Engineering Privacy

Sarah Spiekermann and Lorrie Faith Cranor, *Senior Member*, *IEEE*

Abstract—In this paper, we integrate insights from diverse islands of research on electronic privacy to offer a holistic view of privacy engineering and a systematic structure for the discipline's topics. First, we discuss privacy requirements grounded in both historic and contemporary perspectives on privacy. We use a three-layer model of user privacy concerns to relate them to system operations (data transfer, storage, and processing) and examine their effects on user behavior. In the second part of this paper, we develop guidelines for building privacy-friendly systems. We distinguish two approaches: "privacy-by-policy" and "privacy-by-architecture." The privacy-by-policy approach focuses on the implementation of the notice and choice principles of fair information practices, while the privacy-by-architecture approach minimizes the collection of identifiable personal data and emphasizes anonymization and client-side data storage and processing. We discuss both approaches with a view to their technical overlaps and boundaries as well as to economic feasibility. This paper aims to introduce engineers and computer scientists to the privacy research domain and provide concrete guidance on how to design privacy-friendly systems.

Index Terms—Privacy, security, privacy-enhancing technologies, anonymity, identification.

1 Introduction

While privacy has long been heralded as a dead issue by some [1], [2], it is viewed as a key business requirement by others [3], [4]. New regulatory requirements and consumer concerns are driving companies to consider more privacy-friendly policies, but such policies often conflict with the desire to leverage customer data. The widespread adoption of loyalty card schemes and the rise of social network platforms suggest that some consumers are willing to sacrifice privacy for benefits they value. At the same time, perceived privacy breaches often result in consumer outcries. For example, the social networking website Facebook has repeatedly sparked protest from its users by introducing new services with privacy-invasive features turned on by default [5]. Negative news on privacy issues impact stock market valuation [6] and companies are confronted with expensive fines or settlements for privacy breaches [7], [8]. As a result, companies are increasingly unsure how critical customer privacy really is to their operations and sustainable market success.

Surveys suggest that individuals are deeply concerned about privacy. An increasing majority of US citizens say that existing laws and organizational practices do not provide a reasonable level of consumer privacy protection and that companies share personal information inappropriately [7], [9]. Even in Germany, which has the highest legal data protection standards worldwide [10], 47 percent of people do not believe their personal data is adequately protected [11].

- *S. Spiekermann is with the Institute of Information Systems, Humboldt University Berlin, Spandauer Strasse 1, 10178 Berlin, Germany. E-mail: sspiek@wiwi.hu-berlin.de.*
- *L.F. Cranor is with Carnegie Mellon University, 4720 Forbes Ave., Pittsburgh, PA 15213. E-mail: lorrie@cmu.edu.*

Manuscript received 17 Jan. 2008; revised 3 Sept. 2008; accepted 16 Sept. 2008; published online 15 Oct. 2008.
Recommended for acceptance by P. McDaniel.
For information on obtaining reprints of this article, please send e-mail to: tse@computer.org, and reference IEEECS Log Number TSE-2008-01-0018.
Digital Object Identifier no. 10.1109/TSE.2008.88.

While there is evidence that consumers may not always act on their privacy concerns [12], [13], there is convincing data to suggest that these concerns have some impact on consumer behavior and the acceptability and adoption of new technologies. A 2005 survey conducted by Privacy & American Business found that concerns about the use of personal information led 64 percent of respondents to decide not to purchase something from a company [14]. In many countries, new privacy regulations as well as media attention are increasing public awareness of privacy. For example, a 2004 analysis by the European press on radio frequency identification technology (RFID) revealed that about one-third of media messages about the new technology were related to consumer privacy fears [15]. Laboratory studies have shown that, when privacy information is readily available in search results, some consumers will pay a small premium to shop at websites with good privacy policies [16]. Against this background, privacy is a highly relevant issue in systems engineering today.

Despite increasing consciousness about the need to consider privacy in technology design, engineers have barely recognized its importance. Lahlou et al. [17] found that, when engineers were asked about privacy issues as related to prototype development, the issues were viewed either as "an abstract problem, not an immediate problem, not a problem at all (firewalls and cryptography would take care of it), not their problem (one for politicians, lawmakers, or society), or simply not part of the project deliverables." Conversely, privacy-conscious engineers often strive for extremely high degrees of privacy protection that may lead to mechanisms that undermine system usability [18], [19].

In the privacy research literature, we observe two areas of work with seemingly very different goals. The first area includes research aimed at developing cryptographic privacy protections and systems with provable privacy guarantees (JAP [20], Tor [21], and work on differential privacy [22]). Researchers in this area work under a threat model that assumes sophisticated adversaries who will not be deterred by policies or regulations, or regard states and

their security agencies as potential privacy intruders. The second area includes research aimed at protecting consumer data from accidental disclosure or misuse and facilitating informed choice options [23], [24], [25]. Researchers in this area assume that policies and regulations are generally enforceable and that the role of technology is to aid enforcement, but not necessarily to guarantee it.

We aim to present a holistic view of the privacy field, situating each approach to privacy in a spectrum of system design options. We derive system requirements from accepted privacy definitions as well as from user concerns, and propose a framework that integrates existing research to provide engineers a clear roadmap for building privacy-friendly information systems. While recognizing that engineers must work within the constraints set by their employers, we believe that they hold a major responsibility for privacy engineering because they are the ones devising the technical architecture and creating the code.

Several authors have proposed privacy design frameworks for specific domains. Earp et al. [26] proposed a framework for privacy management and policies that addresses various organizational perspectives, focusing on how organizations should evaluate their own privacy policies. Hong et al. [27] proposed privacy risk models as an approach to the design of privacy-sensitive ubiquitous computing systems. Feigenbaum et al. [23] proposed privacy engineering guidelines for digital rights management systems. Our approach has similarities to these proposals, but applies to a wider variety of systems including e-commerce websites and ubiquitous computing applications.

The remainder of this paper proceeds as follows: In Section 2, we discuss frequently cited definitions of privacy and translate them into a high-level responsibility framework for privacy engineering. The framework serves as an underlying concept for the privacy requirements analysis presented in Section 3, which discusses how computing activities (data collection, storage, and processing) can lead to privacy invasion and describes the types of activities that raise privacy concerns in consumers. In Section 4, we present concrete privacy engineering practices. We begin with an overview of fair information practices (FIPs) as outlined by the Organization for Economic Co-operation and Development (OECD) and the more limited "notice and choice" approach of the US Federal Trade Commission (FTC). We then discuss architectural choices that may serve as an alternative to notice and choice. We argue that notice and choice are needed to implement "privacy-by-policy" only where "privacy-by-architecture" cannot be implemented. We present guidelines for implementing privacy-by-policy in Section 5 and present our conclusions in Section 6.

2 Framing Privacy for Engineering

An often-cited 1890 conceptualization of privacy is the "right to be let alone" popularized by Warren and Brandeis in their seminal *Harvard Law Review* article on privacy [28]. They were the first scholars to recognize that a right to privacy had evolved in the 19th century to embrace not only *physical* privacy—a concept embedded in most European legal systems since the middle ages—but also a potential "injury of the feelings," which could, for example, result from the public disclosure of embarrassing private facts [29].

Efforts to define and analyze the privacy concept evolved considerably in the 20th century. In 1975, Altman conceptualized privacy as a "boundary regulation process whereby people optimize their accessibility along a spectrum of "openness" and "closedness" depending on context" [30]. Similarly, Westin [31] described privacy as a "personal adjustment process" in which individuals balance "the desire for privacy with the desire for disclosure and communication" in the context of social norms and their environment. Privacy thus requires that an individual has a means to exercise selective *control* of access to the self and is aware of the potential consequences of exercising that control [30], [32].

It must be noted that Altman and Westin were referring to nonelectronic environments, where privacy intrusion was typically based on fresh information, referring to one particular person only, and stemming from traceable human sources. The scope of possible privacy breaches was therefore rather limited. Today, in contrast, details about an individual's activities are typically stored over a longer period of time and available from multiple electronic sources. Privacy breaches can therefore also occur indirectly. For example, customer segmentation, a practice where companies divide their potential customers into groups that share similar characteristics, can lead to an exclusion of people from services based on potentially distorted judgments. Often, the sources of personal data are not traceable due to myriad collecting, combining, and processing entities. Solove [33] notes that, as a result, information privacy is now not only about controlling immediate access to oneself but also about reducing the risk that personal information might be used in an unwanted way.

Solove's distinction between *access control* and *risk management* suggests two distinct dimensions to building privacy-friendly technologies and information systems. In context with the historical interpretation of privacy as a boundary regulation process [30], engineers are first responsible for ensuring that users can exercise immediate control over access to themselves and their personal data. Second, they are responsible for minimizing future privacy risks by protecting data after it is no longer under a user's direct control. Given current IT architectures, it can be argued that these responsibilities extend to three distinct technical domains: the user sphere, the recipient sphere, and a joint sphere.

The "user sphere" encompasses a user's device. From a privacy perspective, user devices should be fully controllable by the people who own them. Data should not flow in and out of them without their owners being able to intervene. Additionally, devices should respect their owners' physical privacy, interrupting them only when needed and at appropriate times.

The "recipient sphere" is a company-centric sphere of data control that involves backend infrastructure and data sharing networks. Though this information is less open to public scrutiny, engineers still have a responsibility to minimize the risk of potential privacy breaches due to data

TABLE 1
Three-Layer Privacy Responsibility Framework and Engineering Issues

Privacy Spheres	Where Data is Stored	Engineer's Responsibility	Engineering Issues
User Sphere	Users' desktop personal computers, laptops, mobile phones, RFID chips	• Give users control over access to themselves (in terms of access to data and attention)	• What data is transferred from the client to a data recipient? • Is the user explicitly involved in the transfer? • Is the user aware of remote and/or local application storing data on his system? • Is data storage transient or persistent?
Joint Sphere	Web service provider's servers and databases	• Give users some control over access to themselves (in terms of access to data and attention) • Minimize users' future privacy risks	• Is the user fully aware of how his data is used and can he control this?
Recipient Sphere	Any data recipients: servers and databases of network providers, service providers or other parties with whom data recipient shares data	• Minimize users' future privacy risks	• What data is being shared by the data recipient with other parties? • Can the user expect or anticipate a transfer of his data by the recipient? • Is personal data adequately secured? • Is data storage transient or persistent? • Can the processing of personal data be foreseen by the user? • Are there secondary uses of data that may not be foreseen by the user? • Is there a way to minimize processing? (e.g. by delegating some pre-processing to User Sphere)

leakage or uncontrolled or undocumented access and sharing practices.

