CHAPTER

5

Privacy and Cyberspace

LEARNING OBJECTIVES

Upon completing this chapter, you will successfully be able to:

- Describe the ways in which personal privacy is threatened in the digital era and determine whether cybertechnology has introduced any new or unique privacy issues,
- Explain why *privacy* is a difficult concept to define and describe the key elements of a comprehensive privacy theory that helps us to distinguish between a *loss* of privacy and a *violation* of privacy,
- Explain why privacy is valued and why it is an important social value as well as an individual good,
- Describe how one's privacy is impacted by *data-gathering* techniques such as RFID technologies and (Internet) cookies,
- Assess privacy-related concerns generated by *data-analysis* techniques involving Big Data, such as data mining and Web mining,
- Explain what is meant by the problem of protecting "privacy in public," especially in light of challenges posed by routine uses of online search facilities such as Google and social networking sites such as Facebook,
- Evaluate the debate between proponents of stronger privacy legislation and those who advocate for industry self-regulation practices as an alternative,
- Describe and assess the arguments for whether people should, in certain cases, have a right to have online personal information about them "erased," or at least "delinked" from search engine indexes.

In this chapter, we examine a wide range of privacy issues, including concerns that affect the day-to-day activities of ordinary individuals carried out in both online and offline contexts. We begin by reflecting on a scenario that illustrates some privacy concerns surrounding the opening of a controversial data center by the National Security Agency (NSA) in 2014.

▶ **SCENARIO 5–1:** A New NSA Data Center

The NSA, officially formed in 1952 and headquartered at Ft. Meade, Maryland, is one of the largest intelligence organizations in the United States (and in the world). Although the NSA's original charter was to conduct surveillance on "foreign" sources—that is, non-U.S. governments, organizations, and individuals—many critics, including reputable journalists, claim that the agency's mission has since been significantly expanded to include surveillance on American citizens as well. Sensitive documents that Edward

Snowden leaked to the press in May 2013 allegedly revealed some of the controversial surveillance techniques the NSA uses to collect data on U.S. citizens. These revelations have resulted in increased public awareness of the NSA, as well as closer media scrutiny of that organization's activities, especially in light of some embarrassing details about foreign government officials that were included in the leaked documents. (See Scenario 4–2 in Chapter 4 for more detail about controversies resulting from the documents leaked by Snowden.)

In the aftermath of the Snowden leaks, the NSA not only continued to defend its activities but has also recently opened a controversial data center, or "data farm" as some refer to it, in Utah. The new data center, which is able to gather significantly more information than was previously possible, reportedly collects phone records, text messages, e-mails, and other forms of electronic data. Privacy advocates in the United States are concerned that the new center also provides NSA with enhanced tools to analyze much of the electronic data generated by ordinary American citizens. Many NSA critics believe that the organization would have collected more data in the past but that it was hampered from doing so by limitations affecting both (i) storage capacity and (ii) the technology-related resources needed to conduct more extensive searches. But the new NSA data center (whose contents are officially classified) is reported to have 100,000 square feet of computers (to harvest data) and five zettabytes of storage capacity to house it.[1] So, critics worry that the NSA's activities now pose an even greater threat than before to the privacy of ordinary American citizens. ■

Are privacy advocates justified in their concerns about the NSA's increased ability to "spy" on American citizens and to collect vast amounts of data about them? The NSA has argued that its surveillance and data-collection techniques simply follow the organization's charter, which is to keep Americans safe. That organization's defenders suggest that in an era of global terrorism, U.S. citizens should be less concerned about their privacy and more worried about their safety. Arguably, this tension underlies one of the crucial dilemmas facing American citizens, as well as citizens in many countries around the world—namely, how much individual privacy should be we willing to trade off for increased security? But is the dichotomy between privacy and security, as articulated here, a "genuine dilemma"?[2] (Recall our analysis of the False Dichotomy/Either–Or fallacy in Chapter 3.) We examine key issues at the heart of the privacy-versus-security debate in detail in Chapter 6, and we briefly describe some concerns affecting governmental surveillance and data-collection techniques in Section 5.4.4.

The main purpose of Scenario 5–1 has been to get us to begin thinking about the value of privacy in general, especially in light of the serious challenges it faces in the second decade of the twenty-first century. For example, we will see that our privacy is threatened not only by governmental surveillance but by many of the data-collection and data-analysis techniques currently used in the commercial sector as well. This is especially apparent in the case of major search engine companies like Google, which collect vast amounts of personal information on ordinary users. So, our privacy is currently threatened on many different levels and in many different sectors.

▶ 5.1 PRIVACY IN THE DIGITAL AGE: *WHO* IS AFFECTED AND WHY SHOULD WE WORRY?

Of all the ethical and social issues associated with the digital era, perhaps none has received more media attention than concern about the loss of personal privacy. As we shall see, however, cybertechnology is not the first technology to threaten personal privacy. Nevertheless, that technology now threatens privacy in ways that were not previously possible.

5.1.1 Whose Privacy Is Threatened by Cybertechnology?

Virtually no one living today in a developed nation is exempt or immune from some kind of cybertechnology-related privacy threat. In fact, people who have never owned or never even used a networked computer are still affected. Consider that in carrying out many of our day-to-day activities, we supply information to organizations that use computers and electronic devices to record, store, and exchange those data. These activities can include information we provide in filling out various forms, or they can include information acquired from our commercial transactions in a bank or a store.

Some might assume that using the Internet only for noncommercial activities will help them to avoid many of the typical privacy threats associated with the online world, such as privacy risks associated with shopping at e-commerce sites. However, even users who navigate the Web solely for recreational purposes are at risk with respect to their privacy. For example, if you use a major search engine (such as Google) or interact in a social media site (such as Facebook), personal data about your interests and preferences can be acquired by those organizations, whose methods for collecting this information are not always obvious to ordinary users. Furthermore, personal data about us collected via our online activities can be sold to third parties.

Also consider that applications such as Google Street View (a feature of Google Earth and Google Maps) make use of satellite cameras and GPS software that enable Internet users to zoom in on your house or place of employment and potentially record information about you. Additionally, closed-circuit televisions (CCTVs) located in public places and in shopping malls record many of your daily movements as you casually stroll through those environments. So even if you have never owned or used a computer, cell phone, (Internet-enabled) electronic device, and so forth, your privacy is threatened in ways that were not possible in the past.

Concerns about privacy now affect many aspects of one's life—from commerce to healthcare to work to recreation. So, some analysts organize these concerns into categories such as consumer privacy, medical/healthcare privacy, employee/workplace privacy, etc. Although some cyber-related privacy concerns are specific to one or more spheres or sectors—that is, employment, healthcare, and so forth—others cut across multiple dimensions of our lives and thus affect each of us regardless of our employment or health status.

Unfortunately, we cannot examine all of the current privacy concerns in a single chapter. So, we will have to postpone our analysis of certain kinds of privacy issues until the later chapters in the book. For example, we will examine some cyber-related privacy concerns that conflict with cybersecurity issues and national security interests in Chapter 6, where privacy-related concerns affecting "cloud computing" are also considered. Some specific employee/workplace privacy-related issues are examined in our discussion of workplace surveillance and employee monitoring in Chapter 10. And in our analysis of emerging and converging technologies (such as nanotechnology and ambient intelligence) in Chapter 12, we examine some controversies surrounding a relatively new category of privacy called "location privacy."

Despite the significant impact that cybertechnology has had and continues to have for our privacy, one still might ask: Do any of the current privacy concerns differ in kind from privacy issues that arose in the predigital era? In other words, is there anything new, or even special, about cyber-related privacy issues? We next propose a strategy for addressing that question.

5.1.2 Are Any Privacy Concerns Generated by Cybertechnology Unique or Special?

We begin by noting that concerns about personal privacy existed long before the advent of computers and cybertechnology. Prior to the digital era, for example, technologies such as the camera and the telephone presented challenges for privacy. So we can ask: What, if anything,

is special about the privacy concerns that are associated with cybertechnology? Consider the impact that changes involving this technology have had on privacy with respect to the:

- Amount of personal information that can be collected
- Speed at which personal information can be transmitted
- Duration of time that the information can be retained
- Kind of information that can be acquired and exchanged

Cybertechnology makes it possible to collect and store much more information about individuals than was possible in the predigital era. The *amount* of personal information that could be collected in that era was determined by practical considerations, such as the physical space required to store the data and the time and difficulty involved in collecting the data. Today, of course, digitized information that can be stored electronically in computer databases takes up very little storage space and can be collected with relative ease. As we will see in Section 5.5, many people are now worried about the sheer volume of personal data that can be collected and analyzed by various techniques associated with "big data" and data mining.

Next, consider the *speed* at which information is exchanged and transferred between databases. At one time, records had to be physically transported between filing destinations; the time it took to move them depended upon the transportation systems—for example, motor vehicles, trains, airplanes, and so forth—that carried the records. Now, of course, records can be transferred between electronic databases in milliseconds through wireless technologies, high-speed cable lines, or even ordinary telephone lines.

With so much information being collected and transferred so rapidly, many have expressed concerns about its accuracy as well as the difficulties in tracking down and correcting any inaccuracies that might have been transferred. In an interview conducted for the BBC TV series *The Machine that Changed the World*, Harvard law professor Arthur Miller points out that trying to correct such information is like "chasing a greased pig"—you may get your hands on the pig, but it is very difficult to keep the pig firmly in your grip.[3] Although issues concerning the accuracy of personal information are clearly distinguishable from those concerning privacy per se, accuracy issues are frequently associated with privacy issues, and both are impacted by cybertechnology.

Also, consider the *duration* of information—that is, how long information can be kept. Before the digital era, information was manually recorded and stored in file cabinets and then in large physical repositories; it is unlikely that report cards my parents received as high school students still exist somewhere as physical records in file cabinets, for at that time report cards were not computerized but instead existed, literally, as ink marks on paper. But the report cards my daughter received when she was a high school student were both generated and stored using computer technology. As an electronic record, her report card can be kept indefinitely, and the grades she received as a high school student (as well as the grades she received in elementary school and in college) can follow her throughout her life.

In the past, practices involving the retention of personal data were perhaps more "forgiving." Because of practical limitations, such as physical storage space, that affected how long personal data could be kept on file, much of the personal information collected and stored had to be destroyed after a certain number of years. Since information could not be archived indefinitely, people with blemished records sometimes had the opportunity to start over again by physically relocating. Today, however, one's electronic dossier would likely follow, making it very difficult, if not impossible, for that person to start over with a clean slate. We can argue whether the current means of data retention is a good thing, but it is difficult to dispute the claim that now, because of cybertechnology, most of us have what Arthur Miller calls a "womb-to-tomb dossier." It is also worth noting, however, that a 2014 court ruling by the European Court of Justice (ECJ) gave citizens in European nations the right, in certain cases, to have

some kinds of online personal information about them deleted or "erased." We examine this principle, commonly referred to as the right to "be forgotten" or "to erasure," in detail in Section 5.8.

Cybertechnology has also generated privacy concerns because of the *kind* of personal information that can now be collected. For example, every time you engage in an electronic transaction, such as making a purchase with a credit card or withdrawing money from an ATM, transactional information is collected and stored in several computer databases; this information can then be transferred electronically across commercial networks to agencies that request it. Personal information, retrieved from transactional information that is stored in computer databases, has been used to construct electronic dossiers containing detailed information about an individual's commercial transactions, including purchases made and places traveled—information that can reveal patterns in a person's preferences and habits.

Additionally, we should note that cybertechnology raises privacy concerns because of the myriad ways in which it enables our personal information to be manipulated once it has been collected. For example, unrelated pieces of information about us that reside in separate databases can be combined to construct electronic personal dossiers or profiles. Also, information about us included in one database can be matched against records in other databases that contain information about us. Furthermore, our personal information can be mined (from databases, as well as from our activities on the Web) to reveal patterns in our behavior that would have been very difficult to discern in the predigital era. Of course, our personal data could have been, and in some instances was, manipulated in the predigital era as well. But there were practical limitations to the amount of data collection and analysis that could be done manually by humans.

Although the privacy concerns that we now associate with cybertechnology may not be totally new, or even altogether different in kind, from those we associate with earlier technologies, few would dispute the claim that cybertechnology has exacerbated them. In Sections 5.4–5.5, we examine some controversial uses of cybertechnology that raise specific concerns for personal privacy. First, however, we examine the concept of personal privacy to better understand what privacy is and why we value it.

► 5.2 WHAT IS PERSONAL PRIVACY?

Although many definitions have been put forth, there is no universally agreed upon definition. To understand why this concept has been so difficult to define, consider the diverse range of metaphors typically associated with privacy. Sometimes, we speak of privacy as something that can be lost or diminished, suggesting that privacy can be understood in terms of a repository of personal information that can be either diminished altogether or gradually eroded. Contrast this view with descriptions of privacy as something that can be intruded upon or invaded, where privacy can be understood in terms of a spatial metaphor, such as a zone, that deserves protection. Alternatively, privacy is sometimes described as something that can be violated or breached, when we think of it in terms of either a right or an interest that deserves legal protection. Because of these different conceptions of privacy, we will see that it is useful to distinguish between the notions of one's having privacy (in a descriptive sense) and one's having a (normative) right to privacy. We will say more about this distinction in Section 5.2.4.

Privacy analysts have pointed out that in the United States, the meaning of privacy has evolved since the eighteenth century. Initially, privacy was understood in terms of freedom from (physical) intrusion. Later, it became associated with freedom from interference into one's personal affairs, including one's ability to make decisions freely. Most recently,

privacy has come to be closely identified with concerns affecting access to and control of personal information—a view that is also referred to as "informational privacy." Although the main emphasis in this chapter is on informational privacy, we also briefly describe the other two views.

5.2.1 Accessibility Privacy: Freedom from Unwarranted Intrusion

In a seminal paper on privacy, Samuel Warren and Louis Brandeis suggested that privacy could be understood as "being let alone" or "being free from intrusion." Appearing in the *Harvard Law Review* in 1890, the Warren and Brandeis article made the first explicit reference to privacy as a legal right in the United States. Many Americans are astonished to find out that there is no explicit mention of privacy in either the Constitution or its first ten amendments, the Bill of Rights. However, some legal scholars believe that a right to privacy can be inferred from the Fourth Amendment, which protects citizens against unreasonable searches and seizures of personal affects (i.e., papers, artifacts, etc.) by the government. Some legal scholars suggest that the Fourth Amendment may also provide legal grounds for a right to privacy protection from nongovernmental intrusion as well.

Warren and Brandeis also suggested that our legal right to privacy is grounded in our "right to inviolate personality." In part, they were responding to a certain use of a new technology—not the computer, of course, but rather the camera—which had begun to threaten individual privacy in new ways.[4] Photographs of people began to appear in newspapers, for example, in gossip columns, along with stories that were defamatory and sometimes even false. Warren and Brandeis believed that individuals have a (legal) right not be intruded upon in this manner. Because this definition of privacy as freedom from unwarranted intrusion focuses on the harm that can be caused through physical access to a person or to a person's possessions, DeCew (1997, 2006) and others have described this view as *accessibility privacy*.

