In: P. Brusilovsky, A. Kobsa, W. Nejdl, eds.: The Adaptive Web: Methods and Strategies of Web Personalization. Berlin, Heidelberg, New York: Springer Verlag, 628-670. This is the author's version of the work. The definitive work is DOI 10.1007/978-3-540-72079-9_4, © Springer Verlag.

21

Privacy-Enhanced Web Personalization

Alfred Kobsa

Donald Bren School of Information and Computer Sciences
University of California, Irvine
Irvine, CA 92697-3440, U.S.A.
kobsa@uci.edu
http://www.ics.uci.edu/~kobsa

Abstract. Consumer studies demonstrate that online users value personalized content. At the same time, providing personalization on websites seems quite profitable for web vendors. This win-win situation is however marred by privacy concerns since personalizing people's interaction entails gathering considerable amounts of data about them. As numerous recent surveys have consistently demonstrated, computer users are very concerned about their privacy on the Internet. Moreover, the collection of personal data is also subject to legal regulations in many countries and states. Both user concerns and privacy regulations impact frequently used personalization methods. This article analyzes the tension between personalization and privacy, and presents approaches to reconcile the both.

21.1 Introduction

It has been tacitly acknowledged for many years that personalized interaction and user modeling have significant privacy implications, due to the fact that large amounts of personal information about users needs to be collected to perform personalization. For instance, frequent users of a search engine may appreciate that their search terms become recorded to disambiguate future queries and deliver results that are better geared towards their interests (see Chapter 6 of this book [130]). They may not appreciate though if their search history from the past few years becomes accessible to others. Secretaries may value if the help component of their text editor can give personalized advice based on a model of their individual word-processing skills that it built over time by watching how they interact with the word processor [113]. They are however likely to be concerned if the contents of their model becomes accessible to others, specifically if negative consequences may arise from a disclosure of the skills they lack. Other potential privacy concerns in the context of personalized systems include (see [37]): unsolicited marketing, computer "figuring things out" about the user [197], fear of price discrimination, information being revealed to other users of

the same computer, unauthorized access to accounts, subpoenas by courts, and government surveillance.

Kobsa [104, 105] was arguably the first to point out the tension between personalization and privacy nearly twenty years ago, but without much impact. The only privacy solution of these days was to ascertain that users could store their models on a diskette (R. Oppermann, GMD) or a PCMCIA card (J. Orwant, MIT [143]), and carry them with. The situation changed completely in the late 1990s, for four main reasons:

Personalized systems moved to the web. Web retailers quickly realized the enormous potential of personalization for customer relationship management and made their websites user-adaptive. This had significant privacy implications. While user models were previously confined to stand-alone machines or local networks, people's profiles were now collected on dozens if not hundreds of personalized websites. Widely publicized security glitches and privacy breaches as well as aggressive telemarketing led to a widespread (~60-80%) stated reluctance of Internet users to disclose personal data and being tracked online. This reluctance however endangers the basic foundations of personalization, which highly relies on such data [180].

More sources of user data available. While in the 1980s the main source of user modeling was nearly exclusively textual data entered by the user, assumptions about users can nowadays be drawn from, e.g., their mouse movements, mouse clicks, eye movements, facial expression, physiological data and location data. Completely new privacy threats arose in ubiquitous computing environments where users are no longer merely IP addresses in an abstract online space, but become identified individuals who are being monitored and contacted by their physical surroundings.

More powerful analyses of user data available. More powerful computers, computer networks, sensors and algorithms have made it possible to collect, connect and analyze far more data about users than ever before. Complete digital lifetime archives replete with personal data may soon become reality.

Restrictions imposed by privacy legislation. Many more countries, states and provinces have meanwhile introduced privacy laws, which severely affect not only commercial websites but also experimental research on user modeling (in many cases even when it is done "just with IP numbers", "just with our students", or "just for testing purposes") [107]. Areas that are specifically affected include data mining for personalization purposes (see Chapter 3 of this book [135]), adaptive tutoring systems that build learner models (see Chapter 1 of this book [23]), and adaptation to the needs of people with special needs [58, 174].

Consumer studies demonstrate that online users value personalized content [32, 149, 178]. At the same time, providing personalization on websites seems quite profitable for web vendors [11, 34, 77, 86]. This win-win situation is however marred by privacy concerns since personalizing people's interaction entails gathering considerable amounts of data about them. As a consequence, the topic of "Privacy and Personalization" has received considerable attention from industry and academia in the past few years. Three industry conferences with this title were held in 2000-01 (in New York, London and San Francisco) with a participation of about 150 people. The

Ubiquitous Computing conferences have held privacy workshops in the past four years[1] that address, among other topics, privacy in context-aware systems. In July 2005, twenty researchers met for a first workshop on Privacy-Enhanced Personalization at the 10th International User Modeling Conference in Edinburgh, Scotland [108], and one year later for a second workshop at the CHI-2006 conference in Montréal, Canada [130].

The aim of research on privacy-enhanced personalization is to reconcile the goals and methods of user modeling and personalization with privacy considerations, and to strive for best possible personalization within the boundaries set by privacy. This research field is widely interdisciplinary, with contributions coming from information systems, marketing research, public policy, economics, computer-mediated communication, law, human-computer interaction, and the information and computer sciences. This chapter analyzes how current research results on privacy in electronic environments relate to the aims of privacy-enhanced personalization. It first discusses the impact of Internet users' privacy concerns on the disclosure of personal data. Section 21.3 then reviews the current state of research on factors that contribute to alleviating privacy concerns and to encouraging the disclosure of personal data. Section 21.4 analyzes the impact of privacy regulation on personalized systems (specifically of privacy legislation, but also industry and company self-regulation as well as principles of fair information practices). Section 21.5 finally describes privacy-enhancing technical solutions that are particularly well suited for personalized systems.

As we will see throughout these discussions, there exists no magic bullet for making personalized systems privacy-enhanced, neither technical nor legal nor social/organizational. Instead, numerous small enhancements need to be introduced, which depend on the application domain as well as the types of data, users and personalization goals involved. At the end of most sections and subsections, we will list the lessons for the privacy-minded design of personalized systems that ensue from the research results discussed in the respective section. In a concrete project, though, the applicability of these recommendations will still need to be verified as part of the normal interaction design and user evaluation process [155, 175].

21.2 Individuals' Privacy Concerns

21.2.1 Methodological Preliminaries

This section analyzes empirical results regarding people's privacy-related attitudes, and the subsequent section known motivators and deterrents for people disclosing personal information to websites. Two principal types of empirical methods are available for identifying such attitudes, motivators and deterrents:

1. *Inquiry-based methods*. In this approach, the participants of an empirical study are being asked about their privacy attitudes ("reported/perceived attitudes"), their disclosure behavior in the past ("reported/perceived behavior"), and their antici-

[1] See the links at http://www.cs.berkeley.edu/~jfc/privacy/

pated disclosure behavior under certain privacy-related circumstances ("stated behavioral intentions"). In the third case, these privacy-related circumstances can be merely described to subjects, or one can try to immerse subjects in them as much as possible (e.g. by showing them a website with characteristics that may be important in subjects' disclosure decisions).
2. *Observation-based methods*. In this approach, the privacy-related behavior of participants is being observed during the empirical study. Subjects are put into a situation that resembles the studied circumstances as much as possible (usually a lab experiment, ideally a field experiment), and they have to exhibit privacy-related behavior therein (e.g., disclose their own personal data while purchasing products) rather than merely answer questions about their likely behavior.