Finally, the "joint sphere of privacy control" encompasses companies that host peoples' data and provide (often free of charge) additional services (e.g., e-mail). Strictly speaking, these services are under the full control of the companies providing them. However, users may expect (or even believe they have) "privacy" when they use these services. Because it is *their* e-mail and/or the network space *they* personalized for *themselves* and view in *their* browsers, people expect that their privacy is protected. Google garnered strong criticism for mining its users' gmail accounts for advertisement purposes [34]. Network-based personal service environments therefore call for careful privacy design in which users and providers have a joint say as to the degree of "access" allowed. The risk of personal data abuse should also be minimized through the use of proper security mechanisms.

Table 1 summarizes the privacy spheres and resultant engineering responsibilities in a three-layer privacy responsibility framework.

3 Privacy Requirements Analysis

System requirement analysis typically starts with a detailed understanding of the relevant processes as well as stakeholder needs surrounding these processes [35]. In the context of privacy engineering, we therefore need to understand what user privacy perceptions and expectations exist, and how they might be compromised by IT processes [36]. We also need to understand the level of privacy protection that is required. In this section, we begin with an analysis of privacy sensitive processes, followed by an overview of user perceptions and concerns. We then discuss opposing views in the research community about how much privacy is actually required in different contexts.

3.1 System Activities and How They Can Differentially Impact User Privacy

All information systems typically perform one or more of the following tasks: data transfer, data storage, and data processing. Each of these activities can raise privacy concerns. However, their impact on privacy varies depending on *how* they are performed [37], *what type of data* is involved, *who* uses the data [38], [39], and in *which of the three spheres* they occur.

3.1.1 Data Transfer

Data transfer can occur at three levels. First, data may be transferred from a user's system to a service provider. Second, after the initial transfer, recipients may share the data within their own organizations. Third, data may be transferred to external third parties. External third parties are any data recipients outside the organizational boundaries of the user's direct interaction partner.

The first type of data transfer involves the user sphere. Here, engineers must ensure a controlled transition of data from the user to the selected recipient. From a privacy perspective, two types of transfers can be distinguished: transfers involving *explicit* user involvement and transfers not directly involving a user. When data transfers involve users—for example, when users fill out website forms—users are aware that the transfer is taking place and are likely to understand what benefit they are receiving in return (although they may not necessarily understand the privacy implications). When data transfer occurs without direct user involvement—for example, when Web browsers send cookie information back to websites, when passive RFID tags are read without notice, or when cameras record

activities in an environment—users may not understand why the transfer is taking place or even realize that it is taking place at all. This type of *implicit* data transfer tends to raise greater privacy concerns than those initiated by the user [37].

The design and communication of data transfer occurring in the joint and recipient spheres are a challenge for privacy engineers. Since users are generally not involved, they need to trust in the contextual integrity of the data recipient [40]. They need to trust that data recipients will only transfer their personal data in an appropriate (necessary and secured) way, consistent with their expectations. To earn this trust, engineers need to minimize the risk of inappropriate or uncontrolled transfers and create transparency as to what transfers occur (i.e., through policy communication). The joint sphere may be a particular area of sensitivity for users who believe their online data should not be available to any third party.

3.1.2 Data Storage

Privacy-sensitive storage of personal data can occur in the user, joint, and remote spheres. Generally, data storage occurs at a collecting entity's backend. Data may be stored in databases, transaction records, or log files on primary servers and backup tapes. Ensuring that stored data is adequately protected from unauthorized access is a key engineering responsibility. Privacy law in some countries also dictates that engineers must ensure transparency and some degree of control over personal data stored in backend systems (see Section 3.1.3).

Privacy also becomes an issue when local applications store data on a user's personal system (in the user sphere), sometimes without the user's awareness. For example, word processors embed personally identifiable metadata into documents to describe document creation and change history [41] and Web browsers store users' browsing history and cache Web content. Privacy breaches occur when users, unaware of such client-side storage, have their activities discovered by others. Users may also be uneasy about remote entities storing data on their local systems, especially when they do not understand the purpose of the data storage or are unable to control it. This is typically done in order for the remote application to operate a service, and the information generally consists of identifiers and information state. For example, many websites store cookies on a user's system in order to identify the client on the next visit. In this way, profiles can then be created and stored by data recipients, often without the knowledge or explicit consent of users. Thus, users must be made aware of data storage activities in their sphere so that they are not surprised by them later and do not interpret them as unwanted intrusions [42].

Generally, it is useful to distinguish between persistent and transient storage. *Persistent* storage involves data that is stored indefinitely or for some period of time that goes beyond a single transaction or session. It allows data from multiple transactions or sessions to be accumulated over time and retrieved later upon request. *Transient* storage refers to user data that is stored for the purpose of an immediate transaction and then deleted. Transient data storage has minimal privacy implications, while persistent data storage can raise significant privacy concerns [37]. As a result, the use of transient data storage can reduce privacy hurdles.

3.1.3 Data Processing

Data processing refers to any use or transformation of data. It is typically done outside of the user's sphere of influence. Data processing that is a necessary part of delivering a service or billing for a service is generally anticipated by users and does not typically raise privacy concerns.

However, companies often engage in secondary uses of personal data that may not be foreseen by users. For example, companies may group customers into segments based on their purchases or scan their e-mails to market personalized services. Such secondary use of data can occur with or without explicit user involvement. Under European privacy laws, users must be informed up front of all secondary uses of data and given an opportunity to provide or withhold their consent (in some cases, this can be satisfied by providing an opt-out, in other cases an opt-in is necessary) [43]. In the US, sector-specific legal requirements regulate secondary use of data (e.g., in the healthcare, telecommunications, and financial services sectors) [44].

Typically, the company that collects the data does the data processing. However, data processing may be outsourced to a third-party service provider, raising additional privacy concerns. Steps must be taken to ensure that third parties protect the data they receive and do not use it for their own purposes. There is a growing list of privacy breaches and identity theft incidents that have occurred due to negligence on the part of third-party service providers. One highly publicized data breach involved a data broker, ChoicePoint, that allowed fraudsters to register as legitimate businesses and gain access to consumer databases used by insurance companies, government agencies, and companies who use this information to run background checks [45].

3.2 Understanding User Privacy Expectations and Behavior

It is important for engineers to understand how privacy breaches can occur as a result of data transfer, storage, and processing. It is equally important for them to understand user expectations with regard to the privacy-friendliness of a system. Though people have many concerns about privacy issues, this privacy consciousness does not always fall in line with actual behavior and is highly variable from person to person [46], [47].

3.2.1 How Can Privacy Be Breached from a User's Perspective?

A number of studies have investigated individuals' privacy concerns [50], [51], [52], [53]. In 1996, Smith et al. identified seven areas of activity that cause unease [53]:

1. *collection* and storage of extensive amounts of personal data,
2. *unauthorized secondary use by the collecting organization*,
3. *unauthorized secondary use by an external organization* with whom personal data has been shared,
4. *unauthorized access* to personal data, e.g., identity theft or snooping into records,

5. *errors* in personal data, whether deliberately or accidentally created,
6. *poor judgment* through decisions made automatically based on incorrect or partial personal data, and
7. *combination of personal data* from disparate databases to create a combined and thus more comprehensive profile for a person.

This list shows that users are concerned about the amount of personal data leaving their sphere of influence. But even more (six out of seven) concerns arise from undesired data usage once data has been collected and is no longer under the user's control (joint and recipient sphere).

Since the Smith et al. [53] privacy study, technological advancements have added new issues to this list of privacy concerns. Shorter product lifecycles of digital devices and an increased ubiquity of information services have led to new forms of privacy breach. Garfinkel and Shelat [54] demonstrated the issue of uncovering personal information on used hardware. The unauthorized execution of operations on a personal device, taking advantage of increased processing power in personal devices, is also becoming more common. Unauthorized operations such as spyware may cause a computer system to become unstable, inundate the user with unwanted advertising, or trigger unauthorized collection of personal data [55].

Furthermore, pervasive computing environments have the potential to magnify privacy concerns as they multiply the number of interfaces people have with the network. As technology continues to progress, data transfer, storage, and processing volume is expected to increase substantially. Early consumer studies on RFID technology reveal that people are aware of and feel threatened by the new volume of data that is passively collected about them [56], [57]. The historic concept of "the right to be let alone" [28] may be expanded to include the right not to be addressed by or forced to view digital services in pervasive computing spaces. Pop-ups, e-mail, and SMS spam are now regularly considered to be privacy intrusions [58].

Finally, the past decade of technological advancement, in particular growth in bandwidth and connections to the Internet, has led to a new breed of service platforms that typify the joint sphere. These platforms offer the management and display of personal information along with communication services. Examples are online blogs, social network platforms, online e-mail services, or media-file repositories. Depending on the nature of these services, privacy breaches can either be triggered by users themselves or by the platforms' operations. Undesired "exposure" [33] is probably the most frequent privacy breach reported on in this context. For example, people publicly described unfavorably by other online users may feel hurt or disgraced. The frequent practice whereby service providers mine users' content (e.g., e-mails) may also lead to a feeling of undesired exposure [34].

Table 2 relates the three-sphere framework to the user privacy concerns described in this chapter.