5.2.2 Decisional Privacy: Freedom from Interference in One's Personal Affairs

Privacy is also sometimes conceived of as freedom from interference in one's personal choices, plans, and decisions; some refer to this view as decisional privacy. This kind of privacy has also been associated with reproductive technologies having to do with contraception. In *Griswold v. Connecticut* (1965), the court ruled that a person's right to get counseling about contraceptive techniques could not be denied by state laws. The view of privacy as freedom from external interference into one's personal affairs has since been appealed to in legal arguments in a series of controversial court cases, such as those involving abortion and euthanasia. For example, this view of privacy was appealed to in the landmark Supreme Court decision on abortion (*Roe v. Wade* 1973), as well as in a state court's decision involving Karen Ann Quinlan's right to be removed from life-support systems and thus her "right to die."[5] Because it focuses on one's right not to be interfered with, decisional privacy can be distinguished from both accessibility privacy and informational privacy.

5.2.3 Informational Privacy: Control over the Flow of Personal Information

Because of the increasing use of technology to gather and exchange personal information, many contemporary analysts view privacy in connection with one's ability to restrict access to and control the flow of one's personal information. Privacy concerns are now often framed in terms of questions such as: Who should have access to one's personal information? To what extent can individuals control the ways in which information about them can be gathered,

TABLE 5-1 Three Views of Privacy

Accessibility privacy	Privacy is defined as one's (physically) being let alone, or being free from intrusion into one's physical space
Decisional privacy	Privacy is defined as freedom from interference in one's choices and decisions
Informational privacy	Privacy is defined as control over the flow of one's personal information, including the ways in which that information is collected and exchanged

stored, mined, combined, recombined, exchanged, and sold? These are our primary concerns in this chapter, where we focus on informational privacy.

Table 5-1 summarizes the three views of privacy.

5.2.4 A Comprehensive Account of Privacy

James Moor has put forth a privacy framework that incorporates important elements of the nonintrusion, noninterference, and informational views of privacy. According to Moor,

> An individual [has] privacy in a *situation* with regard to others if and only if in that situation the individual [is] protected from intrusion, interference, and information access by others.[6]

An important element in this definition is Moor's notion of "situation," which he deliberately leaves broad so that it can apply to a range of contexts, or zones, that can be "declared private." For example, a situation can be an "activity" or a "relationship," or it can be the "storage and access of information" in a computer (Moor 2000).

Central to Moor's theory is a distinction between naturally private and normatively private situations, enabling us to differentiate between the conditions required for (i) having privacy and (ii) having a right to privacy. This distinction, in turn, enables us to differentiate between a loss of privacy and a violation of privacy. In a naturally private situation, individuals are protected from access and interference from others by natural means, for example, physical boundaries such as those one enjoys while hiking alone in the woods. In this case, privacy can be lost but not violated, because there are no norms—conventional, legal, or ethical—according to which one has a right, or even an expectation, to be protected. In a normatively private situation, on the other hand, individuals are protected by conventional norms (e.g., formal laws and informal policies) because they involve certain kinds of zones or contexts that we have determined to need normative protection. The following two scenarios will help us to differentiate between normative and natural (or descriptive) privacy.

▶ **SCENARIO 5–2:** Descriptive Privacy

Mary enters her university's computer lab at 11:00 P.M. to work on a research paper that is due the next day. No one else is in the lab at the time that Mary arrives there, and no one enters the lab until 11:45 P.M., when Tom—the computer lab coordinator—returns to close the lab for the evening. As Tom enters, he sees Mary typing on one of the desktop computers in the lab. Mary seems startled as she looks up from her computer and discovers that Tom is gazing at her. ■

Did Mary lose her privacy when Tom entered the lab and saw her? Was her privacy violated? Before Tom noticed her in the lab, we could say that Mary had privacy in the descriptive, or natural, sense of the term because no one was physically observing her while she was in the lab. When Tom entered and noticed that Mary was typing on a computer, Mary lost her natural (or descriptive) privacy in that situation. However, we should not infer that her privacy was violated in this incident, because a university's computer lab is not the kind of situation or zone that is declared normatively private and thus protected.

▶ **SCENARIO 5–3:** Normative Privacy

Tom decides to follow Mary, from a distance, as she leaves the computer lab to return to her (off-campus) apartment. He carefully follows her to the apartment building, and then stealthily follows Mary up the stairway to the corridor leading to her apartment. Once Mary is safely inside her apartment, Tom peeps through a keyhole in the door. He observes Mary as she interacts with her laptop computer in her apartment. ■

Has Mary's privacy been violated in this scenario? In both scenarios, Tom observes Mary interacting with a computer. In the first scenario, the observation occurred in a public place. There, Mary may have lost some privacy in a descriptive or natural sense, but she had no expectation of preserving her privacy in that particular situation. In the second scenario, Mary not only lost her privacy but her privacy was violated as well, because apartments are examples of zones or "situations" that we, as a society, have declared normatively private.

We have explicit rules governing these situations with respect to privacy protection. Note that it was not merely the fact that Tom had observed Mary's interactions with a computer that resulted in her privacy being violated in the second scenario. Rather, it was because Tom had observed her doing this in a normatively protected situation. So, there was nothing in the information per se that Tom acquired about Mary that threatened her privacy; it was the situation or context in which information about Mary was acquired that caused her privacy to be violated in the second scenario.

5.2.5 Privacy as "Contextual Integrity"

We have seen the important role that a situation, or context, plays in Moor's privacy theory.[7] But some critics argue that the meaning of a situation or context is either too broad or too vague. Helen Nissenbaum elaborates on the notion of a context in her model of privacy as "contextual integrity," where she links adequate privacy protection to "norms of specific contexts." She notes that the things we do, including the transactions and events that occur in our daily lives, all take place in some context or other. In her scheme, contexts include "spheres of life" such as education, politics, the marketplace, and so forth (Nissenbaum 2004a, 2010).

Nissenbaum's privacy framework requires that the processes used in gathering and disseminating information (i) are "appropriate to a particular context" and (ii) comply with norms that govern the flow of personal information in a given context.[8] She refers to these two types of informational norms as follows:

1. Norms of appropriateness
2. Norms of distribution

Whereas norms of appropriateness determine whether a given type of personal information is either appropriate or inappropriate to divulge within a particular context, norms of distribution restrict or limit the flow of information within and across contexts. When either norm has been "breached," a violation of privacy occurs; conversely, the contextual integrity of the flow of personal information is maintained when both kinds of norms are "respected."[9]

As in the case of Moor's privacy model, Nissenbaum's theory demonstrates why we must always attend to the context in which information flows, and not to the nature of the information itself, in determining whether normative protection is needed. To illustrate some of the nuances in her framework of privacy as contextual integrity, consider the following scenario in which a professor collects information about students in his seminar.

▶ SCENARIO 5–4: Preserving Contextual Integrity in a University Seminar

Professor Roberts teaches a seminar on social issues in computing to upper-division, undergraduate students at his university. Approximately half of the students who enroll in his seminar each semester are computer science (CS) students, whereas the other half are students majoring in humanities, education, business, etc. At the first class meeting for each seminar, Professor Roberts asks students to fill out an index card on which they include information about their major, their year of study (junior, senior, etc.), the names of any previous CS courses they may have taken (if they are non-CS majors), their preferred e-mail address, and what they hope to acquire from the seminar. Professor Roberts then records this information in his electronic grade book. ■

Has Professor Roberts done anything wrong in requesting and collecting this information? For the most part, it is information that he could have gathered from the registrar's office at his university—for example, information about which CS courses and which general education courses the students have previously taken, and so forth. But Roberts finds it much more convenient to collect information in the classroom, and he informs the students that he uses that information in determining which kinds of assignments he will decide to give to the class in general and which kinds of criteria he will use to assign students to various group projects.

Because Professor Roberts has informed the students about how the information they provided to him will be used in the context of the classroom, and because the students have consented to give him the information, no privacy violation seems to have occurred. In fact, the process used by Professor Roberts satisfies the conditions for Nissenbaum's norm of appropriateness with respect to contextual integrity.

Next, suppose that Professor Roberts has lunch a few weeks later with a former student of his, Phil, who recently graduated and now has a job as a software engineer for a publishing company. Phil's company plans to release its first issue of a new magazine aimed at recent CS graduates, and it has launched an advertising campaign designed to attract undergraduate CS majors who will soon graduate. Phil asks Professor Roberts for the names of the CS majors in the seminar he is teaching. Professor Roberts is initially inclined to identify some students that Phil would likely know from classes that he had taken the previous year at the university. But should Professor Roberts reveal those names to Phil?

If he did, Professor Roberts would violate the privacy norm of distribution within the context of the seminar he is teaching. Consider that the students gave information about themselves to Professor Roberts for use in the context of that seminar. While his use of that information for purposes of the seminar is context appropriate, passing on (i.e., distributing) any of that information to Phil is not, because it would violate the integrity of that context. Even though the information about the students that Professor Roberts has collected is neither sensitive nor confidential in nature, it was given to him for use only in the context of the seminar he is teaching. Insofar as Professor Roberts uses the information in that context, he preserves its integrity. But if he elects to distribute the information outside that context, he violates its integrity and breaches the privacy of his students.

▶ 5.3 WHY IS PRIVACY IMPORTANT?

Of what value is privacy? Why does privacy matter and why should we care about it? In 1999, Scott McNealy, then CEO of Sun Microsystems, uttered his now famous remark to a group of reporters: "You have zero privacy anyway. Get over it." And in 2013, Facebook CEO Mark Zuckerberg proclaimed that privacy "is no longer a social norm." So, should we infer from these remarks that the idea of personal privacy is merely a relic of the past? Although Froomkin (2000), Garfinkel (2000), and others speak of the "death of privacy," not everyone has been

willing to concede defeat in the battle over privacy. Some privacy advocates staunchly believe that we should be vigilant about retaining what little privacy we may still have. Others note that we do not appreciate the value of privacy until we lose it, and by then, it is usually too late. They point out that once privacy has been lost, it is difficult, if not impossible, to get back. So perhaps, we should heed their warnings and try to protect privacy to the degree that we can.

We might also question whether the current privacy debate needs to be better understood in terms of differences that reflect generational attitudes. For many so-called Millennials, who are now college aged, privacy does not always seem to be of paramount importance. Most Millennials, as well as many members of Generations X and Y, seem all too eager to share their personal information widely on social networking services such as Facebook, and many also seem willing to post "away messages" on AIM or Skype that disclose their whereabouts at a given moment to a wide range of people. But for many older Americans, including Baby Boomers, privacy is something that is generally still valued. So the relative importance of privacy may vary considerably among the generations; however, we will proceed on the assumption that privacy has value and thus is important.

Is privacy universally valued? Or is it valued mainly in Western, industrialized societies where greater importance is placed on the individual? Solove (2008) notes that privacy is a "global concern," which suggests that it is valued universally. However, it has also been argued that some non-Western nations and cultures do not value individual privacy as much as we do in the West. Alan Westin believes that countries with strong democratic political institutions consider privacy more important than do less democratic ones.[10] Nations such as Singapore and the People's Republic of China seem to place less importance on individual privacy and greater significance on broader social values, which are perceived to benefit the state's community objectives. Even in countries such as Israel, with strong democratic systems but an even stronger priority for national security, individual privacy may not be as important a value as it is in most democratic nations. So, even though privacy has at least some universal appeal, it is not valued to the same degree in all nations and cultures. As a result, it may be difficult to get universal agreement on privacy laws and policies in cyberspace.

5.3.1 Is Privacy an Intrinsic Value?

Is privacy something that is valued for its own sake—that is, does it have intrinsic value? Or is it valued as a means to an end, in which case it has only instrumental worth? Recall our discussion of intrinsic and instrumental values in Chapter 2. There, we saw that happiness has intrinsic value because it is desired for its own sake. Money, on the other hand, has instrumental value since it is desired as a means to some further end or ends.

While few would argue that privacy is an intrinsic value, desired for its own sake, others, including Fried (1990), argue that privacy is not merely an instrumental value or instrumental good. Fried suggests that unlike most instrumental values that are simply one means among others for achieving a desired end, privacy is also essential, that is, necessary to achieve some important human ends, such as trust and friendship. We tend to associate intrinsic values with necessary conditions and instrumental values with contingent, or non-necessary conditions; so while privacy is instrumental in that it is a means to certain human ends, Fried argues that it is also a necessary condition for achieving those ends. Solove also believes that privacy has aspects that cut across the intrinsic–instrumental divide, and he argues "intrinsic and instrumental value need not be mutually exclusive."[11]

Although agreeing with Fried's claim that privacy is more than merely an instrumental value, and with Solove's insight that privacy is a value that spans the intrinsic–instrumental divide, Moor (2004) takes a different approach to illustrate this point. Like Fried, Moor argues that privacy itself is not an intrinsic value. But Moor also believes that privacy is an articulation, or "expression" of the "core value" security, which in turn is essential across cultures, for

human flourishing. (We examine the concept of security as it relates to privacy in Chapter 6.) And like Fried, Moor shows why privacy is necessary to achieve certain ends. Moor further suggests that as information technology insinuates itself more and more into our everyday lives, privacy becomes increasingly important for expressing (the core value) security.

Does privacy play a key role in "promoting human well-being," as Spinello (2010) claims? Perhaps, one way it does is by serving as a "shield" that protects us from interference. DeCew (2006), who believes that the value of privacy lies in the "freedom and independence" it provides for us, argues that privacy shields us from "pressures that preclude self-expression and the development of relationships."[12] She claims that privacy also acts as a shield by protecting us from coercion and the "pressure to conform." In her view, the loss of privacy leaves us vulnerable and threatened because we are likely to become more conformist and less individualistic.

5.3.2 Privacy as a Social Value

Based on the insights of DeCew and others, one might infer that privacy is a value that simply benefits individuals. However, some authors have pointed out the social value that privacy also provides, noting that privacy is essential for democracy. Regan (1995) points out that we often frame the privacy debate simply in terms of how to balance privacy interests as individual goods against interests involving the larger social good; in such debates, Regan notes that interests benefiting the social good will generally override concerns regarding individual privacy. If, however, privacy is understood as not solely concerned with individual good but as contributing to the broader social good, then in debates involving the balancing of competing values, individual privacy might have a greater chance of receiving equal consideration.

Solove (2008) also believes that privacy has an important social dimension, when he notes that the value of privacy is both communal *and* individual. Employing an argument similar to Regan's, Solove points out that privacy becomes "undervalued," when it is viewed as an overly individualistic concept. Arguing instead for what he calls a "pragmatic approach," Solove believes that it is important to assess the value of privacy in terms of its "contribution to society."[13]

Since privacy can be of value for greater social goods, such as democracy, as well as for individual autonomy and choice, it would seem that it is important and worth protecting. But privacy is increasingly threatened by new cyber- and cyber-related technologies. In Sections 5.4 and 5.5, we examine how privacy is threatened by two different kinds of practices and techniques that use cybertechnology:

a. *Data-gathering* techniques used to collect and record personal information, often without the knowledge and consent of users

b. *Data-analysis* techniques, including data mining, used to manipulate large data sets of personal information to discover patterns and generate consumer profiles (also typically without the knowledge and consent of users)

▶ 5.4 GATHERING PERSONAL DATA: SURVEILLANCE, RECORDING, AND TRACKING TECHNIQUES

Collecting and recording data about people is hardly new. Since the Roman era, and possibly before then, governments have collected and recorded census information. Not all data-gathering and data recording practices have caused controversy about privacy. However, cybertechnology makes it possible to collect data about individuals without their knowledge and consent. In this section, we examine some controversial ways in which cybertechnology is used to gather and record personal data, as well as to monitor and track the activities and locations of individuals.

