family. For such a system, privacy is not an issue. There would be a temptation, however, to include demographic data and possibly identifying data on the tracker.

Who's at the Bar?

Hundreds of bars installed cameras with a face recognition system to provide data to a website and smartphone app. The app tells users the number of people at a particular bar, the male/female ratio, and the approximate age range. Each bar gets summary statistics on its patrons that could be useful for advertising or other business planning. The system does not identify individual people and does not store the video. So this is not a privacy issue. Or is it?

The point is that such an application can remain utterly unthreatening, or it can drift over the boundary into location tracking and privacy infringement. The bar owners do not control the system, so they cannot be certain that what they tell their customers about it is true. (There are many examples of systems collecting and storing data without the knowledge of the businesses that use the system.) The developer and operator of the system might exercise great care to protect patrons' privacy, or they might succumb to temptation to add new features that require storing video or identifying individuals. Awareness of potential risks and understanding of good privacy practices are essential for both the software developers who invent and upgrade such systems and the managers who make decisions about what features to implement.

Tools for parents

Many technologies help parents track their children's physical location. Cellphone services enable parents to check a child's location from the parent's mobile device. Devices installed in a car tell parents where their teens are and how fast they are driving. A company sells wireless watchband transmitters for children, so parents can monitor them. RFID tags in shoes and clothes can be monitored hundreds of feet away. These might be very helpful with young children who wander off in a crowded place.

Tracking children can increase safety, but there are parenting issues and risks involved in using tracking tools. At what age does tracking become an invasion of the child's privacy? Should parents tell children about the tracking devices and services they are using? Informed consent is a basic principle for adults. At what age does it apply to children? Will intense tracking and monitoring slow the development of a child's responsible independence?

A monitoring system that sends easily read or easily intercepted signals could decrease rather than increase the safety of a child. Child molesters and identity thieves could collect personal data. Parents need to be aware of potential for false alarms and for a false sense of security. For example, a child might lose a phone or leave a tagged article of clothing somewhere. Older kids might figure out how to thwart tracking systems. Clearly, how and when to use surveillance tools should involve thoughtful decisions for families.

Pets, prisoners, and people with Alzheimer's disease can wear devices that locate them if they wander off. Veterinarians implant ID chips under the skin of pets and farm animals.

Foiling poachers, following turtles, tracking guitars

Owners tag very valuable and extremely rare plants, both in the wild and in gardens, with tracking chips so they can locate them if stolen.

Satellite technology and microprocessors enormously improved animal tracking. Scientists now attach tiny transmitters to rare birds and other animals to study their behavior and learn how to protect their food sources. Researchers learned that some animals travel much farther than previously thought: Sea turtles swim from the Caribbean to Africa. A nesting albatross flew from Hawaii to the San Francisco Bay, a weeklong round-trip, to get food for its young. To encourage interest from the public, researchers set up websites where we can follow the animals' movements.[47]

These are valuable services. What happens when the same technologies track people?

I recently toured a guitar factory. The tour guide showed us a partially complete guitar neck. And there, on the front of the neck, was an RFID chip. The fret board, when attached to the neck, covers the chip. The guide explained how useful the chip was for tracking guitars through production and for finding a specific guitar in the stock room. The chip remains in the guitar when a customer buys it. Manufacturers put RFID tags in many other products, in addition to guitars, to track them through the manufacturing and sales processes. What is the potential for tracking people via the products they buy? Does it matter?

Some people have suggested doing this for prisoners and children. Does the suggestion of implanting tracking chips in people make you wonder if that is such a good idea? After heavy opposition from parents, a school dropped its proposal to require that all students wear an RFID-equipped device while on school grounds. The constant surveillance and the risks of misuse were enough, in the minds of many parents, to outweigh the benefits of a removable tracking device.

2.3.4 A Right to Be Forgotten

People sometimes want to remove information about themselves from the Internet or from a company's records. It could be an offensive comment made in anger, a photo on one's own social network page or a photo-sharing site, information in online directories, or personal data posted by others (e.g., on a genealogy site). It could be the profile an advertising company developed by tracking the person's Web activity, a collection of data gleaned from the person's smartphone use, or the collection of the person's search queries that a search engine stores. It could be unflatering images or information that other people posted. It could be a search engine's links to such material. Legislators and privacy advocates in the United States and the European Union are promoting a legal right to demand that websites remove material about oneself. The right to have material removed, as a legal or ethical right, has come to be called the "right to be forgotten." The wide range of material a person might want to

The right to be forgotten in the EU: Section 2.5.3

remove suggests many practical, ethical, social, and legal questions and criticisms about such a right.[48]

The policies of various websites about removing material vary. Some sites with members, such as social networks, respond to a member's request to delete material the user posted and to delete a member's material when the member closes the account. When the material is not in a user's account, the situation is more complicated. Some sites, such as directories, collect information automatically; thus, deleted information can reappear. A filter system to prevent reposting for a particular person has the problem of correctly distinguishing that person from others with the same or similar names.

Should a company or website always comply with a request to delete a particular item or a person's record any time a person makes such a request? We understand that people do foolish things and regret them later. It is reasonable to let many of them be forgotten. If a person wants to delete something he or she posted on a website, it is reasonable, courteous, good-spirited, and perhaps a good business policy to comply. If someone else posts compromising photos or information from a person's past, removing it raises issues of free speech and truth. If the person is not a public figure and the information has no broad social value, removing it might be the reasonable, courteous thing to do. Complying with the request could be ethically acceptable and admirable but not ethically obligatory. In some cases, it could be a bad idea. The information might matter to people in a particular community. The person who posted it might have a good reason. The appropriate decision in specific cases might be difficult.

What about the data that advertisers and search engines collect about us? Must they, from an ethical standpoint, comply with a request from a person who wants his or her record deleted? If the companies collected the data secretly, without permission, or in violation of their stated privacy policies and terms of use, then there are good reasons to require its deletion independent of any right to be forgotten. Suppose the information is the set of a person's search queries or something similar that a free website collects, and suppose the site makes its collection and use of the data clear in its terms of use. The company's use of the data is, in part, our payment for the free service it provides. If the company agrees to delete people's records upon request, it is providing its service to those people for free (or at a "discount" if they continue to view ads on the site). If a relatively small number of individuals request deletion of their data, a large company can probably afford to comply without significant inconvenience or reduction in the value it gets from analysis of user data. Many companies give some products and services for free. Again, complying with deletion requests could be ethically and socially admirable, good-spirited, and perhaps a good business policy. On the other hand, a person might make a deletion request to hide some illegal or offensive behavior or to remove evidence in a dispute of some kind.

If the right to be forgotten is a negative right (a liberty), it could mean that we may choose to stay off the Internet and become a recluse, but we cannot force someone else to remove a photo that we are in. As a positive right (a claim right), it is akin to requiring

that others erase their minds, as well as their photos, blogs, and links. It can mean that others may not write about a person or exchange specified information about the person—information gained without violating any of the person's rights. This can infringe freedom of speech. In some applications, the right would mean that a person may break agreements (e.g., terms of use for a Web service) at will. There seems to be little if any basis for such an ethical right.

Are there contexts in which it makes sense to enforce a legal requirement to remove material when a person requests it? Perhaps for special populations, such as children (where parents might make the request or a young adult might want to remove seminude sexting photos sent to friends while in high school). Perhaps in other special situations. Legislators must carefully craft any such legal requirement to avoid conflict with free speech, free flow of information, and contractual agreements. A legal requirement to honor removal requests will be more of a burden to small sites than to large ones, which can develop software to help automate the process and have legal staffs to defend against complaints.

Sexting: Section 3.2.3

2.4 Government Systems

2.4.1 Databases

Federal and local government agencies maintain thousands of databases containing personal information. Examples include tax, property ownership, medical, travel, divorce, voter registration, bankruptcy, and arrest records. Others include applications for government grant and loan programs, professional and trade licenses, and school records (including psychological testing of children). And there are many, many more. Government databases help government agencies perform their functions, determine eligibility for government benefits programs, detect fraud in government programs, collect taxes, and catch people who are breaking laws. The scope of government activities is enormous, ranging from catching violent criminals to licensing flower arrangers. Governments can arrest people, jail them, and seize assets from them. Thus, the use and misuse of personal data by government agencies pose special threats to liberty and personal privacy. It seems reasonable to expect governments to meet an especially high standard for privacy protection and adherence to their rules.

The Privacy Act of 1974 is the main law about the federal government's use of personal data. A summary of the provisions of the Act appears in Figure 2.3. Although this law was an important step in attempting to protect our privacy from abuse by federal agencies, it has problems. The Privacy Act has, to quote one expert on privacy laws, "many loopholes, weak enforcement, and only sporadic oversight."[49] The E-Government Act of 2002 added some privacy regulations for electronic data and services—for example, requiring agencies

- Restricts the data in federal government records to what is "relevant and necessary" to the legal purpose for which the government collects it
- Requires federal agencies to publish a notice of their record systems in the Federal Register so that the public may learn about what databases exist
- Allows people to access their records and correct inaccurate information
- Requires procedures to protect the security of the information in databases
- Prohibits disclosure of information about a person without his or her consent (with several exceptions)

Figure 2.3 Provisions of the Privacy Act of 1974.

to conduct privacy impact assessments for electronic information systems and to post privacy policies on agency websites used by the public.