TABLE 2
Three-Layer Privacy Responsibility Framework and Associated User Privacy Concerns

Sphere of Influence	User privacy concerns
User Sphere	• Unauthorized collection • Unauthorized execution • Exposure • Unwanted inflow of data
Joint Sphere	• Exposure • Reduced Judgment • Improper access • Unauthorized secondary use
Recipient sphere	• Internal unauthorized use • External unauthorized use • Improper access • Errors • Reduced judgment • Combining data

3.2.2 Privacy Attitudes and Privacy Behavior

While privacy preservation seems to be of growing concern, empirical studies, as well as observations of actual user behavior, suggest just the opposite: People appear to be unconcerned about privacy until it is actually breached [48], [49] and they do not necessarily act according to the privacy preferences they claim to have. Spiekermann et al. [12], [59] showed that regardless of a user's expressed privacy concerns, they are willing to reveal the most intimate details of their personal preferences if deemed appropriate. Social network platforms, including online blogs and other intimate presentations of personal lives, are flourishing on the Net and suggest a "new exhibitionism" [60]. Millions of people are regularly using loyalty cards through the use of which they reveal most of their consumption preferences. Industry professionals who observe this phenomenon are tempted to interpret this behavior as a decreasing interest in privacy in modern society. However, such a conclusion could be short sighted.

Studies show that users differ in the degree and focus of their privacy concerns [52], [61]. One group of people continuously identified in privacy studies is referred to as "unconcerned" or "marginally concerned" [31], [52], [49], [47]. This group of people, estimated to represent around 20-25 percent of the population, may contribute to the impression that "data intensive" services (such as loyalty cards) are accepted across the population and that privacy is becoming less important. Yet, other privacy opinion clusters identified in the same studies include privacy "fundamentalists" and "pragmatists," denoting very high and medium degrees of privacy concern [31], [52]. Among pragmatists, consumers can be grouped further into those who are more "identity aware" and those who are more "profile aware" [46], [47]. Identity aware users are people who worry more about sharing identifying information such as e-mail addresses, physical address, or phone numbers than profiling practices. Profile aware users are more concerned about sharing characterizing profile information such as hobbies, age, interests, or preferences. A study by one of the authors suggests that attitudes toward RFID services are consistent with this cluster affiliation. Also, people who are unconcerned or less concerned about privacy tend to be those with a lower level of education [49], a finding that is mirrored in technology acceptance studies [57]. It has been observed that

people who know little about data processing are twice as likely to use data-intensive services, such as loyalty cards, than those who have a realistic perception of data use [62].

Finally, a research community working under the term "privacy economics" argues that economic rationale can explain why people do not display privacy protection behavior that is consistent with their privacy attitudes. Varian [65] argues that "a transaction is made more efficient if detailed information about the consumer's tastes is available to the seller" and thus posits that it is rational for people to reveal personal data in sales contexts. Others draw on the theory of immediate gratification and comment that consumers probably give higher value to immediate benefits from data-intensive services (such as advice from an e-commerce website) than to the long-term desire to maintain privacy [63]. People may overvalue the immediate benefits they obtain from revealing information and underestimate the cumulative risks associated with the cost of privacy loss [64].

Complementary to privacy economics research, other studies investigate privacy behavior from a more psychological perspective. Strandburg [66] argues that people have a willpower problem and cannot resist the temptation to reveal. Huberman et al. [67] found that people are most restrictive about those personal data points where they diverge from the average of their peer group. Spiekermann [62] argues that peoples' privacy behavior may be driven by offline evolutionary experience: For example, if their data goes to a huge database, as is the case with loyalty cards, they believe that their data is drowned in and protected by the mass of others' data (just as individual behavior stands out less in a crowd). Generally, people seem to have difficulty grasping that "the Internet does not forget" [69].

In addition, much existing online privacy behavior simply goes unobserved leading to the false conclusion that people do not care. Gumbrecht [70] showed that blog authors use ambiguous language and references in order to protect their and others' privacy. Viégas [68] found that the majority of bloggers carefully consider whether certain topics are too personal to write about. They often develop sophisticated rules on how to write about others and whether to identify their subjects. These findings make it plain that privacy protection is an inherent part of how people act online even though it cannot be observed electronically.

It is also difficult for marketers to monitor comparative and relative behavior in the e-commerce world. Participants in a laboratory study who were provided with easy-to-understand information about website privacy policies were more likely to make purchases from sites with better privacy policies than those who did not receive this information [71], [16]. At the same time, it is impossible for marketers to tell how many customers are lost due to a lack of privacy sensitivity.

In summary, findings on user behavior suggest that privacy is an issue for the majority of people despite the fact that they engage in data-intensive services and do not protect their personal data sufficiently. Personal propensity, group behavior, irrational decision making, a lack of IT education, long-standing information-sharing habits, unobserved privacy behavior, as well as some economic calculus may explain this behavior. What seems an acceptable and tempting service-data exchange to some engineers or marketers may be unacceptable to many users, lead to an unpredictable market [72] and lead to customer backlash over privacy issues [5]. In order to protect companies from such volatility in customer perceptions, shown to be relevant to stock-market valuation [6], it may be advisable to build systems and follow privacy policies based on some baseline privacy protection, as described in Sections 4 and 5.

3.3 Privacy Expectations and Threat Model

When designing a privacy-friendly system, engineers must consider customer expectations and the extent that privacy-enhancing technologies (PETs) are needed to address users' privacy concerns and to meet legal requirements.

Primary and secondary data recipients therefore play a crucial role in the degree of privacy protection and information collection [39], [73]. Application service providers, network providers, software vendors, social network providers, location service providers, etc., all have to ask themselves *"what do our customers expect from us in the context of their transactions?"* How much and what data do they expect will be collected in each instance according to the norm? Which of this data do they expect will be distributed further? Government requirements also come into play. Service providers are now asked by governments to store transaction data much longer than needed for billing purposes in order to facilitate criminal investigations [74]. If there is no such regulatory mandate, the degree of data parsimony and privacy built into a system is at the discretion of the system designers.

How much privacy should be built into a system that can be called privacy protective or privacy enhancing? Little consensus has been reached in the privacy research community. PET researchers have different opinions on who the privacy "attacker" is. Is it the government, with potentially unlimited resources able to systematically reconstruct an individual's transactions? Are privacy-friendly systems for protecting people from each other (e.g., malicious hackers with limited resources but the desire to intrude on others' privacy)? Or, are they supposed to protect people from becoming digits in a commercial "database nation" [75], where companies accumulate extensive individual profiles for profit maximization? Cryptography researchers and privacy rights organizations tend to favor systems that prevent access to individuals and their information at all cost. The goal is to make access to the individual tamper-proof and to build a technological infrastructure based on nonidentifiability of users even vis-à-vis governments. Often, unfortunately, achieving this ambitious goal undermines system usability and drives system cost to a point where marketability and adoption of the solution becomes difficult. However, recent technological advances, for example, in the area of privacy-preserving data mining [76] and differential privacy [22], may lead to deployable solutions with strong privacy guarantees.

Other groups in the privacy technology community care less about making access theoretically and cryptographically tamper-proof and acknowledge that information may

TABLE 3
The Fair Information Practices Proposed by the US Federal Trade Commission in Their 2000 Report to Congress [79]

(1) Notice	Websites would be required to provide consumers clear and conspicuous notice of their information practices, including what information they collect, how they collect it (e.g., directly or through non-obvious means such as cookies), how they use it, how they provide Choice, Access, and Security to consumers, whether they disclose the information collected to other entities, and whether other entities are collecting information through the site.
(2) Choice	Websites would be required to offer consumers choices as to how their personal identifying information is used beyond the use for which the information was provided (e.g., to consummate a transaction). Such choice would encompass both internal secondary uses (such as marketing back to consumers) and external secondary uses (such as disclosing data to other entities).
(3) Access	Websites would be required to offer consumers reasonable access to the information a Website has collected about them, including a reasonable opportunity to review information and to correct inaccuracies or delete information.
(4) Security	Websites would be required to take reasonable steps to protect the security of the information they collect from consumers.

be collected for useful purposes such as personalized services. For them, the threat model is what is commercially feasible to do and not what is theoretically doable. This group's goal is to give people control through informed consent to personal data use. They are willing to forego cryptographically sophisticated solutions in favor of "good enough," easy-to-use, and affordable privacy protection, while recognizing that such solutions may offer insufficient protection against well-funded attackers.

The choice of threat model around which to design a PET may be dictated by customer requirements or legal requirements. However, for consumer-focused systems that will be used by a broad range of users, there may not be one threat model that is appropriate for all users. In some cases, it may be possible to design a privacy-friendly system that users can customize based on their personal threat model. For example, Cranor et al. [77] designed a graphical user interface for the Tor anonymity system that would allow users to select configuration options based on their privacy needs.

In order to scale the degree of privacy built into systems, engineers need to consider customer expectations, government regulations, and the threat model they believe to be viable for the majority of their customers—generally in consultation with other decision-makers in their company—and determine the architectural design options that are appropriate for various protection levels.

4 Engineering Privacy-Friendly Systems

In this section, we propose a methodology for systematically engineering privacy friendliness. First, we introduce the "notice and choice" approach based on FIPs. We discuss how this approach can be supported through "privacy-by-policy." We then discuss an alternative approach, "privacy-by-architecture."

4.1 Principles of Fair Information Practice

In 1980, the OECD published eight Guidelines on the Protection of Privacy and Trans-Border Flows of Personal Data [78], which have since served as the basis for privacy legislation in Europe [43] and many other countries. Often referred to as Fair Information Practices (FIPs), these principles emphasize the need to minimize the collection and use of personal data, to inform individuals about data collection, and to adequately maintain and protect collected data. US regulatory and self-regulatory efforts supported by the Federal Trade Commission (FTC) have over the past decade focused on a subset of these principles, tailored to the e-commerce context—notice, choice, access, and security—as shown in Table 3 [79]. Because the FTC principles focus on notice and choice rather than minimizing data collection or use limitation, they are sometimes referred to as a "notice and choice" approach to privacy. This is a pragmatic approach that recognizes that companies are reluctant to stop collecting or using data, but also recognizes that individuals expect to retain control over how their data is used.