5.4.1 "Dataveillance" Techniques

Some believe that the greatest threat posed to personal privacy by cybertechnology lies in its capacity for surveillance and monitoring. Others worry less about the monitoring per se and more about the vast amounts of transactional data recorded via cybertechnology. Roger Clarke uses the term *dataveillance* to capture both the surveillance (data monitoring) and data recording techniques made possible by computer technology.[14] There are, then, two distinct controversies about dataveillance: one having to do with surveillance as a form of data monitoring and one having to do with the recording and processing of data once the data are collected. We examine both controversies, beginning with a look at data monitoring aspects of surveillance.

First, we should note the obvious, but relevant, point that privacy threats associated with surveillance are by no means peculiar to cybertechnology. Long before the advent of cybertechnology, individuals (e.g., private investigators and stalkers) as well as organizations, including governmental agencies all over the world, have used the latest technologies and techniques available to them to monitor individuals and groups.

Telephone conversations have been subject to government surveillance by wiretapping, but phone conversations have also been monitored in the private sector as well; for example, telephone conversations between consumers and businesses are frequently monitored, sometimes without the knowledge and consent of the consumers who are party to them. So surveillance is neither a recent concern nor one that should be associated exclusively with the use of cybertechnology to monitor and record an individual's online activities. However, surveillance has clearly been exacerbated by cybertechnology. Consider that video cameras now monitor consumers' movements while they shop at retail stores, and scanning devices used by "intelligent highway vehicle systems," such as E-ZPass, subject motorists to a type of surveillance while they drive through tollbooths. Sue Halpern notes that, as of 2011, approximately 500 companies monitor and track all of our movements online.[15]

In the past, it was not uncommon for companies to hire individuals to monitor the performance of employees in the workplace. Now, however, there are "invisible supervisors," that is, computers, that can continuously monitor the activities of employees around the clock without failing to record a single activity of the employee. We will examine workplace monitoring in detail, including some arguments that have been used to defend and to denounce computerized monitoring, in Chapter 10, where we consider some impacts that cybertechnology has for the contemporary workplace. In the remainder of this section, we consider surveillance techniques that involve nonworkplace-related monitoring and recording of personal data in both off- and online activities.

Although users may not always realize that they are under surveillance, their online activities are tracked by Web site owners and operators to determine how frequently users visit their sites and to draw conclusions about the preferences users show while accessing their sites. We next consider some controversies associated with a type of online surveillance technology known as cookies.

5.4.2 Internet Cookies

Cookies are text files that Web sites send to and retrieve from the computer systems of Web users, enabling Web site owners to collect information about a user's online browsing preferences whenever that user visits a Web site. The use of cookies by Web site owners and operators has generated considerable controversy, in large part because of the novel way that information about Web users is collected and stored. Data recorded about the user are stored on a file placed on the hard drive of the user's computer system; this information can then be

retrieved from the user's system and resubmitted to a Web site the next time the user accesses that site.

Those who defend the use of cookies tend to be owners and operators of Web sites. Proprietors of these sites maintain that they are performing a service for repeat users of a Web site by customizing the user's means of information retrieval. They also point out that, because of cookies, they are able to provide a user with a list of preferences for future visits to that Web site. Privacy advocates, on the other hand, see the matter quite differently. They argue that information gathered about a user via cookies can eventually be acquired by online advertising agencies, which can then target that user for online ads. For example, information about a user's activities on different Web sites can, under certain circumstances, be compiled and aggregated by online advertising agencies. The information can then be combined and cross-referenced in ways that enable a marketing profile of that user's online activities to be constructed and used in more direct advertisements. Also consider that Google now integrates information gathered from cookies with its wide array of applications and services, which include Gmail, Google+, Google Chrome, and others. As Zimmer (2008) notes, Google's ability to integrate this information provides the search engine company with a "powerful infrastructure of dataveillance" in which it can monitor and record users' online activities.

Some critics have also argued that because cookie technology both (a) monitors and records a user's activities while visiting Web sites (often without the user's knowledge and consent) and (b) stores that information on a user's computer or device, it violates the user's privacy. To assist Internet users who may be concerned about cookies, a number of privacy-enhancing tools (PETs), which are briefly described in Section 5.7, are available. Also, most current Web browsers provide users with an option to disable cookies. So with these browsers, users can either opt-in or opt-out of (accepting) cookies, assuming that they (i) are aware of cookie technology and (ii) know how to enable/disable that technology on their Web browsers. However, some Web sites will not grant users access unless they accept cookies.

Many privacy advocates also object to the fact that the default status for most Web browsers is such that cookies will automatically be accepted unless explicitly disabled by the user. So, cookie technology has raised a number of privacy-related concerns because of the controversial methods it uses to collect data about users who visit Web sites.

5.4.3 RFID Technology

Another mode of surveillance made possible by cybertechnology involves the use of radio frequency identification (RFID) technology. In its simplest form, RFID technology consists of a tag (microchip) and a reader. The tag has an electronic circuit, which stores data, and an antenna that broadcasts data by radio waves in response to a signal from a reader. The reader also contains an antenna that receives the radio signal, and it has a demodulator that transforms the analog radio information into suitable data for any computer processing that will be done (Lockton and Rosenberg 2005).

Although the commercial use of RFIDs was intended mainly for the unique identification of real-world objects (e.g., items sold in supermarkets), the tags can also be used to monitor those objects after they are sold. This relatively new mode of (continuous or "downstream") tracking of consumers' purchases has caused concern among some privacy advocates; for example, Nissenbaum (2004a) worries that consumers may not realize how RFID tags now make it possible for store managers to record, track, and share information about their purchases well beyond the initial point of sale.[16]

In one sense, the use of these tags in inventory control in retail contexts would seem uncontroversial. For example, Garfinkel (2002) notes that a company such as Playtex could place an RFID tag in each bra it manufactures to make sure that shipments of bras headed for Asia are not diverted to New York. He also points out, however, that a man with a handheld (RFID) reader in his pocket who is standing next to a woman wearing such a bra can learn the make and size of her bra. Additionally, and perhaps more controversially, RFID technology can be used for tracking the owners of the items that have these tags. So, on the one hand, RFID transponders in the form of "smart labels" make it much easier to track inventory and protect goods from theft or imitation. On the other hand, these tags pose a significant threat to individual privacy. Critics of this technology, which include organizations such as the Electronic Privacy Information Center (EPIC) and the American Civil Liberties Union (ACLU), worry about the accumulation of RFID transaction data by RFID owners and how those data will be used in the future.

RFID technology is already widely used—as Garfinkel notes, it has been incorporated into everything from automobile keys to inventory control systems to passports. If you have an E-ZPass (or some other intelligent highway systems) transponder in your car, for example, you already possess a wireless tag; E-ZPass uses the serial number on it to debit your account when your car passes through a tollbooth. Garfinkel notes that these tags now also appear in some of our clothing.

Many ranchers in the United States now track their cattle by implanting RFID tags in the animals' ears. In the future, major cities and municipalities might require RFID tags for domestic animals and pets; in Taiwan, for example, owners of domesticated dogs are required to have a microchip containing an RFID tag inserted in their pet dog's ear. In this case, the tag identifies the animal's owner and residence. Policies requiring RFID tags for some humans, especially for the elderly, may also be established in the near future. In the United States, some nursing homes now provide their patients with RFID bracelets. And chips (containing RFID technology) can now be implanted in children so that they can be tracked if abducted. On the one hand, this use of RFID technology may seem to empower parents of young children; on the other hand, however, Adam (2005) fears that we may come to rely too heavily on these technologies to care for children.

Like Internet cookies and other online data-gathering and surveillance techniques, RFID clearly threatens individual privacy. But unlike surveillance concerns associated with cookies, which track a user's habits while visiting Web sites, RFID technology can be used to track an individual's location in the offline world. We examine some specific privacy-and-surveillance concerns affecting RFID in connection with "location privacy" and "pervasive surveillance" issues in Chapter 12 in our discussion of ambient intelligence.

5.4.4 Cybertechnology and Government Surveillance

So far, we have examined surveillance techniques involving cybertechnology that are used mainly in the business and commercial sectors to monitor the activities of consumers and to record data about them. Another mode of surveillance that is also associated with cybertechnology involves governments and government agencies that monitor the activities of citizens, a practice that is sometimes referred to as "domestic spying."

Some cybertechnologies, despite their initial objectives and intent, can facilitate government surveillance. Consider, for example, that cell phone companies in the United States are required by law to include a GPS locator chip in all cell phones (manufactured after December 2005). This technology, which assists "911 operators" in emergency situations, also enables any cell phone user to be tracked within 100 meters of his or her location; so some privacy advocates worry that this information can also be used by the government to spy on individuals.

Government agencies currently use a variety of technologies that enable them to intercept and read private e-mail messages. In Chapter 6, we will see that this practice, initiated by the George W. Bush administration to monitor e-mail between U.S. residents and people living outside the United States, has been controversial. And in Section 5.6.1, we will see why the U.S. government's decision to subpoena the records of online search requests made by users of search engines such as Google, which are recorded and archived in computer databases, has also been controversial. In Chapter 7, we describe in detail some of the specific technologies (such as keystroke monitoring and biometric technologies) that have been used by government agencies in the United States to conduct surveillance on individuals. There, we will also see why these technologies, which have been used to combat terrorism and crime in cyberspace, have been controversial from the point of view of privacy and civil liberties.

While few would object to the desirable ends that increased security provides, we will see that many oppose the means—that is, the specific technologies and programs supporting surveillance operations, as well as legislation such as the USA Patriot Act—that the U.S. government has used to achieve its objectives. Our main purpose in this section has been to briefly describe how government surveillance of citizens illustrates one more way that cybertechnology both contributes to and enhances the ability of organizations to gather and record data about individuals.

In concluding this section, you may wish to revisit Scenario 5–1, where we noted that NSA's original charter was to conduct surveillance on entities (countries, organizations, and individuals) outside the United States. We also noted, however, that there is now compelling evidence to suggest that the NSA's mission has been broadened to include as surveillance on U.S. citizens as well. For more details on NSA-related surveillance in connection with the Edward Snowden controversy, see Scenario 4–2 in Chapter 4.

► 5.5 ANALYZING PERSONAL DATA: BIG DATA, DATA MINING, AND WEB MINING

In the previous section, we examined some ways in which personal data could be gathered using surveillance techniques. Other tools, however, have been devised to manipulate and analyze that (collected) data before it is transferred across, and exchanged between, electronic databases. Our focus in this section is on data analysis, as opposed to data collection. Simply collecting and recording personal data, per se, might not seem terribly controversial if, for example, the data were never manipulated (e.g., combined, recombined, matched, transferred, and exchanged) in preparation for further analysis. Some would argue, however, that the mere collection of personal data without someone's knowledge and consent is in itself problematic from a privacy perspective. Others assume that if data are being collected, there must be some motive or purpose for its collection. Of course, the reason, as many now realize, is that transactions involving the sale and exchange of personal data are a growing business.

Much of the personal data gathered or collected electronically by one organization is later exchanged with other organizations; indeed, the very existence of certain institutions depends on the exchange and sale of personal information. Some privacy advocates believe that professional information-gathering organizations, such as Equifax, Experian (formerly TRW), and TransUnion (credit reporting bureaus), as well as the Medical Information Bureau (MIB), violate the privacy of individuals because of the techniques they use to transfer and exchange personal information across and between databases. Many also believe that this process has been exacerbated by the phenomenon of *big data*.

5.5.1 Big Data: What, Exactly, Is It, and Why Does It Threaten Privacy?

Ward and Barker (2013) note that while the term "big data" has become "ubiquitous," it has no precise or "unified single" meaning. They also point out that the definitions of big data put forth thus far are not only "diverse," but are often "contradictory" as well.[17] However, Ward and Barker also propose a working definition based on an "extrapolation" of key factors that cut across various definitions of big data, which they describe as an "analysis of large and/or complex data sets using a series of techniques."[18] So, in their view, big data can be understood in terms of one or more of three key factors: size, complexity, and technologies/tools (used to analyze the data).

Initially, one might assume that the concept of big data simply refers to the size or scale of the data being analyzed. For example, Boyd and Crawford (2012) suggest that big data can be understood mainly in terms of its "capacity to search, aggregate and cross-reference *large data sets*."[19] Definitions that focus on capturing the large size of the data sets involved often view big data primarily in terms of its *volume*. Other definitions, however, include factors sometimes referred to as the "three Vs": *variety*, *velocity*, and *veracity*.[20] The "variety" component or element describes the wide range of sources involved in the data analysis, which include social media, scientific applications, business transactions, Internet search indexing, medical records, and Web logs. Whereas "velocity" captures the speed ("fast data in/out") involved in the process, "veracity," refers to the notion of trust in the (big) data analysis that needs to be established for business decision making. So the concept of big data is a far more complex phenomenon than merely the size, or volume, of the data involved. As Poskanzer (2015) points out, in the case of big data, "more isn't just more—more is different." She further suggests that big data can be better understood as a "new mode of knowledge production."[21]

Others believe that the concept of big data can be understood in terms of certain "insights" it purportedly provides into new and emerging types of data and content. Some also suggest that big data can be viewed as an "emerging paradigm," or perhaps even as providing a "paradigm shift" for analyzing data. But has a genuinely new or an emerging paradigm been provided or has the expression "big data" simply become a "new buzzword"—one that is now so ambiguous, as well as ubiquitous, that it is no longer effective? Perhaps the real "shift" (paradigm or otherwise) in data analysis was ushered in when the technique commonly known today as *data mining* (defined in the following section) first became available—that is, when certain kinds of pattern-matching algorithms made possible by AI research were first used to analyze terabytes of data to "discover" information that otherwise would not have been obvious. For example, some data mining techniques led to the "discovery" of new (and mostly nonobvious) groups and of "new facts" about people. This technique or process has also been referred to as "knowledge discovery in databases" (KDD).

Regardless of which expression we use to describe this phenomenon—big data, data mining, or KDD—serious privacy concerns have been generated by it. Some believe that these kinds of concerns justify the need for a new legal category of privacy, which some call "group privacy."[22] However, we will see that many, if not most, of the kinds of privacy concerns currently associated with big data had already been introduced by the use of various data mining techniques, beginning in the 1990s. We next examine the concept of data mining in detail to understand how these specific privacy issues arose and why they are problematic.

5.5.2 Data Mining and Personal Privacy

Data mining can be defined as a technique that involves the manipulation of information, including personal information, through an analysis of implicit patterns discoverable in large data sets. (In this respect, it is very similar to many definitions of big data.) Also, data mining can generate new and sometimes nonobvious classifications or categories of persons (which is

again similar to some definitions of big data); as a result, individuals whose data are mined can become identified with or linked to certain newly created groups that they might never have imagined to exist. This is further complicated by the fact that current privacy laws offer individuals relatively little protection with respect to how information about people acquired through data mining activities is subsequently used, even though important decisions can be made about those individuals based on the patterns found in the mined personal data. So, data mining technology can be used in ways that raise special concerns for personal privacy.[23]

But what is so special about the privacy concerns raised by data mining? For example, how do they differ from privacy issues introduced by more traditional data retrieval and exchange techniques? For one thing, privacy laws as well as informal data protection guidelines have been established for protecting personal data that are:

- Explicit in databases (in the form of specific electronic records)
- Confidential in nature (e.g., data involving medical, financial, or academic records)
- Exchanged between databases

However, relatively few legal or normative protections apply to personal data manipulated in the data mining process, where personal information is typically:

- Implicit in the data
- Nonconfidential in nature
- Not necessarily exchanged between databases

Unlike personal data that reside in explicit records in databases, information acquired about persons via data mining is often derived or inferred from implicit patterns in the data. The patterns can suggest "new" facts, relationships, or associations about a person, placing that person in a "newly discovered" category or group. Also, because most personal data collected and used in data mining applications is considered neither confidential nor intimate in nature, there is a tendency to presume that such data must, by default, be public data. And unlike the personal data that are often exchanged between or across two or more databases in traditional database retrieval processes, in the data mining process, personal data are often manipulated within a single database or within a large "data warehouse."