Both approaches have complementary strengths and weaknesses, and mixes of both approaches are therefore customary. Inquiry-based methods do not directly unveil people's actual privacy-related attitudes and disclosure behavior, but only their perception thereof, which may not be in sync with reality. Observation-based methods on the other hand often do not allow one to recognize people's higher-level behavioral patterns or rationale, which in return can be more easily accessed through inquiries. In addition, both approaches are equally subject to various potential biases that must be eliminated through careful experimental design (see e.g. [141]).

In the area of privacy, other factors also seem to come into play that may skew the results of empirical studies. Such known or suspected factors include the following:

1. *Biased self-selection*. It may be the case that predominantly those people volunteer to participate in a privacy study, or take the pains to complete it until the very end, for whom privacy is a personal concern. This may bias the responses towards higher concerns.
2. *Socially desirable responses*. It may be the case in privacy studies that subjects tend to respond and act in ways that are deemed socially desirable. For instance, in times of ever-increasing identity theft this bias may skew responses towards higher concerns since not having privacy concerns might be viewed as displaying a lack of prudence and responsibility.
3. *Discrepancies between stated attitudes and observed behavior*. In several privacy studies in e-commerce contexts, discrepancies have already been observed between users stating high privacy concerns but subsequently disclosing personal data carelessly [17, 128, 172]. Several authors therefore challenge the genuineness of such reported privacy attitudes [80, 90] and emphasize the need for experiments that allow for an observation of actual online disclosure behavior [128, 167].

It seems possible to eliminate the first two sources of bias through careful experimental design and post-hoc recalibration of socially desirable responses. The discrepancies between stated privacy attitudes and observed disclosure behavior will be discussed in more detail in Sections 21.2.5 and 21.3.5. For the time being, it seems useful though to clearly distinguish whether an experimental finding stems from the observation of actual human disclosure behavior in an experiment, or is based on subjects' reports of attitudes, past behavior or behavioral intentions. We will therefore

introduce the convention of marking findings of the first kind with an asterisk (*) in the remainder of this chapter.[2]

21.2.2 Potential Effects of Privacy Concerns on Personalized Systems

Numerous consumer surveys and research studies have revealed that Internet users harbor considerable privacy concerns regarding the disclosure of their personal data to websites, and the monitoring of their Internet activities. These studies were primarily conducted between 1998 and 2003, mostly in the United States (see [156] for an incomplete listing). In the following, we summarize a few important findings (the percentage figures indicate the ratio of respondents who adopted the respective view). For a more detailed discussion we refer to [180].

Personal data

1. Internet users who are concerned about the privacy or security of their personal information online: 70% [15], 83% [196], 89.5% [187], 84% [147];
2. People who have refused to give personal information to a web site at one time or another: 95% [87], 83% [82], 82% [44];
3. Internet users who would never provide personal information to a web site: 27% [63];
4. Internet users who supplied false or fictitious information to a web site when asked to register: 40% [87]; 34% [44]; 24% [63]; 15% more than half of the time [167]; 6% always, 7% often, 17% sometimes [163]; 48.9% never, 24.1% a quarter or less of the time, 18% between ¼ and over ¾ of the time [76]; 19.4% in an experiment (half of them multiple times) [127]; 39.6% in an experiment (2-3 items on average, and the likelihood of falsification was correlated with the stated sensitivity of the item).
5. People who are concerned if a business shares their data for a purpose that is different from the one for which they were originally collected: 90% [163], 89% [147];
6. Online users who believe that sites that share personal information with other sites invade privacy: 83% [45].

Significant concern about the use of personal data is visible in these results, which may cause problems for those personalized systems that depend on users disclosing data about themselves. More than a quarter of respondents stated that they would never consider providing personal information to a web site. Quite a few users indicated having supplied false or fictitious information to a web site when asked to register, which makes all personalization based on such data dubious, and may also jeopardize cross-session identification of users as well as all personalization based thereon. Furthermore, 80-90% of the respondents are concerned if a business shares their information for a different than the original purpose. This may have severe impacts on central user modeling servers that collect data from, and share them with, different user-adaptive applications (see Chapter 4 of this book [130]).

---

[2] Note that a cited article may both describe observation-based findings (which will be marked with an asterisk) and findings that are based on subjects' reports (which will not).

User tracking and cookies

1. People who are concerned about being tracked on the Internet: 54% [63], 63% [82], 62% [147];
2. People who are concerned that someone might know what web sites they visited: 31% [63];
3. Users who feel uncomfortable being tracked across multiple web sites: 91% [82];
4. Internet users who generally accept cookies: 62% [148];
5. Internet users who set their computers to reject cookies: 25% [44], 10% [63]; and
6. Internet users who delete cookies periodically: 53% [148].

These results reveal significant user concerns about tracking and cookies, which may have effects on the acceptance of personalization that is based on usage logs. Observations 3–6 directly affect machine-learning methods that operate on user log data since without cookies or registration, different sessions of the same user can no longer be linked. Observation 3 may again affect the acceptance of the above-mentioned user modeling servers which collect user information from several websites (see Chapter 4 of this book [130]).

These survey results indicate that privacy concerns may indeed severely impede the adoption of personalized web-based systems. As a consequence, personalized systems may become less used, personalization features may become switched off if this is an option, fewer personal information may become disclosed, and escape strategies may be adopted such as submitting falsified data, maintaining multiple accounts/identities, deleting cookies, etc. However, developers of personalized web-based systems should not feel completely discouraged by the abundance of stated privacy concerns in consumer surveys. As we will see, privacy concerns are only one of many factors that influence whether and to what extent people disclose data about themselves and utilize personalized systems. In Section 21.3 we will discuss numerous factors that can seemingly mitigate users' privacy concerns and prompt them to nevertheless disclose personal data about themselves. Designers of personalized systems will have to carefully analyze users' privacy concerns in their application domain and address them, but also consider those mitigating factors and ascertain that as many of them as possible are present in the design of their systems.

21.2.3 Effect of Information Type

Not surprisingly, many surveys indicate that users' willingness to disclose personal information also depends on the kind of information in question. For instance,

- Ackerman et al. [1] found that the vast majority of their respondents always or usually felt comfortable providing information about their own preferences, including favorite television show (82%) and favorite snack food (80%). In contrast, only a very small number said they would usually feel comfortable providing their credit card number (3%) or social security number (1%). The figures decreased in all categories if the data was about subjects' children and not about themselves.

- Phelps [151] found that consumers are more willing to provide marketers with demographic and lifestyle information than with financial, purchase-related, and personal identifier information. The vast majority of respondents were always or somewhat willing to share their two favorite hobbies, age, marital status, occupation or type of job, and education.
- Metzger [127] found that participants of her experiment "were most willing to provide basic demographic information (e.g., sex, age, education level, marital status), and slightly less willing to provide information about their actual online behavior (past purchases time spent online), religion, political party identification, race, hobbies/interests, and occupation. Respondents were by far most protective of their personal contact information (telephone number and email address) and financial information (credit card number, social security number, and income)."*
- In a different experiment, Metzger [126] found that participants were most likely to withhold their credit card and social security numbers. The next-most withheld items included email address, telephone number, favorite website, hobbies/interests, and last purchase made online, income and political party affiliation. Participants were least likely to withhold general demographic information about themselves, for example, their sex, race, education, marital status, time spent online, number of people in their household, and age. Name and address were given out most frequently, but those were required for receiving a free CD.