The Government Accountability Office (GAO) is Congress' "watchdog agency." Over the past 25 years, the GAO has released numerous studies showing lack of compliance with the Privacy Act and other privacy risks and breaches. The GAO reported in 1996 that White House staffers used a "secret" database with records on 200,000 people (including ethnic and political information) without adequate access controls. A GAO study of 65 government websites found that only 3% of the sites fully complied with the fair information standards for notice, choice, access, and security established by the Federal Trade Commission (FTC) for commercial websites. (The FTC's site was one that did not comply.) The GAO reported that the Internal Revenue Service (IRS), the Federal Bureau of Investigation (FBI), the State Department, and other agencies that use data mining to detect fraud or terrorism did not comply with all rules for collecting information on citizens. The GAO found dozens of weaknesses in the operation of the government's communication network for transmitting medical data in the Medicare and Medicaid programs—weaknesses that could allow unauthorized access to people's medical records.[50]

The IRS is one of several federal government agencies that collects and stores information on almost everyone in the country. It is also a major secondary user of personal information. Year after year, hundreds of IRS employees are investigated for unauthorized snooping in people's tax files. (An IRS employee who was a Ku Klux Klan member read tax records of members of his Klan group looking for income information that would indicate that someone was an undercover agent.) These abuses led to a law with tough penalties for government employees who snoop through people's tax information without authorization. However, a GAO report a few years later found that while the IRS had made significant improvements, the tax agency still failed to adequately protect people's financial and tax information. IRS employees were able to alter and delete data without authorization. Employees disposed of disks with sensitive taxpayer information without

erasing files. Hundreds of tapes and diskettes were missing. A report by the Treasury's Inspector General said that the IRS did not adequately protect taxpayer information on more than 50,000 laptops and other storage media. Personal financial information that taxpayers provide to the IRS is "at risk" from hackers and disgruntled employees because many of the 250 state and federal agencies to which the IRS provides taxpayer information do not have adequate safeguards.[51]

Various reviews of compliance with the Privacy Act and the E-Government Act have highlighted weaknesses in these laws. The GAO advocated modifying the Privacy Act to cover all personally identifiable information collected and used by the federal government, thus closing gaping loopholes that exempt much government use of personal information from the law's provisions. The GAO advocated stricter limits on use of personal information. Recognizing that most people do not read the *Federal Register,* the GAO suggested better ways of informing the public about government databases and privacy policies. The Information Security and Privacy Advisory Board (a government advisory board) pointed out: "The Privacy Act does not adequately cover government use of commercially-compiled databases of personal information. The rules about the federal government's use of commercial databases, and even use of information gleaned from commercial search engines, have been vague and sometimes non-existent." Thus, agencies can bypass the protections of the Privacy Act by using private-sector databases and searches, rather than collecting the information itself.[52]

> *Quis custodiet ipsos custodes? (Who will guard the guards themselves?)*
>
> —Juvenal

Database example: tracking college students

The U.S. Department of Education proposed establishing a database to contain the records of every student enrolled in a college or university in the United States. The proposal would require colleges and universities to provide and regularly update the records including each student's name, gender, Social Security number, major, courses taken, courses passed, degrees, loans, and scholarships (public and private). The government would keep the data indefinitely. The department has not yet implemented the proposal because of intense opposition. The government already has similar databases, and proposals for massive government databases of personal information appear regularly. We discuss this one as an example for analysis; the issues and questions we raise here apply in many other situations.

The student database would have many beneficial uses: The federal government spends billions of dollars each year on federal grants and loans to students but has no good way to measure the success of these programs. Do students who get aid graduate? What majors do they pursue? The database would help evaluate federal student aid programs

and perhaps lead to improvements in the programs. The database would provide more accurate data on graduation rates and on actual college costs. The ability to track the number of future nurses, engineers, teachers, and so on, in the educational pipeline can help shape better immigration policy and business and economic planning.

On the other hand, the collection of so much detail about each student in one place generates a variety of privacy risks. Several of the points in the list in Section 2.1.2 are relevant here. It is very likely that the government would find new uses for the data that are not part of the original proposal. Such a database could be an ideal target for identity thieves. Leaks of many sorts are possible and likely. There is potential for abuse by staff members who maintain the data; for example, someone might release college records of a political candidate. And there would undoubtedly be errors in the database. If the department limits the data's use to generalized statistical analysis, errors might not have a big impact, but for some potential uses, the errors could be quite harmful.


More about identity theft: Section 5.3

Some educators worry that a likely eventual link between the database and public school databases (on children in kindergarten through high school) would contribute to "cradle-to-grave" tracking of childhood behavior problems, health and family issues, and so on.[53]

The planned uses of the database do not include finding or investigating students who are breaking laws, but it would be a tempting resource for law enforcement agencies. A Virginia state law requires colleges to provide the names and other identifying information for all students they accept. State police then check if any are in sex-offender registries. What else might they check for? What other government agencies might want access to a federal student database? Would the Defense Department use the database for military recruiting? What potential risks arise if employers get access? All such uses would be secondary uses, without the consent of the students.


Risks from errors in sex-offender registries: Section 8.1.2

It makes sense for the government to monitor the effectiveness of the grants and loans it gives to college students. It is therefore reasonable to require data on academic progress and graduation from students who receive federal money or loan guarantees. But what justifies demanding the data on all other students? For statistics and planning, the government can do voluntary surveys, just as businesses and organizations, without the government's power of coercion, must do. Are the benefits of the database central enough to the fundamental responsibilities of government to outweigh the risks and to justify a mandatory reporting program of so much personal data on every student?*

* Critics of the proposal, including many universities, point out other risks and costs besides privacy. Colleges fear that collection of the data would lead to increased federal control and interference in management of colleges. The reporting requirements would impose a high cost on the schools. The whole project would have high costs to taxpayers.

The U.S. Census

The U.S. Constitution authorizes and requires the government to count the people in the United States every 10 years, primarily for the purpose of determining the number of Congressional representatives each state will have. Between 1870 and 1880, the U.S. population increased by 26%. It took the government nine years to process all the data from the 1880 census. During the 1880s, the population increased by another 25%. If the Census Bureau used the same methods, it would not complete processing data from the 1890 census until after the 1900 census was to begin. Herman Hollerith, a Census Bureau employee, designed and built punch-card processing machines—tabulators, sorters, and keypunch machines—to process census data.* Hollerith's machines did the complete 1890 population count in only six weeks—an amazing feat at the time. The Bureau completed the rest of the processing of the 1890 census data in seven years. It could have been done sooner, but the new machines allowed sophisticated and comprehensive analysis of the data that was not possible before. Here is an early example of computing technology enabling increased processing of data with the potential for good and bad effects: better use of information and invasion of privacy.

The Census Bureau requires everyone to provide name, gender, age, race, and relationship to people one lives with. It requires three million households a year to fill out a longer form that contains questions about marital history, ancestry, income, details about one's home, education, employment, disabilities, expenditures, and other topics.

Census information is supposed to be confidential. Federal law says that "in no case shall information furnished . . . be used to the detriment of any respondent or other person to whom such information relates."[54]

During World War I, the Census Bureau provided names and addresses of young men to the government to help find and prosecute draft resisters. During World War II, the Census Bureau assisted the Justice Department in using data from the 1940 census to find U.S. citizens of Japanese ancestry; the army rounded up Japanese-Americans and put them in internment camps. With the introduction of electronic computers and the advances in computing technology, using the data "to the detriment of any respondent" is easier. Some cities used census data to find poor families who violated zoning or other regulations by doubling up in single-family housing. They evicted the families. A few years after the 9/11 terrorist attacks, at the request of the Department of Homeland Security, the Census Bureau prepared lists showing the number of people of Arab ancestry in various zip codes throughout the United States. A government spokesperson said they needed the data to determine which airports should have signs in Arabic. Privacy and civil liberties organizations were skeptical.[55]

* The company Hollerith formed to sell his machines later became IBM.

When considering each new system or policy for personal data use or data mining by government, we should ask many questions: Is the information it uses or collects accurate and useful? Will less intrusive means accomplish a similar result? Will the system inconvenience ordinary people while being easy for criminals and terrorists to thwart?

How significant are the risks to innocent people? Are privacy protections built into the technology and into the rules controlling usage?

Fighting terrorism

Before the terrorist attacks on the United States on September 11, 2001, law enforcement agencies lobbied regularly for increased powers that conflicted with privacy. Sometimes they got what they wanted; sometimes they did not. Generally, people resisted privacy intrusion by government. After the attacks on the World Trade Center and the Pentagon, more people became willing to accept uses of personal data and forms of search and surveillance that would have generated intense protest before. Two examples are the intrusive searches at airports and the Transportation Security Administration's (TSA) requirement that airlines provide the name and birth date of every passenger to the TSA so that it can match people against its watch list. In 2012, the government extended to five years the amount of time the National Counterterrorism Center may store data on Americans with no known connection to terrorism or criminal activity.

Errors in terrorism watch lists: Section 8.1.2

Proposals for new data mining programs to find terrorists and terrorist plots continue to appear. We summarize an interesting point Jeff Jonas and Jim Harper present about the suitability of data mining for this purpose.[56] Marketers make heavy use of data mining. They spend millions of dollars analyzing data to find people who are likely to be customers. How likely? In marketing, a response rate of a few percent is considered quite good. In other words, expensive, sophisticated data mining has a high rate of false positives. Most of the people whom data mining identifies as potential customers are not. Many targeted people will receive ads, catalogs, and sales pitches they do not want. Junk mail and pop-up ads annoy people, but they do not significantly threaten civil liberties. A high rate of false positives in data mining for finding terrorist suspects does. Data mining might be helpful for picking terrorists out of masses of consumer data, but appropriate procedures

Reducing privacy intrusions for air travel

Travelers are familiar with x-ray scanning machines at airports. The machines display on a computer screen the image of a person's body and any weapons and packets of drugs hidden under clothing and wigs. The American Civil Liberties Union (ACLU) describes the scan as "a virtual strip search." In response to strong objections from the public and privacy advocates, the TSA modified the software to display a generic line drawing of a body, instead of the x-ray image of the actual person scanned.[57]

Why didn't the TSA build in this obvious privacy-protecting feature at the beginning? There might be technical problems, but perhaps they did not because no law or regulation requires such privacy protection.

are essential to protect innocent but mistakenly selected people. Jonas and Harper argue that other methods for finding terrorists are more cost-effective and less threatening to the privacy and civil liberties of large numbers of people.

2.4.2 Public Records: Access versus Privacy

Governments maintain "public records," that is, records that are available to the general public. Examples include bankruptcy records, arrest records, marriage license applications, divorce proceedings, property-ownership records (including mortgage information), salaries of government employees, and wills. These have long been public, but by and large they were available only on paper in government offices. Lawyers, private investigators, journalists, real estate brokers, neighbors, and others use the records. Now that it is so easy to search and browse through files on the Web, more people access public records for fun, for research, for valid personal purposes—and for purposes that can threaten the peace, safety, and personal secrets of others.

Public records include sensitive information such as Social Security numbers, birth dates, and home addresses. Maricopa County in Arizona, the first county to put numerous and complete public records on the Web, had the highest rate of identity theft in the United States.[58] Obviously, certain sensitive information should be withheld from public-record websites. That requires decisions about exactly what types of data to protect. It requires revisions to government software systems to prevent display of specified items. Because of the expense and lack of accountability, incentives within government agencies to do this are weak. A few have adopted policies to block display of sensitive data in files posted online, and some states have laws requiring it. Several software companies produced software for this purpose, using a variety of techniques to search documents for sensitive data and protect them. Until new systems—in which such security is part of the basic design—replace older systems, the patches and add-ons, while helpful, are likely to miss a lot of sensitive data.


More about identity theft: Section 5.3

To illustrate more issues about public records and potential solutions, we describe a few kinds of specialized information (political contributions, flight information for private airplanes, and the financial statements of judges), then raise some questions.

Political campaign committees must report the name, address, employer, and donation amount for every donor who contributes more than $100 to a candidate for president. This information is available to the public. In the past, primarily journalists and rival campaigns examined it. Now it is on the Web and easy to search. Anyone can find out what candidate their neighbors, friends, employees, and employers support. We can also find the addresses of prominent people who might prefer to keep their address secret to protect their peace and privacy.