While the notice and choice approach is useful, it is *not* clear that it should serve as the golden rule for privacy design since notice, choice, access, and security only come into play when a system collects personal data. These principles could be largely irrelevant in systems built with privacy-friendly architectures in which little or no personal data is collected in the first place. If a company decided to abstain from collecting personal data and base its business model entirely on pseudonymous and nonidentifiable user tokens, it should not be required to provide complex privacy notifications and choices to its customers. However, in some jurisdictions, opt-out choices may be legally required even when data is used in anonymous form. Even without legal mandates, companies may find it beneficial to provide a simple, unobtrusive notice to let customers know that they are not actually collecting any personally identifiable information. For example, if a company was able to build or market its products and services entirely through a client-controlled architecture combined with nonidentified transaction mechanisms [80], [81], notice and choice may not be needed. We call this approach *privacy-by-architecture*. On the other hand, if a company opted to implement just enough privacy mechanisms to let users feel comfortable and perceive an adequate level of protection, then notice and choice would play an important role, providing users with some degree of control over their personal data. We call this approach *privacy-by-policy*. In a hybrid approach, privacy-by-policy can be enhanced through technical mechanisms that audit or enforce policy compliance. The decision to use any one of these system designs may be

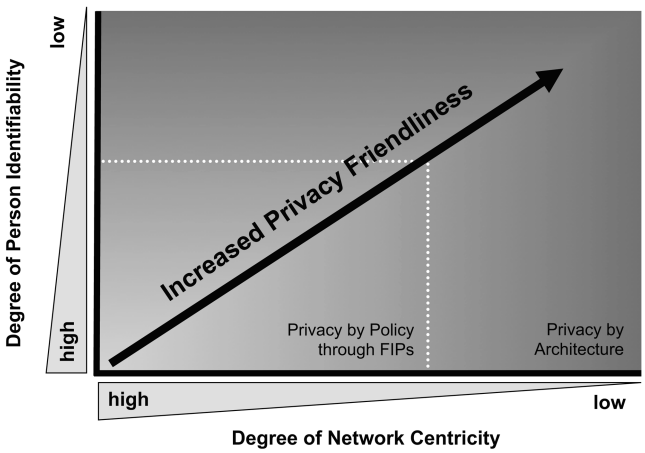

Fig. 1. Privacy friendliness of architectural choices.

based on customers' concerns and the relevant privacy threat model, as well as on technological capabilities, business needs, or regulatory requirements. The next section provides a more detailed description of these architectural options.

4.2 Architectural Choices

When building a new system from scratch, we argue that engineers typically can make architectural choices on two dimensions: network centricity and identifiability of data. "Network centricity" is the degree to which a user's system relies on a network infrastructure to provide a service, as well as the degree of control a network operator can exercise over a client's operations. More network centricity means potentially less privacy for clients. The more network-centric a system is, the more the network operator knows about the client and the more he can control the client. "Identifiability" can be defined as the degree to which data can be directly attributed to an individual. Personal data can be entered into a system anonymously (e.g., e-voting) or by identifying oneself (e.g., when conducting online banking transactions). Naturally, anonymous transactions imply a higher degree of privacy for the data provider [82], [83].

Recently, systems have been developed offering more client-centric architectures and anonymous transactions. These systems embed privacy features and create privacy-by-architecture, providing higher levels of privacy friendliness than systems that collect personally identifiable data and adhere to a FIP policy [84]. Applications with client-centric architectures minimize the need for personal information to leave the user sphere. For example, Place Lab is a software framework for location-based services that allows devices to locate themselves without revealing their location to a central server [85]. Furthermore, by using anonymous or pseudonymous credentials that attest to a relevant fact rather than to a person's identity, secure transactions can take place outside the user sphere without the transfer of personal information [80], [81] (Fig. 1).

4.2.1 Network- versus Client-Centric Architectures

Network-centric architectures facilitate the use of inexpensive client devices with minimal storage and processing capabilities. As clients become more powerful, it may be possible in many instances for data processing to occur on a user's computer, eliminating the need for data transfer and remote storage, minimizing unwanted secondary data use, and improving service quality. For example, while most recommendation systems rely on a central database of user preferences or behaviors, Canny [86] has proposed a collaborative filtering system architecture in which individual participants store their data preferences on their own systems and compute an "aggregate" of their data to share with other members of their community. Alternative designs for location-based services (LBSs) illustrate well how client-centric architectures are more privacy friendly: LBSs rely on knowing the exact location of a user's device. Mobile operators who offer LBSs typically calculate a device's location through triangulation based on network information. They use location information to customize the information they send to the mobile device. This network-centric architecture stands in sharp contrast to client-centric mobile location solutions such as Place Lab [85] or those based on GPS. A GPS-enabled smart phone can use satellite data to calculate its own position and provide that location information to an application running directly on the phone. Since no information is sent back upstream, the user's location remains completely private.

One way to provide some of the privacy protections associated with a client-centric architecture while allowing for the use of inexpensive clients is to deploy a system in which clients communicate with a trusted intermediary that makes anonymized location requests on their behalf [87]. Another approach involves using clients that frequently change their network identifiers to reduce the ability of service providers to track a client over time [88].

However, a company's decision to choose a more client-centric architecture may have important strategic implications for its business model and position in the value chain. Greater network control offers authority over who accesses the customer base and thus a controllable competitive landscape. Mobile operators, for example, have a genuine interest in building network-centric architectures. Knowing where a client is allows them to sell this information or generate extra revenue through LBSs sold over the network. Pure client-centric GPS systems leave operators out of higher margin content business. A similar dynamics has been observed for DRM systems where more network control would be in the interest of copyright owners and distributors while undermining users' privacy [23]. The examples show that architectural decisions in favor of privacy entail trade-offs in terms of profit and system dependability.

4.2.2 User Identifiability

A company's technical and business strategy does not always allow implementation of a more privacy-friendly client-centric system. However, privacy concerns can be reduced in network-centric systems if data is not stored in a form that facilitates identification of a unique individual.

System designers should consider the extent to which users can remain unidentified during electronic transactions. Indeed, many service providers on the Internet acknowledge this factor already: they offer their users a

TABLE 4
Framework for Privacy-Friendly System Design

Privacy stages	identifiability	Approach to privacy protection	Linkability of data to personal identifiers	System Characteristics
0	identified	privacy by policy (notice and choice)	linked	• unique identifiers across databases • contact information stored with profile information
1	pseudonymous		linkable with reasonable & automatable effort	• no unique identifies across databases • common attributes across databases • contact information stored separately from profile or transaction information
2		privacy by architecture	not linkable with reasonable effort	• no unique identifiers across databases • no common attributes across databases • random identifiers • contact information stored separately from profile or transaction information • collection of long term person characteristics on a low level of granularity • technically enforced deletion of profile details at regular intervals
3	anonymous		unlinkable	• no collection of contact information • no collection of long term person characteristics • *k*-anonymity with large value of *k*

pseudonymous self-representation. Thus, the service provider can store extensive customer profiles and offer personalized services or products with reduced privacy risks [89], [90], [91]. It should be noted that the pseudonyms provided by many service providers (e.g., AOL screen names) only protect a user's identity from other users. They do not provide users with privacy vis-à-vis the service provider who can typically reidentify them knowing the link between the pseudonym and the real identity. Consequently, customer pseudonyms do not automatically provide privacy-by-architecture.

Reidentification typically occurs in one of two ways: First, pseudonymous profiles can be reidentified by linking them with identity information stored in another database within the same company. Most companies collect identification data from customers in order to bill them or ship products. Reasonably easy linkage can be achieved if both customer profile and billing/shipping databases share a common attribute such as an e-mail address. A second means of reidentification is to apply data mining techniques to pseudonymous transaction logs. Since users often reveal personal data as a part of their pseudonymous transactions, their identities may be derived from the data traces they create. When AOL released logs of search queries identified only by pseudonyms, some users were identified because their names or contact information appeared in some of their search queries [92]. This created a public relations nightmare for AOL and several AOL employees involved in the incident resigned or were fired [93]. The next section discusses how engineers can actively specify the degree of user identifiability.

4.3 Degrees of Identifiability

The framework for privacy-friendly system design presented in Table 4 shows that the degree of privacy-friendliness of a system is inversely related to the degree of user data identifiability [94]. The more personally identifiable data that exists about a person, the less she is able to control access to information about herself, and the greater the risk of unauthorized use, disclosure, or exposure of her personal data.

The ability to link personal information to create a comprehensive and identified profile is the key to determining the degree of privacy a person has. Linkage can occur directly by joining database information or it can be achieved indirectly by pattern matching [95]. Table 4 shows what measures can be taken by engineers to reduce the risk of profile linkage and pattern matching and thus embed more or less privacy into their systems.

In stage 0 of the framework, unique identifiers (social security numbers, stable IP addresses, etc.) and contact or other information that can be used to readily identify a unique individual or household are stored in a user's profile. Such data sets can be characterized as identified. For example, if an online store records its customers' purchases in combination with their names and addresses or telephone numbers, then

the purchases are linked to identified individuals. In such a system, privacy can be protected only through policies that restrict the store's use and disclosure of customer data and that provide notice and choice to customers.

In stage 1 of the framework, contact information is required from a customer, but immediate links between a person's profile and her identified self are avoided by storing the contact information and the profile information in separate databases, prohibiting the use of unique identifiers across these databases, and storing the profile information under a pseudonym. Yet, as outlined above, reidentifiability is an issue of concern when choosing this technical privacy strategy. This is because common identifiers across the contact and profile databases may still exist and could be used to resolve the pseudonym. For example, both databases may contain common e-mail addresses or unique identifiers such as those stored in cookies. Users might also select common pseudonyms or passwords across multiple systems. As a result, the probability of reidentifying individuals registered under a pseudonym is reasonably high. Reidentification can sometimes even be done in an automated fashion, rendering the reidentification process cost efficient [96]. As a result, this type of system design strategy only provides a medium degree of privacy. Since reidentification is still technically possible, policies should be put in place to prohibit it, and users should be informed of these policies, as well as of the steps they can take to protect their privacy (for example, choosing unique pseudonyms and passwords).