An experiment by Huberman et al. [88] suggests that not only different data categories, but also different values within the same category may have different privacy valuations*. A group of experimental subjects participated in a reverse auction for the disclosure of certain personal information to all others (namely of individuals' age, weight, salary, spousal salary, credit rating and amount of savings). The anonymously submitted asking prices for this personal data turned out to be a (largely linear) function of the deviance of the data values from the socially desirable standard (this holds true both for individually perceived and actual deviance)*. The results seem to indicate that the more undesirable a trait is with respect to the group norm, the higher is its privacy valuation.

The lesson from these findings for the design of personalized web-based systems seems that highly sensitive data categories should never be requested without the presence of some of the mitigating factors that will be discussed later. To lower privacy concerns for data values that are possibly highly deviant, one-sided open intervals should be considered whose closed boundary does not deviate too much from the expected norm (such as "weight: 250 pounds and above" for male adults).

* An asterisk indicates that the data is based on an observation of human privacy-related behavior in an experiment rather than a survey of stated attitudes, reported past behavior, or stated behavioral intentions (see Section 24.2.1 for a more detailed explanation).

21.2.4 Interpersonal Differences in Privacy Attitudes

Various studies established that age [52, 132], education [151] and income [3] are positively associated with the degree of stated Internet privacy concern. Smith et al. [170] also found that people who were victims of a perceived privacy invasion or had heard of one had higher privacy concerns. Gender effects on Internet privacy concerns could not be clearly established so far.

In a broad privacy survey that was first conducted in 1991 [81] and since then repeated several times, Harris Interactive and Alan Westin clustered respondents into three groups, namely privacy fundamentalists, the privacy unconcerned, and privacy pragmatists. Privacy fundamentalists generally express extreme concern about any use of their data and unwillingness to disclose them, even when privacy protection mechanisms would be in place. In contrast, the privacy unconcerned tend to express mild concern for privacy only, and also mild anxiety about how other people and organizations use information about them. Privacy pragmatists as the third group are generally concerned about their privacy as well. In contrast to the fundamentalists though, their privacy concerns are lower and they are far more willing to disclose personal information, e.g. when they understand the reasons for its use, when they see benefits for doing so, or when they see privacy protections in place.

In the latest edition of this survey in 2003[3], privacy fundamentalists comprise about 26% of all adults, the privacy unconcerned about 10%, and the privacy pragmatists 64% [179]. Previous editions and other studies yield slightly different figures and/or clusters. For instance, the clustering of the responses in [1] resulted in 17% privacy fundamentalists, 27% "marginally concerned", and 56% members of the "pragmatic majority". Acquisti and Grossklags [3] found four different clusters: "privacy fundamentalists with high concern toward all collection categories (26.1 percent), two medium groups with concerns either focused on the accumulation of data belonging to online or offline identity (23.5 percent and 20.2 percent, respectively), and a group with low concerns in all fields (27.7 percent)." Spiekermann et al. [172] also identified privacy fundamentalists (30%) and marginally concerned users (24%). In addition, the authors were able to split the remaining respondents into two distinct groups, namely ones who are concerned about revealing information such as their names, email or mailing addresses ("identity concerned, 30%) and others who are rather more concerned about the profiling of their interests, hobbies, health and other personal information ("profiling averse", 25%).

21.2.5 Stated Attitudes versus Reported and Observed Behavior

What are the effects of high privacy concerns? If one looks at people's reported past behavior or intended future behavior, the effects seem straightforward:

- Sheehan and Hoy [167] found that people's stated concern for privacy correlates negatively with the reported frequency of registering with websites in the past, and positively with providing incomplete information when they do register.

[3] See [114] for a more detailed comparison of privacy concern indicators over different years.

- Metzger [127] more generally found that stated concern for online privacy negatively predicted reported past online information disclosure (i.e., those who expressed high privacy concerns also tended to report less information disclosure in the past, and vice versa).
- Smith et al. [170] developed and validated a survey instrument for determining individuals' level of privacy concerns, which is composed of four subscales that measure concerns about inappropriate collection, unauthorized secondary use, improper access, and errors in storing. Research by Xu et al. [195] that used this instrument indicates that if people's individual privacy "sub-concerns" are addressed, their intended data disclosure rose significantly (concerns regarding improper access and unauthorized secondary use had particularly high regression coefficients).
- Finally, Chellappa and Sin [31] found that users' stated intention to use personalization services (which necessitates their willingness to disclose information about themselves) is also negatively influenced by their individual level of privacy concern.

Other survey results however shed doubts on whether Internet users always follow through on their stated concerns. A large majority of people buy online (and thereby give out personal data) despite professing privacy concerns [16, 76, 179]. More paradoxically, Behrens [15] found that 20 percent of adults who say they have placed an order on the Internet in the past three months also say they won't put personal information such as their name and address on the Web.

If we look at observable user behavior, the discrepancy to stated privacy concerns becomes even more apparent. The experiment of Metzger [128] did not confirm the hypothesis that individuals' level of concern about online privacy and data security is negatively related to the *observed* amount of personal information they disclosed to a commercial Web site*. The experiment by Spiekermann et al. [172] showed that privacy fundamentalists in particular did not live up to their expressed attitudes*. They only answered 10 percentage points fewer questions than marginally concerned participants.

As mentioned above, a lesson from the apparent discrepancy between intended and actual disclosure behavior of highly privacy-concerned individuals is that developers in the area of personalized web-based systems should not feel completely discouraged by the abundance of stated privacy concerns in consumer surveys. User experiments and daily web practice prove that people do disclose their personal data, since other factors are in effect at the same time that override or alleviate their privacy concerns. Such factors will be discussed in the next few sections. Moreover, based on the abovementioned results of Xu et al. [195], it seems worthwhile to address people's individual privacy "sub-concerns". Section 21.5.4 will discuss methods for dealing with privacy in a more personalized manner.

21.3 Factors Fostering the Disclosure of Personal Information

This section describes factors that have been shown to influence people's willingness to disclose personal data about themselves on the Internet. Those factors include the value that people assign to personalization, their knowledge of and control over how personal information is used, users' trust in a website (and known antecedents thereof, namely positive past experience, the design, operation and reputation of a website, and the presence of privacy statements and privacy seals), as well as data disclosure benefits other than personalization. The section also discusses consequences of these findings for the design of web-based personalized systems. In Section 21.3.5 we will describe how users consider these factors in a situation-specific cost-benefit analysis when deciding on whether or not to disclose individual personal data.

21.3.1 Value of Personalization

Chellappa and Sin [31] found that the value which Internet users assign to personalization is a very important factor with regard to their stated intention to use personalized websites, and that it can "override" privacy concerns: "the consumers' value for personalization is almost two times […] more influential than the consumers' concern for privacy in determining usage of personalization services. This suggests that while vendors should not ignore privacy concerns, they are sure to reap benefits by improving the quality of personalized services that they offer" [31]. A study by White [194] also confirmed that users are more likely to provide personal information when they receive personalization benefits (the opposite seems to hold true however in the case of potentially embarrassing information in combination with a deep relationship between consumer and business, as will be explained in more detail in Section 21.3.3.1).

How much value, then, do Internet users assign to personalized services? Consumer surveys from the turn of the century (i.e. from the time when personalization features became first visible on the web) suggest that a slight majority of respondents value personalization, but that about a quarter sees no value in personalization or is not willing to disclose personal data to receive it:

1. Online users who see / do not see personalization as a good thing: 59% / 37% [82];
2. People who are willing to give information to receive a personalized online experience: 51% (15% not) [148], 43% (39% not) [163];
3. Types of information users would provide to a web site that used it to personalize/ customize their experience, compared to one that does not provide any personalization: hobbies 76% vs. 51%, address 81% vs. 60%, job title 50% vs. 32%, phone number 45% vs. 29%, income 34% vs. 19%, name 96% vs. 85%, mother's maiden name 22% vs. 14%, e-mail address 95% vs. 88%, credit card number 22% vs. 19%, social security number 6% vs. 7%.
4. Online users who find it useful if a site remembers basic information (name, address): 73% (9% not) [148];
5. Online users who find it useful if a site remembers information (preferred colors, music, delivery options etc.): 50% (20% not) [148];

6. People who are bothered if a web site asks for information one has already provided (e.g., mailing address): 62% [148].