The pilots of the roughly 10,000 company airplanes in the United States file a flight plan when they fly. A few businesses have combined this flight information, obtained

from government databases, with aircraft registration records (also public government records) to provide a service telling where a particular plane is, where it is going, when it will arrive, and so on. Who wants this information? Competitors can use it to determine with whom top executives of another company are meeting. Terrorists could use it to track movements of a high-profile target. The information was available before, but not so easily and anonymously.

Federal law requires federal judges to file financial disclosure reports.[59] The public can review these reports to determine whether a particular judge might have a conflict of interest in a particular case. The reports were available in print but not online. When an online news agency sued to make the reports available online, judges objected that information in the reports can disclose where family members work or go to school, putting them at risk from defendants who are angry at a judge. Ultimately, the reports were provided for posting online, with some sensitive information removed.[60]

The change in ease of access to information changes the balance between the advantages and disadvantages of making some kinds of data public. Whenever access changes significantly, we should reconsider old decisions, policies, and laws. Do the benefits of requiring reporting of small political contributions outweigh the privacy risks? Do the benefits of making all property ownership records public outweigh the privacy risks? Maybe. The point is that such questions should regularly be raised and addressed.

How should we control access to sensitive public records? Under the old rules for the financial statements of judges, people requesting access had to sign a form disclosing their identity. This is a sensible rule. The information is available to the public, but the record of who accessed it could deter most people intent on doing harm. Can we implement a similar system online? Technologies for identifying and authenticating people online are developing, but they are not yet widespread enough for use by everyone accessing sensitive public data on the Web. We might routinely use them in the future, but that raises another question: How will we distinguish data that requires identification and a signature for access from data the public should be free to view anonymously, to protect the viewer's privacy?[61]

2.4.3 National ID Systems

In the United States, national identification systems began with the Social Security card in 1936. In recent decades, concerns about illegal immigration and terrorism provided the most support for a more sophisticated and secure national ID card. Opposition, based on concerns about privacy and potential abuse (and cost and practical problems), prevented significant progress on a variety of national ID proposals made by many government agencies. In this section, we review Social Security numbers, various issues about national ID systems, and the REAL ID Act, a major step toward turning driver's licenses into national ID cards.

Social Security numbers[62]

The history of the Social Security number (SSN) illustrates how the use of a national identification system grows. When SSNs first appeared in 1936, they were for the exclusive use of the Social Security program. The government assured the public at the time that it would not use the numbers for other purposes. Only a few years later, in 1943, President Roosevelt signed an executive order requiring federal agencies to use the SSN for new record systems. In 1961, the IRS began using it as the taxpayer identification number. So employers and others who must report to the IRS require it. In 1976, state and local tax, welfare, and motor vehicle departments received authority to use the SSN. A 1988 federal law requires that parents provide their SSN to get a birth certificate for a child. In the 1990s, the Federal Trade Commission encouraged credit bureaus to use SSNs. A 1996 law required that states collect SSNs for occupational licenses, marriage licenses, and other kinds of licenses. Also in 1996, Congress required that all driver's licenses display the driver's SSN, but it repealed that law a few years later due to strong protests. Although the government promised otherwise, the SSN has become a general identification number.

We use our Social Security number for identification for credit, financial services, and numerous other services, yet its insecurity compromises our privacy and exposes us to fraud and identity theft. For example, a part-time English teacher at a California junior college used the Social Security numbers of some of her students, provided on her class lists, to open fraudulent credit card accounts. Because the SSN is an identifier in so many databases, someone who knows your name and has your SSN can, with varying degrees of ease, get access to your work and earnings history, credit report, driving record, and other personal data. SSNs appear on public documents and other openly available forms. Property deeds, which are public records (and now online), often require SSNs. For decades, SSNs were the ID numbers for students and faculty at many universities; the numbers appeared on the face of ID cards and on class rosters. The state of Virginia included SSNs on published lists of voters until a federal court ruled that its policy of requiring the SSN for voter registration was unconstitutional. Some employers used the SSN as an identifier and put it on badges or gave it out on request. Many companies, hospitals, and other organizations to which we might owe a bill request our SSN to run a credit check. Some routinely ask for an SSN and record it in their files, although they do not need it.

More than 30 years ago, the U.S. Department of Agriculture (USDA) began including the SSN as part of the ID number for farmers who received loans or grants. In 2007, the USDA admitted that since 1996 it had inadvertently included the SSNs of more than 35,000 farmers on the website where it posted loan details.[63] This example illustrates how practices begun well before the Web have continuing repercussions. It also illustrates the importance of careful and thorough evaluation of decisions to put material on the Web. There are likely many similar examples that no one has yet noticed.

SSNs are too widely available to securely identify someone. Social security cards are easy to forge, but that hardly matters, because those who request the number rarely ask for

the card and almost never verify the number. The Social Security Administration itself used to issue cards without verification of the information provided by the applicant. Criminals have little trouble creating false identities, while innocent, honest people suffer disclosure of personal information, arrest, fraud, destruction of their credit rating, and so on, because of problems with the SSN.

Gradually, governments and businesses began to recognize the risks of careless use of the SSN and reasons why we should not use it so widely. It could take a long time to undo the damage its widespread use has already done to privacy and financial security.

A new national ID system

> *Places like Nazi Germany, the Soviet Union, and apartheid South Africa all had very robust identification systems. True, identification systems do not cause tyranny, but identification systems are very good administrative systems that tyrannies often use.*
>
> —Jim Harper, Director of Information Policy Studies, Cato Institute[64]

Various national ID card proposals in recent years would require citizenship, employment, health, tax, financial, or other data, as well as biometric information such as fingerprints or a retina scan, depending on the specific proposal and the government agency advocating it. In many proposals, the cards would also access a variety of databases for additional information.

More about biometrics: Section 5.3.3

Advocates of national ID systems describe several benefits: You would need the actual card, not just a number, to verify identity. The cards would be harder to forge than Social Security cards. A person would need to carry only one card, rather than separate cards for various services as we do now. The authentication of identity would help reduce fraud both in private credit card transactions and in government benefit programs. Use of ID cards for verifying work eligibility would prevent people from working in the United States illegally. Criminals and terrorists would be easier to track and identify.

Opponents of national ID systems argue that they are profound threats to freedom and privacy. "Your papers, please" is a demand associated with police states and dictatorships. In Germany and France, identification papers included the person's religion, making it easy for the Nazis to capture and remove Jews. Under the infamous pass laws of South Africa, people carried passes, or identification papers, that categorized them by race and controlled where they could live and work. Cards with embedded chips or magnetic strips and the large amount of personal information they can carry or access have even more potential for abuse. Most people would not have access to the machinery that reads the cards. Thus, they would not always know what information they are giving others about themselves. Theft and forgery of cards would reduce some of the potential benefits. Peter Neumann and Lauren Weinstein warned of risks that arise from the databases

and communication complexes that would support a national ID card system: "The opportunities for overzealous surveillance and serious privacy abuses are almost limitless, as are opportunities for masquerading, identity theft, and draconian social engineering on a grand scale."[65]

A woman in Canada could not get her tax refund because the tax agency insisted she was dead. Her identification number had been mistakenly reported in place of her mother's when her mother died. She would still have been able to get a new job, withdraw money from her bank account, pay her rent, send email, and go to her doctor while she was resolving the problem with the tax agency. What if the worker verification database connected to the death records database? Or what if a mistake cancelled the one ID card required for all these transactions? A critic of a proposal for a national identification card in Australia described the card as a "license to exist."[66]

The REAL ID Act attempts to develop a secure national identification card by setting federal standards for driver's licenses (and state-issued ID cards, for people without driver's licenses). Licenses must meet the federal standards for use for identification by the federal government. Such purposes include airport security and entering federal facilities. By implication, they likely include working for the federal government and obtaining federal benefits. It is likely that the government will add many new uses, as it did with the Social Security number. Businesses and state and local governments are likely to require the federally approved ID card for many transactions and services. The federal government pays for approximatley half the medical care in the United States (for example, Medicare, benefits for veterans, and numerous federally funded programs). It is not hard to envision requiring the driver's license for federal medical services and eventually it becoming a de facto national medical ID card.

The REAL ID Act requires that, to get a federally approved driver's license or ID card, each person must provide documentation of address, birth date, Social Security number, and legal status in the United States. Motor vehicle departments must verify each person's information, in part by accessing federal databases such as the Social Security database. The departments must scan documents submitted by drivers and store them in transferable form, for at least 10 years (making motor vehicle records a desirable target for identity thieves). The licenses must satisfy various requirements to reduce tampering and counterfeiting, and they must include the person's photo and machine-readable information to be determined by the Department of Homeland Security.

The REAL ID Act puts the burden of verifying identity on individuals and the state motor vehicle departments. Errors in federal databases used for verification could prevent people from getting their driver's licenses. Many states object to the mandate and its high costs (estimated in billions of dollars). More than 20 states passed resolutions refusing to participate. Residents in states without a federally approved driver's license could experience serious inconvenience. Congress passed REAL ID in 2005, and it was originally to take effect in 2008. The Department of Homeland Security extended the deadline for

Accuracy of worker verification database: Section 6.3.1

compliance several times, while some members of Congress have been working to modify or repeal REAL ID. As I write this, the deadline remains in the future, and Congress has not repealed the law.

Many European and Asian countries require national ID cards. An unpopular plan for an expensive mandatory national ID card in the United Kingdom stalled when emails about weaknesses of the plan leaked from government offices. The government of Japan implemented a national computerized registry system that included assigning an ID number to every citizen of the country. The system is for government purposes, initially with approximately 100 applications, but eventually its uses will probably be in the thousands. The intention is to simplify administration procedures and make them more efficient. Privacy advocates and protesters have complained of insufficient privacy protection, potential abuse by government, and vulnerability to hackers. The Indian government is building a national ID database for its 1.2 billion people. The database will include each person's photo, fingerprints, iris scan, birth date, and other information. Its stated purposes include improving provision of government services and catching illegal immigrants.

> *As soon as you are willing to put your home, your office, your safe deposit box, your bike lock, your gym key, and your desk key all onto one and ask the government to issue that one key, you will be okay with the national ID. But until then, we need to think more in terms of diversification of identification systems.*
>
> —Jim Harper, Director of Information Policy Studies, Cato Institute[67]

2.5 Protecting Privacy: Technology, Markets, Rights, and Laws

2.5.1 Technology and Markets

Many individuals, organizations, and businesses help meet the demand for privacy to some degree: Individual programmers post free privacy-protecting software on the Web. Entrepreneurs build new companies to provide technology-based privacy protections. Large businesses respond to consumer demand and improve policies and services. Organizations such as the Privacy Rights Clearinghouse provide excellent information resources. Activist organizations such as the Electronic Privacy Information Center inform the public, file lawsuits, and advocate for better privacy protection.