In stage 2, systems are actively designed for nonidentifiability of users, creating what we denote as "privacy-by-architecture." Separate databases for profile and contact information must be created in such a way that common attributes are avoided. In addition, steps should be taken to prevent future databases from reintroducing common identifiers. Identifiers should therefore be generated at random and any information that is highly specific to an individual (e.g., birth dates or contact data) should be avoided whenever possible. For example, if the age of a customer matters for a business's marketing purposes, the year of birth should be registered without the precise day and month. If the birthday matters for a business's marketing, then day and month should be recorded without the year of birth. The general guideline here is to *minimize the granularity of long-term personal characteristics collected about an individual*.

Even so, it may still be possible to individually identify a person based on transaction patterns. Pattern matching exploits the notion that users can be reidentified based on highly similar behavior or on specific items they carry over time and across settings. Reidentification based on pattern matching also relies on the existence of one identified pattern or a way to add identifying data to an existing pattern. For example, mobile operators may be able to reidentify a customer even if he uses an unidentified prepaid phone. This can be done by extracting the pattern of location movements over a certain time span and extracting the endpoints of the highly probable home and work locations. Typically, only one individual will share one home and work location. Researchers have also demonstrated that a relatively small amount of information about an individual's tastes in movies is sufficient to identify them in an anonymized movie rating database [95].

Pattern matching does not always result in the identification of a unique individual. Often, a pattern may match multiple individuals. In some cases, a unique match can be obtained with some additional effort by contacting the individuals, observing their behavior, or enhancing their profiles with information about them from other sources. k-Anonymity is a concept that describes the level of difficulty associated with uniquely identifying an individual [97]. The value k refers to the number of individuals to whom a pattern of data, referred to as *quasi-identifiers*, may be attributed. If a pattern is so unique that k equals one person ($k = 1$), then the system is able to uniquely identify an individual. Detailed data tends to lower the value of k (for example, a precise birth date including day, month, and year will match fewer people than a birthday recorded without year of birth). Long-term storage of profiles involving frequent transactions or observations also tends to lower the value of k because unique patterns will emerge based on activities that may reoccur at various intervals. The values of k associated with a system can be increased by storing less detailed data and by purging stored data frequently. The frequency of observations or transactions will dictate the frequency of purging necessary to maintain a high value of k. Some privacy laws, such as the German data protection law, require that identified data be deleted after its purpose has been fulfilled. However, there is value in purging nonidentified data as well, to minimize the risk of reidentification based on pattern matching.

In some cases, large values of k may be insufficient to protect privacy because records with the same quasi-identifiers do not have a diverse set of values for their sensitive elements. For example, a table of medical records may use truncated zip code and age range as quasi-identifiers, and may be k-anonymized such that there are at least k records for every combination of quasi-identifiers. However, if for some sets of quasi-identifiers, all patients have the same diagnosis or a small number of diagnoses, privacy may still be compromised. The l-diversity principle can be used to improve privacy protections by adding the requirement that there be at least l values for sensitive elements that share the same quasi-identifiers [98].

In stage 2 of our privacy framework, privacy-by-architecture does not guarantee unlinkability; rather, it ensures that the process of linking a pseudonym to an individual will require an extremely large effort. The degree of effort required may change over time due to technical advances or the availability of new data sources. Reidentification that previously required manual effort might become automatable, and thus might become more cost-effective. Therefore, it is important that privacy claims be made with a view to the future and that they be periodically reevaluated.

An extreme form of privacy-by-architecture is denoted in our framework where users remain anonymous. Anonymity can be provided if no collection of contact information or of long-term personal characteristics occurs. Moreover, profiles collected need to be regularly deleted and anonymized to achieve k-anonymity with large values for k or l-diversity with large values for l.

TABLE 5
Providing Notice to Address User Privacy Concerns

User concerns	Notice should be given about…
Marketing Practices	
Combining Data	Notice about data combination practices • external data purchases? • linking practices?
Reduced Judgment	Notice about segmentation practices • type of judgments made? • personalization done? • what does personalization lead to for the customer? • sharing of segmentation information?
Future attention consumption	• contact plans (i.e. through newsletters, SMS)
IS Practices	
External unauthorized transfer	• is data shared outside the initial data recipient? • if yes, with whom is data shared?
External unauthorized processing	• is data processed externally for other purposes than initially specified? • if yes, for what purposes?
Internal unauthorized transfer	• is data transferred within a company conglomerate? • if yes with whom within the comglomerate?
Internal unauthorized processing	• is data processed internally for other purposes than initially specified? • if yes, for what purposes?
Unauthorized collection of data from client	• use of re-identifiers (i.e. cookies, stable IP address, phone number, EPC) • collection of information about device nature (i.e. browser, operating system, phone type) • collection of information from the device (i.e. music library, cache information)
Unauthorized execution of operations on client	• installation of software? • updates?
Exposure	• cached information (i.e browser caches, document histories) • collection of information from the device (i.e. music library, cache information)

To conclude, we argue that if a company pursues a privacy-by-architecture approach, it should be allowed to forgo any further notice and choice communication with customers. Since no personally identifiable data is technically created or recreatable with reasonable effort, no real threat to a person's privacy is established. Consequently, companies opting for this privacy strategy should be relieved of the duty to engage in complex privacy policy exchanges. Nonetheless, notices and opt-out opportunities may be required in some jurisdictions, even for anonymous data use. Furthermore, unobtrusive communications that explain how privacy is being protected may help build trust and allow a company to promote their privacy protective architecture. Such notices can also allow independent parties to assess the risk that data might be reidentifiable in the future. In contrast, if companies do not opt for a privacy-by-architecture approach, then a privacy-by-policy approach must be taken where notice and choice will be essential mechanisms for ensuring adequate privacy protection. The next section details this approach.

5 Implementing Privacy-by-Policy

If a company opts to collect identified or reidentifiable personal information in accordance with stages 0 and 1 of the framework presented in Table 4, then it pursues a strategy we characterize as privacy-by-policy. In this section, we provide guidelines on how companies can implement the notice, choice, and access FIPs, including ways to meaningfully inform users about data practices without being overly disruptive. The security FIP can be implemented by adhering to security best practices, which are covered extensively elsewhere [99], [100].

5.1 Providing Notice and Choice

Companies can instill trust by providing information about what personal data they collect and how they use and protect it. Information can be provided in the form of a comprehensive privacy policy, a short or layered notice [101], or brief notifications and opportunities to make choices at the time that data is collected, stored, or processed. Users should be given the opportunity to make choices about secondary uses of their personal information —those uses that go beyond the original purpose for which the data was provided. Privacy policies should cover the type of data collected, how that data will be used, the conditions under which it will be shared, how it will be secured, and how individuals can access their own data and provide or withhold their consent to data processing. Discussions of how to create usable privacy policies are covered elsewhere [102], [103].

What should users be informed about? Based on users' privacy concerns discussed in Section 3, Table 5 proposes best practices for providing meaningful privacy notices.

Depending on the technological and business environment, some information may not apply (i.e., if there are no sharing practices or no use of data for marketing purposes). Even though customers typically do not read all of the information, it still serves as a signal. The mere fact that it is available generates trust and motivates companies' internal compliance.

An additional challenge for engineers is to design privacy interfaces that give users appropriate ad hoc notices concerning data collection and use choices. Ad hoc notices may take the form of pop-up windows, short notices incorporated into online forms, or alerts issued by handheld devices. Because this type of notice interrupts the user's workflow, users may consider them a nuisance and ignore them. Meaningful and timely information can be offered with minimal disruption by positioning notices at the point in an interaction where they are most relevant, by providing information in a format that succinctly conveys the most important information, and by limiting notices to situations that are most likely to raise privacy concerns. Such notices may provide a link to a more comprehensive privacy statement.

Decisions about when to interrupt users with privacy-related information can be difficult to make [104] and should be based on the extent to which privacy concerns are raised by data collection or processing. Generally, interruptions are less disturbing if they do not force the user to pay attention and are presented between tasks [105]. While there is a risk of burdening users with too many ad hoc notices, users who are not informed about data collection or processing may lose trust in a company if they discover later that their data has been collected or used.

Systems that allow users to specify privacy preferences up front and have them applied in future situations, as well as systems that learn users' privacy preferences over time, may further minimize the need for user interruption. Websites may use the Platform for Privacy Preferences (P3P) to provide privacy information in a computer-readable format, allowing user agents to make automated privacy choices based on a user's stored privacy preferences [32], [25]. Instant messaging and chat clients may allow users to set up privacy preferences indicating who can see their presence information or other personal information so that they need not be prompted every time someone requests their information [106]. Current research on privacy in location-based services seeks to find technical mechanisms for allowing users to retain fine-grained control over the conditions under which their location information may be released, without requiring them to explicitly authorize every release [107], [108].

5.2 Providing Access

Companies often provide access mechanisms by supplying a point of contact for customers, generally in the form of a phone number for customer service. To ensure that updates made through the customer service desk are propagated back to the systems where the data is stored and processed, engineers need to ensure efficient, yet secure access to a current and complete view of customer information. It is also necessary to minimize the risk that the access mechanism itself will open up additional privacy vulnerabilities. Privacy vulnerabilities may occur due to customer service employees making unauthorized use of personal data, or by people calling customer service and pretending to be someone they are not in order to obtain personal data —a practice known as pretexting [109], [110]. To prevent and detect unauthorized access, access to a full customer profile view should be auditable.