Next, consider a scenario involving data mining practices at a lending institution in determining whether or not to grant mortgages to its customers. As you consider the privacy issues raised in the following scenario, keep in mind Nissenbaum's distinction between "norms of appropriateness" and "norms of distribution" for determining contextual integrity (described in Section 5.2.5).

▶ **SCENARIO 5–5:** Data Mining at the XYZ Credit Union

Jane, a senior real estate professional at CBA Real Estate, wishes to purchase a condominium, and she has recently applied for a mortgage at the XYZ Credit Union. To be considered for this loan, Jane is required to fill out a number of mortgage-related forms, which she willingly completes. For example, on one form, she discloses that she has been employed by CBA for more than seven years and that her current annual salary is $95,000. On another form, Jane discloses that she has $50,000 in her savings account at a local bank (much of which she plans to use for the down payment on the house she hopes to purchase). Additionally, she discloses that she has $1,000 of credit card debt and still owes $3,000 on an existing car loan. The amount of the loan for the mortgage she hopes to secure is for $100,000 over a 30-year period.

After Jane has completed the forms, the credit union's computing center runs a routine data mining program on information in its customer databases and discovers a number of patterns. One reveals that real estate professionals earning more than $80,000 but less than $120,000 annually are also likely to leave their current employers and start their own businesses after 10 years of employment. A second data

mining algorithm reveals that the majority of female real estate professionals declare bankruptcy within two years of starting their own businesses. The data mining algorithms can be interpreted to suggest that Jane is a member of a group that neither she nor possibly even the mortgage officers at the credit union had ever known to exist—namely, the group of female real estate professionals likely to start a business and then declare bankruptcy within two years. With this newly inferred information about Jane, the credit union determines that Jane, because of the newly created category into which she fits, is a long-term credit risk. So, Jane is denied the mortgage. ■

Does the credit union's mining of data about Jane raise any significant privacy concerns? At one level, the transaction between Jane and the credit union seems appropriate. To secure the mortgage from XYZ Credit Union, Jane has authorized the credit union to have the information about her, that is, her current employment, salary, savings, outstanding loans, and so forth, that it needs to make an informed decision as to whether or not to grant her the mortgage. So, if we appeal to Nissenbaum's framework of privacy as contextual integrity, it would seem that there is no breach of privacy in terms of norms of appropriateness.

However, Jane gave the credit union information about herself for use in one context, namely, to make a decision about whether or not she should be granted a mortgage for her condominium. She was also assured that the information given to the credit union would not be exchanged with a third party, without first getting Jane's explicit consent. So, no information about Jane was either exchanged or cross-referenced between external databases—that is, there is no breach of the norms of distribution (in Nissenbaum's model, described in Section 5.2.5). However, it is unclear whether the credit union had agreed not to use the information it now has in its databases about Jane for certain in-house analyses.

Although Jane voluntarily gave the credit union information about her annual salary, previous loans, and so forth, she gave each piece of information for a specific purpose and use, in order that the credit union could make a meaningful determination about Jane's request for a mortgage. However, it is by no means clear that Jane authorized the credit union to use disparate pieces of that information for more general data mining analyses that would reveal patterns involving Jane that neither she nor the credit could have anticipated at the outset. Using Jane's information for this purpose would now raise questions about "appropriateness" in the context involving Jane and the XYZ Credit Union.

The mining of personal data in Jane's case is controversial from a privacy perspective for several reasons. For one thing, the information generated by the data mining algorithms suggesting that Jane is someone likely to start her own business, which would also likely lead to her declaring bankruptcy, was not information that was "explicit" in any of the data (records) about Jane per se; rather, it was "implicit" in patterns of data about people similar to Jane in certain respects but also vastly different from her in other respects. For another thing, Jane's case illustrates how data mining can generate new categories and groups such that the people whom the data mining analysis identifies with those groups would very likely have no idea that they would be included as members. And we have seen that, in the case of Jane, certain decisions can be made about members of these newly generated groups simply by virtue of those individuals being identified as members. For example, it is doubtful that Jane would have known that she was a member of a group of professional individuals likely to start a business and that she was a member of a group whose businesses were likely to end in bankruptcy. The "discovery" of such groups is, of course, a result of the use of data mining tools.

Even though no information about Jane was exchanged with databases outside XYZ, the credit union did use information about Jane internally in a way that she had not explicitly authorized. And it is in this sense—unauthorized internal use by data users—that many believe data mining raises serious concerns for personal privacy. Note also that even if Jane had been granted the mortgage she requested, the credit union's data mining practices would still have raised privacy concerns with respect to the *contextual integrity* of her personal information.

Jane was merely one of many credit union customers who had voluntarily given certain personal information about themselves to the XYZ for use in one context—in this example, a mortgage request—and subsequently had that information used in ways that they did not specifically authorize.

Controversial Consumer Profiles Generated by Data Mining Techniques

The scenario involving Jane is, of course, hypothetical. But there is now empirical evidence to suggest that banks and consumer credit organizations are using data mining techniques to determine an individual's "credit worthiness" in ways that are not so different from the process described in Scenario 5–5. So, in some cases, a consumer's credit rating is actually determined via profiling schemes that can suggest "guilt by association." For example, a consumer could be denied a credit card, or have one revoked, merely because of where she shops or where she lives. Also, consider that people living in neighborhoods where there have been high rates of home foreclosures, or people holding mortgages with certain banks or lending institutions that have experienced high rates of home foreclosures, may now be considered credit risks by virtue of their association with either a certain neighborhood or bank, even though they have been responsible in paying their mortgages and other loans on time.

Similarly, if individuals shop at a certain kind of retail store, say, Walmart, information about their purchases at such a store can associate them with other individuals who shop there, and who may have a higher-than-average default rate on their credit cards. For example, Stuckey (2008) describes an incident where a 37-year-old computer consultant had two of his American Express cards canceled and the limit on a third card reduced based on criteria having to do with (i) where he shopped and (ii) the financial institution with whom he held his mortgage. When this person questioned American Express's decision, he was informed that the criteria it uses to decide to reduce the spending limit on someone's credit card include:

> credit experience with customers who have made purchases at establishments where you have recently used your card.
>
> analysis of the credit risk associated with customers who have residential loans from the creditor(s) indicated in your credit report.

While there had been suspicion for many years that credit card companies did indeed engage in the kind of profiling scheme used by American Express, consumer advocates and credit analysts believe that this may have been the first time that a major credit company admitted to using such criteria. In its defense, however, American Express claimed that it needed to analyze its exposure to risk as it reviews its cardholder's credit profiles in light of the economic turndown in the United States (in 2008–2009) that severely affected the credit industry at that time (Stuckey 2008).

Can Data Mining Techniques also Be Used in Ways that Protect Consumer Privacy?

We have seen how data mining can be used to threaten consumer privacy. But can it also be used to protect consumers against fraudulent activities? Perhaps not surprisingly, data mining, like other technologies, can be viewed as a "double-edged sword" with respect to consumers' interests, as the following story suggests. One day, to my surprise, I received a telephone call from my credit card company informing me that a purchase, which the company apparently viewed as suspicious, had been charged earlier that day to my credit card account. When asked about the purchase, I informed the company's representative that it had not been made by me, and I also thanked the person for notifying me so promptly about this transaction. The company representative then immediately canceled my existing credit card and issued me a new card with a new account number.

Why did the company suspect that the purchase made that day with my credit card was questionable? It would seem that the data mining algorithms used by the credit card company to determine the patterns of my purchases—which kinds of purchases and credit card transactions I typically make, with whom and where I make them, and when—generated suspicion about the questionable purchase made that day with my credit card. So in this instance, data mining appeared to have been used in a way that protected the interests of a consumer.

5.5.3 Web Mining: Analyzing Personal Data Acquired from Our Interactions Online

Initially, the mining of personal data depended on large (offline) commercial databases called *data warehouses*, which stored the data, consisting primarily of transactional information. Data mining techniques are now also commonly used by commercial Web sites to analyze data about Internet users. This process is sometimes referred to as *Web mining*, which has been defined as the application of data mining techniques to discover patterns from the Web.[24] The various kinds of patterns discovered via Web mining are often used by marketers, especially in their online advertisements and promotional campaigns.

A now classic case of Web mining involved Facebook's Beacon initiative in 2007, which enabled Facebook friends to share information about their online activities, including online purchases they made. This initiative was controversial from the outset, however, because it also allowed targeted advertisements by the Web sites sending the data to Facebook. In response to the outpouring of criticism Facebook received for collecting more user information for advertisers than it had originally admitted, the popular social networking service decided to cancel Beacon in December 2007.[25] However, critics worry that Facebook and other social networking services still engage in various forms of Web mining.

Because the amount (or volume) of data currently available on the Web is so vast, one might assume that it is impossible to mine those data in ways that could be useful. However, current data mining tools employ sophisticated techniques that can "comb" through the massive amounts of data on the Web; collecting and analyzing this volume of data would not have been possible using earlier kinds of information gathering/analysis techniques. (Recall our brief examination of "big data" in Section 5.5.1, where we saw how easy it is now to analyze extremely large data sets.) Also, sophisticated search engines have programs (called "spiders") that "crawl" through the Web in order to uncover general patterns in information across multiple Web sites. Halpern (2011) points out that approximately 500 companies now mine the "raw material of the Web" and then sell it to data mining companies. And Pariser (2011) notes that one of these companies, Acxiom, has managed to accumulate 1,500 pieces of data, on average, for each person in its database; this personal data ranges from people's credit scores to the kinds of medications they use.

Pariser also notes that Google and other major search engine companies use "prediction engines" to construct and refine theories about us and the kinds of results we desire from our search queries. In Section 5.7.1, we examine some specific ways in which the use of Internet search engines raise privacy concerns, even though the kind of personal information about us that is acquired by search engine companies might not initially seem to warrant explicit privacy protection. To see why such protection might indeed be needed in these cases, however, we first examine some questions underlying a concern that Nissenbaum (2004b) calls the "problem of privacy in public."

► 5.6 PROTECTING PERSONAL PRIVACY IN PUBLIC SPACE

So far, we have examined how cybertechnology can be used to gather, exchange, and mine personal information. With the exception of data mining, which manipulates personal, but nonconfidential information, the kind of personal information gathered and exchanged was

often confidential and sensitive in nature. For example, financial and medical records could be exchanged between two or more databases using computerized merging and matching techniques. This confidential/sensitive personal information is sometimes referred to as nonpublic personal information (NPI). Privacy analysts are now concerned about a different kind of personal information—public personal information (PPI), which is neither confidential nor sensitive and which is also being gathered, exchanged, and mined via cybertechnology.

5.6.1 PPI vs. NPI

PPI includes information about you, such as where you work or attend school or what kind of car you drive. Even though it is information about you as a particular person, PPI has not enjoyed the privacy protection that has been granted to NPI. Until recently, most concerns about personal information that was gathered and exchanged electronically were limited to NPI, and because of the attention it has received, privacy laws and policies were established to protect NPI. But now, privacy advocates are extending their concern to PPI; they argue that PPI deserves greater legal and normative protection than it currently has. As noted previously, Nissenbaum refers to this challenge as the problem of protecting privacy in public.

Why should the collection and exchange of PPI raise privacy concerns? Suppose that I discover some of the following information about you: you are a junior at Technical University, you frequently attend your university's football games, and you are actively involved in your university's computer science club. In one sense, the information that I have discovered about you is personal, because it is about you (as a person), but it is also public, because it pertains to things that you do in the public sphere. Should you be worried that this information about you is so easily available?

In the past, the public availability of such seemingly harmless and uncontroversial information about you was no cause for concern. Imagine that 80 years ago a citizen petitioned his or her congressperson to draft legislation protecting the privacy of each citizen's movements in public places. It would have been difficult then to make a strong case for such legislation; no one would have seen any need to protect that kind of personal information. But today, many argue that we need to protect privacy in public, pointing out that our earlier assumptions are no longer tenable. Nissenbaum (2004b) believes that many in the commercial sector proceed from an assumption that she believes is "erroneous"—namely, "There is a realm of public information about persons to which no privacy norms apply."[26] Keep this assumption in mind as you consider the following two scenarios.

▶ **SCENARIO 5–6:** Shopping at SuperMart

On your way home from class, you decide to stop at SuperMart to shop for groceries. If I also happen to shop there and see you enter or leave SuperMart, or if we are both shopping in this store at the same time, I now have information that you shop (or, at least, have once shopped) at SuperMart. (This information could be considered "public" because it was acquired in a public forum and because it is neither intimate nor confidential in nature.) If I also happen to pass by you in one of the aisles at SuperMart, I can observe the contents of your shopping basket; I may notice, for example, that your cart contains several bottles of wine but relatively little food. Again, I have acquired this information about you by observing your activity in a public forum. ■

Because the information I have acquired about you in the above scenario can be considered public information, it would not warrant any legal privacy protection. And even though this information is about you as a person, it is not the kind of personal information to which we, as a society, would typically grant normative privacy protection. What, exactly, is the privacy problem regarding the kind of personal information about your public activities in shopping at SuperMart? Why should you be concerned about information that is gathered about

what you do at SuperMart or, for that matter, in any public place? Let us continue the shopping metaphor, but this time, we consider shopping that takes place in an online forum.

▶ **SCENARIO 5–7:** Shopping at Nile.com

Imagine that you visit an online bookstore called Nile.com to view a particular book that you are considering purchasing. Because you are visiting this bookstore via a computer or electronic device, you cannot be physically observed by other users who also happen to be visiting Nile's Web site at that time. However, from the moment you enter that site, information about you is being intentionally gathered and carefully recorded—that is, information about the exact time that you entered Nile, as well as the exact time that you leave. As you make your initial contact with the Nile Web site, Nile requests a cookie file from your device to determine whether you have previously visited this site. If you have visited this site before and have clicked on items that interested you, Nile can find a record of these items. The information stored in that cookie file can also be used by Nile to alert you to newly released books that it believes might interest you, based on an analysis of the data Nile collected from your previous visits to its site. ■

The information that Nile now has about you does not seem categorically different from the information that SuperMart might also have about you (assuming, for example, that you used that store's "courtesy card" or discount card in making your purchases). However, there are significant differences in the ways that information about you can be gathered, recorded, and then used as a result of your shopping at each store.

When you shopped in physical space at SuperMart, only a list of your actual purchases could be recorded and stored in SuperMart's databases. Items that might have only caught your attention and items that you might also have picked up or even placed in your cart at one point while shopping but did not eventually purchase at the checkout register are not recorded by SuperMart's data-collection system. However, as you shop, or even browse, at Nile, there is a record of virtually every move you make—every book that you search, review, etc., as well as the one(s) you purchase. Yet, just like the information gathered about your shopping habits in physical space at SuperMart, this personal information that Nile has gathered about your browsing and shopping habits online is considered and treated as public information (i.e., not treated as NPI).