New applications of technology can often solve problems that arise as side effects of technology. Soon after "techies" became aware of the use of cookies by Web sites, they wrote cookie disablers and posted them on the Web. Software to block pop-up ads appeared soon after the advent of such ads. People figured out how to prevent ads from appearing in their Gmail and told the world. Companies sell software to scan for spyware;

some versions are free. We can install free add-ons to our browsers that block Web activity trackers. Several companies provide services, called anonymizers, with which people can surf the Web anonymously, leaving no record that identifies them or their computers. Some search engines do not store user search queries in a way that allows linking them together to one person.[68] Companies offer products and services to prevent forwarding, copying, or printing email. (Lawyers are among the major customers.) There are services that fully erase email or text messages (on both the sender's and recipient's phones) after a user-specified time period. They can be helpful for doctors, who must follow very strict medical privacy regulations. Some tracking systems for laptops, tablets, and phones include a feature that allows the owner of a stolen or lost laptop to encrypt, retrieve, and/or erase files remotely.


More about anonymizers: Section 3.4

These are a very few examples of the many products and technology applications that protect privacy. They illustrate that individuals, businesses, and organizations are quick to respond and make privacy-protecting tools available. They have advantages and disadvantages; they do not solve all problems. Learning about, installing, and using privacy tools might be daunting to nontechnical, less educated users—a large part of the public—hence the importance of designing systems with privacy protection in mind, building in protective features, and having privacy-protecting policies.


Protections against identity theft: Section 5.3.2

Encryption

> *Cryptography is the art and science of hiding data in plain sight.*
>
> —Larry Loen[69]

It is possible to intercept email and data in transit on the Internet and to pick wireless transmissions out of the air. Someone who steals a computer or hacks into one can view files on it. Most eavesdropping by private citizens is illegal. Hacking and stealing laptops are crimes. The law provides for punishment of offenders when caught and convicted, but we can also use technology to protect ourselves.

Encryption is a technology, often implemented in software, that transforms data into a form that is meaningless to anyone who might intercept or view it. The data could be email, business plans, credit card numbers, images, medical records, cellphone location history, and so on. Software at the recipient's site (or on one's own computer) decodes encrypted data so that the recipient or owner can view the messages or files. Software routinely encrypts credit card numbers when we send them to online merchants. People are often not even aware that they are using encryption. The software handles it automatically.

Many privacy and security professionals view encryption as the most important technical method for ensuring the privacy of messages and data sent through computer networks. Encryption also protects stored information from intruders and abuses by

employees. It is the best protection for data on laptops and other small data storage devices carried outside an office.

Encryption generally includes a coding scheme, or cryptographic algorithm, and specific sequences of characters (e.g., digits or letters), called *keys*, used by the algorithm. Using mathematical tools and powerful computers, it is sometimes possible to "break" an encryption scheme—that is, to decode an encrypted message or file without the secret key.

Modern encryption technology has a flexibility and variety of applications beyond protecting data. For example, it is used to create digital signatures, authentication methods, and digital cash. Digital signature technology allows us to "sign" documents online, saving time and paper for loan applications, business contracts, and so on. In one specialized authentication application, aimed at reducing the risk of unauthorized access to medical information online, the American Medical Association issues digital credentials to doctors that a laboratory website can verify when a doctor visits to get patient test results. There are likely to be thousands of applications of this technology.

Digital cash and other encryption-based privacy-protected transaction methods can let us do secure financial transactions electronically without the seller acquiring a credit card or checking account number from the buyer. They combine the convenience of credit card purchases with the anonymity of cash. With such schemes, it is not easy to link records of different transactions to form a consumer profile or dossier. These techniques can provide both privacy protection for the consumer with respect to the organizations he or she interacts with and protection for organizations against forgery, bad checks, and credit card fraud. However, cash transactions make it harder for governments to detect and prosecute people who are "laundering" money earned in illegal activities, earning money they are not reporting to tax authorities, or transferring or spending money for criminal purposes. Thus, most governments would oppose and probably prohibit a truly anonymous digital cash system. Some digital cash systems include provisions for law enforcement and tax collection. The potential illegal uses of digital cash have long been possible with real cash. It is only in recent decades, with increased use of checks and credit cards, that we lost the privacy we had from marketers and government when we used cash for most transactions.

> *The technologies of anonymity and cryptography may be the only way to protect privacy.*
>
> —Nadine Strossen, president of the American Civil Liberties Union[70]

Policies for protecting personal data

The businesses, organizations, and government agencies that collect and store personal data have an ethical responsibility (and in many cases a legal one) to protect it from misuse. Responsible data holders must anticipate risks and prepare for them. They must continually update security policies to cover new technologies and new potential threats.

Encryption Policy

For centuries before the Internet, governments, their military agencies, and their spies were the main users of codes. For decades, most of the cryptographers in the United States worked for the National Security Agency (NSA). The NSA almost certainly could break virtually any codes that were in use until the early 1970s.[71] The NSA worked hard to keep everything about encryption secret. In the 1970s, a private-sector breakthrough called public key cryptography produced encryption that was relatively easy to use and very difficult to crack. Keeping encryption as an exclusive tool of governments and spies was no longer an option.

More about the NSA: Section 2.6.3

Throughout the 1990s, when people began using encryption for email and other purposes, the U.S. government battled the Internet community and privacy advocates to restrict the availability of secure encryption (that is, encryption that is so difficult and expensive to crack that it is not practical to do so.) It maintained a costly and ultimately futile policy of prohibiting export of powerful encryption software. The government interpreted anything posted on the Internet as effectively exported. Thus, even researchers who posted encryption algorithms on the Net faced possible prosecution. The government argued that the export prohibition was necessary to keep strong encryption from terrorists and enemy governments. The U.S. policy was strangely out of date. The stronger encryption schemes were available on Internet sites all over the world.

The National Research Council (the research affiliate of the National Academy of Sciences) strongly supported the use of powerful encryption and the loosening of export controls. It argued that strong encryption provides increased protection against hackers, thieves, and terrorists who threaten our economic, energy, and transportation infrastructures.[72] The need for strong encryption in electronic commerce was becoming obvious as well.

Concurrently with the ban on export of strong encryption, the government attempted to ensure its access to encryption keys (or to the unencrypted content of encrypted messages) for encryption used within the United States. Pedophiles and child molesters encrypt child pornography on their computers. Other criminals encrypt email and files to hide their contents from law enforcement agents. The FBI supported a bill requiring a loophole, or "backdoor," in all encryption products made, sold, or used in the United States to permit immediate decryption of the encrypted data upon the receipt of a court order.[73] The FBI argued that authority to intercept telephone calls or email or seize computers meant nothing if agents could not read what they seized. Technical experts argued that such a law would be extraordinarily difficult to implement because encryption is now part of Web browsers and many other common computing tools. Implementation of an immediate decryption mechanism would threaten privacy and seriously weaken security of electronic commerce and communications.

During the same time, courts considered legal challenges to the export restrictions based on the First Amendment. The question is whether cryptography algorithms, and computer programs in general, are speech and hence protected by the First Amendment. The government argued that software is not speech and that control of cryptography was a national security issue, not a freedom-of-speech issue. The federal judge who heard the case thought otherwise. She said:

> This court can find no meaningful difference between computer language . . . and German or French. . . . Like music and mathematical equations, computer language is just that, language, and it communicates information either to a computer or to those who can read it. . . . For the purposes of First Amendment analysis, this court finds that source code is speech.[74]

The U.S. government removed almost all export restrictions on encryption in 2000. Congress did not pass a law requiring all encryption to have a mechanism for law enforcement access. Among thousands of wiretaps approved for criminal investigations in 2010, law enforcement agents encountered encryption only six times and were able to obtain the plain text of the messages.[75]

Employers must train those who carry around personal data about the risks and proper security measures.

A well-designed database for sensitive information includes several features to protect against leaks, intruders, and unauthorized employee access. Each person with authorized access to the system should have a unique identifier and a password. A system can restrict users from performing certain operations, such as writing or deleting, on some files. User IDs can be coded so that they give access to only specific parts of a record. For example, a billing clerk in a hospital does not need access to the results of a patient's lab tests. The computer system keeps track of information about each access, including the ID of the person looking at a record and the particular information viewed or modified. This is an *audit trail* that can later help trace unauthorized activity. The knowledge that a system contains such provisions will discourage many privacy violations.

Databases with consumer information, Web-activity records, or cellphone location data are valuable assets that give businesses a competitive advantage. The owners of such data have an interest in preventing leaks and unlimited distribution. That includes providing security for the data and developing modes of operation that reduce loss. Thus, for example, mailing lists are usually not sold; they are "rented." The renter does not receive a copy (electronic or otherwise). A specialized firm does the mailing. The risk of unauthorized copying is thus restricted to a small number of firms whose reputation for honesty and security is important to their business. Other applications also use this idea of trusted third parties to process confidential data. Some car rental agencies access a third-party service to check the driving record of potential customers. The service examines the motor vehicle department records; the car rental company does not see the driver's record.

Website operators pay thousands, sometimes millions, of dollars to companies that do *privacy audits*. Privacy auditors check for leaks of information, review the company's privacy policy and its compliance with that policy, evaluate warnings and explanations on its website that alert visitors when the site requests sensitive data, and so forth. Hundreds of large businesses have a position called *chief privacy officer*. This person guides company privacy policy. Just as the Automobile Association of America rates hotels, the Better Business Bureau and similar organizations offer a seal of approval, an icon companies that comply with their privacy standards can post on websites.

Large companies use their economic influence to improve consumer privacy. IBM and Microsoft removed Internet advertising from websites that do not post clear privacy policies. Walt Disney Company and Infoseek Corporation did the same and, in addition, stopped accepting advertising on their websites from sites that do not post privacy policies. The Direct Marketing Association adopted a policy requiring its member companies to inform consumers when they will share personal information with other marketers and to give people an opt-out option. Many companies agreed to limit the availability of sensitive consumer information, including unlisted telephone numbers, driving histories, and all information about children.

There continue, of course, to be many businesses without strong privacy policies, as well as many that do not follow their own stated policies. The examples described here represent a trend, not a privacy utopia. They suggest actions responsible companies can take. As some problems are addressed, new ones continually arise.

2.5.2 Rights and Law

In Section 2.2, we considered some aspects of law and Fourth Amendment principles related to protection of privacy. The Fourth Amendment protects the negative right (a liberty) against intrusion and interference by government. This section focuses mainly on discussion of principles related to rights and legal protections for personal data collected or used by other people, businesses, and organizations.

We separate legal remedies from technical, management, and market solutions because they are fundamentally different. The latter are voluntary and varied. Different people or businesses can choose from among them. Law, on the other hand, is enforced by fines, imprisonment, and other penalties. Thus, we should examine the basis for law more carefully. Privacy is a condition or state we can be in, like good health or financial security. To what extent should we have a legal right to it? Is it a negative right or a positive right (in the sense of Section 1.4.2)? How far should law go, and what should be left to the voluntary interplay of markets, educational efforts of public interest groups, consumer choices and responsibilities, and so forth?