Another approach to information access is for companies to provide customers with direct online or automated telephone system access to their personal information. Companies that sell products and services exclusively online often allow users to set up accounts to store their billing and shipping information, as well as other information such as clothing sizes or product-related preferences. Sometimes companies allow users to use their account to access information about all of their past transactions with that company. Users can also edit or delete their contact information and preferences using the online interface. The number of companies, including telecommunications providers, banks, and others, that offer customers the ability to manage their accounts online is steadily increasing. Online account management interfaces offer an effective way of providing individuals with access to their personal data. However, they can also open up privacy vulnerabilities if inadequate authentication mechanisms are used. Authentication may be particularly problematic when offline retailers offer customers the opportunity to open online accounts. For example, when the US grocery chain Stop & Shop offered its loyalty cardholders the ability to access their grocery purchase records online, cardholders initially authenticated themselves by typing in their loyalty card number. Someone who had obtained a grocery receipt of another customer or had access to their card will be able to access that customer's records [111]. Likewise, users of shared computers may inadvertently provide other users of their computer with access to their personal information through online account mechanisms—for example, if they set up accounts to allow access with a cookie rather than a password.

A discussion of how to properly authenticate account holders is beyond the scope of this paper, and best practice in this area is constantly changing as vulnerabilities are discovered in commonly used authentication mechanisms.

5.3 Responsibility for Informing Users about the Data Sharing Network

When data is shared among multiple recipients, the question arises as to whose responsibility it is to inform users of data collection and usage practices. Engineers first need to determine to what extent their systems have direct relationships with users. If users have a direct relationship with a company X, it is this company X upon which they base their trust to handle their personal date. Therefore, companies that run systems that interact directly with users—service providers (e.g., iTunes or facebook.com), network providers (e.g., Vodafone), or system providers (e.g., Microsoft)—have a responsibility to notify users and provide them with data collection and processing choices. In addition, companies may also need to inform users of the data handling policies of their partners if they share data. This is because data sharing raises significant privacy

concerns and is often not otherwise known to users. Ideally, a company should provide a maximum amount of information about their data sharing network and take responsibility for its conduct. A good industry example of taking such "sharing-network responsibility" is the mobile operator Vodafone. Vodafone has established a Privacy Management Code of Practice that establishes privacy rules for all third parties who want to provide location services to Vodafone customers [112]. Breach of the code can lead to serious consequences for service providers, such as service contract termination, cost recovery, and payment withholding. Vodafone thus takes responsibility not only for its own practices but also informs its customers about its data sharing network and enforces privacy standards on this network.

5.4 Technical Mechanisms to Audit and Enforce Compliance

Companies may adopt an approach that is essentially privacy-by-policy, yet obtain some of the benefits of a privacy-by-architecture approach by also adopting technologies that can aid in auditing or enforcing policy compliance. A number of systems have been proposed that include a compliance engine that evaluates all data access requests according to a set of privacy rules [113], [114]. Thus, data requests may be examined to determine whether the requester is allowed to access that data, whether the purpose specified by the requester is permitted, or other policy requirements. Once a request is determined to be policy compliant and data is released, there is no guarantee that the requester will not misuse the data or disclose it inappropriately. However, this approach helps protect against unintentional privacy violations. Furthermore, associated auditing mechanisms can provide evidence as to which employees accessed a particular data set and thus who may be responsible should a breach occur. Digital rights management and digital watermarking techniques have also been proposed as mechanisms to allow individuals to track the flow of their personal information, and even to discover when digital photographs of themselves have been taken without their knowledge and released publicly [115]. Others have argued that the combination of privacy notices and auditing is a satisfactory and practical solution to addressing privacy concerns that is likely to be more palatable to businesses than privacy-by-architecture approaches [23].

6 Conclusion

In this paper, we have presented an overview of state-of-the-art privacy research and derived concrete guidelines for building privacy-friendly systems. We have introduced a three-sphere model of user privacy concerns and related it to system operations (data transfer, storage, and processing). We then described two types of approaches to engineering privacy: "privacy-by-policy" and "privacy-by-architecture." The privacy-by-policy approach focuses on implementation of the notice and choice principles of FIPs, while the privacy-by-architecture approach minimizes collection of identifiable personal data and emphasizes anonymization and client-side data storage and processing. Systems built such that profiles can be linked with a reasonable or automatable effort do not employ technical privacy-by-architecture, but instead rely on privacy by policy. These systems need to integrate notice, choice and access mechanisms in order to make users aware of privacy risks and offer them choices to exercise control over their personal information. However, if linkability is rigorously minimized—creating privacy-by-architecture—notice and choice may not need to be provided. Because privacy is technically enforced in such systems, little reasonable threat remains to a user's privacy and hence additional warnings, choices, and interruptions may be more confusing than productive.

Today, the privacy-by-policy approach has been embraced by many businesses because it does not interfere with current business models that rely on extensive use of personal information. In the absence of enforced legal restrictions on the use of personal data, the privacy-by-policy approach relies on companies to provide accessible privacy information and meaningful privacy choices so that users can do business with the companies that meet their privacy expectations. However, as Kang [116] notes: "For numerous reasons, such as transaction costs, individuals and information collectors do not generally negotiate and conclude express privacy contracts before engaging in each and every cyberspace transaction. Any proposed market-based solution which does not acknowledge this economic reality is deficient." P3P user agents offer the potential to reduce such transaction costs, but they have yet to gain widespread use. Hybrid approaches may offer a practical solution that satisfies business needs while minimizing privacy risk.

The privacy-by-architecture approach generally provides higher levels of privacy to users more reliably and without the need for them to analyze or negotiate privacy policies. In some cases, it can provide better privacy protection while still offering a viable business plan.

Acknowledgments

This work was supported in part by US Army Research Office contract no. DAAD19-02-I-0389 ("Perpetually Available and Secure Information Systems") to Carnegie Mellon University's CyLab and by the Transcoop programme of the Humboldt-Shiftung.

References

[1] A. Etzioni, *The Limits of Privacy*, 1999.
[2] P. Sprenger, "Sun on Privacy: 'Get over It'," *Wired News*, http://www.wired.com/news/politics/0,1283,17538,00.html, 26 Jan. 1999.
[3] "The Coming Backlash in Privacy," *The Economist*, 2000.
[4] A. Cavoukian and T.J. Hamilton, *The Privacy Payoff: How Successful Businesses Build Customer Trust.* McGraw-Hill, 2002.
[5] J. Guynn, "Facebook Hangs Its Head over Ad System," *Los Angeles Times*, http://www.latimes.com/business/printedition/la-fi-facebook6dec06,0,1006420.story?coll=la-headlines-pe-business, 6 Dec. 2007.
[6] A. Acquisti, A. Friedman et al. "Is There a Cost to Privacy Breaches? An Event Study Analysis," *Proc. Third Int'l Conf. Intelligent Systems*, 2006.
[7] Ernst & Young LLP, *Privacy: What Consumers Want*, 2002.

[8] Int'l Assoc. Privacy Professionals (IAPP), "US Privacy Enforcement Case Studies Guide," http://www.privacyassociation.org/images/stories/pdfs/IAPP_Privacy_Enforcement_Cases_07.05.07.pdf, 2007.

[9] CBS News, "Poll: Privacy Rights under Attack," http://www.cbsnews.com/stories/2005/09/30/opinion/polls/main894733.shtml, Oct. 2005.

[10] Privacy Int'l, *National Privacy Ranking 2006—European Union and Leading Surveillance Societies,* 2006.

[11] *TAUCIS—Technikfolgenabschätzungsstudie Ubiquitäres Computing und Informationelle Selbstbestimmung,* J. Bizer et al., eds. 2006.

[12] S. Spiekermann et al., "E-Privacy in 2nd Generation E-Commerce," *Proc. Third ACM Conf. Electronic Commerce,* 2001.

[13] A. Acquisti and J. Grossklags, "Privacy and Rationality in Individual Decision Making," *IEEE Security & Privacy,* vol. 2, pp. 24-30, 2005.

[14] Privacy & American Business, "New Survey Reports an Increase in ID Theft and Decrease in Consumer Confidence," http://www.pandab.org/deloitteidsurveypr.html, May 2005.

[15] S. Spiekermann, "Acceptance of Ubiquitous Computing Services: About the Importance of Human Control," presentation at the Carnegie Mellon Univ. Heinz School of Public Policy and Management, Pittsburgh, 2006.

[16] J. Tsai, S. Egelman, L. Cranor, and A. Acquisti, "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study," *Proc. Workshop Economics of Information Security,* June 2007.

[17] S. Lahlou, M. Langheinrich, and C. Röcker, "Privacy and Trust Issues with Invisible Computers," *Comm. ACM,* vol. 48, no. 3, pp. 59-60, 2005.

[18] A. Whitten and J.D. Tygar, "Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0," *Proc. Eighth USENIX Security Symp.,* Aug. 1999.

[19] R. Dingledine and N. Mathewson, "Anonymity Loves Company: Usability and the Network Effect," *Security and Usability: Designing Secure Systems that People Can Use,* L. Cranor and S. Garfinkel, eds., pp. 547-559, 2005.

[20] O. Berthold et al., "Web MIXes: A System for Anonymous and Unobservable Internet Access," *Proc. Int'l Workshop Design Issues in Anonymity and Unobservability.* 2001.

[21] R. Dingledine et al., "Tor: The Second-Generation Onion Router," *Proc. 12th USENIX Security Symp.,* 2004.

[22] P. Golle, F. McSherry, and I. Mironov, "Data Collection with Self-Enforcing Privacy," *Proc. 13th ACM Conf. Computer and Comm. Security,* pp. 69-78, http://doi.acm.org/10.1145/1180405.1180416, Oct./Nov. 2006.

[23] J. Feigenbaum, M.J. Freedman, T. Sander, and A. Shostack, "Privacy Engineering for Digital Rights Management Systems," *Revised Papers from the ACM CCS-8 Workshop Security and Privacy in Digital Rights Management,* pp. 76-105, T. Sander, ed., 2002.

[24] B. Friedman, I.E. Smith et al. "Development of a Privacy Addendum for Open Source Licenses: Value Sensitive Design in Industry," *Proc. Eighth Int'l Conf. Ubiquitous Computing,* 2006.

[25] L.F. Cranor, P. Guduru, and M. Arjula, "User Interfaces for Privacy Agents," *ACM Trans. Computer-Human Interaction,* vol. 13, no. 2, pp. 135-178, http://doi.acm.org/10.1145/1165734.1165735, June 2006.