Now, we can see why some people worry about having their movements online tracked and recorded. The information Nile gathered about you is, in effect, Nile's information, even though it pertains to you as a person; Nile now owns that information about you, as well as the information it has about its other customers, and is, in principle at least, free to do with that information whatever it chooses (so long as it is consistent with any consumer privacy policies it may happen to have). On the one hand, the information seems fairly innocuous—after all, who really cares which books you happen to browse or purchase? On the other hand, however, this information can be combined with other information about your online transactions at additional Web sites to create a consumer profile of you, which can then be sold to a third party.

One argument that online entrepreneurs might advance to defend these business practices is that if a user puts information about him- or herself into the public domain of the Internet, then that information is no longer private. Of course, one response to this line of reasoning could be to question whether users clearly understand the ways that data they submit might subsequently be used.

In Scenario 5–7, Nile used information about you in ways that you neither explicitly authorized nor likely intended—an example of the kind of practice that Nissenbaum (2004a, 2010) describes as violating "contextual integrity" (see Section 5.2.5). Also, we can question whether businesses, such as Nile.com, should be able to "own" the information about us that they collect and then do with that information whatever they please and for as long as they want?

Fulda (2004) questions whether the old legal rule that states, "Anything put by a person in the public domain can be viewed as public information," should still apply. He admits that such a rule may have served us well, but only before data were "mined" to produce profiles and other kinds of patterns about individuals.[27]

5.6.2 Search Engines and the Disclosure of Personal Information

Internet search engines are valuable for directing us to available online resources for academic research, commerce, recreation, and so forth; so it might be surprising to find that search engine technology, too, can be controversial from the perspective of personal privacy. How can search engine technology conflict with personal privacy? At least two different kinds of concerns affecting privacy arise because of practices involving search engines: (i) search engine companies such as Google record and archive each search request made by users and (ii) search engines enable users to acquire a wealth of personal information about individuals, with relative ease. We begin with a brief examination of (i).

Google and Its Practice of Collecting Records of Users' Searches

Google creates a record of every search made on its site, which it then archives. The topic searched for, as well as the date and time the specific search request is made by a user, are included in the record. These data can be linked to the IP address and the ISP of the user requesting the search. So individual searches made by a particular user could theoretically be analyzed in ways that suggest patterns of that individual's online behavior, and, perhaps more controversially, these records could later be subpoenaed in court cases. Yet, until relatively recently, many (if not most) Google users were unaware of the company's policy regarding the recording and archiving of users' search requests.

On the one hand, this information might seem relatively innocuous—after all, who would be interested in knowing about the kinds of searches we conduct on the Internet, and who would want to use this information against us? On the other hand, however, consider the case of a student, Mary, who is writing a research paper on Internet pornography. Records of Mary's search requests could reveal several queries that she made about pornographic Web sites, which in turn might suggest that Mary was interested in viewing pornography. Following a controversial decision by the George W. Bush administration in 2005, Google users discovered that any worries they may have had about the lack of privacy protection concerning their Internet searches were justified. That year, the Bush administration informed Google that it would be required to turn over a list of all users' queries entered into its search engine during a one-week period. Initially, Google refused to comply with the subpoena on the grounds that the privacy rights of its users would be violated. Yahoo, however, which also had its search records subpoenaed, complied with the government's initial request.[28]

The Bush administration's decision to seek information about the search requests of ordinary users has since drawn significant criticism from many privacy advocates. Critics argued that although the Bush administration claimed that it had the authority to seek electronic information in order to fight the "war on terror" and to prevent another September 11-like attack, the records at issue in this particular case had to do with the number of users requesting information about, or inadvertently being sent to, pornographic Web sites. Some critics further argued that the Bush administration was interested in gathering data to support its stance on the Children's Internet Protection Act (CIPA), which had been challenged in a U.S. District Court (see Chapter 9). So, many critics were quick to point out that the Bush administration's rationale for obtaining records of search requests made by ordinary citizens seemed politically and ideologically motivated and may have had nothing to do with protecting national security.

Using Search Engines to Acquire Information about People

It is not only the fact that an individual's search requests are recorded and archived by major companies such as Google that make Internet search engines controversial from the perspective of personal privacy. Search engine-related privacy issues also arise because that technology can be used for questionable purposes such as stalking. In fact, one search facility—Gawker-Stalker (www.gawker.com/stalker)—has been designed specifically for the purpose of stalking famous people, including celebrities. For example, suppose that Matt Damon is spotted ordering a drink at an upscale café in Boston. The individual who spots Damon can send a "tip" via e-mail to Gawker-Stalker, informing the site's users of Damon's whereabouts. The Gawker site then provides its users, via precise GPS software, with information about exactly where, and at what time, Damon was sighted. Users interested in stalking Damon can then follow his movements electronically, via the Gawker site, or they can locate and follow him in physical space, if they are in the same geographical vicinity as Damon.

But it is not just celebrities who are vulnerable to information about them being acquired by others via search engines. Consider the amount and kind of personal information about ordinary individuals that is now available to search engines. In some cases, that information may have been placed on the Internet inadvertently, without the knowledge and consent of those affected. Yet information about those persons can be located by an Internet user who simply enters their names in a search engine program's entry box. The fact that one can search the Internet for information about someone might not seem terribly controversial. After all, people regularly place information about themselves on Web sites (or perhaps they authorize someone else to do it for them) and on social networking services such as Facebook. And it might seem reasonable to assume that any online personal information that is currently available to the public should be viewed simply as public information. But should such information about persons be unprotected by privacy norms merely because it is now more easily accessible for viewing by the public? (In Section 5.8, we consider whether users should have a "right" to have some kinds of online personal information about them either deleted or "de-linked" from search engine indexes.)

We have seen how the use of search engines can threaten the privacy of individuals in two distinct ways: (i) by recording and archiving records of a user's search queries that reveal the topic of the search and the time the request was made by the user and (ii) by providing users of search engines with personal information about individuals who may have no idea of the wealth of personal information about them that is available online (and have no control over how it is accessed and by whom it is accessed). The latter concern is further complicated by the fact that individuals who are the subject of online searches, including celebrities who can be stalked (as we saw in the case of Gawker), enjoy no legal protection because of the presumed "public" nature of the personal information about them that is available via online searches.

So far, we have seen how our personal information can be collected and then manipulated by search engines in ways that are controversial.[29] A variation of this privacy-related controversy involves access to personal information that resides in public records made available online via online searches. In Section 5.1.2, we saw that once information is converted to digital form, it can live on indefinitely; so there is no time limit or expiration date for most public records. As we will see in our analysis of a recent European privacy principle called "the right to be forgotten" in Section 5.8, links to documents about an unfortunate incident in one's distant past, which may no longer be "relevant," can continue to be available online and thus haunt that person indefinitely. So one might ask: Do we need stricter privacy laws, especially in the United States, to protect us in the digital era?

▶ 5.7 PRIVACY LEGISLATION AND INDUSTRY SELF-REGULATION

Many privacy advocates believe that stronger privacy laws are needed to protect the interests of online consumers, as well as ordinary users. Others, however, especially those in the commercial sector, argue that additional privacy legislation is neither necessary nor desirable. Instead, they suggest the use of voluntary controls regulated by industry standards. Generally, privacy advocates have been skeptical of voluntary controls, including most industry standards affecting "self-regulation," arguing instead for stricter privacy/data protection frameworks backed by explicit legislation. We begin this section with an examination of some industry-initiated, self-regulatory schemes designed to protect consumer privacy.

5.7.1 Industry Self-Regulation and Privacy-Enhancing Tools

Some who advocate for the use of (voluntary) self-regulatory controls point out that various privacy enhancing tools (PETs), designed to protect a user's privacy while navigating the Internet, are already available. For example, some PETs enable users to navigate the Web anonymously; perhaps, one of the best-known tools of this type is the *Anonymizer* (available from Anonymizer.com).[30] Another useful tool is TrackMeNot (http://cs.nyu.edu/trackmenot/), which was designed to work with the Firefox Web browser to protect users against surveillance and data profiling by search engine companies. Rather than using encryption or concealment tools to accomplish its objectives, TrackMeNot instead uses "noise and obfuscation." In this way, a user's Web searches become "lost in a cloud of false leads." By issuing randomized search queries to popular search engines such as Google and Bing, TrackMeNot "hides users' actual search trails in a cloud of 'ghost' queries." This technique makes it difficult for search engine companies to aggregate the data it collects into accurate user profiles.

Although some users have found anonymity tools (and other kinds of PETs) helpful, many question their overall effectiveness in protecting the privacy of online consumers, as well as ordinary Internet users. In fact, even many industry self-regulation proponents would likely concede that PETs alone are not sufficient. But they still oppose the idea of any additional privacy legislation, arguing instead for better enforcement of industry standards that have already been accepted and implemented. Some of these standards are similar to PETs in their intended objective, that is, to protect an online consumer's privacy, but are also unlike PETs in that they cannot be classified as "tools" in the strict (or technological) sense of the term.

One industry-backed (self-regulatory) framework, designed to help ensure that commercial Web sites adhere to the privacy policies they advertise, is TRUSTe. This framework uses a branded system of "trustmarks" (i.e., graphic symbols) to represent a Web site's privacy policy regarding personal information. Trustmarks provide consumers with the assurance that a Web site's privacy practices accurately reflect its stated policies. If a Web site bearing its trust seal does not abide by the stated policies, users can file a complaint to TRUSTe. Any Web site that bears the TRUSTe mark and wishes to retain that seal must satisfy several conditions: The Web site must clearly explain in advance its general information-collecting practices, including which personally identifiable data will be collected, what the information will be used for, and with whom the information will be shared. Web sites that bear a trust seal but do not conform to these conditions can have their seal revoked. And Web sites displaying trust seals, such as TRUSTe, are subject to periodic and unannounced audits of their sites.

Critics have pointed out some of the difficulties that users encounter interacting with frameworks like TRUSTe. For example, the amount of information users are required to provide can easily discourage them from carefully reading and understanding the agreement. Also, the various warnings displayed may appear unfriendly and thus might discourage users; "friendlier" trustmarks, on the contrary, might result in users being supplied with less direct

information that is important for protecting their privacy. But advocates of self-regulatory frameworks such as TRUSTe argue that, with them, users will be better able to make informed choices regarding online commercial transactions.

Some critics also worry that schemes like TRUSTe do not go far enough in protecting consumers. Consider, for example, a now classic incident involving Toysmart.com, an e-commerce site that once operated in the state of Massachusetts. Consumers who purchased items from Toysmart were assured, via an online trust seal, that their personal data would be protected. The vendor's policy stated that personal information disclosed to Toysmart would be used internally but would not be sold to or exchanged with external vendors. So, users who dealt with Toysmart expected that their personal data would remain in that company's databases and not be further disclosed or sold to a third party. In the spring of 2000, however, Toysmart was forced to file for bankruptcy.

In the bankruptcy process, Toysmart solicited bids for its assets, which included its databases containing the names of customers.[31] One question that arose was whether the parties interested in purchasing that information were under any obligation to adhere to the privacy policy that Toysmart had established with its clients? If they were not, then whoever took over Toysmart's site or purchased its databases, would, in principle, be free to do whatever they wished with the personal information in the databases. They would conceivably be able to do this, despite the fact that such information was given to Toysmart by clients in accordance with an explicit privacy policy that guaranteed that personal information about them would be protected indefinitely.

A slightly different, but related, kind of privacy policy concern arises in the context of search engine companies. Unlike e-commerce sites, which users can easily avoid if they wish, virtually every Internet user depends on search engines to navigate the Web. In Section 5.6.1, we saw how major search engine companies such as Google record and keep a log of users' searches. This practice, as we also saw, has generated privacy-related concerns for ordinary users and was further complicated by the fact that Google offers many other kinds of services in addition to its well-known search engine. These include Gmail, Google Maps, Google+, Google Calendar, Google Chrome, Picasa, AdSense/AdWords, YouTube (which was acquired by Google), and so forth. So, Google had developed separate privacy policies for its services, and these policies varied from service to service.

In 2012, Google announced a new comprehensive privacy policy, which replaced its individual privacy policies for each of its services. The new policy, however, also allowed the sharing of user account data across all its services, subsidiary services, and Web-based applications. When Google implemented its new privacy policy, critics noted that a user's search engine history could now be shared with YouTube, or vice versa, and that a user's Google+ account data might be shared with AdWords to generate more targeted advertising.[32]

Google's 2012 privacy policy, while explicit and transparent, has nonetheless been controversial for several reasons. For one thing, it is not clear how Google will use all of the personal information that it can now access so easily. For another, no one outside Google fully understands how the search engine company uses that information to manipulate (i.e., tailor or personalize) the search results a user receives for his or her search queries. Additionally, it is not clear whether one's personal information collected from the various Google services will be used only internally or will also be available to advertisers and information merchants outside the company (e.g., those Web sites that include embedded Google ads to generate revenue).

Other critics worry whether users can trust Google—a company that officially embraces the motto: "do not be evil"—to abide by its new privacy policy. Some note, for example, that many people who used Apple's Safari Web browser on their computers and iPhones were under the impression that Google was not able to track their browsing activities. In 2012, however, it was discovered Google had used software code that tricked the Safari browser, thus

enabling Google to track the activities of those using that browser. Google disabled the controversial software code shortly after the incident was reported in *The Wall Street Journal*, and Safari users were informed by Google that they could rely on Safari's privacy settings to prevent tracking by Google in the future (Anguin and Valentino-DeVries 2012). But some critics remain skeptical.

Because of concerns involving distrust of major search engine companies like Google, as well as commercial Web sites in general, to regulate themselves, many privacy advocates believe that the only plausible alternative for protecting users is to enact better, and more explicit, privacy laws. We next briefly examine some existing privacy legislations in the North America (mainly the United States) and Europe.

5.7.2 Privacy Laws and Data Protection Principles

Many nations, especially in the West, have enacted strong privacy legislation. The United States, however, has not taken the lead on legislation initiatives in this area; in fact, some would argue that the United States is woefully behind Canada and the European nations when it comes to protecting its citizens' privacy. For example, in the United States, there is currently very little privacy protection provided in legal statutes. In 1974, Congress passed the Privacy Act, which has been criticized both for containing far too many loopholes and for lacking adequate provisions for enforcement. Also, it applies only to records in federal agencies and thus is not applicable in the private sector.

Critics also point out that there is virtually no explicit legal protection for private e-mail communications in the United States. Julian Sanchez notes that the Electronic Communications Privacy Act (ECPA) of 1986, which was "tweaked in the early 1990s," was written before most people had even heard of the Internet. Some U.S. citizens might assume that the Fourth Amendment (prohibiting government search and seizure) also applies to the protection of e-mail communications. However, Sanchez points out that it was not until 2010 that a court in the United States finally ruled in favor of privacy protection for an e-mail communication—and he notes that this ruling was handed down only at the circuit court level of one federal court.[33]

In 2003, the Health Insurance Portability and Accountability Act (HIPAA), which provides protection for "individually identifiable" medical records from "inappropriate use and disclosure," was enacted into law in the United States. But the kind of privacy protection provided by HIPAA does not apply to an individual's nonmedical/health records such as consumer data, or even to one's genetic data. Enactment of the Genetic Information Nondiscrimination Act (GINA), in 2008, explicitly extended privacy protection to personal genetic information. So, some federal privacy laws have successfully targeted specific contexts such as healthcare and genetic information. However, critics argue that privacy legislation in the United States has resulted mostly in a "patchwork" of individual state and federal laws that are neither systematic nor coherent.