Until the late 19th century, courts based legal decisions supporting privacy in social and business activities on property rights and contracts. There was no recognition of an independent right to privacy. In 1890, a crucial article called "The Right of Privacy," by Samuel Warren and Louis Brandeis[76] (later a Supreme Court Justice), argued that privacy was distinct from other rights and needed more protection. Judith Jarvis Thomson, an MIT philosopher, argued that the old view was more accurate, that in all cases where infringement of privacy is a violation of someone's rights, that violation is of a right distinct from privacy.[77] We present some of the claims and arguments of these papers. Then we consider a variety of other ideas and perspectives about laws to protect privacy.

One purpose of this section is to show the kinds of analyses that philosophers, legal scholars, and economists perform in trying to elucidate underlying principles. Another is to emphasize the importance of principles, of working out a theoretical framework in which to make decisions about particular issues and cases.

Warren and Brandeis: The inviolate personality

The main target of criticism in the 1890 Warren and Brandeis article is newspapers, especially the gossip columns. Warren and Brandeis vehemently criticize the press for "overstepping . . . obvious bounds of propriety and decency." The kinds of information of most concern to them are personal appearance, statements, acts, and interpersonal relationships (marital, family, and others).[78] Warren and Brandeis take the position that

people have the right to prohibit publication of facts about themselves and photographs of themselves. Warren and Brandeis argue that, for example, if someone writes a letter in which he says he had a fierce argument with his wife, the recipient of the letter cannot publish that information. They base this claim on no property right or other right except privacy. It is part of the right to be let alone. Warren and Brandeis base their defense of privacy rights on, in their often-quoted phrase, the principle of "an inviolate personality."

Laws against other wrongs (such as slander, libel, defamation, copyright infringement, violation of property rights, and breach of contract) can address some privacy violations, but Warren and Brandeis argue that there remain many privacy violations that those other laws do not cover. For example, publication of personal or business information could constitute a violation of a contract (explicit or implied), but there are many cases in which the person who discloses the information has no contract with the victim. The person is not violating a contract but is violating the victim's privacy. Libel, slander, and defamation laws protect us when someone spreads false and damaging rumors about us, but they do not apply to true personal information whose exposure makes us uncomfortable. Warren and Brandeis say privacy is distinct and needs its own protection. They allow exceptions for publication of information of general interest (news), use in limited situations when the information concerns another person's interests, and oral publication. (They were writing before radio and television, so oral publication meant a quite limited audience.)

Judith Jarvis Thomson: Is there a right to privacy?

Judith Jarvis Thomson argues the opposite point of view. She gets to her point after examining a few scenarios.

Suppose you own a copy of a magazine. Your property rights include the right to refuse to allow others to read, destroy, or even see your magazine. If someone does anything to your magazine that you did not allow, that person is violating your property rights. For example, if someone uses binoculars to see your magazine from a neighboring building, that person is violating your right to exclude others from seeing it. It does not matter whether the magazine is an ordinary news magazine (not a sensitive privacy issue) or some other magazine you do not want people to know you read. The right violated is your property right.

You may waive your property rights, intentionally or inadvertently. If you absent-mindedly leave the magazine on a park bench, someone could take it. If you leave it on the coffee table when you have guests at your home, someone could see it. If you read a pornographic magazine on a bus, and someone sees you and tells other people that you read dirty magazines, that person is not violating your rights. The person might be doing something impolite, unfriendly, or cruel, but not something that violates a right.

Our rights to our person and our bodies include the right to decide to whom we show various parts of our bodies. By walking around in public, most of us waive our right to prevent others from seeing our faces. When a Muslim woman covers her face, she is

exercising her right to keep others from viewing it. If someone uses binoculars to spy on us at home in the shower, they are violating our right to our person.

If someone beats on you to get some information, the beater is violating your right to be free from physical harm done by others. If the information is the time of day, privacy is not at issue. If the information is more personal, then they have compromised your privacy, but the right violated is your right to be free from attack. On the other hand, if a person peacefully asks whom you live with or what your political views are, they have violated no rights. If you choose to answer and do not make a confidentiality agreement, the person is not violating your rights by repeating the information to someone else, though it could be inconsiderate to do so. However, if the person agreed not to repeat the information, but then does, it does not matter whether or not the information was sensitive; the person is violating the confidentiality agreement.

In these examples, there is no violation of privacy without violation of some other right, such as the right to control our property or our person, the right to be free from violent attack, or the right to form contracts (and expect them to be enforced). Thomson concludes, "I suggest it is a useful heuristic device in the case of any purported violation of the right to privacy to ask whether or not the act is a violation of any other right, and if not whether the act really violates a right at all."[79]

Criticisms of Warren and Brandeis and of Thomson

Critics of the Warren and Brandeis position[80] argue that it does not provide a workable principle or definition from which to conclude that a privacy right violation occurs. Their notion of privacy is too broad. It conflicts with freedom of the press. It appears to make almost any unauthorized mention of a person a violation of the person's right.

Critics of Thomson present examples of violations of a right to privacy (not just a desire for privacy), but of no other right. Some view Thomson's notion of the right to our person as vague or too broad. Her examples might (or might not) be a convincing argument for the thesis that considering other rights can resolve privacy questions, but no finite number of examples can prove such a thesis.

Neither article directly refutes the other. Their emphases are different. Warren and Brandeis focus on the use of the information (publication). Thomson focuses on how it is obtained. This distinction sometimes underlies differences in arguments by those who advocate strong legal regulations on use of personal data and those who advocate more reliance on technical, contractual, and market solutions.

Applying the theories

How do the theoretical arguments apply to privacy and personal data today?

Throughout Warren and Brandeis, the objectionable action is publication of personal information—its widespread, public distribution. Many court decisions since the appearance of their article have taken this point of view.[81] If someone published information from a consumer databases (in print or by making it public on the Web), that would

violate the Warren and Brandeis notion of privacy. A person might win a case if someone published his or her consumer profile. But intentional publication is not the main concern in the current context of consumer databases, monitoring of Web activity, location tracking, and so on. The amount of personal information collected nowadays might appall Warren and Brandeis, but their article allows disclosure of personal information to people who have an interest in it. By implication, they do not preclude, for example, disclosure of a person's driving record to a car rental company from which he or she wants to rent a car. Similarly, it seems Warren and Brandeis would not oppose disclosure of information about whether someone smokes cigarettes to a life insurance company from whom the person is trying to buy insurance. Their view does not rule out use of (unpublished) consumer information for targeted marketing, though they probably would disapprove of it.

The content of social networks would probably shock and appall Warren and Brandeis. Their position would severely restrict the sharing of photos that include other people and of the location and activities of friends.

An important aspect of both the Warren and Brandeis paper and the Thomson paper is that of consent. They see no privacy violation if a person consented to the collection and use of the information.

Transactions

We have another puzzle to consider: how to apply philosophical and legal notions of privacy to transactions, which automatically involve more than one person. The following scenario will illustrate the problem.

One day in the small farm community of Friendlyville, Joe buys five pounds of potatoes from Maria, who sells him the five pounds of potatoes. (I describe the transaction in this repetitious manner to emphasize that there are two people involved and two sides to the transaction.)

Either Joe or Maria might prefer the transaction to remain secret. The failure of his own potato crop might embarrass Joe. Or Joe might be unpopular in Friendlyville, and Maria fears the townspeople will be angry at her for selling to him. Either way, we are not likely to consider it a violation of the other's rights if Maria or Joe talks about the purchase or sale of the potatoes to other people in town. But suppose Joe asks for confidentiality as part of the transaction. Maria has three options. (1) She can agree. (2) She can say no; she might want to tell people she sold potatoes to Joe. (3) She can agree to keep the sale confidential if Joe pays a higher price. In the latter two cases, Joe can decide whether to buy the potatoes. On the other hand, if Maria asks for confidentiality as part of the transaction, Joe has three options. (1) He can agree. (2) He can say no; he might want to tell people he bought potatoes from Maria. (3) He can agree to keep the purchase confidential if Maria charges a lower price. In the latter two cases, Maria can decide whether to sell the potatoes.

Privacy includes control of information about oneself. Is the transaction a fact about Maria or a fact about Joe? There does not appear to be a convincing reason for either party to have more right than the other to control information about the transaction. Yet this problem is critical to legal policy decisions about use of consumer information. If we are to assign control of the information about a transaction to one of the parties, we need a firm philosophical foundation for choosing which party gets it. (If the parties make a confidentiality agreement, then they have an ethical obligation to respect it. If the agreement is a legal contract, then they have a legal obligation to respect it.)

Philosophers and economists often use simple two-person transactions or relationships, like the Maria/Joe scenario, to try to clarify the principles involved in an issue. Do the observations and conclusions about Maria and Joe generalize to large, complex societies and a global economy, where, often, one party to a transaction is a business? All transactions are really between people, even if indirectly. So if a property right or a privacy right in the information about a transaction goes to one of the parties, we need an argument showing how the transaction in a modern economy is different from the one in Friendlyville. Later in this section, we describe two viewpoints on the regulation of information about consumer transactions: the free market view and the consumer protection view. The consumer protection view suggests treating the parties differently.

Ownership of personal data

Some economists, legal scholars, and privacy advocates propose giving people property rights in information about themselves. The concept of property rights can be useful even when applied to intangible property (intellectual property, for example), but there are problems in using this concept for personal information. First, as we have just seen, activities and transactions often involve at least two people, each of whom would have reasonable but conflicting claims to own the information about the transaction. Some personal information does not appear to be about a transaction, but there still can be problems in assigning ownership. Do you own your birthday? Or does your mother own it? After all, she was a more active participant in the event.

The second problem with assigning ownership of personal information arises from the notion of owning facts. (Copyright protects intellectual property such as computer programs and music, but we cannot copyright facts.) Ownership of facts would severely impair the flow of information in society. We store information on electronic devices, but we also store it in our minds. Can we own facts about ourselves without violating the freedom of thought and freedom of speech of others?

Although there are difficulties with assigning ownership in individual facts, another issue is whether we can own our "profiles," that is, a collection of data describing our activities, purchases, interests, and so on. We cannot own the fact that our eyes are blue, but we do have the legal right to control some uses of our photographic image. In almost all states, we need a person's consent to use his or her image for commercial purposes. Should the law treat our consumer profiles the same way? Should the law treat the collection of

our search queries the same way? How can we distinguish between a few facts about a person and a "profile"?