More recent surveys found the percentage of respondents who value personalization to be significantly higher. In a 2005 study by ChoiceStream [32], 80% of respondents stated that they are interested in receiving personalized content (news, books, search results, TV/movie, music). This number is consistent with the 2004 edition of the same survey in which 81% expressed their interest in personalized content. Young people are slightly more interested in personalization than older people. No figures are available on those who are not interested. 60% indicated that they would spend at least 2 minutes answering questions about themselves and their interests in order to receive personalized content, versus 56% in 2004. 26% agreed that they would spend at least 6 minutes answering such questions, compared with 21% in 2004. Moreover, 59% (2004: 65%) of respondents indicated a willingness to provide information about their personal preferences, and 46% (2004: 57%) to provide demographics. The authors of the study attribute these decreases in people's willingness to provide personal data to a surge in societal privacy concerns during the intermittent year.

These findings suggest that developers of personalized web-based systems need to make the personalization benefits of their system very clear to users, and ascertain that those benefits are ones that people want.[4] If users perceive value in the personalization services offered, they are considerably more likely to intend to use them and provide the required information about themselves.

21.3.2. Knowledge of and Control over the Use of Personal Information

Many privacy surveys indicate that Internet users find it important to know how their personal information is being used, and to have control over this usage. In a survey of Roy Morgan Research [163], 68% of respondents indicate that it was very important (and 25% that it was important) to know how their personal info may be used. In a survey by Turow [182], 94% even agree that they should have a legal right to know everything that a web site knows about them. In a 1997 survey by Harris Interactive [119], 63% of people who had provided false information to a website or declined to provide information said they would have supplied the information if the site provided notice about how the information would be used prior to disclosure, and if they were comfortable with these uses.

As far as control is concerned, 69% of subjects said in a 2003 Harris poll [179] that "controlling what information is collected about you" is extremely important, and 24% still regarded it as somewhat important. Likewise in a direct marketing study of Phelps et al. [151], the vast majority of respondents desire more control over what companies do with their information. Sheehan and Hoy [168] even found that control

[4] Time savings and monetary savings, and to a lesser extent pleasure, received the highest approval in surveys conducted by Tan et al. [89, 177] on benefits that businesses collecting personal information should offer. In a survey by Cyber Dialogue [122], customized content provision and the remembering of preferences were quoted as the main reasons for users to personalize websites.

(or lack of control) over the collection and usage of information is the most important factor for people's stated privacy concerns, explaining 32.8% of the variance.

Some empirical evidence also exists that people are more willing to disclose their personal data if they possess knowledge of and/or control over the use of this data. In the above-mentioned survey by Roy Morgan Research [163], 59% said they'd be more likely to trust an organization if it gave them more control over how their personal information was used (as we will see in Section 21.3.3, trust in turn is an important factor for people's willingness to disclose their personal data). In a 1998 survey [76], 73.1% indicated that they would give demographic information to a Web site if a statement was provided regarding how the information was going to be used. In a survey by Hoffman [87], 69% of Web users who do not provide data to Web sites say it is because the sites provide no information on how the data will be used. In an experiment by Kobsa and Teltzrow [111], users disclosed significantly more information about themselves when, for every requested piece of personal information, a website explained the user benefits and the site's privacy practices in connection with the requested data* (the effects of these two factors were not separated in this study).

These findings suggest that personalized systems should be able to explain to users what facts and assumptions are stored about them and how these are going to be used.[5] Moreover, users should be given ample control over the storage and usage of this data. This is likely to increase users' data disclosure and at the same time complies with the rights of data subjects accorded by many privacy laws, industry and company privacy regulations, and Principles of Fair Information Practices (see Section 21.4).

Extensive work in this direction has been carried out by Judy Kay and her team under the notion of "scrutability" [47, 99-101]. According to Kay, "this means that the user can scrutinise the model to see what information the system holds about them. In addition, it means that the user can scrutinise the processes underlying the user modelling. These include the processes used to collect data about the user. It also includes the processes that made inferences based on that data." [102]. A qualitative evaluation was carried out which showed that "participants in the evaluation could, generally, understand how the material was adapted and how to control that adaptation" [48]. It was challenging for them though to determine what content was included/excluded on a page and what caused the adaptation, and to understand how to change their profiles to control the inclusion or exclusion of content.

[5] In Section 24.3.3.4 we will see that "privacy statements" (aka "privacy policies"), which constitute the current best practice for privacy disclosures, are not an effective medium for providing such explanations.

21.3.3 Trust in a Website

Trust in a website is a very important motivational factor for the disclosure of personal information.[6] In a survey by Hoffman et al. [87], nearly 63% of consumers who declined to provide personal information to web sites stated as the reason that they do not trust those who are collecting the data (similar responses can be found in Milne and Boza [132]). Conversely, Schoenbachler and Gordon [165] found a positive relationship between trust in an organization and stated willingness to provide personal information. In the experiment of Metzger [127], Internet users' trust in a company's Web site positively influenced their information disclosure to the site*. Trust was also found to positively affect the intended use of an e-commerce website [68].

Several antecedents to trust have been empirically established, and for many of them effects on disclosure have also been verified. Such trust-inducing factors include[7]

- positive experiences in the past,
- the design of a website,
- the reputation of the website operator,
- the presence of a privacy seal, and
- the presence of a privacy statement (but not necessarily its content).

These factors will be discussed in the following subsections.

21.3.3.1 Positive Experiences in the Past

Positive experience in the past is an established factor for trust. Almost half (47%) of the respondents in a consumer survey of the Australian privacy commissioner [163] agreed that their trust in an organization with their personal information would be based on good past experience. Pavlou [145] found a highly significant positive effect of good prior experience on trust for a number of existing websites.

The impact of positive experience in the past on the disclosure of personal information is well supported. In an open-ended questionnaire by Culnan and Boza [132], the number one reason that consumers gave for trusting organizations with personal information was past experience with the company. Culnan and Armstrong [43] found that people who agree that their personal data be used for targeted marketing purposes are more likely to have prior experience with direct marketing than people who do not agree. Metzger [127] observed what types of information subjects disclosed to an experimental website and also asked them what types of information they had disclosed in the past. The author found the total amount of past information disclosure to be a good predictor for the current amount of information disclosure*.

---

[6] A number of different definitions and conceptualizations of trust in online environments have been proposed or used in the literature. For a discussion and critical analysis of those we refer to [69, 73, 124].

[7] Telling users how their personal data will be used and giving them control over this usage (see Section 24.3.2) may also increase users' trust [87]. We refrain from listing it as a factor for trust though since the empirical support for this claim seems insufficient to date.