Generally, U.S. lawmakers have resisted requests from privacy advocates and consumer groups for stronger consumer privacy laws. Instead, they have sided with business interests in the private sector, who believe that such legislation would undermine economic efficiency and thus adversely impact the overall economy. Critics point out, however, that many American businesses that have subsidiary companies or separate business operations in countries with strong privacy laws and regulations, such as nations in Western Europe, have found little difficulty in complying with the privacy laws of the host countries; furthermore, profits for those American-owned companies have not suffered because of their compliance. In any event, there has been increased pressure on the U.S. government, especially from Canada and countries in the European Union (EU), to enact stricter privacy laws (as well as pressure on American businesses to adopt stricter privacy policies and practices to compete in e-commerce at the global level).

EU nations, through the implementation of strict "data protection" principles, have been far more aggressive than the United States in both anticipating and addressing privacy concerns raised by cybertechnology. In the early 1990s, the European community began to synthesize the "data protection" laws of the individual European nations.[34] The European community has since instituted a series of "directives," including the EU Directive 95/46/EC of the European Parliament and of the Council of Europe of 24 October 1995.[35] The latter, also sometimes referred to simply as the EU Directive on Data Protection, was designed to protect the individual rights of citizens who reside within the EU, while also facilitating the flow of data beyond the EU nations. As such, the EU Directive on Data Protection prohibits the "transborder flow" of personal data to countries that do not provide adequate protection of personal data. Elgesem (2004) has pointed out that a central focus of this directive, unlike earlier privacy legislation in Europe that focused simply on the recording and the storage of personal data, is on the "processing and flow" of that data.

Several principles make up the European Directive on Data Protection; among them are the principles of Data Quality and Transparency. Whereas the Data Quality Principle is concerned with protecting the data subject's reasonable expectations concerning the processing of data about that subject (ensuring that the personal data being processed is true, updated, and properly kept), the Transparency Principle grants the data subject the rights to be informed, to contest, to correct, and "to seek judicial redress."[36] What helps to ensure that each of these principles is enforced on behalf of individuals, or "data subjects," is the presence of privacy protection commissions and boards in the various European nations. As in the case of Canada, which has also set up privacy oversight agencies with a Privacy Commissioner in each of its provinces, every member state of the EU is required to institute a Data Protection Authority (DPA). (This "authority" can consist of a board, a commission, or an individual commissioner.) DPAs are empowered to check to see that all of the laws are being followed in the processing of personal data, and they can impose very severe sanctions when personal data is processed illegally. In Europe, willful data protection breaches may also be criminal offenses, and can even rise to the level of felonies in certain circumstances.

A recent challenge for the EU Directive on Data Protection—and one that has international implications because of the flow of personal information across the porous boundaries of cyberspace—has involved the question of whether users should have a right to have certain kinds of personal information about them deleted, or at least "delinked" from search engine indexes. This right would apply mainly to personal information in digital form that is shown to be either inaccurate or no longer "relevant."

▶ 5.8 A RIGHT TO "BE FORGOTTEN" (OR TO "ERASURE") IN THE DIGITAL AGE

In our discussion of privacy issues affecting online public records in Section 5.6.3, we saw that a record about an unfortunate incident in one's past can now live on indefinitely—that is, once it has been converted into digital form and identified with a digital link or universal resource locator (URL). Is this necessarily a bad thing? One might argue that our being able to access information about a person's past convictions for crimes such as child molestation or pedophilia is very important; for example, a community's residents would be able to view information concerning past criminal records of prospective home buyers wishing to move into their neighborhood. But do we always need access to an online public record about someone's past to accomplish this specific objective? In the United States, and possibly in other countries well, some explicit laws are already in place requiring that the names of past offenders of various kinds of child- and sex-related crimes be included on a "list" or index, and also requiring these

people register with police departments in communities where they wish to live. So, in these instances, individuals who have been convicted of certain kinds of crimes are required to self-report, even though information about their past may also be readily available via online public records as well. A more interesting challenge, however, can arise in the case of access to online records about a person's past arrest for a less serious offense, such as underage drinking. This kind of situation is illustrated in the following scenario.

▶ **SCENARIO 5–8:** An Arrest for an Underage Drinking Incident 20 Years Ago

Philip Clark is a 39-year-old resident of Amityville, where he lives with his wife and two school-age children. He is a respectable member of his community, active in his local church as well as in several civic organizations. Philip is employed by the DEF Corporation in Amityville, where he has worked for several years and is a highly valued employee. However, when Philip attempted to change jobs a few years ago, he was unsuccessful. He strongly suspects that this may be due to an online document about an incident in Philip's past that shows up whenever someone searches his name: a newspaper story describing Philip Clark's arrest (along with the arrests of two of his friends) in an underage drinking incident that occurred 20 years ago, when Philip was a sophomore in college. Philip had pleaded guilty to the charge and received a reduced sentence of 30 hours of work-related service in his community; the judge presiding over Philip's case informed Philip that because it was his first offense, the conviction would not be included in his permanent record. So, Philip believed that the incident was behind him and that he would not have to worry about any official public record affecting his future.

Since the time of his arrest as a teenager, Philip has not violated any laws; in fact, he is viewed by many in his community as a "model citizen" because of his volunteer work with youth groups and his various contributions to neighborhood initiatives. Yet Philip continues to be haunted by the unfortunate incident in his past because of the online link to the (20-year-old) newspaper story about his arrest, which cannot be expunged in the same way that a public record can. And since a prospective employer who searches for "Philip Clark" will almost certainly discover the link to the newspaper article describing Philip's arrest, which is featured prominently in the list of search returns, Philip believes that his future employment prospects are not very promising. In fact, during the past few years, Philip has been turned down by a number of prospective employers who initially seemed very interested in hiring him. Unable to change jobs, and feeling locked out of potential career opportunities, Philip concludes that he will be stuck in his current employment position as long as the information about his past underage-drinking arrest continues to be available online. So Philip decides to contact Google, Bing, and other major search engine companies with requests to have their links to that 20-year-old newspaper story removed from their indexes. ■

Does, or should, Philip have a right to make this request? If not, what alternative recourse, if any, does/should Philip have to get the link to this old, and arguably now "irrelevant," information removed? While some might be sympathetic to Philip's request, others oppose any legislation that would give people a right to have any online personal information about them deleted, or even "de-linked" or deindexed, from search engines. The questions raised in Scenario 5–8 reflect some of the key issues at stake in the current debate in Europe about the Right to Be Forgotten (RTBF), sometimes also referred to as the "right to erasure." Whereas the scenario depicting Philip Clark is hypothetical, an actual case in Spain, involving Mario Costeja González, triggered a controversial debate in Europe (and elsewhere), the ramifications of which are still being sorted out.

In 2010, González, a Spanish citizen, sought to have some unflattering (and arguably no longer "relevant") information about him removed from Google's list of returns for searches of his name. Specifically, he wanted Google to delete a link to an article in a Spanish newspaper about his home foreclosure that occurred 16 years earlier. So González appealed to Spain's National Data Protection Agency to have the link to the report about his foreclosure removed, arguing that because the information was no longer relevant, it should not be prominently

featured in Google's list of returns on searches of his name. The Spanish court ruled in González's favor. While this court's decision arguably set a precedent that is favorable to Spanish citizens with respect to RTBF requests, it was not clear whether this ruling should apply in other EU countries as well.

Google Inc. challenged the Spanish court's ruling, and the European Court of Justice (ECJ), which presides over all of the EU countries, agreed to consider the case. Initially, it seemed that the European court might side with Google. However, when the ECJ formally considered the case in May 2014, it upheld the Spanish court's ruling.[37] Two important qualifications affecting the ECJ's ruling on RTBF are worth mentioning: (i) the right is not absolute (but instead needs to be balanced against other rights, including freedom of expression) and (ii) the right does not apply in the same way to "public figures" in Europe, including politicians and celebrities. Google (in Europe) agreed to comply with the ECJ's decision, which affected only its European users (e.g., those using a service like Google.co.uk in England, but not to the Google.com users outside Europe). However, many of those outside (as well as inside) Europe have been critical of the ECJ's ruling. We briefly examine some of their arguments.

5.8.1 Arguments Opposing RTBF

Major search engine companies and journalists/publishers have been among RTBF's staunchest opponents. Search engine companies generally make two different kinds of claims, arguing that they:

a. Do not control *content* on the Internet (and thus cannot be held responsible for the relevance, or even the accuracy, of the content on sites to which they provide links)

b. Cannot be expected to respond to all of the links requested by users (even if the information being linked to is either inaccurate or no longer relevant, because doing so would be too *impractical*, if not impossible)

Regarding (a), search engine companies tend to view themselves as "services" that provide links to online content, and not as "content providers." In the United States, search engine companies are not held legally liable for the content to which they link, as long as they comply with official legal requests to remove links to sites whose content explicitly violates the law—for example, sites that willingly and intentionally violate U.S. laws involving copyright, child pornography, and so forth. So if search engine companies make a good faith effort to remove those links, they are immune from legal liability (in accordance with Section 512 of the Digital Millennium Copyright Act). In Europe, however, search engine companies are viewed as "controllers of personal data" that are responsible/liable for the content that is accessible through their services. So, American search engine companies, as well as all non-European companies, operating in Europe are required to comply with RTBF.

According to (b), Google and other major search engine companies have argued that it would be extremely difficult, as well as very time consuming, for them to have to respond to every RTBF-like request made by users. For example, some critics have noted that in the time period between the ECJ's ruling on RTBF, in May 2014 and April 2015, Google received more than 244,000 requests for delinking.[38] So, these critics might also argue that Google's obligation to sort through these requests would not only be a daunting task but that being required to respond meaningfully to all of those requests in a timely manner would seem virtually impossible. However, we can ask whether these factors in themselves would be sufficient for someone or some company not to comply with a law. As Bottis (2014) points out, we do not cease to enact and comply with laws simply because their enforcement could not possibly eliminate certain crimes. She notes, for example, that even though it has not been possible to eliminate crimes like prostitution, drug dealing, and so forth, we do not "de-legislate" those

crimes. Bottis further points out that in the digital world, protecting privacy and copyright has sometimes seemed impossible to achieve, but we still propose, enact, and enforce laws to protect both as best we can.[39]

Those who defend (b) might seem to have stronger case when Google's situation is examined from the vantage point of RTBF requests it receives to remove personal information solely on grounds that the information is embarrassing (e.g., one of the arguments made in the Mario González case). However, one could respond to this objection by noting that there is a critical difference between a company being required to comply with requests to delete or delink to personal information that is merely "unflattering" or "embarrassing" versus requests to delink to personal information that is either inaccurate or no longer relevant. So, perhaps a different set of standards could apply in the case of requests to delete/delink from personal information of the latter type, as opposed to the former.

We next examine two kinds of arguments typically used by journalists and publishers against RTBF. Essentially, these groups believe that being required to comply with RTBF is:

c. Tantamount to "Internet censorship" (because it violates "freedom of expression")

d. Harmful to the general public (because it interferes with a citizen's "right to know")

Regarding (c), many critics believe that requiring publishers to delete some kinds of online personal information (but not other kinds), or requiring search engine companies to remove links to that information, is a step toward censoring the Internet. Some American journalists and publishers also worry that RTBF, and principles like it, violate the First Amendment of the U.S. Constitution. For example, they argue that the RTBF principle interferes with the free flow of information (as well as freedom of expression), which is essential to their news reporting and journalistic investigations. But RFTB's supporters counter by claiming that "the spirit" of this privacy principle is "to empower individuals to manage their personal data" while also "explicitly protecting the freedom of expression and of the media."[40]

With respect to (d), many journalists and publishers also argue that RTBF threatens the public's right to access information and thus their right to know. So, these critics believe that the general public is harmed by RTBF and principles like it. Some of these critics also suggest that since RTBF would contribute to making the Internet less robust and would "degrade" its (overall) quality, because of the deleted online information and/or the removal of links to it.

5.8.2 Arguments Defending RTBF

In making their case for RTBF, many European supporters begin by pointing out that personal privacy is a human right (as stated in the United Nation's Declaration of Human Rights). Some of these supporters also see RTBF as a subset of a (principle in the) EU Data Protection Directive that already exists—namely, Article 12, which enables a user to request the deletion of personal data that is "no longer necessary."[41] But some also believe that Article 12 needs to be "updated and clarified for the digital age," and that this would include an explicit right to delete online personal information that is no longer *relevant*, as well as online personal information that is "inadequate" or "excessive."[42] Arguments supporting RTBF generally fall into two broad categories, claiming that this privacy principle is needed to:

i. Prevent innocent people from being harmed

ii. Protect people whose personal identity evolves over time

Regarding (i), supporters argue that without a principle like RTBF, many people are at "risk" and thus vulnerable to "harm" in a variety of ways. For example, Bottis (2014) notes that people can easily be "defamed, humiliated, and degraded" by the kinds of inappropriate personal information about them that is readily accessible on the Internet. Consider some relatively

recent incidents involving "revenge porn" sites, where victims have unfairly suffered significant psychological harm.[43] Arguably, these victims should have the right to have this kind of inappropriate information about them deleted, or at least delinked.

In addition to the kinds of psychological harm caused to these victims, however, Bottis notes that without RTBF, some people may even be at risk of being put in serious physical danger. To illustrate this claim, she points to an actual case involving a rape victim from the past whose name was later revealed to the public via a newspaper article in the *Florida Star*. (While this newspaper had legally acquired the name of the victim via a court record, its decision to print the name violated a Florida state statute.) In this incident, the past victim, whose real identity had been exposed, was "*threatened again* with rape, forced to move and change jobs [and suffered] from deep distress and public humiliation."[44] This particular incident involved personal information that had been published in a physical newspaper. But many newspaper articles have since been converted to digital form and made available online, which means that these articles are now discoverable by, and accessible to, a much wider audience. So, the rape victim whose name was disclosed by the *Florida Star* (several years ago) could now potentially be at an even greater risk regarding her physical safety. However, with an RFTB-like privacy principle in place, that victim would have some legal recourse in being able to request the removal of any links to the online version of that newspaper article (and possibly to having the online version of the article deleted altogether).

With respect to (ii), some RTBF supporters worry about a person's ability to protect his or her *personal identity* and *autonomy*—for example, to develop one's identity in an autonomous way in a digital world. Earlier (in Section 5.3), we noted some connections between autonomy and privacy and showed how the latter can be essential for the former. Some also believe that privacy is essential for one's personal identity. For example, Floridi (2014) has argued that a society in which "no informational privacy is possible . . . is one in which no personal identity can be maintained."[45] So, some RTBF advocates argue that people who are continually "stigmatized" by the presence of online personal information about their distant past would not have the level of privacy that is essential to protect their personal identities. These advocates believe that because the Internet "never forgets," a principle like RTBF is needed to protect those whose personal identities may evolve over time.