Judge Richard Posner, a legal scholar who has extensively studied the interactions between law and economics, gives economic arguments about how to allocate property rights to information.[82] Information has both economic and personal value, he points out. It is valuable to us to determine if a business, customer, client, employer, employee, and so on, is reliable, honest, and so on. Personal and business interactions have many opportunities for misrepresentation and therefore exploitation of others. Posner's analysis leads to the conclusion that, in some cases, individuals or organizations should have a property right to information, while in other cases, they should not. That is, some information should be in the public domain. A property right in information is appropriate where the information has value to society and is expensive to discover, create, or collect. Without property rights to such information, the people or businesses that make investments in discovering or collecting the information will not profit from it. The result is that people will produce less of this kind of information, to the detriment of society. Thus, the law should protect, for example, trade secrets, the result of much expenditure and effort by a business. A second example is personal information, such as the appearance of one's naked body. It is not expensive for a person to obtain, but virtually all of us place value on protecting it, and concealment is not costly to society. So it makes sense to assign the property right in this information to the individual. Some privacy advocates want to protect information that can lead to denial of a job or some kind of service or contract (e.g., a loan). They advocate restrictions on sharing of information that might facilitate negative decisions about people—for example, landlords sharing a database with information about tenant payment histories. Posner argues that a person should not have a property right to negative personal information or other information whose concealment aids people in misrepresentation, fraud, or manipulation. Such information should be in the public domain. That means a person should not have the right to prohibit others from collecting it, using it, and passing it on, as long as they are not violating a contract or confidentiality agreement and do not obtain the information by eavesdropping on private communications or by other prohibited means.

In recent decades, the trend in legislation has not followed Posner's position. Some critics of Posner's point of view believe that moral theory, not economic principles, should be the source of property rights.

A basic legal framework

A good basic legal framework that defines and enforces legal rights and responsibilities is essential to a complex, robust society and economy. One of its tasks is enforcement of agreements and contracts. Contracts—including freedom to form them and enforcement of their terms by the legal system—are a mechanism for implementing flexible and diverse economic transactions that take place over time and between people who do not know each other well or at all.

We can apply the idea of contract enforcement to the published privacy policies of businesses and organizations. The Toysmart case is an example. Toysmart, a Web-based seller of educational toys, collected extensive information on about 250,000 visitors to its website, including family profiles, shopping preferences, and names and ages of children. Toysmart had promised not to release this personal information. When the company filed for bankruptcy, it had a large amount of debt and virtually no assets—except its customer database, which had a high value. Toysmart's creditors wanted the database sold to raise funds to repay them. Toysmart offered the database for sale, causing a storm of protest. Consistent with the interpretation that Toysmart's policy was a contract with the people in the database, the bankruptcy-court settlement included destruction of the database.[83]

A second task of a legal system is to set defaults for situations that contracts do not explicitly cover. Suppose a website posts no policy about what it does with the information it collects. What legal rights should the operator of the site have regarding the information? Many sites and offline businesses act as though the default is that they can do anything they choose. A privacy-protecting default would be that they can use the information only for the direct and obvious purpose for which they collected it. The legal system can (and does) set special confidentiality defaults for sensitive information, such as medical and financial information, that tradition and most people consider private. If a business or organization wants to use information for purposes beyond the default, it would have to specify those uses in its policies, agreements, or contracts or request consent. Many business interactions do not have written contracts, so the default provisions established by law can have a big impact.

A third task of a basic legal structure is to specify penalties for criminal offenses and breach of contracts. Thus, law can specify penalties for violation of privacy policies and negligent loss or disclosure of personal data that businesses and others hold. Writers of liability laws must strike a balance between being too strict and too lenient. If too strict, they make some valuable products and services too expensive to provide. If too weak, they provide insufficient incentive for businesses and government agencies to provide reasonable security for our personal data.


More about liability issues: Section 8.3.3

Regulation

Technical tools, market mechanisms, and business policies for privacy protection are not perfect. Is that a strong argument for regulatory laws? Regulation is not perfect either. We must evaluate regulatory solutions by considering effectiveness, costs and benefits, and side effects, just as we evaluate other kinds of potential solutions to problems caused by technology. The pros and cons of regulation fill entire books. We briefly make a few points here. (We will see similar problems in Section 8.3.3 when we consider responses to computer errors and failures.)

There are hundreds of privacy laws. When Congress passes laws for complex areas like privacy, the laws usually state general goals and leave the details to government agencies

that write hundreds or thousands of pages of regulations, sometimes over many years. It is extremely difficult to write reasonable regulations for complex situations. Laws and regulations often have unintended effects or interpretations. They can apply where they do not make sense or where people simply do not want them.

Regulations often have high costs, both direct dollar costs to businesses (and, ultimately, consumers) and hidden or unexpected costs, such as loss of services or increased inconvenience. For example, regulations that prohibit broad consent agreements and instead require explicit consent for each secondary use of personal information have an attribute economists call "high transaction cost." The consent requirement could be so expensive and difficult to implement that it eliminates most secondary uses of information, including those that consumers find desirable.

Although regulations have disadvantages, we should remember that businesses sometimes overestimate the cost of privacy regulations. They also sometimes underestimate the costs, to themselves and to consumers, of not protecting privacy.[84]

Contrasting Viewpoints

> *When asked "If someone sues you and loses, should they have to pay your legal expenses?" more than 80% of people surveyed said "yes." When asked the same question from the opposite perspective: "If you sue someone and lose, should you have to pay their legal expenses?" about 40% said "yes."*

The political, philosophical, and economic views of many scholars and advocates who write about privacy differ. As a result, their interpretations of various privacy problems and their approaches to solutions often differ, particularly when they are considering laws and regulation to control collection and use of personal information by businesses.* We contrast two perspectives. I call them the free market view and the consumer protection view.

The free market view

People who prefer market-oriented solutions for privacy problems tend to emphasize the freedom of individuals, as consumers or in businesses, to make voluntary agreements; the diversity of individual tastes and values; the flexibility of technological and market solutions; the response of markets to consumer preferences; the usefulness and importance of contracts; and the flaws of detailed or restrictive legislation and regulatory solutions. They emphasize the many voluntary organizations that provide consumer education, develop guidelines, monitor the activities of business and government, and pressure

* There tends to be more agreement among privacy advocates when considering privacy threats and intrusions by government.

businesses to improve policies. They may take strong ethical positions but emphasize the distinction between the role of ethics and the role of law.

A free market view for collection and use of personal information emphasizes informed consent: Organizations collecting personal data (including government agencies and businesses) should clearly inform the person providing the information if they will not keep it confidential (from other businesses, individuals, and government agencies) and how they will use it. They should be legally liable for violations of their stated policies. This viewpoint could consider truly secret forms of invisible information gathering to be theft or intrusion.

A free market view emphasizes freedom of contract: People should be free to enter agreements (or not enter agreements) to disclose personal information in exchange for a fee, services, or other benefits according to their own judgment. Businesses should be free to offer such agreements. This viewpoint respects the right and ability of consumers to make choices for themselves based on their own values. Market supporters expect consumers to take the responsibility that goes with freedom—for example, to read contracts or to understand that desirable services have costs. A free market view includes free flow of information: the law should not prevent people (or businesses and organizations) from using and disclosing facts they independently or unintrusively discover without violating rights (e.g., without theft, trespass, or violation of contractual obligations).

We cannot always expect to get exactly the mix of attributes we want in any product, service, or job. Just as we might not get cheeseless pizza in every pizza restaurant or find a car with the exact set of features we want, we might not always be able to get both privacy and special discounts—or free services. We might not be able to get certain websites—or magazines—without advertising, or a specific job without agreeing to provide certain personal information to the employer. These compromises are not unusual or unreasonable when interacting with other people.

Market supporters prefer to avoid restrictive legislation and detailed regulation for several reasons. Overly broad, poorly designed, and vague regulations stifle innovation. The political system is a worse system than the market for determining what consumers want in the real world of trade-offs and costs. It is impossible for legislators to know in advance how much money, convenience, or other benefits people will want to trade for more or less privacy. Businesses respond over time to the preferences of millions of consumers expressed through their purchases. In response to the desire for privacy many people express, the market provides a variety of privacy protection tools. Market supporters argue that laws requiring specific policies or prohibiting certain kinds of contracts violate the freedom of choice of both consumers and business owners.

This viewpoint includes legal sanctions for those who steal data and those who violate confidentiality agreements. It holds businesses, organizations, and government agents responsible for loss of personal data due to poor or negligent security practices. To encourage innovation and improvement, advocates of this viewpoint are more likely to prefer penalties when a company loses, inappropriately discloses, or abuses the data, rather

than regulations that specify detailed procedures that holders of personal information must follow.

The free market viewpoint sees privacy as a "good," both in the sense that it is desirable and that it is something we can obtain varying amounts of by buying or trading in the economy, like food, entertainment, and safety. Just as some people choose to trade some safety for excitement (bungee jumping, motorcycle riding), money (buying a cheaper but less safe product), or convenience, some choose different levels of privacy. As with safety, law can provide minimum standards, but it should allow the market to provide a wide range of options to meet the range of personal preferences.

The consumer protection view

Advocates of strong privacy regulation emphasize the unsettling uses of personal information we have mentioned throughout this chapter, the costly and disruptive results of errors in databases (which we discuss in Chapter 8) and the ease with which personal information leaks out, via loss, theft, and carelessness. They argue for more stringent consent requirements, legal restrictions on consumer profiling, prohibitions on certain types of contracts or agreements to disclose data, and prohibitions on businesses collecting or storing certain kinds of data. They urge, for example, that the law require companies to have opt-in policies for secondary uses of personal information, because the opt-out option might not be obvious or easy enough for consumers who would prefer it. They would prohibit waivers and broad consent agreements for secondary uses.

The focus of this viewpoint is to protect consumers against abuses and carelessness by businesses and against their own lack of knowledge, judgment, or interest. Advocates of the consumer protection view emphasize that people do not realize all the ways others may use information about them. They do not understand the risks of agreeing to disclose personal data. Those who emphasize consumer protection are critical of programs to trade free devices and services for personal information or consent for monitoring or tracking. Many support laws prohibiting collection or storage of personal data that could have negative consequences, if they believe the risks are more important than the value of the information to the businesses that want to collect it. Consumer advocate and privacy "absolutist" Mary Gardiner Jones objected to the idea of consumers consenting to dissemination of personal data. She said, "You can't expect an ordinary consumer who is very busy trying to earn a living to sit down and understand what [consent] means. They don't understand the implications of what use of their data can mean to them."[85] She said this roughly 20 years ago. Understanding the implications of the ways data are collected and used now is more difficult. A former director of the ACLU's Privacy and Technology Project expressed the view that informed consent is not sufficient protection. She urged a Senate committee studying confidentiality of health records to "re-examine the traditional reliance on individual consent as the linchpin of privacy laws."[86]

Those who emphasize the consumer protection point of view would argue that the Joe/Maria scenario in Friendlyville, described earlier in this section, is not relevant in a

complex society. The imbalance of power between the individual and a large corporation is one reason. Another is that in Friendlyville the information about the transaction circulates to only a small group of people, whom Joe and Maria know. If someone draws inaccurate or unfair conclusions, Joe or Maria can talk to the person and present his or her explanations. In a larger society, information circulates among many strangers, and we often do not know who has it and what decisions about us they base on it.