Of specific importance are established, long-term relationships. Sheehan and Hoy [168] prompted subjects for their privacy concerns in 15 hypothetical scenarios. They identified three factors in these scenarios that explain the stated level of privacy concerns, one of them including "items that suggest that the online user has an established relationship with the online entity, in which some level of communication and interaction has already been established between the two parties." Schoenbachler and Gordon [165] found a positive relationship between respondents' perception of a relationship with an organization and their stated willingness to provide personal information. The same was the case in a study by White [194] for two pieces of information that were determined to be specifically private (namely one's address and telephone number). Interestingly enough, the author also found that a deeper relationship with the customer lead to a *decreased* willingness to disclose two other pieces of personal information that were determined to be specifically embarrassing and might cause a loss of face when disclosed, namely one's purchase history of Playboy/ Playgirl magazine and of condoms.

The lesson for the design of personalized systems is not to regard the disclosure of personal information as a one-time matter. Users of personalized websites can be expected to become more forthcoming with personal details over time if they obtain positive experiences with the same site or comparable sites. Personalized websites should be designed in such a way that they can deliver satisfactory user experiences with any amount of personal data that users chose to disclose, and allow users to add more personal detail incrementally at later times.

21.3.3.2 Design and Operation of a Website

Various interface design elements and operational characteristics of a website have been found to increase users' trust in the website, such as

- the absence of errors, such as wrong information or incorrect processing of inputs and orders [12],
- the (professional) design of a site [59, 61],
- the usability of a site [49, 59, 162], specifically for information-rich sites such as sports, portal, and e-commerce sites [12],
- the presence of contact information, namely physical address, phone number or email address [59, 60],
- links from a believable website [59],
- links to outside sources and materials [59],
- updated since last visit [59],
- quick responses to customer service questions [59, 60],
- email confirmation for all transactions [59],
- the presence of an interactive communication channel with a site, specifically instant messaging or voice communication [13], and
- the presence of a photo of a "customer care person" (positive effect for sites with low reputation, negative effect for sites with high reputation)* [159].

While there do not seem to be studies yet that measure directly the effect of website design and operational characteristics on users' willingness to disclose personal data, the established effect of trust on user disclosure behavior (see Section 21.3.3) makes

the existence of such an effect very plausible. The lesson from the above findings for the design of personalized websites is therefore to use personalization preferably in professionally designed and easy-to-use websites that also possess some other of the above-mentioned trust-increasing design elements and operational characteristics.

21.3.3.3 Reputation of the Website Operator

The reputation of the organization that operates a website is an important factor for users' trust in the website. Schoenbachler and Gordon [165] found a positive relationship between the perceived reputation of a company and stated trust in the company. Likewise, Metzger [127] established a positive correlation between individuals' subjective regard for a non-existing company whose fictitious website they saw, and their stated trust in this website. Jarvenpaa et al. [95] and Pavlou [145] also found an effect of reputation on trust for several existing websites (Jarvenpaa et al. [94, 95] moreover determined that perceived company size is positively associated with consumers' trust in these websites, though size and reputation are highly related). Metzger [128] varied the reputation of a website between subjects and found that the one with higher reputation was deemed more trustworthy than the one with lower reputation.

Not surprisingly then, reputation is positively correlated with users' willingness to disclose personal information. In a Canadian consumer survey [30], 74% indicate that a company's reputation would make them more comfortable with providing personal information. In a paper-based experiment, Andrade et al. [5] found an effect of perceived reputation on stated concern about the disclosure of personal information (this effect only approached statistical significance though). In the online survey of Earp and Baumer [53], subjects were randomly shown one of 30 web pages from higher traffic and lower traffic websites in different sectors. Subjects were significantly less willing to provide personally identifiable information (specifically their phone numbers, home and email addresses, and social security and credit card numbers) to the lower-traffic sites (which were presumably less known to them).[8]

The lesson for the design of personalized systems seems to be that everything else being equal, users' information disclosure at sites of well-reputed companies is likely to be higher than at sites with lower reputation. Personalization is therefore likely to be more successful at sites with higher reputation. It may of course be possible to compensate for the lack of reputation by putting more emphasis on other factors that foster the disclosure of personal data. Designers should however clearly refrain from using personalization features as a "gimmick" to increase the popularity of websites with low reputation since based on the aforesaid, it is unlikely that users will take much advantage of the personalization features if they have to disclose personal data to a low-reputation website.

---

[8] Metzger [128] found that regard for the company had a stronger relationship with disclosure than did trust, which is somewhat contradictory to the current view that reputation effects disclosure indirectly via fostering trust.

21.3.3.4 Presence of a Privacy Statement

Privacy statements on websites (which are often also called "privacy policies") describe the privacy-related practices of these sites. Most countries that have privacy laws enacted require that users be informed about the data being collected and the purposes for which they are used. And even in jurisdictions where omnibus privacy legislation does not exist, special provisions at the federal or state level or simply public relation motives prompt many companies to publish privacy statements at their websites.[9] The comprehensibility of these disclosures for normal Internet users is however fairly low [96, 121].

There exists weak empirical evidence that the mere *presence* of a privacy statement at a website fosters trust.[10] For instance, 55% of the respondents in a survey of the Australian Privacy Commissioner [163] indicated that having a privacy statement would help build trust. This leads to the expectation that the presence of a privacy statement would also foster purchases and disclosure, namely via increased trust (see Section 21.3.3), which already received some empirical confirmation. In the study of Jensen et al. [97], the presence of a privacy statement proved to be one of the two best predictors for subjects' stated intent to buy from a website. In an experiment with Singaporean students, Hui et al. [90] found an effect of the presence of a privacy statement on subjects' willingness to completely fill in an online questionnaire with personal information*, but this effect only approached statistical significance ($p<0.1$).[11] Metzger [126] however found the opposite effect. In her experiment, 43.7% of subjects who bought CDs from a fake online music store, or completed a questionnaire to receive a free CD, withheld information when a privacy policy was present. In contrast, only 15% withheld information when the privacy policy was *not* present, and the difference was statistically significant* [129].

Not too many people seem to view and read privacy policies. As far as self-reported past behavior is concerned, the percentage of respondents who indicated having looked at privacy policies varies between 3% ("most of the time, carefully") [83], 4.5% ("always") [133], 14.1% ("frequently") [133], 31.8% ("sometimes") [133], 33% ("sometimes, carefully") [83], 23.7% ("likely, at first visit") [97], and 43% ("likely, e-commerce site, before buying") [97]. Milne and Culnan [133] found that stated concern for privacy is positively associated with stated tendency to read online privacy notices.

Observing user behavior in experiments and real life portrays a somewhat different picture though. Jensen et al. [97] found that subjects read privacy statements in 25.9% of cases where they were available* (the authors believe though that this number is

---

[9] For example, as far as the U.S. is concerned, a 2004 survey of more than 1,000 websites across a spectrum of industries found that 93% of them featured a privacy statement [40].

[10] As in the case of privacy seals, Internet users seem to be confused about what protection the presence of a privacy policy affords to them. For instance, Turow [182] found that 57% of U.S. adults who use the Internet at home agree or strongly agree with the statement "When a web site has a privacy policy, I know that the site will not share my information with other websites or companies."

[11] Interestingly, the authors did not find the same effect when the tested website featured both a privacy policy and a privacy seal.

inflated since subjects knew they were being observed and what the purpose of the experiment was and therefore took more care in their decision-making than usual). In the experiment of Kobsa and Teltzrow [111], only two out of 52 subjects accessed the privacy statement*. The most reliable figures are presumably real-world server-side observations: only one percent of users or less click at links to a website's privacy statement according to Reagan [158], and less than 0.5% according to Kohavi [112]. In contrast to the above-mentioned survey results of Milne and Culnan [133], Jensen et al. [97] also found that those subjects whom they classified as privacy fundamentalists were no more likely to read privacy policies than the privacy unconcerned*.