RTBF supporters further argue that since one's personal identity can evolve significantly over one's lifetime, certain kinds of information about a person's characteristics in the distant past may no longer be relevant or appropriate. For example, Bottis notes that some kinds of information revealed about one's past religious beliefs, sexual orientation, and so forth, may no longer reflect that person's present identity. Consider the case of a person who had a sex change 40 years ago. Does information about that incident need to be available online? Unless the person who had the sex change is a public figure, in which case he or she would not be protected by RTBF anyway, information about that person's past sexual identity could be viewed as neither relevant nor something that the general public has a need to know.

We have briefly examined some arguments for and against RTBF. While this privacy principle may never extend beyond the EU countries, RTBF is nevertheless the law in Europe; so, any search engine company, or any other kind of company, doing business in Europe must comply with it. But a difficult challenge facing those companies is determining the appropriate criteria for whether or not the digital links to a specific incident in one's past warrant removal.

5.8.3 Establishing "Appropriate" Criteria

Although the ECJ ruled in favor of RTBF in May 2014, it did not provide precise criteria for search engine companies to comply with the new privacy principle. Google has since established an advisory council to come up with appropriate criteria. It would seem that at least two important factors need to be taken into consideration: (i) the nature of the *personal information*

itself and (ii) the *context(s)* in which this information flows. With regard to (i), Floridi raises an interesting point by asking: "Is the information this person would like to see de-linked, or even perhaps removed, *constitutive* of that person . . . [o]r is it something completely irrelevant?"[46] Initially, at least, Floridi's distinction would seem to provide a very helpful criterion for search engine companies and online content providers to consider. Recall, for example, the hypothetical incident involving "Philip Clark" (in Scenario 5–8). Is the information about Philip's arrest 20 years ago for underage drinking "constitutive" of Philip's identity? How relevant is that information about Philip? Based simply on what has been disclosed about him in Scenario 5–8, the information would not seem very relevant at all. So, Floridi's (constitutive) criterion would seem to work well in this scenario.

Suppose, however, we were to alter that scenario slightly, that is, in such a way that Philip is considering running for a political office in Amityville (e.g., as a city alderman). Arguably, that information about Philip's past would now seem more *relevant*, even if it is nor more *constitutive* of Philip's identity than it was before. It is one thing if Philip, as a private citizen of Amityville, wishes to have a link to a story about a 20-year-old underage drinking arrest removed because it causes problems for him in trying to change jobs. But if Philip decides to run for public office, even at a very low level, it is reasonable to argue that this information about his past arrest may indeed be relevant. So, it would seem that in this case, the question of whether such-and-such personal information is constitutive of Philip's identity does not play a key role in our decision. (Of course, once someone becomes a "public figure," the RTBF principle no longer applies to that person in the same way that it does to an ordinary citizen.)

While there may indeed be many clear-cut cases where Floridi's "constitutive" criterion can be applied fairly easily, borderline cases will also likely arise where its application might be less effective. Consider again the case of Philip Clark, but this time suppose Philip's arrest had been made when he was 21 years of age (i.e., when he was legally an adult in the U.S.) and that it involved driving while intoxicated (DWI) instead of underage drinking in a home or dorm room. Information about Philip in the DWI incident may be no more constitutive of Philip's identity than the information about his underage drinking. Yet, there may be compelling reasons not to delete the former information about Philip, even if there is agreement that information about his underage drinking arrest does warrant removal.

We now turn to (ii)—namely, criteria affecting the *context* in which ones online personal information flows. Recall Helen Nissenbaum's framework of "privacy as contextual integrity," which we examined in Section 5.4. There we saw that Nissenbaum's privacy framework requires that the processes used in disseminating personal information must not only be "appropriate" to a particular context but must also comply with "norms that govern the flow of personal information" within that context. We saw that in Nissenbaum's framework, "norms of distribution" can restrict or limit the flow of information both within and across various contexts. We also saw that when that norm is "breached," a violation of privacy occurs. So, in Nissenbaum's scheme, it is not necessarily the information itself—for example, whether it happens to be constitutive of one's identity—that is germane; rather it is the context and the "norms" that govern the flow of the personal information in that context.

We can now see why attempts at resolving the question about which *kinds* of personal information should be eligible for deletion, and which should not, may be more difficult than anticipated. But even if clear-cut criteria could, in principle, be established, questions still remain about the process involved for removing and deindexing the personal information. For example, should users contact the search engine companies that provide the links to the information (or content) or should they instead contact the publishers who make the online content available? Floridi (2014) notes that while a search engine company has "no creative power with respect to the personal information it indexes," a publisher "has both creative and controlling power over the personal information in question." So, he argues that a publisher, unlike a search engine, can "block access to personal information quite easily." In light of this

distinction, Floridi argues for a procedure whereby a user would first make a request to a publisher to remove the information in question. If that fails, then the user could next request the search engine company to delink it. If that still does not work, the user could then appeal to the national Data Protection Authority (DPA) in his or her country (in Europe). And, finally, if that does not work, the user could appeal to the ECJ.[47]

An additional RTB-related question to consider has to do with the principle's scope: how widely should it apply? As already noted, the ECJ's ruling affects only those search engines operating in Europe. (We should also note that the Google search engine in Europe displays the "removal notification" at the bottom of the search page in the case of "name searches" it has delinked.) So, people living outside Europe can still access information that has been removed (e.g., information about Mario Gonzalez) via Google.com, but not through Google.co.es (in Spain). Finally, it is worth noting that the ECJ allows Google, as well as other search engine companies operating in Europe, to assess RTBF requests on a case-by-case basis in determining which requests must be honored and which can be rejected.

▶ 5.9 CHAPTER SUMMARY

We began this chapter by examining some ways that cybertechnology has exacerbated privacy concerns introduced by earlier technologies. We then briefly examined the concept of privacy and some theories that have attempted to explain and defend the need for privacy protection. We saw that "informational privacy" could be distinguished from "accessibility privacy" and "decisional privacy," and that Moor's privacy theory was able to integrate key components of three traditional theories into one comprehensive theory of privacy. We also saw that privacy is an important value, essential for human ends such as friendship and autonomy.

Next, we saw how personal privacy is threatened by data-gathering techniques such as RFID technologies and Internet cookies and by data-analysis techniques such as those associated with Big Data. We then saw the impact that data mining technologies have for privacy, especially for many forms of "public personal information" that have no explicit normative protection. In our analysis of the problem of "protecting privacy in public," we examined ways in which contemporary search engines pose some significant challenges. We then examined the debate between those who advocate for stricter privacy laws and those who champion (industry) self-regulation standards as an alternative to additional privacy legislation. Finally, we examined the current dispute involving an individual's alleged "right to be forgotten" (or "right to erasure") in a digital world.

We also noted at the outset that not all computer-related privacy concerns could be examined in this chapter. For example, specific kinds of privacy issues pertaining to employee monitoring in the workplace are examined in Chapter 10, while surveillance concerns affecting "location privacy" made possible by pervasive computing and ambient intelligence are examined in Chapter 12. Although some privacy concerns affecting personal information collected by governmental organizations were briefly identified and considered in this chapter, additional privacy issues in this area are examined in Chapter 6 in the context of our discussion of computer/cyber security.

▶ REVIEW QUESTIONS

1. Describe four ways in which the privacy threats posed by cybertechnology differ from those posed by earlier technologies.
2. What is personal privacy and why is it difficult to define?
3. Describe some important characteristics that differentiate "accessibility privacy," "decisional privacy," and "informational privacy."
4. How does James Moor's theory of privacy combine key elements of these three views of privacy? What

does Moor mean by a "situation," and how does he distinguish between "natural privacy" and "normative privacy"?

5. Why is privacy valued? Is privacy an intrinsic value or is it an instrumental value? Explain.
6. Is privacy a social value or is it simply an individual good?
7. What does Roger Clarke mean by "dataveillance"? Why do dataveillance techniques threaten personal privacy?
8. What are Internet cookies and why are they considered controversial from the perspective of personal privacy?
9. What is RFID technology and why is it a threat to privacy?
10. Describe some surveillance techniques that the U.S. government has used to collect data on its citizens. Why are they considered controversial?
11. What is meant by "Big Data"? Why is this notion difficult to define?
12. What is data mining and why is it considered controversial?
13. What is Web mining and how is it similar to and different from traditional data mining?
14. What is the difference between public personal information (PPI) and nonpublic personal information (NPI)?
15. What is meant by "privacy in public"? Describe the problem of protecting personal privacy in public space.
16. Why are certain uses of Internet search engines problematic from a privacy perspective?
17. Describe some of the voluntary controls and self-regulation initiatives that have been proposed by representatives from industry and e-commerce.
18. Why do many privacy advocates in the U.S. believe that industry self-regulation and voluntary controls are not adequate and that stronger privacy legislation is needed?
19. What are some of the criticisms of U.S. privacy laws such as HIPAA and the Privacy Act of 1974?
20. Describe some principles included in the EU Directive on Data Protection. What do you believe to be some of the strengths and weaknesses of those principles when compared to privacy laws in the United States?
21. What is the meant by the "Right to Be Forgotten"? Why is this "right" so controversial?

▶ DISCUSSION QUESTIONS

22. Review Helen Nissenbaum's framework of privacy in terms of "contextual integrity." What are the differences between what she calls "norms of appropriateness" and "norms of distribution"? Give an example of how either or both norms can be breached in a specific context.
23. Through the use of currently available online tools and search facilities, ordinary users can easily acquire personal information about others. In fact, anyone who has Internet access can, via a search engine such as Google, find information about us that we ourselves might have had no idea is publicly available there. Does this use of search engines threaten the privacy of ordinary people? Explain.
24. In debates regarding access and control of personal information, it is sometimes argued that an appropriate balance needs to be struck between individuals and organizations: individuals claim that they should be able to control who has access to their information and organizations, including government and business groups, claim to need that information in order to make appropriate decisions. How can a reasonable resolution be reached that would satisfy both parties?
25. Reexamine the arguments made by the U.S. government and by Google regarding the government's requests for information about users' search requests made during the summer of 2005. Are the government's reasons for why it should have access to that information reasonable? Does Google have an obligation to protect the personal information of its users, with respect to disclosing information about their searches? Could this obligation be overridden by certain kinds of national defense interests? If, for example, the government claimed to need the information to prevent a potential terrorist attack, would that have changed your analysis of the situation? Or does the government have the right, and possibly an obligation to the majority of its citizens, to monitor the searches if doing so could positively affect the outcome of child pornography legislation?
26. Initially, privacy concerns involving computer technology arose because citizens feared that a strong centralized government could easily collect and store data about them. In the 1960s, for example, there was talk of constructing a national computerized database in the United States, and many were concerned that George Orwell's prediction of Big Brother in his classic book 1984 had finally arrived. The centralized database, however, never materialized. Prior to September 11, 2001, some privacy advocates suggested that we have fewer reasons to be concerned about the federal government's role in privacy intrusions (Big Brother) than we do about privacy threats from the commercial sector (Big Bucks and Big Browser). Is that assessment still accurate? Defend your answer.

Scenarios for Analysis

1. In the days and weeks immediately following the tragic events of September 11, 2001, some political leaders in the United States argued that extraordinary times call for extraordinary measures; in times of war, basic civil liberties and freedoms, such as privacy, need to be severely restricted for the sake of national security and safety. Initially, the majority of American citizens strongly supported the Patriot Act, which passed by an overwhelming margin in both houses of Congress and was enacted into law on October 21, 2001. However, between 2001 and 2005, support for this act diminished considerably. Many privacy advocates believe that it goes too far and thus erodes basic civil liberties. Some critics also fear that certain provisions included in the act could easily be abused. Examine some of the details of the Patriot Act (which can be viewed on the Web at www.govtrack.us/congress/bills/107/hr3162/text) and determine whether its measures are as extreme as its critics suggest. Are those measures also consistent with the value of privacy, which many Americans claim to embrace? Do privacy interests need to be reassessed, and possibly recalibrated, in light of ongoing threats from terrorists? To what extent does the following expression, attributed to Benjamin Franklin, affect your answer to this question: "They who can give up essential liberty to obtain a little temporary safety, deserve neither liberty nor safety."
2. At the beginning of this chapter, we suggested that concerns about the loss of privacy may have a generational dimension or element—that is, younger people may be less concerned about privacy loss involving cybertechnology than older people. To further explore this possibility, conduct a series of informal interviews with individuals that represent three generations: Millennials, Gen X/Y, and Baby Boomers. Ask members of each group how much they value their privacy and how much of it they are willing to trade off for the convenience of cybertechnology. Compare the results of the answers you get from the three groups. Are their respective views about the importance of privacy as far apart as some might expect? Explain.

▶ ENDNOTES

1. See, for example, Berkes (2014). Berkes notes that a zettabyte of data is roughly equivalent to the amount of data that would "fill 250 billion DVDs" and that it is estimated that five zettabytes of data could store approximately "100 years worth of [all] worldwide communications." For additional information about the controversy surrounding the opening of this new NSA data facility, see http://www.npr.org/2013/06/10/190160772/amid-data-controversy-nsa-builds-its-biggest-data-farm. See also the description of the new NSA facility in Bamford (2012).
2. See, for example, Solove (2011).
3. See the interview with Arthur Miller in the video, "The World at Your Fingertips," in the BBC/PBS Series, *The Machine that Changed the World*, 1990.
4. See Warren and Brandeis (1890) for more detail.
5. For a discussion of the right to privacy in the Quinlan case, see "Court at the End of Life—The Right to Privacy: Karen Ann Quinlan" at http://www.libraryindex.com/pages/582/Court-End-Life-RIGHT-PRIVACY-KAREN-ANN-QUINLAN.html.
6. Moor (2000, p. 207). [Italics added]
7. Key aspects of Moor's privacy framework are more fully developed in Tavani and Moor (2001).
8. Nissenbaum (2004a, p. 137).
9. *Ibid*, p. 135. For examples of ways in which Nissenbaum's contextual integrity model of privacy can be applied to the blogosphere and to "the Cloud," see Grodzinsky and Tavani (2010, 2011), respectively.
10. See Westin (1967) for more detail on this point.
11. Solove (2008, p. 84).
12. DeCew (2006, p. 121). Moor (2006, p. 114) also describes privacy as a kind of "shield" that protects us.
13. Solove, p. 91.
14. See Clarke's account of dataveillance, available at http://www.rogerclarke.com/DV/.
15. See Halpern (2011) for more detail.
16. Nissenbaum, p. 135.
17. Ward and Barker (2013, p. 1). Available at Ar Xiv:1309.5821lvl [cs.DB]. In their view, this lack of consistency both "introduces ambiguity" and "hampers discourse" about issues affecting big data.
18. *Ibid*, p. 2. [Italics Ward and Barker] Although their definition also includes the "collection," as well the analysis of data, we focus solely on the analysis aspect of big data in this section.
19. Boyd and Crawford (2012, p. 663). [Italics added]
20. Two of these categories—variety and velocity (along with "volume")—were first articulated in the now classic Gartner Report (2001).
21. Poskanzer (2015, p. 210).
22. Vedder (2004) refers to the kind of privacy protection needed for groups as "Categorial Privacy."
23. In composing this section on data mining, I have drawn from and expanded upon some concepts and distinctions introduced in Tavani (1999, 2007).
24. See "Web Mining." In *Wikipedia*. Available at http://en.wikipedia.org/wiki/Web_mining.
25. See, for example, the account of Facebook Beacon in http://en.wikipedia.org/wiki/Facebook.
26. Nissenbaum (2004b, p. 455).