A consumer cannot realistically negotiate contract terms with a business. At any specific time, the consumer can only accept or reject what the business offers. And the consumer is often not in a position to reject it. If we want a loan for a house or car, we have to accept whatever terms lenders currently offer. If we need a job, we are likely to agree to disclose personal information against our true preference because of the economic necessity of working. Individuals have no meaningful power against large companies like Google and Apple. They have to use search engines whether or not they know or accept a company's policy about use of their search queries.

In the consumer protection view, self-regulation by business does not work. Business privacy policies are weak, vague, or difficult to understand. Businesses sometimes do not follow their stated policies. Consumer pressure is sometimes effective, but some companies ignore it. Instead, we must require all businesses to adopt pro-privacy policies. Software and other technological privacy-protecting tools for consumers cost money, and many people cannot afford them. They are far from perfect anyway and hence not good enough to protect privacy.

The consumer protection viewpoint sees privacy as a right rather than something we bargain about. For example, a website jointly sponsored by the Electronic Privacy Information Center and Privacy International flashes the slogans "Privacy is a right, not a preference" and "Notice is not enough."[87] The latter indicates that they see privacy as a positive right, or claim right (in the terminology of Section 1.4.2). As a negative right, privacy allows us to use anonymizing technologies and to refrain from interacting with those who request information we do not wish to supply. As a positive right, it means we can stop others from communicating about us. A spokesperson for the Center for Democracy and Technology expressed that view in a statement to Congress, saying that we must incorporate into law the principle that people should be able to "determine for themselves when, how and to what extent information about them is shared."[88]

2.5.3 Privacy Regulations in the European Union

The European Union (EU) has a comprehensive Data Protection Directive (passed in 1995).[89] It covers processing of personal data, including collection, use, storage, retrieval, transmission, destruction, and other actions. The directive sets forth Fair Information Principles that EU member nations must implement in their own laws. Several are similar to the first five principles in Figure 2.2 (in Section 2.1.3). The EU has some additional or stronger rules. They include:

- Processing of data is permitted only if the person has consented unambiguously or if the processing is necessary to fulfill contractual or legal obligations or is needed for tasks in the public interest or by official authorities to accomplish their tasks (or a few other reasons).
- Special categories of data—including ethnic and racial origin, political and religious beliefs, health and sex life, and union membership—must not be processed without the person's explicit consent. Member nations may outlaw processing of such data even if the subject does consent.
- Processing of data about criminal convictions is severely restricted.

The EU's rules are stricter than those in the United States, as the next few examples illustrate.

Google modified its privacy policy in 2012 to allow the company to combine information it collects on members from its various services. The EU argued that average users could not understand how Google uses their data under the new policy and that that violates the EU's privacy regulations. A court in Germany said that some of Facebook's policies in its member agreement (for example, granting Facebook a license to use material a member posts or stores at Facebook) are illegal there. The German government told Facebook to stop running face recognition applications on German users; it violates German privacy laws.

The EU devised legal guidelines for social networking sites. The guidelines say the sites should set default privacy settings at a high level, tell users to upload a picture of a person only if the person consents, allow the use of psuedonyms, set limits on the time they retain data on inactive users, and delete accounts that are inactive for a long time.

The European Commission proposed granting a legal "right to be forgotten." It would, among other things, require that a website remove information, photos, and so on, of a particular person if that person requests it, whether that person posted the material or someone else did. It appears also to require that search engines remove links to material a person wants removed. Such a "right" clearly conflicts with freedom of speech in cases where another person posted the material and does not want it removed.


More about a right to be forgotten: Section 2.3.4

A Spanish government agency ordered Google to remove links from its search results to dozens of articles that have sensitive information about individual people. (Google fought the demand in European court, arguing that the order violated freedom of expression and that the government did not require news media to remove the articles.) Because of Germany's strict privacy laws, Google's Street View allowed anyone to request that their home or office be blurred out on its street images. Google won a lawsuit about Street View violating a homeowner's privacy, but the company discontinued taking photos for Street View in Germany.[90]

While the EU has much stricter regulations than the United States on collection and use of personal information by the private sector, some civil libertarians believe that the

regulations do not provide enough protection from use of personal data by government agencies. Although the directive says that data should not be kept longer than necessary, European countries require that ISPs and telephone companies retain records of customer communications (date, destination, duration, and so on) for up to two years and make them available to law enforcement agencies. The EU said it needs this requirement to fight terrorism and organized crime.[91]

The EU's strict privacy directive does not prevent some of the same abuses of personal data that occur in the United States. In Britain, for example, the Information Commissioner reported that data brokers use fraud and corrupt insiders to get personal information. As in the United States, customers of illegal services include journalists, private investigators, debt collectors, government agencies, stalkers, and criminals seeking data to use for fraud.[92]

The EU Data Privacy Directive prohibits transfer of personal data to countries outside the European Union that do not have an adequate system of privacy protection. This part of the directive caused significant problems for companies that do business both in and outside Europe and might normally process customer and employee data outside the EU. The EU determined that Australia, for example, did not have adequate privacy protection. Australia allows businesses to create their own privacy codes consistent with the government's National Privacy Principles. The United States has privacy laws covering specific areas such as medical information, video rentals, driver's license records, and so on, but does not have comprehensive privacy laws covering all personal data. The EU agreed to the "Safe Harbor" plan, under which companies outside the EU that agree to abide by a set of privacy requirements similar to the principles in the Data Protection Directive may receive personal data from the EU.[93] After the terrorist attacks in 2001, screening of air travel passengers from Europe to the United States raised problems. The U.S. government wanted more information about the passengers than the EU wanted to provide.

Many privacy advocates describe U.S. privacy policy as "behind Europe" because the United States does not have comprehensive federal legislation regulating personal data collection and use. Others point out that the United States and Europe have different cultures and traditions. European countries tend to put more emphasis on regulation and centralization, especially concerning commerce, whereas U.S. tradition puts more emphasis on contracts, consumer pressure, flexibility and freedom of the market, and penalties for abuses of personal information by enforcement of existing laws (such as those against deceptive and unfair business practices).

2.6 Communications

Law enforcement agencies intercept communications to collect evidence of criminal activities. Intelligence agencies intercept communications to collect information about

the activities and plans of hostile governments and terrorists. The Fourth Amendment to the U.S. Constitution and various laws put restraints on their activities to protect innocent people and reduce the opportunity for abuses. In this section, we consider how changing technologies and government policies affect the ability of law enforcement agencies to intercept the contents of communications and to obtain other information about communications. We begin with background on wiretapping of telephone conversations and laws about the privacy of telephone and email. We consider the Communications Assistance for Law Enforcement Act (CALEA), which requires that the technology used in communications systems be designed or modified to ensure the ability of law enforcement agencies to intercept communications. Then we consider interception of communications for national security.

2.6.1 Wiretapping and Email Protection

Telephone

Within 10 years of the invention of the telephone, people (in and out of government) were wiretapping them.[94] Before that, people intercepted telegraph communications. Throughout the years when human operators made telephone connections and most people had party lines (one telephone line shared by several households), operators and nosy neighbors sometimes listened in on telephone conversations.

Increased wealth and new technology eliminated party lines and human operators, but telephones were still vulnerable to wiretapping. The legal status of wiretapping was debated throughout most of the 20th century. Federal and state law enforcement agencies, businesses, private detectives, political candidates, and others widely used wiretapping. In 1928, the Supreme Court ruled that wiretapping by law enforcement agencies was not unconstitutional, although Congress could ban it. In 1934, Congress passed the Communications Act. This law states that, unless authorized by the sender, no person could intercept and divulge a message; there is no exception for law enforcement agencies. A 1937 Supreme Court decision ruled that wiretapping violated this law.[95] Federal and state law enforcement agencies and local police ignored the ruling and continued to wiretap regularly for decades, sometimes with the approval of the U.S. Attorney General. In one well-publicized case, the FBI monitored the telephone calls between a defendant and her attorneys during her trial. Evidence obtained by illegal wiretapping is inadmissable in court, so the FBI kept a separate, secret file system. The FBI bugged and wiretapped members of Congress and the Supreme Court. Although there was publicity about extensive use of wiretapping by police, no prosecutions resulted. In many cases, of course, law enforcement agencies were wiretapping people suspected of crimes, but in many other cases, they tapped people with unconventional views, members of civil rights groups, and political opponents of powerful government officials.

A fierce debate on the wiretap issue continued. Congress repeatedly rejected proposals to allow wiretapping and electronic surveillance. In 1967 (in *Katz v. United States*,

discussed in Section 2.2.2), the Supreme Court ruled that intercepting telephone conversations without a court order violated the Fourth Amendment to the U.S. Constitution. In 1968, as part of the Omnibus Crime Control and Safe Streets Act, Congress explicitly allowed wiretapping and electronic surveillance by law enforcement agencies, with a court order, for the first time in U.S. history. The main argument given for this change was the necessity to combat organized crime.

The government needs a court order to (legally) intercept or record the content of a telephone call for a criminal investigation.* Law enforcement agents must justify the request, and the wiretap permission is granted for a limited time period. Government agents may determine the telephone numbers called from a particular telephone and the number from which someone made a call with less court scrutiny and justification.

Senator Sam Ervin commented in 1968, "The mere fact of passing a law never resolves a controversy as fierce as this one."[96] He was right. Debate continued about whether the privacy protections in the Omnibus Crime Act were strong enough to be constitutional. Supreme Court justices disagreed. Wiretapping by government and politicians that was illegal or of questionable legality continued, most notably during the Vietnam War. Journalists and government employees were victims of unconstitutional wiretaps during the Nixon administration. In 1998, Los Angeles police officers admitted using wiretaps improperly in a large number of cases.

Most other countries have constitutional and legal protections for communications privacy, but police and intelligence agencies in many countries routinely perform illegal monitoring of political opponents, human rights workers, and journalists.[97]

Email and other new communications

Old laws did not explicitly cover email and cellphone conversations, and interception was common when email and cellphones were new. Driving around Silicon Valley eavesdropping on cellphone conversations was, reportedly, a popular form of industrial spying in the 1980s. Snoops intercepted cellphone conversations of politicians and celebrities. The Electronic Communications Privacy Act of 1986 (ECPA), with amendments in 1994, extended the 1968 wiretapping restrictions to electronic communications, including electronic mail, cordless and cellular telephones, and paging devices. This was a significant step toward protecting privacy in cyberspace from private and governmental snooping. It requires that the government get a court order to legally intercept email.† Controversy continued about the standard law enforcement agencies must meet to obtain copies of stored email. The government argued that people give up their expectation of privacy by allowing ISPs to store their email on the ISP's computers; thus, the strict requirements of the Fourth Amendment would


Expectation of privacy: Section 2.2.2

* The government may intercept the content of communications without a court order in some emergencies.