When users do read privacy statements, the effect on users' behavior is unclear as yet. In an experiment by Metzger [126], 62.5% of the participants who clicked on the strong version of the privacy policy disclosed some information to the website and only 37.5% of those who clicked on the weak version, but the difference was not statistically significant*. Likewise in the experiment of Spiekermann et al. [172], the privacy protection that was promised in privacy statements did not have a statistically significant effect on subjects' willingness to disclose personal data* (subjects had to sign that they had read and accepted this statement prior to shopping at the experimental website). In contrast to these negative results, Andrade et al. [5] did find an effect of the length or level of detail of privacy statements: subjects who saw a 12-word statement professed considerable higher concern about the disclosure of personal information than subjects who saw a 88-word statement (a 22-word example statement was initially presented to all subjects as being "typical"). The ecological relevance of this experiment is however unclear since real-life privacy statements usually comprise several pages of text and not just a few words.

The preliminary lesson for the design of personalized systems seems to be that *traditional* privacy statements should not be posted in the expectation of increasing users' trust and/or disclosure of personal information, even when the statement describes good company privacy practices. There may of course be other reasons for posting such statements, such as legal or self-regulatory requirements (see Section 21.4), or demonstration of good will. Evidence is mounting though that privacy-minded company practices can have such a positive effect if they are communicated to web users in comprehensible forms, such as the following:

- Kobsa and Teltzrow [111] found that subjects disclose significantly more information about themselves if every website does not only display a link to a privacy policy, but if additionally every entry field for personal information is accompanied by a short summary of the website's privacy practices regarding specifically the solicited piece of information (and an explanation of why it is needed)*.
- Gideon at al. [70] asked subjects to search for vendors of a given product in a search engine and to buy the product with their own credit cards. For every site in the result list, the color of an appended "Privacy Bird" [9, 38] indicated whether the P3P [193] encoded privacy policy of the site matches typical medium-level privacy expectations, does not match them, or could not be parsed. If the website had no P3P policy posted, no bird would appear. The authors of the study found that when subjects were asked to buy a pack of condoms, they patronized websites with conforming privacy policies significantly more often than a control group that saw no privacy birds*. No such difference could be found when subjects had to buy a surge protector rather than condoms.

21.3.3.5 Presence of a Privacy Seal

Privacy seals are logos of certifying agencies such as consumer organizations, data commissioner's offices or private companies. These agencies assert to web visitors that websites that display their seals respect privacy to some extent. The amount of assured privacy protection varies from seal to seal and also over time. U.S. privacy seals originally merely asserted that a website abides to its published privacy statement, no matter how privacy-friendly this policy actually was. Meanwhile, trust organization require minimum privacy standards such as the observance of the FTC principles of notice, choice and consent [65].

A number of recent studies uncovered several problems with at least some privacy seals though:

Insufficient scrutiny of trust organizations: Using webbots that analyze websites' privacy practices, Edelman [54] found that sites that used practices most Internet users would find objectionable nevertheless received a privacy seal from TRUSTe, the leading US trust mark. The percentage of untrustworthy sites certified by an TRUSTe seal even significantly increased over time (to nearly 3.5% as of January 2006). Various privacy breaches at websites that carried the TRUSTe seal, and to a much smaller extent also at sites with the BBBOnLine seal, have been reported as well [115, 137, 169].

Negative self-selection of seal-bearing websites. Several studies came to the conclusion that websites that decide to "pay up" for certain privacy seals seem to have more questionable privacy practices than ones that don't. Larose and Rifon [115] found that sealed sites requested significantly more personal information from users than unsealed sites. Miyazaki and Krishnamurthy [134] reviewed 60 high-traffic websites and found no support for the hypothesis that participation in a seal program is an indicator of better privacy practices (Larose and Rifon [115] made similar findings). While these studies were all performed manually, Edelman [54] analyzed more than 500,000 websites with web bots. He found that the ratio of untrustworthy vs. trustworthy sites certified by TRUSTe (5.4%) is more than twice as high as for non-certified sites (2.5%). In a regression model with several site characteristics, the presence of an TRUSTe privacy seal turns out to be a statically significant *negative* coefficient for site trustworthiness. In contrast, the much less frequent BBBOnLine privacy seal that comes with a more cumbersome and restrictive certification process does not seem to suffer from such an adverse self-selection, and seal-bearing websites are slightly more likely to be trustworthy than a random cross-section of sites.

Seals not understood by web users: The results of a study by Portz et al. [154] on a specific privacy seal, WebTrust, "were mixed in terms of potential customers correctly understanding what WebTrust signifies." In a study by Moores [136], 42% recognized the TRUSTe logo and 29% the BBBOnline logo as a privacy seal. A whopping 15% however also mistook an officially looking fake graphic for a genuine privacy seal.

The presence of privacy seals clearly does have an effect on web users though, despite this confusion about what assurances they actually afford. Rifon et al. [160] found a positive effect on the perception of trust in a website. Miyazaki and Krishnamurthy

[134] found that the presence of a privacy seal resulted in more favorable consumer perceptions regarding the privacy policies of a website. In the study of Jensen et al. [97], the presence of a privacy seal turned out to be one of the two best predictors for subjects' stated intent to buy from a website.

There is also empirical evidence for an effect of the presence of a privacy seal on users' stated willingness to disclose personal data to the website [116]. Other studies found that this effect was moderated by other factors, namely

- *Perceived self-efficacy* (i.e. confidence in one's ability to protect one's privacy): Rifon et al. [160] found that for individuals with lower self-efficacy, the presence of a privacy seal had a positive effect on anticipated disclosure of personal data. No such effect on subjects with high self-efficacy could be found.
- *Perceived online shopping risk* (when compared to transactions made at traditional brick and mortar stores): Miyazaki and Krishnamurthy [134] found a positive effect of privacy seal presence on anticipated disclosure of personal information for those subjects who experience relatively high levels of online shopping risk. No effect on subjects with low-risk experience could be found.

It remains to be seen whether these moderating factors are independent of each other, or rather correlated (which seems more likely). For designers of web-based personalized systems, the pragmatic conclusion at this point is to display privacy seals as long as web users associate trust with them since doing so is likely to foster users' disclosure behavior.

21.3.4 Benefits other than Personalization

Financial Rewards. In a consumer survey by Turow [182], 16% of respondents agreed or even strongly agreed with the statement "I will give out information to a website only if I am paid or compensated in some way". Hann et al. [78] found that a financial reward of 20 Singapore dollars[12] (but not of 10 or 5 dollars) had a statistically significant positive effect on intended disclosure behavior. However, this economic benefit turned out to be relatively less important by a considerable margin than three privacy concerns measured by the survey instrument of Smith [170], namely unauthorized internal/external secondary usage, unauthorized access, and errors in the data. The authors calculated that monetary compensations between about S$15.00 and S$50.00 would be needed to motivate subjects to overcome these concerns. Financial rewards also had a statistically significant effect on observed disclosure behavior* in an experiment by Hui et al. [90], even though rewards only ranged from S$1 to S$9.

Social Adjustment Benefits. A study by Lu et al. [120] demonstrated that social adjustment benefits, i.e. the opportunity of establishing social identity by integrating into desired social groups [14], can also have an effect on intended disclosure behavior. The three experimental conditions were (a) no benefits (control group), (b) opportu-

[12] One Singapore dollar equaled 0.54 U.S. dollars in 2002.

nity of face-to-face interaction with other people (namely meetings with people having similar interests, participation in focus groups, membership in downtown clubs of the Internet business), and (c) opportunity of online interaction (namely access to online chat-rooms with similar interests, exclusive membership in the online clubs of the Internet business, access to online forums featuring focus groups). For extrovert subjects, both treatment conditions had a statistically significant effect on their intended disclosure of personal data, while for introvert subjects this was only the case when online interaction was offered.