[26] J.B. Earp, A.I. Antón, and O. Jarvinen, "A Social, Technical and Legal Framework for Privacy Management and Policies," *Proc. Americas Conf. Information Systems,* 2002.

[27] J.I. Hong, J.D. Ng, S. Lederer, and J.A. Landay, "Privacy Risk Models for Designing Privacy-Sensitive Ubiquitous Computing Systems," *Proc. Fifth Conf. Designing Interactive Systems: Processes, Practices, Methods, and Techniques,* pp. 91-100, http://doi.acm.org/10.1145/1013115.1013129, Aug. 2004.

[28] D. Warren and L. Brandeis, "The Right to Privacy," *Harvard Law Rev.,* vol. 45, 1890.

[29] R.A. Posner, "Privacy—A Legal Analysis," *Philosophical Dimensions of Privacy,* F.D. Schoeman, ed., 1984.

[30] I. Altman, *The Environment and Social Behavior: Privacy, Personal Space, Territory, Crowding.* Brooks/Cole, 1975.

[31] A.F. Westin, *Privacy and Freedom.* Atheneum, 1967.

[32] L.F. Cranor, "Privacy Policies and Privacy Preferences," *Security and Usability: Designing Secure Systems that People Can Use,* L. Cranor and S. Garfinkel, eds., 2005.

[33] D.J. Solove, "A Taxonomy of Privacy," *Univ. of Pennsylvania Law Rev.,* vol. 154, 2005.

[34] Heise Online, "Datenschützer: Google's Mail-Service in Deutschland unzulässig,"http://www.heise.de/newsticker/meldung/46383, 2004.

[35] J.A. Hoffer et al., *Modern Systems Analysis and Design.* Prentice Hall, 2002.

[36] J.C. Cannon, *Privacy: What Developers and IT Professionals Should Know.* Addison-Wesley Professional, 2004.

[37] L.F. Cranor, "I Didn't Buy It for Myself," *Designing Personalized User Experiences in E-Commerce,* C.-M. Karat, J.O. Blom, and J. Karat, eds., Kluwer Academic Publishers, 2004.

[38] A. Adams and A. Sasse, "Taming the Wolf in Sheep's Clothing: Privacy in Multimedia Communications," *Proc. Seventh ACM Int'l Multimedia Conf.,* 1999.

[39] A. Adams and A. Sasse, "Privacy in Multimedia Communications: Protecting Users, Not Just Data," *People and Computers XV—Interaction without Frontiers,* J. Blandford, J. Vanderdonkt, and P. Gray, eds., pp. 49-64, Springer, 2001.

[40] H. Nissenbaum, "Privacy as Contextual Integrity," *Washington Law Rev.,* vol. 791, pp. 119-158, 2004.

[41] S. Byers, "Information Leakage Caused by Hidden Data in Published Documents," *IEEE Security & Privacy,* vol. 2, no. 2, pp. 23-27, 2004.

[42] G.J. Nowag and J. Phelps, "Direct Marketing and the Use of Individual-Level Consumer Information: Determining How and When Privacy Matters," *J. Direct Marketing,* vol. 93, pp. 46-60, 1995.

[43] "Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data," *J. European Communities,* vol. 281, no. 31, 1995.

[44] *The Privacy Law Sourcebook 2004: United States Law, International Law, and Recent Developments,* M. Rotenberg, ed. EPIC, 2004.

[45] P.N. Otto, A.I. Anton, and D.L. Baumer, "The ChoicePoint Dilemma: How Data Brokers Should Handle the Privacy of Personal Information," *IEEE Security & Privacy,* vol. 5, no. 5, pp. 15-23, Sept./Oct. 2007, doi: http://doi.ieeecomputersociety.org/10.1109/MSP.2007.126.

[46] S. Spiekermann et al., "Stated Privacy Preferences versus Actual Behaviour in EC Environments: A Reality Check," *Proc. Fifth Internationale Tagung Wirtschaftsinformatik,* 2001.

[47] B. Berendt et al., "Privacy in E-Commerce: Stated Preferences versus Actual Behavior," *Comm. ACM,* vol. 484, pp. 101-106, 2005.

[48] P.M. Regan, *Legislating Privacy: Technology, Social Values, and Public Policy.* Univ. of North Carolina Press, 1995.

[49] K.B. Sheehan, "Toward a Typology of Internet Users and Online Privacy Concerns," *The Information Soc.,* vol. 1821, pp. 21-32, 2002.

[50] M. Brown and R. Muchira, "Investigating the Relationship between Internet Privacy Concerns and Online Purchase Behavior," *J. Electronic Commerce Research,* vol. 5, no. 1, pp. 62-70, 2004.

[51] N. Malhotra, S.S. Kim, and J. Agarwal, "Internet Users' Information Privacy Concerns IUIPC: The Construct, the Scale, and a Causal Model," *Information Systems Research,* vol. 15, no. 4, pp. 336-355, 2004.

[52] M.S. Ackerman, L.F. Cranor, and J. Reagle, "Privacy in E-Commerce: Examining User Scenarios and Privacy Preferences," *Proc. First ACM Conf. Electronic Commerce,* pp. 1-8, http://doi.acm.org/10.1145/336992.336995, Nov. 1999.

[53] J.H. Smith et al., "Information Privacy: Measuring Individuals' Concerns about Organizational Practices," *MIS Quarterly,* vol. 202, pp. 167-196, 1996.

[54] S. Garfinkel and A. Shelat, "Remembrance of Data Passed: A Study of Disk Sanitization Practices," *IEEE Security & Privacy,* Jan./Feb. 2003.

[55] N.F. Awad and K. Fitzgerald, "The Deceptive Behaviors that Offend Us Most about Spyware," *Comm. ACM,* vol. 48, pp. 55-60, http://doi.acm.org/10.1145/1076211.1076240, Aug. 2005.

[56] O. Berthold et al., "RFID Verbraucherängste und Verbraucherschutz," *Wirtschaftsinformatik Heft,* vol. 6, 2005.

[57] O. Guenther and S. Spiekermann, "RFID and Perceived Control—The Consumer's View," *Comm. ACM,* vol. 489, pp. 73-76, 2005.

[58] S.M. Edwards, H. Li et al., "Forced Exposure and Psychological Reactance: Antecedents and Consequences of the Perceived Intrusiveness of Pop-Up Ads," *J. Advertising,* vol. 313, pp. 83-96, 2002.