† The ECPA allows businesses to read the email of employees on the business system. We discuss this issue of employee privacy in Chapter 6.

not apply. A federal appeals court ruled that people *do* have an expectation of privacy for email stored at their ISP and that police need a search warrant to get it.[98]

The USA PATRIOT Act (passed soon after the terrorist attacks in 2001) weakened the ECPA and loosened restrictions on government surveillance and wiretapping activities. It allows law enforcement agents to more easily get header information (such as destination and time) for email. Law enforcement agents use the looser standards to get cellphone location records without a search warrant. This practice is controversial. Some judges and many privacy advocates argue that the loosened provisions of the ECPA violate the Fourth Amendment.

2.6.2 Designing Communications Systems for Interception

New technologies, market competition, and varied customer needs have generated a great diversity of telecommunications services, equipment, protocols, algorithms, and companies. Law enforcement agencies argue that communication technologies developed in the past few decades have made their job of intercepting communications and obtaining communication records (for example, phone numbers called) more difficult. Internet telephone calls and email travel in small pieces (called packets) mingled with packets from other communications. Packets from one message might follow different routes to the destination. Thus, intercepting Internet communications is more difficult than attaching a clip to an old analog telephone wire. When people use call forwarding, the first number called—the number law enforcement agents can legally get fairly easily (without a search warrant)—does not give information about the actual recipient of the call.

The Communications Assistance for Law Enforcement Act (CALEA) requires that the design of telecommunications equipment ensure that the government can intercept telephone calls (with a court order or other authorization). CALEA passed in 1994, so it does not explicitly cover many newer ways of communicating. The law was controversial, and repeated attempts by the government to extend it continue to be controversial. In 2010, the government proposed legislation to require social networking sites and Internet phone services to modify their systems so that law enforcement agents can monitor the communications of users.[99]

In the past, engineers designed communications equipment for its communications purpose. The FBI developed its tools for interception, and communications providers had to assist. The significance of CALEA is that, previously, the government could not require the design and modification of communications equipment to meet the interception needs of law enforcement.

The essential argument in favor of CALEA (and other government programs to intercept communications) is to maintain the ability of law enforcement agencies to protect us from drug dealers, organized crime, other criminals, and terrorists in a changing technological environment. "The prospect of trying to enforce laws without a nationwide standard for surveillance would turn enforcement into a nightmare," according to the

The security of BlackBerrys

Research in Motion (RIM) provides encrypted, highly secure communications on its BlackBerrys. They are popular among business people, government agencies, and ordinary users around the world. Several governments (including China, India, Kuwait, and the United Arab Emeriates) pressured RIM to provide access to BlackBerry communications by government agents. Several threatened to ban BlackBerry (and other) services they could not monitor. RIM made agreements with some governments to allow access to users' communications when government requests comply with the country's laws.

FBI.[100] The problems with CALEA and other programs to intercept communications, according to critics, include threats to privacy and civil liberties, the potential for abuse by government, and the side effects of "backdoor" access that threaten the security of communications systems. Critics of CALEA also argue that requirements for determining the physical location of cellphone users and ensuring that the government could intercept Internet communications go beyond the scope of the law and extend the government's surveillance power beyond what Congress intended when it passed the law. The idea of designing communications technology for a "nationwide standard for surveillance" is a nightmare to those who place high value on privacy and civil liberties. Should the designers of communications systems be free to use the best technology available for achieving speed, convenience, low cost, and privacy?

More than 80% of the wiretaps courts authorize for criminal investigations are for drug cases.[101] Critics claim that wiretaps are a less useful law enforcement tool than informants, detective work, witnesses, and so on. Supporters of CALEA argue that wiretaps are essential for catching and/or convicting dangerous criminals. The focus of criminal wiretaps on drug crimes raises the question of whether the government really needs such extreme, system-wide controls on the communication systems used by 300 million Americans. If drug prohibition were to end, as alcohol prohibition did in the 1930s, would we find ourselves with a costly and privacy-threatening infrastructure of intrusion and relatively little legitimate need for it?

2.6.3 The NSA and Secret Intelligence Gathering[102]

The purpose of the National Security Agency (NSA) is to collect and analyze foreign intelligence information related to national security and to protect U.S. government communications and sensitive information related to national security. Because governments encrypt their sensitive material, the NSA has long devoted a huge amount of resources to cryptology and has the most advanced code-breaking capabilities. A secret presidential order formed the NSA

The NSA and encryption policy: Section 2.5.1

in 1952. Its budget is still secret, although its website says the NSA/CSS (NSA and Central Security Service) is about the size of one of the larger Fortune 500 companies.[103] The NSA builds and uses enormously powerful supercomputers.* It collects and stores huge masses of information.

Some of the NSA's activities go well beyond its purpose—or interpret the purpose to an extraordinarily broad degree—and potentially threaten the privacy and freedoms of U.S. citizens. Because the NSA uses methods that do not satisfy the Fourth Amendment, it was legally restricted to intercepting communications outside the United States (with some exceptions). Through its history, the agency generated much controversy by secretly violating restrictions on surveillance of people within the United States. In the 1960s and 1970s, the NSA monitored communications of specific American citizens (including civil rights leader Martin Luther King Jr. and entertainers who opposed the Vietnam War). A Congressional committee chaired by Senator Church found that the NSA had been secretly and illegally collecting international telegrams, including telegrams sent by American citizens, since the 1950s and searching them for foreign intelligence information. As a result, Congress passed the Foreign Intelligence Surveillance Act of 1978 (FISA) establishing oversight rules for the NSA. The law prohibited the agency from collecting masses of telegrams without a warrant and from compiling lists of Americans to watch without a court order. The law set up a secret federal court, the Foreign Intelligence Surveillance Court, to issue warrants to the NSA to intercept communications of people it could show were agents of foreign powers or involved in terrorism or espionage.

Secret access to communications and communications records

The NSA collects information by intercepting communications. While some newer technologies (such as fiber optic cable) made wiretapping more difficult, other technological changes in the past few decades make communications of ordinary people more vulnerable. Satellite communications were a boon to the NSA; it could pick messages out of the air. Increased wealth, travel, and trade generated more international communication—cluttering communications channels and potentially making it harder for the NSA to detect messages of interest. Then, vastly increased processing power of computer systems enabled the NSA to filter and analyze huge quantities of communications of innocent people instead of targeting only specific suspects. In cyberspace, our email, cellphone conversations, tweets, searches, purchases, financial information, legal documents, and so on, mix with military, diplomatic, and terrorist communications. The NSA sifts through it all. It does "deep packet inspection," analyzing the packets of information traveling through the Internet, and collects whatever is of interest. It collects all communications to and from approximately a million people on its watch lists. Its interception activity is

* As of 2012, the government had an IBM supercomputer that operated at 16.32 petaflops (16.32 million billion operations per second), officially the fastest in the world.[104]

extremely controversial because the NSA processes and collects data on Americans with no court order and no approval from the FISA court.

In 2006, an AT&T employee described (under oath) a secret, secure room the NSA set up at an AT&T switching facility. From this room, the NSA had access to email, telephone, and Web communications of AT&T users.[105] The NSA built a database of telephone and email records of millions of Americans. The government argued that the NSA was not intercepting or listening to telephone calls and was not collecting personal identifying information. It used sophisticated data mining technology to analyze calling patterns to learn how to detect communications of terrorist cells. The agency analyzes calling patterns because the sources of terrorism are diffuse and require broader means of detection and surveillance than old-time spy work. The NSA can no longer rely on monitoring only the telephone traffic of a few hostile governments and a small number of known suspects. It can no longer monitor just those specific physical telephone lines or communications links that connect specific military facilities or other sites of interest. Analysis of communications traffic helps the NSA determine what is suspicious. Opponents of the monitoring program argued that it was a huge intrusion on privacy. Even if the NSA did not collect customer names, it is quite easy to re-identify people from their phone records. Opponents said the warrantless collection of the records by the NSA was illegal, and it was illegal for a telephone company to provide them. Several groups filed suits against AT&T for violating its stated privacy policies and communications privacy law by assisting the NSA.

Congress passed the FISA Amendments Act in 2008. This law retroactively protects AT&T (and other entities that assist the NSA) from lawsuits. Although it includes provisions to restrict domestic surveillance, overall it reduces previous protections. The FISA Amendments Act is controversial, and a lawsuit challenging its constitutionality is ongoing.[106] In the meantime, it became clear that the NSA installed and continues to operate secret monitoring rooms at other major U.S. telecommunications company facilities, where it can filter and collect whatever domestic communications it chooses. The NSA built an enormous new data center to store, decrypt, and analyze billions of gigabytes of communications and files.[107] What it cannot decrypt now, it stores to decrypt later when it develops faster computers or better algorithms. Civil libertarians are concerned that the NSA is collecting huge quantities of ordinary business and personal encrypted data that have nothing to do with terrorism or foreign intelligence.

Before the USA PATRIOT Act, there was a sharp boundary between legal rules for terrorism investigations (involving foreigners) and criminal investigations (involving people within the United States). The PATRIOT Act allows information obtained in terrorism investigations under FISA warrants to be used in criminal cases. Government officials do not follow the normal protections and rules for search warrants in terrorism cases. In addition, prosecutors normally provide defense attorneys in a criminal case with recordings of intercepted messages. When obtained as part of a terrorism case, the

government does not have to provide transcripts. Thus, the broader powers to violate privacy of communications in terrorism investigations can have serious impacts on people accused of ordinary crimes.

How can we evaluate the NSA's programs of massive collection of communications? How should we react when powerful government agencies break laws that protect privacy of communications? Accessing data on specific suspects is a reasonable, essential, and responsible part of criminal and terrorist investigations. Broad access and data mining are more questionable because they threaten the safety and freedom of innocent people if investigators mistakenly decide someone's transactions look suspicious. They provide the mechanisms for totalitarian control. Is the secrecy justifiable? Is the secrecy essential? Exposure of monitoring programs leads terrorists to take new measures to hide their activities, communications, and transactions. Temporary secrecy is essential for many criminal and terrorist investigations, but secret programs to monitor and collect communications present a huge potential for abuse, as we have seen often in the past. We have also seen the hideous effects of terrorism. Is the fact that there were no successful terrorist attacks in the United States after 2001 (up to the time I write this) due in part to the secret communications monitoring and analysis programs of the NSA and other intelligence agencies?

EXERCISES

Review Exercises

2.1 What does the term *personal information* mean?

2.2 What does the term *secondary use* mean? Give an example.

2.3 What does the term *re-identification* mean? Give an example.

2.4 Explain the difference between *opt-in* and *opt-out* policies for secondary uses of personal information.

2.5 Describe one application of face recognition that infringes privacy.

2.6 Describe two tools people can use to protect their privacy on the Web.

2.7 Describe two methods a business or agency can use to reduce the risk of unauthorized release of personal information by employees.

General Exercises

2.8 A company in the Netherlands that makes navigation devices collects location data from the devices to provide real-time services to its customers. It also provides anonymous statistical data to government agencies to improve roads and traffic flow. Unknown to the company and its customers, the police used the data to choose sites for traffic cameras to catch speeders. Was this a privacy violation? Why or why not?[108]