Both results seem only marginally relevant for personalized web-based systems since those normally do not offer such benefits. In special application scenarios though, the provision of personal data might open an opportunity for financial benefits (e.g., targeted advertising with special discounts) or social adjustment benefits (e.g., participation in discussion groups with people who have similar goals or interests). Designers should consider taking advantage of the increase in trust and disclosure that these benefits may entail.

21.3.5 Disclosure Behavior as the Result of a Cost-Benefit Analysis

Current privacy theory regards people's disclosure behavior to websites as the result of a situation-specific cost-benefit analysis, in which the potential risks of disclosing one's personal data are weighed against potential benefits of the data disclosure.[13] Trust thereby is an important risk-mitigating factor [31, 87, 95, 124, 128].

This cost-benefit tradeoff explains the discrepancies between stated privacy concerns and observed "inconsequent" data disclosure behavior that was discussed in Section 21.2.5. While it seems true that many Internet users are privacy-concerned, it is also a fact that most are willing to "trade off" their concerns against benefits that they value (see Sections 21.3.1 and 21.3.4) [17, 78, 173, 179], and become even more swayed to do so by the presence of trust-evoking signals such as those discussed in Section 21.3.3.

Acquisti and Grossklags [2, 3] point out however that Internet users often lack sufficient information to be able to make educated privacy-related decisions (for instance, they underestimate the probability with which they can be identified if they disclose certain data, or are unfamiliar with a site's privacy practices since they hardly ever read privacy statements (see Section 21.3.3.4). Like all complex probabilistic decisions, privacy-related decisions are moreover affected by systematic deviations from rationality [98]. For instance, Acquisti and Grossklags [3] present evidence of hyperbolic temporal discounting, which may lead to an overvaluation of small but immediate benefits and an undervaluation of future negative privacy impacts.

An implication of users' cost-benefit analysis for personalized systems is that developers can work in four, and possibly even five directions to encourage more liberal disclosure behavior, and thereby enhance the quality of the system's personalized services. They can

[13] Culnan [42] coined the term "privacy calculus" to refer to this cost-benefit comparison (the term dates back to Laufer et al.'s [117, 118] notion of "calculus of behavior").

1. address the privacy concerns directly, as explained in Sections 21.2.5 and 21.5.4,
2. ensure that the user values the personalization benefits of the system (see Section 21.3.1),
3. ascertain that the user trusts the website (which mitigates privacy concerns), e.g. by establishing the trust-enhancing factors described in Sections 21.3.3.1 – 21.3.3.5,
4. ascertain that the user is made aware of, and can control, how personal information is being used (see Section 21.3.2), and
5. if meaningful, ascertain that financial rewards and social adjustment benefits are provided (see Section 21.3.4).

Interaction effects between these factors have not been established as yet. From the experiment of Chellappa et al. [31] (see Section 21.3.1) and the work of Acquisti and Grossklags [2, 3] we can conclude that *instant personalization benefits* will be a very important factor in the outcomes of users' cost-benefit analyses.

21.4 Privacy Laws, Industry and Company Regulations, and Principles of Fair Information Practices

To date, more than forty countries and numerous states have privacy laws enacted [161, 190]. Many companies and a few industry sectors additionally or alternatively adopted self-regulatory privacy guidelines. These laws and self-regulations are often based on more abstract principles of fair practices regarding the use of personal information. In this section, we will analyze the effects that these regulatory instruments have, specifically on personalization in web-based systems. We will uncover some deficits in current personalized systems, which open avenues for interesting and challenging future research. Privacy laws, industry and company regulations and Principles of Fair Information Practices may also impose requirements that are not directly related to personalization but affect any system that collects personal data. These more general implications cannot be discussed here. Readers are advised to consult their national privacy literature.

21.4.1 Privacy laws

Since personalized systems collect personal data of individual people, they are also subject to privacy laws and regulations if the respective individuals are in principle identifiable. To date, more than forty countries and numerous states have privacy laws enacted. They lay out procedural, organizational and technical requirements for the collection, storage and processing of personal data, in order to ensure the protection of these data as well as the data subjects to whom the data apply. These requirements include disclosure duties (e.g. about the purpose of data processing), and conditions for legitimate data acquisition, data transfer (e.g., to third parties or across national borders) and the processing of personal data (e.g., their storage, modification and deletion). Other requisites include user opt-in (e.g., asking for their consent before collecting their data), opt-out (e.g., of data collection or data processing), and users'

right to be informed (e.g., about what personal information has been collected and possibly how it is processed and used). Other legal stipulations establish adequate security mechanisms (e.g., access control), and the supervision and audit of personal data processing.

Some requirements imposed by privacy laws directly or indirectly affect the permissibility of personalization methods. Here are some examples:

1. *Value-added (e.g. personalized) services based on traffic or location data require the anonymization of such data or the user's consent* [56][14]. This clause requires the user's consent for any personalization based on interaction logs if the user can be identified.
2. *Users must be able to withdraw their consent to the processing of traffic and location data at any time* [56]. In a strict interpretation, this stipulation requires personalized systems to immediately honor requests for the termination of all traffic or location based personalization, i.e. even during the current session. A case can probably be made that users should not only be able to make all-or-none decisions, but also decisions with regard to individual aspects of traffic or location based personalization (such as agreeing to be informed about nearby sights but declining to receive commercial offers from nearby businesses).
3. *The personalized service provider must inform the user of the type of data which will be processed, of the purposes and duration of the processing, and whether the data will be transmitted to a third party, prior to obtaining her consent* [56]. It is sometimes fairly difficult for personalized service providers to specify beforehand the particular personalized services that an individual user would receive. The common practice is to collect as much data about the user as possible, to lay them in stock, and then to apply those personalization methods that "fire" based on the existing data (see, e.g., rule-based personalization or stereotype activation [109]). Also, internal inference mechanisms may augment the available user information by additional assumptions about the user, which in return may trigger additional personalization activities. For meeting the disclosure requirements of privacy laws in such cases of low ex-ante predictability, it should suffice to list a number of *typical* personalization examples (preferably those that entail the most severe privacy consequences) [79].
4. *Personal data that were obtained for different purposes may not be grouped* [46]. This limitation affects centralized user modeling servers (see Chapter 4 of this book [130]), which store user information from, and supply this data to, different personalized applications. Such servers must not return data to requesting personalized applications that was collected for a different purpose than the one for which the data is now being sought.
5. *Usage data must be erased immediately after each session* (except for very limited purposes) [50]. This requirement could affect the use of machine learning methods that derive additional assumptions about users (see Chapter 3 of this book [135]), when the learning takes place over several user sessions.
6. *No fully automated individual decisions are allowed that produce legal effects concerning the data subject or significantly affect him and which are based solely*

[14] EU directives are "Europe-wide minimum standards" in the sense that all European Union member states have to implement them in their national legislation, but are free to go beyond them.

on automated processing of data intended to evaluate certain personal aspects relating to him, such as his performance at work, creditworthiness, reliability, conduct, etc. [55]. These provisions could affect, for example, personalized tutoring applications (see Chapter 22 of this book [84]), if they assign scores to users that significantly affect them.