27. Fulda (2004, p. 472).
28. See, for example, Nissenbaum (2010).
29. For an extended discussion of privacy issues generated by search engine technology, see Tavani (2012).
30. Critics point out that tools like *Anonymizer* are not effective in e-commerce contexts; they also note that even in non-e-commerce contexts, users' online activities can still be tracked via their IP addresses.
31. For more information about the Toysmart case, see Morehead (2000), who notes that Toysmart made the mistake of separating its customer list as a "separate asset" instead of grouping it together with other aspects of its "corporate package."
32. See Werner (2012) for a more detailed analysis of this controversy.
33. See Julian Sanchez (2013) interviewed in *Online Privacy: How Did We Get Here?* PBS/Digital Studios. Available at http://www.theatlantic.com/video/archive/2013/07/what-we-talk-about-when-we-talk-about-privacy/278134/.
34. See http://www.oecd.org/document/18/0,3343,en_2649_34255_1815186_1_1_1_1,00.html.
35. See http://www.cdt.org/privacy/eudirective/EU_Directive_.html.
36. For more detail about the various principles and how each works, see http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML.
37. According to the revisions included in Article 17 of the EU Directive, individuals have the right—under certain conditions—to ask search engines to remove links with personal information about them.
38. See "A New Ethics Case Study." Available at http://www.scu.edu/ethics-center/ethicsblog/internet-ethics.cfm?c=22135&utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+EthicalIssuesInTheOnlineWorld+%28Ethical+Issues+in+the+Online+World%29. See also Herritt (2014).
39. See Bottis (2014). I am grateful to Professor Bottis for some additional points she raised in a series of e-mail exchanges with me on the topic of RTBF, which are included here.
40. See the *Factsheet on the "Right to Be Forgotten" Ruling C131/12*. Available at http://ec.europa.eu/justice/data-protection/files/factsheets/factsheet_data_protection_en.pdf.
41. Article 12 (Right of Access) states "the rectification, erasure, or blocking of data the processing of which does not comply with the provisions of the Directive because the information is incomplete or inaccurate and a notification to whom the data has been disclosed of any rectification, erasure, or blocking . . . unless this proves impossible or involves disproportionate effort."
42. These criteria are in included in Article 17 ("Right to be forgotten and to erasure") of the EU Directive.
43. Of course, the topic of "revenge porn" is one that warrants separate discussion because of the cluster of ethical issues it raises. But it is mentioned here because it makes available one kind of online personal information that an injured party should clearly be able to have deleted.
44. Bottis (2014, p. 3). [Italics Bottis]
45. See the interview with Floridi in Herritt (2014, p. 2).
46. *Ibid*, p. 4. [Italics added]
47. *Ibid*. This procedure would follow one that is already in place for requesting the deletion of "unnecessary personal information," in accordance with Article 12 of the EU Privacy Directive.

▶ REFERENCES

Adam, Alison. 2005. "Chips in Our Children: Can We Inscribe Privacy in a Machine?" *Ethics and Information Technology* 7, no. 4: 233–42.

Anguin, Julia and Jennifer Valentino-DeVries. 2012. "Google's iPhone Tracking: Web Giant, Others Bypassed Apple Browser Settings for Guarding Privacy." *Wall Street Journal*, February 17. Available at http://online.wsj.com/article_email/SB10001424052970204880404577225380456599176-lMyQjAxMTAyMDEwNjExNDYyWj.html

Bamford, James. 2012. "The NSA Is Building the Country's Biggest Spy Center (Watch What You Say)." *Wired*, March 15. Available at http://m.wired.com/threatlevel/2012/03/ff_nsadatacenter/all/1.

Berkes, Howard. 2014. "Amid Data Controversy NSA Builds Its Biggest Data Farm." Available at http://www.npr.org/2013/06/10/190160772/amid-data-controversy-nsa-builds-its-biggest-data-farm.

Bottis, Maria. 2014. "Allow Me to Live the Good Life, Let me Forget: Legal and Psychological Foundations of the Right to Be Forgotten and the New Developments in the European Union Laws." In *Well-Being, Flourishing, and ICTs: Proceedings of the Eleventh International Conference on Computer Ethics–Philosophical Enquiry*. Menomonie, WI: INSEIT, Article 10.

Boyd, Danah and Kate Crawford. 2012. "Critical Questions for Big Data: Provocations for a Cultural, Scholarly, and Technological Phenomenon." *Information, Communication, and Society* 15, no. 5: 62–79.

DeCew, Judith W. 1997. *In Pursuit of Privacy: Law, Ethics, and the Rise of Technology*. Ithaca, NY: Cornell University Press.

DeCew, Judith W. 2006. "Privacy and Policy for Genetic Research." In H. T. Tavani, ed. *Ethics, Computing, and Genomics*. Sudbury, MA: Jones and Bartlett, pp. 121–35.

Elgesem, Dag. 2004. "The Structure of Rights in Directive 95/46/EC on the Protection of Individuals with Regard to the Processing of Personal Data and the Free Movement of Such Data." In R. A. Spinello and H. T. Tavani, eds. *Readings in CyberEthics*. 2nd ed. Sudbury, MA: Jones and Bartlett Publishers, pp. 418–35.

Floridi, Luciano. 2014. "The Right to Be Forgotten – The Road Ahead." *The Guardian*, October 8. Available at http://www.theguardian.com/technology/2014/oct/08/the-right-to-be-forgotten-the-road-ahead.

Fried, Charles. 1990. "Privacy: A Rational Context." In M. D. Ermann, M. B. Williams, and C. Gutierrez, eds. *Computers, Ethics, and Society*. New York: Oxford University Press, pp. 51–67.

Froomkin, Michael. 2000. "The Death of Privacy?" *Stanford Law Review* 52. Available at www.law.miami.edu/~froomkin/articles/privacy-deathof.pdf.

Fulda, Joseph S. 2004. "Data Mining and Privacy." In R. A. Spinello and H. Tavani, eds. *Readings in CyberEthics*. 2nd ed. Sudbury, MA: Jones and Bartlett Publishers, pp. 471–5.

Garfinkel, Simson. 2000. *Database Nation: The Death of Privacy in the 21st Century*. Cambridge, MA: O'Reilly and Associates.

Garfinkel, Simson. 2002. "RFID Bill of Rights." *Technology Review*, October. Available at http://www.technologyreview.com/article/401660/an-rfid-bill-of-rights/.

Grodzinsky, Frances S. and Herman T. Tavani. 2010. "Applying the 'Contextual Integrity' Model of Privacy to Personal Blogs in the Blogosphere," *International Journal of Internet Research Ethics* 3, no. 1, pp. 38–47.

Grodzinsky, Francis, S. and Herman T. Tavani. 2011. "Privacy in 'the Cloud': Applying Nissenbaum's Theory of Contextual Integrity." *Computers and Society* 41, no. 1: 38–47.

Halpern, Sue. 2011. "Mind Control and the Internet." *New York Review of Books*, June 23. Available at http://www.nybooks.com/articles/archives/2011/jun/23/mind-control-and-internet/.

Herritt, Robert. 2014. "Google's Philosopher." *Nature and Technology*, December 30. Available at http://www.psmag.com/navigation/nature-and-technology/googles-philosopher-technology-nature-identity-court-legal-policy-95456/.

Lockton, Vance and Richard S. Rosenberg. 2005. "RFID: The Next Serious Threat to Privacy." *Ethics and Information Technology* 7, no. 4: 221–31.

Moor, James H. 2000. "Towards a Theory of Privacy for the Information Age." In R. M. Baird, R. Ramsower, and S. E. Rosenbaum, eds. *Cyberethics: Moral, Social, and Legal Issues in the Computer Age*. Amherst, NY: Prometheus Books, pp. 200–12.

Moor, James H. 2004. "Reason, Relativity, and Responsibility in Computer Ethics." In R. A. Spinello and H. T. Tavani, eds. *Readings in CyberEthics*. 2nd ed. Sudbury, MA: Jones and Bartlett Publishers, pp. 40–54.

Moor, James H. 2006. "Using Genetic Information While Protecting the Privacy of the Soul." In H. T. Tavani, ed. *Ethics, Computing, and Genomics*. Sudbury, MA: Jones and Bartlett, pp. 109–19.

Morehead, Nicholas. 2000. "Toysmart: Bankruptcy Litmus Test." *Wired* 7, no. 12. Available at http://archive.wired.com/techbiz/media/news/2000/07/37517.

Nissenbaum, Helen. 2004a. "Privacy as Contextual Integrity." *Washington Law Review* 79, no. 1: 119–57.

Nissenbaum, Helen. 2004b. "Toward an Approach to Privacy in Public: Challenges of Information Technology." In R. A. Spinello and H. T. Tavani, eds. *Readings in CyberEthics*. 2nd ed. Sudbury, MA: Jones and Bartlett, pp. 450–61. Reprinted from *Ethics and Behavior* 7, no. 3 (1997): 207–19.

Nissenbaum, Helen. 2010. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Palo Alto, CA: Stanford University Press.

Pariser, Eli. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. New York: Penguin.

Poskanzer, Deborah. 2015. "Big Data." In J. Britt Holbrook and Carl Mitcham, eds. *Ethics, Science, Technology, and Engineering: A Global Resource*. Vol. 1, 2nd ed. Farmington Hills, MI: Macmillan Reference, pp. 210–12.

Regan, Priscilla M. 1995. *Legislating Privacy: Technology, Social Values, and Public Policy*. Chapel Hill, NC: The University of North Carolina Press.

Solove, Daniel J. 2008. *Understanding Privacy*. Cambridge, MA: Harvard University Press.

Solove, Daniel J. 2011. *Nothing to Hide: The False Tradeoff between Privacy and Security*. New Haven, CT: Yale University Press.

Spinello, Richard A. 2010. "Informational Privacy." In G. Brenkert and T. Beauchamp, eds. *The Oxford Handbook of Business Ethics*. Oxford, UK: Oxford University Press, pp. 366–87.

Stuckey, Mike. 2008. "Amex Rates Credit Risk by Where You Live, Shop." *MSNBC.Com*. Available at http://www.msnbc.msn.com/id/27055285/.

Tavani, Herman T. 1999. "Informational Privacy, Data Mining and the Internet." *Ethics and Information Technology* 1, no. 2: 137–45.

Tavani, Herman T. 2007. "Philosophical Theories of Privacy: Implications for an Adequate Online Privacy Policy." *Metaphilosophy* 38, no. 1, pp. 1–22.

Tavani, Herman T. 2012. "Search Engines and Ethics." *Stanford Encyclopedia of Philosophy*. Available at http://plato.stanford.edu/entries/ethics-search/.

Tavani, Herman T. and James H. Moor. 2001. "Privacy Protection, Control over Information, and Privacy-Enhancing Technologies." *Computers and Society* 31, no. 1: 6–11.

Vedder, Anton. 2004. "KDD, Privacy, Individuality, and Fairness." In R. A. Spinello and H. T. Tavani, eds. *Readings in CyberEthics*. 2nd ed. Sudbury, MA: Jones and Bartlett, pp. 462–70.

Ward, Jonathan Stuart and Adam Barker. 2013. "Undefined by Data: A Survey of Big Data Definitions." Cornell University Library. Available at http://arxiv.org/abs/1309.5821.

Warren, Samuel and Louis Brandeis. 1890. "The Right to Privacy." *Harvard Law Review* 4, no. 5: 193–220.

Werner, Jeff. 2012. "Should You Be Worried about Google's New Privacy Policy?" *NWFDailyNews.com*, March 25. Available at http://www.nwfdailynews.com/articles/google-48355-new-policy.html.

Westin, Alan. 1967. *Privacy and Freedom*. New York: Atheneum Press.

Zimmer, Michael 2008. "The Gaze of the Perfect Search Engine: Google as an Institution of Dataveillance." In A. Spink and M. Zimmer, eds. *Web Search: Multidisciplinary Perspectives*. Berlin: Springer-Verlag, pp. 77–99.

► FURTHER READINGS

Alfino, Mark. 2001. "Misplacing Privacy." *Journal of Information Ethics* 10, no. 1: 5–8.

Floridi, Luciano. 2014. "Right to Be Forgotten: Who May Exercise Power, over which kind of Information." *The Guardian*. Available at http://www.theguardian.com/technology/2014/oct/21/right-to-be-forgotten-who-may-exercise-power-information?CMP=twt_gu. October 22.

Lyon, David. 2014. "Surveillance, Snowden, and Big Data: Capacities, Consequences, Critique." Big Data and Society 1:1–13.

Schneier, Bruce. 2015. *Data and Goliath: The Hidden Battles to Collect your Data and Control Your World*. New York: W. W. Norton.

Shoemaker, David W. 2010. "Self Exposure and Exposure of the Self: Informational Privacy and the Presentation of Identity." *Ethics and Information Technology* 12, no. 1: 3–15.

Spinello, Richard A. 2011. "Privacy and Social Networking." *International Review of Information Ethics* 16: 42–6.

Zimmer, Michael. 2005. "Surveillance, Privacy, and the Ethics of Vehicle Safety Communication Technologies." *Ethics and Information Technology* 7, no. 4: 201–10.

CHAPTER

6

Security in Cyberspace

LEARNING OBJECTIVES

Upon completing this chapter, you will successfully be able to:

- Articulate what is meant by *security* in the context of cybertechnology and differentiate issues in cybersecurity from both *cyberprivacy*-related issues and *cybercrime*-related issues,
- Distinguish among three distinct categories of security affecting cybertechnology: *data* security, *system* security, and *network* security,
- Describe key challenges that *cloud computing* poses for cybersecurity,
- Explain what is meant by the terms *hacking*, *Hacker Ethic*, and *hacktivism,*
- Describe the parameters of *cyberterrorism* and show how it can be distinguished both from hacktivism and information warfare,
- Explain what is meant by *information warfare* and show how it is both similar to and different from cyberterrorism.

In this chapter, we examine a wide range of issues affecting cybersecurity. Among them is the question whether cyber intrusions can ever be justified on ethical grounds? For example, would it ever be morally permissible for governmental organizations in (sovereign) nation states to engage in cyberattacks and computer break-ins? The following scenario, illustrating an alleged intrusion involving three nations, briefly addresses that question.

▶ **SCENARIO 6–1:** The "Olympic Games" Operation and the Stuxnet Worm

In June 2012, the *New York Times* reported that the United States and Israeli governments had been cooperating on an initiative code-named *Olympic Games*. Originally conceived and developed during the George W. Bush administration, the Olympic Games operation aimed at disrupting Iran's uranium enrichment program and thus damaging that nation's nuclear capability. At the core of this joint operation was a computer worm known as Stuxnet, a "cyberweapon" that targeted "electronic program controllers" developed by Siemens Corporation (in Germany) for industrial controlled computers (ICCs) that were installed in Iran. The Stuxnet worm was allegedly responsible for (i) sending misleading data to computer monitors in Iran and (ii) causing several of that nation's centrifuges—that is, fast-spinning machines that enrich uranium—to spin out of control. The Stuxnet attack was estimated to have destroyed approximately 1,000 of Iran's (then) 6,000 centrifuges.[1] ■

Was the Olympic Games operation a justified breach of cybersecurity? If it is wrong for ordinary individuals and nongovernmental actors/organizations to break into and disrupt someone's computer system, is it also wrong for sovereign nation states to do this as well? Or,