[59] S. Spiekermann, "The Desire for Privacy: Insights into the Views and Nature of the Early Adopters of Privacy Services," *Int'l J. Technology and Human Interaction,* vol. 11, 2004.
[60] D. Spiegel, "Exhibitionismus—leichtgemacht," *Der Spiegel,* vol. 29, 2006.
[61] P. Kumaraguru and L. Cranor, "Privacy Indexes: A Survey of Westin's Studies," ISRI Technical Report CMU-ISRI-05-138, http://reports-archive.adm.cs.cmu.edu/anon/isri2005/abstracts/05-138.html, 2005.
[62] S. Spiekermann, "Auswirkungen der UC-Technologie auf Verbraucher: Chancen und Risiken," *Technikfolgenabschätzung Ubiquitäres Computing und Informationelle Selbstbestimmung TAUCIS,* J. Bizer, O. Guenther, and S. Spiekermann, eds. Berlin, Bundesministerium für Bildung und Forschung BMBF, pp. 153-196, 2006.
[63] G. Löwenstein and D. Prelect, "Anomalies in Intertemporal Choice: Evidence and an Interpretation," *Choices, Values, and Frames,* D. Kahneman and A. Tversky, eds., pp. 578-596, Cambridge Univ. Press, 2000.
[64] A. Acquisti, "Privacy in Electronic Commerce and the Economics of Immediate Gratification," *Proc. Fifth ACM Conf. Electronic Commerce,* pp. 21-29, http://doi.acm.org/10.1145/988772.988777, May 2004.
[65] H. Varian, "Economic Aspects of Personal Privacy," *Privacy and Self-Regulation in the Information Age,* 1996.
[66] K. Strandburg, *Privacy, Rationality, and Temptation: A Theory of Willpower Norms.* College of Law, DePaul Univ., 2005.
[67] B. Huberman et al., "Valuating Privacy," *IEEE Security & Privacy,* vol. 1, pp. 22-25, 2004.
[68] F.B. Viégas, "Bloggers' Expectations of Privacy and Accountability: An Initial Survey," *J. Computer-Mediated Comm.,* vol. 103, 2005.
[69] V. Mayer-Schönberger, *Useful Void: The Art of Forgetting in the Age of Ubiquitous Computing.* John F. Kennedy School of Government, Harvard Univ., 2007.
[70] M. Gumbrecht, "Blogs as 'Protected Space'," *Proc. Workshop Weblogging Ecosystem: Aggregation, Analysis, and Dynamics at the World Wide Web Conf.,* 2004.
[71] J. Gideon, L. Cranor, S. Egelman, and A. Acquisti, "Power Strips, Prophylactics, and Privacy, Oh My," *Proc. Second Symp. Usable Privacy and Security,* vol. 149, pp. 133-144, http://doi.acm.org/10.1145/1143120.1143137, July 2006.
[72] N.F. Awad and M.S. Krishnan, "The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to be Profiled Online for Personalization," *MIS Quarterly,* vol. 301, pp. 13-28, 2006.
[73] S. Lederer and A. Dey, *A Conceptual Model and a Metaphor of Everyday Privacy in Ubiquitous Computing Environments.* Univ. of California, Berkeley, 2002.
[74] *Richtlinie Des Europäischen Parlaments Und Des Rates über die Vorratsspeicherung von Daten, die bei der Bereitstellung öffentlich zugänglicher elektronischer Kommunikationsdienste oder öffentlicher Kommunikationsnetze erzeugt oder verarbeitet werden, und zur Änderung der Richtlinie 2002/58/EG,* Brussels, European Parlament2005/0182 COD, EU, 2006.
[75] S. Garfinkel, *Database Nation—The Death of Privacy in the 21st Century.* O'Reilly & Assoc., 2000.
[76] R. Agrawal and R. Srikant, "Privacy-Preserving Data Mining," *SIGMOD Record,* vol. 29, no. 2, pp. 439-450, June 2000, doi: http://doi.acm.org/10.1145/335191.335438.
[77] L. F. Cranor, S. Egelman, J. Hong, P. Kumaraguru, C. Kuo, S. Romanosky, J. Tsai, and K. Vaniea, "FoxTor: A Tor Design Proposal," http://cups.cs.cmu.edu/pubs/TorGUIContest113005.pdf, 2005.
[78] OECD, "OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data," http://www.oecd.org/document/18/0,2340,en_2649_201185_1815186_1_1_1_1,00.html, 1980.
[79] US Federal Trade Commission, *Privacy Online: Fair Information Practices in the Electronic Marketplace, A Report to Congress,* http://www.ftc.gov/reports/privacy2000/privacy2000.pdf, 2000.
[80] D. Chaum, "Security without Identification: Transaction Systems to Make Big Brother Obsolete," *Comm. ACM,* vol. 28, no. 10, 1030-1044, http://doi.acm.org/10.1145/4372.4373, Oct. 1985.
[81] J. Camenisch and E. Van Herreweghen, "Design and Implementation of the 'Idemix' Anonymous Credential System," *Proc. Ninth ACM Conf. Computer and Comm. Security,* V. Atluri, ed., pp. 21-30, http://doi.acm.org/10.1145/586110.586114, Nov. 2002.
[82] B. Pfitzmann, M. Waidner et al. "Rechtssicherheit Trotz Anonymität in Offenen Digitalen Systemen; Datenschutz und Datensicherung," *Datenschutz und Datensicherheit DuD,* vol. 14, pp. 5-6, 1990.
[83] M.K. Reiter and A.D. Rubin, "Anonymous Web Transactions with Crowds," *Comm. ACM,* vol. 42, no. 2, pp. 32-48, http://doi.acm.org/10.1145/293411.293778, Feb. 1999.
[84] J.I. Hong and J.A. Landay, "An Architecture for Privacy-Sensitive Ubiquitous Computing," *Proc. Second Int'l Conf. Mobile Systems, Applications, and Services,* pp. 177-189, http://doi.acm.org/10.1145/990064.990087, June 2004.
[85] A. LaMarca, Y. Chawathe, S. Consolvo, J. Hightower, I. Smith, J. Scott, T. Sohn, J. Howard, J. Hughes, F. Potter, J. Tabert, P. Powledge, G. Borriello, and B. Schilit, "Place Lab: Device Positioning Using Radio Beacons in the Wild," *Proc. Pervasive 2005,* 2005.
[86] J. Canny, "Collaborative Filtering with Privacy," *Proc. IEEE Symp. Security and Privacy,* pp. 45-57, 2002.
[87] J. Zibuschka, L. Fritsch, M. Radmacher, T. Scherner, and K. Rannenberg, "Enabling Privacy in Real-Life LBS: A Platform for Flexible Mobile Service Provisioning," *New Approaches for Security, Privacy and Trust in Complex Environments,* H. Venter, M. Eloff, L. Labuschagne, J. Eloff, and R. von Solms, eds., IFIP Int'l Federation for Information Processing, vol. 232, pp. 325-336, Springer, 2008.
[88] M. Gruteser and D. Grunwald, "Enhancing Location Privacy in Wireless LAN through Disposable Interface Identifiers: A Quantitative Analysis," *Mobile Networks and Applications,* vol. 10, no. 3, pp. 315-325, June 2005, doi: http://doi.acm.org/10.1145/1145911.1145917.
[89] S. Spiekermann et al., "User Agents in E-Commerce Environments: Industry versus Consumer Perspectives on Data Exchange," *Proc. 15th Conf. Advanced Information Systems Eng.,* 2003.
[90] A. Kobsa and J. Schreck, "Privacy through Pseudonymity in User-Adaptive Systems," *ACM Trans. Internet Technology,* vol. 3, no. 2, pp. 149-183, http://doi.acm.org/10.1145/767193.767196, May 2003.
[91] A. Kobsa, "Privacy-Enhanced Web Personalization," *The Adaptive Web: Methods and Strategies of Web Personalization,* P. Bruslikovsky, A. Kobsa, and W. Nejdl, eds., Springer Verlag, 2007.
[92] M. Barbaro and T. Zeller, "A Face is Exposed for AOL Searcher No. 4417749," *The New York Times,* 9 Aug. 2006.
[93] E. Mills and A. Broache, "Three Workers Depart AOL after Privacy Uproar," *CNET News.com,* 21 Aug. 2006.
[94] A. Pfitzmann and M. Hansen, *Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management—A Consolidated Proposal for Terminology Version v0.30,* http://dud.inf.tu-dresden.de/Anon_Terminology.shtml, Nov. 2007.
[95] A. Narayanan and S. Vitaly, "Robust De-Anonymization of Large Sparse Datasets," *Proc. IEEE Symp. Security and Privacy,* 2008.
[96] B. Malin, "Betrayed by My Shadow: Learning Data Identity Via Trail Matching," *J. Privacy Technology,* http://www.jopt.org/publications/20050609001_malin_abstract.html, 2005.
[97] L. Sweeney, "k-Anonymity: A Model for Protecting Privacy," *Int'l J. Uncertainty, Fuzziness and Knowledge-Based Systems,* vol. 10, no. 5, pp. 557-570, 2002.
[98] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, "L-Diversity: Privacy Beyond k-Anonymity," *ACM Trans. Knowledge Discovery from Data,* vol. 1, no. 3, Mar. 2007, doi: http://doi.acm.org/10.1145/1217299.1217302.
[99] A.K. Ghosh, *Security and Privacy for E-Business.* John Wiley & Sons, 2001.
[100] R.J. Anderson, *Security Engineering: A Guide to Building Dependable Distributed Systems.* John Wiley & Sons, 2001.
[101] "The Center for Information Policy Leadership," *Multi-Layered Notices Explained,* http://www.hunton.com/files/tbl_s47Details/FileUpload265/1303/CIPL-APEC_Notices_White_Paper.pdf, Jan. 2007.
[102] L.F. Cranor, *Web Privacy with P3P.* O'Reilly, 2002.
[103] I. Pollach, "What's Wrong with Online Privacy Policies?" *Comm. ACM,* vol. 50, no. 9, pp. 103-108, http://doi.acm.org/10.1145/1284621.1284627, Sept. 2007.
[104] Microsoft, *Privacy Guidelines for Developing Software Products and Services—Version 1.0,* 2006.
[105] D. McFarlane, "Comparison of Four Primary Methods for Coordinating the Interruption of People in Human-Computer Interaction," *Human-Computer Interaction,* vol. 173, pp. 63-139, 2002.

[106] G. Hsieh, K.P. Tang, W.Y. Low, and J.I. Hong, "Field Deployment of 'IMBuddy': A Study of Privacy Control and Feedback Mechanisms for Contextual IM," *Proc. Ninth Int'l Conf. Ubiquitous Computing*, pp. 91-108, 2007.
[107] J. Cornwell, I. Fette, G. Hsieh, M. Prabaker, J. Rao, K. Tang, K. Vaniea, L. Bauer, L. Cranor, J. Hong, B. McLaren, M. Reiter, and N. Sadeh, "User-Controllable Security and Privacy for Pervasive Computing," *Proc. Eighth IEEE Workshop Mobile Computing Systems and Applications*, 2007.
[108] M. Prabaker, J. Rao, I. Fette, P. Kelley, L. Cranor, J. Hong, and N. Sadeh, "Understanding and Capturing People's Privacy Policies in a People Finder Application," *Proc. Workshop Ubicomp Privacy*, Sept. 2007.
[109] Information Commissioner's Office, "What Price Privacy? The Unlawful Trade in Confidential Personal Information," http://www.ico.gov.uk/upload/documents/library/corporate/research_and_reports/what_price_privacy.pdf, 2006.
[110] D. McCullough and A. Broache, "HP Scandal Reviving Pretexting Legislation," *CNET News.com*, 15 Sept. 2006.
[111] "Your Grocery Purchases on the Web for All to See?" *Privacy J.*, vol. 27, no. 5, pp. 1-7, Mar. 2001.
[112] Vodafone, *Vodafone Location Services—Privacy Management Code of Practice*, 2003.
[113] A. Barth and J.C. Mitchell, "Enterprise Privacy Promises and Enforcement," *Proc. Workshop Issues in the Theory of Security*, pp. 58-66, http://doi.acm.org/10.1145/1045405.1045412, Jan. 2005.
[114] C. Powers and M. Schunter, "Enterprise Privacy Authorization Language EPAL 1.2," *W3C Member Submission 10*, http://www.w3.org/Submission/EPAL/, Nov. 2003.
[115] M. Deng, L. Fritsch, and K. Kursawe, "Personal Rights Management," *Proc. Sixth Workshop Privacy-Enhancing Technologies*, G. Danezis and P. Golle, eds., 2006.
[116] J. Kang, "Information Privacy in Cyberspace Transactions," *Stanford Law Rev.*, vol. 50, pp. 1194-1294, 1998.

Sarah Spiekermann is a professor of information systems at Humboldt University Berlin and at the European Business School (Schloss Reichartshausen). Her research focus is value sensitive and behaviorally informed system design. Before she joined academia in 2003, she worked as a business consultant for A.T. Kearney and led the European business intelligence for Openwave Systems. http://amor.rz.hu-berlin.de/~spiekers/Homepage.htm.

Lorrie Faith Cranor is an associate professor in the School of Computer Science and the Department of Engineering and Public Policy at Carnegie Mellon University, where she is the director of the CMU Usable Privacy and Security Laboratory (CUPS). She was previously a researcher at AT&T-Labs Research. She is a senior member of the IEEE and the ACM. http://lorrie.cranor.org.