Besides "omnibus" privacy laws at the national or state level, there also exist various sectorial laws. Examples in the U.S. include the Health Insurance Portability and Accountability Act (HIPAA [183]) for the privacy of medical data, the Gramm-Leach-Bliley Act (GLB [185]) for the privacy of financial data, and the Children's Online Privacy Protection Act (COPPA [184]) for protecting the privacy of children aged 13 and younger. The HIPPA and GLB Acts would affect personalized systems that collect or process users' medical or financial information, and COPPA those that have children among their users.

21.4.2 Industry and company regulations

Many companies have internal guidelines in place for dealing with personal data. Several industry associations also developed privacy standards to which their members must subject themselves (e.g., the Direct Marketing Association, the Online Privacy Alliance, and the Personalization Consortium). Both company and supra-company self-regulations may affect the aims and methods of personalized systems, as is the case for privacy legislation. For instance, the privacy principles of the members of the U.S. Network Advertising Initiative [139] prohibit the use of "personally identifiable information ("PII") […] collected offline merged with PII collected online for online preference marketing unless the consumer has been afforded robust notice and choice about such merger before it occurs." This stipulation thus restricts the merger of clickstream data with data from legacy customer databases, which is a frequently-found functionality of commercial user modeling servers (see Chapter 4 of this book [130]).

21.4.3 Principles of Fair Information Practices

Over the past three decades, several collections of basic principles have been defined for ensuring privacy when dealing with personal information. So-called Principles of Fair Information Practices have been drafted by several countries as a foundation of their national privacy laws [10, 41], by supra-national organizations as a guidance for their member states [6, 142], and by professional societies as recommendations for policy makers and as guidance for the professional conduct of their members [186].

Developers of personalized systems should also take such privacy principles into account if those are not already indirectly considered through applicable privacy laws and industry or company guidelines. Many guidelines have direct implications on personalized systems. As an example, let us consider excerpts from the recommendations of the U.S. Public Policy Committee of the Association for Computing Machinery (ACM) [186], the largest computer science association worldwide. These

recommendations have been strongly shaped by the 1980 OECD guidelines [142][15] but are more modern and concrete in their technical demands.

Minimization principles.

1. *Collect and use only the personal information that is strictly required for the purposes stated in the privacy policy.*
2. *Store information for only as long as it is needed for the stated purposes.*
3. *Implement systematic mechanisms to evaluate, reduce, and destroy unneeded and stale personal information on a regular basis, rather than retaining it indefinitely.*
4. *Before deployment of new activities and technologies that might impact personal privacy, carefully evaluate them for their necessity, effectiveness, and proportionality: the least privacy-invasive alternatives should always be sought.*

Somewhat in contradiction to these requirements, a current tacit paradigm of personalized systems seems to collect as much data as possible and lay them in stock, and to let personalization being triggered by the currently available personal data (data-driven personalization). Applications in several personalization areas[16] have now sufficiently progressed that it should be possible to determine in hindsight which of the collected data hardly ever trigger personalization, and to forego storing these less needed data in the future even when they would be readily available.

Consent principles.

5. *Unless legally exempt, require each individual's explicit, informed consent to collect or share his or her personal information (opt-in); or clearly provide a readily-accessible mechanism for individuals to cause prompt cessation […] including when appropriate, the deletion of that information (opt-out).*

One implication of this requirement for personalized systems is that personalization based on the users' personal data must be an option that can be switched on and off at any time.

Openness principles.

8. *Whenever any personal information is collected, explicitly state the precise purpose for the collection and all the ways that the information might be used […].*
10. *Explicitly state how long this information will be stored and used, consistent with the "Minimization" principle.*

---

[15] See [37] for a discussion of the effects of the OECD principles on personalized e-commerce systems.

[16] For instance, student-adaptive tutoring systems (see Chapter 22 of this book [84]), customer relationship management on the web (see [109] and Chapter 16 of this book [72]), and recommender systems (see Chapters 9-12 of this book [24, 146, 164, 171]).

11. *Make these privacy policy statements clear, concise, and conspicuous to those responsible for deciding whether and how to provide the data.*

The likely positive effect of such explanations on users' willingness to disclose personal data was discussed in Section 21.3.2. Some difficulties in providing a full explanation of the personalization purposes were discussed in Section 21.4.1 (3).

Access principles.

14. *Establish and support an individual's right to inspect and make corrections to her or his stored personal information, unless legally exempted from doing so.*

This principle calls for online inspection and correction mechanisms for personal data, as discussed in Section 21.3.2.

Accuracy principles.

17. *Ensure that personal information is sufficiently accurate and up-to-date for the intended purposes.*
18. *Ensure that all corrections are propagated in a timely manner to all parties that have received or supplied the inaccurate data.*

So far, allowing users to verify their data seems to be the only solution for assuring data accuracy that has been adopted in the personalization literature. Little attention has been paid to recognizing the obsoleteness of data, and to recording the provenance of data and propagating error and change notifications to the data sources.

Security principles.

19. *Use appropriate physical, administrative, and technical measures to maintain all personal information securely and protect it against unauthorized and inappropriate access or modification.*

This principle not only entails that user information must be protected when it is stored in a repository, but also while it is in transit (e.g. by only using secure channels between authenticated senders and receivers). In the case of personalized systems, the latter is currently not often considered.

21.5 Privacy-Enhancing Technology for Personalized Systems

In this section, we describe and analyze several technical approaches that may reduce privacy risks and make privacy compliance easier. They are by no means complete "technical solutions" to the privacy risks of personalized systems, and their presence is also unlikely to "charm away" users' privacy concerns. Rather, these technologies should only be employed as additional privacy protections in the context of a user-oriented system design that also takes normative aspects into account (see Section 21.4). This analysis will be restricted to technologies that are specifically intended for

personalized web-based systems. For an overview of more general privacy-enhancing technologies that can be applied to wider classes of systems (including personalized web-based systems in many cases), we refer to [21, 25, 71, 188].

21.5.1 Pseudonymous users and user models

It is possible for users of personalized systems to enjoy anonymity and at the same time receive full personalization [110, 166]. In an anonymization infrastructure that supports personalization, users would need to have the following characteristics (using the terminology of [93, 150]):

- *Unidentifiable*. Neither the personalized system nor third parties should be able to determine the identity of pseudonymous users;
- *Linkable for the personalized system*. The personalized system can link every interaction with a specific user, even across sessions (users maintain a persistent identity);
- *Unlinkable for third parties*. Third parties cannot link two interaction steps of the same user;
- *Unobservable for third parties*. Third parties cannot recognize that a personalized system is being used by a given user.

To ensure their linkability, users would need to employ a "pseudonym" in all their transactions, i.e. a unique and persistent identifier that differentiates them from all other users. The personalized system may allow users to freely define their pseudonyms (or pick them from a list of available pseudonyms) without disclosing their true identities. Users may however also be required to reveal their identities to a registrar who assigns pseudonyms to them ("escrowed identity" [103], "initially nonpublic pseudonym" [150]). In the latter case, the pseudonym may be revoked at a later time, by an act of the registrar alone or in tandem with the website operator and/or user. This revocation of pseudonyms may be desirable in cases of misuse or when the identification of the user becomes necessary for other reasons, such as non-anonymous payment and delivery scenarios.

A number of authors proposed infrastructures for pseudonymous yet personalized user interaction with websites based on some or all of the above properties [8, 66, 85, 92, 110, 166]. Protecting the identity of users may not be enough, however. If user data is stored on a user modeling server on the Internet (see Chapter 4 of this book [130]), not only the user but also the user modeling server may need to remain anonymous. User models may reside anywhere on the network, like on the user's platform (as is envisaged, e.g., in the P3P framework [193]) or on a remote server (such as in Microsoft's Passport architecture [144]). A location close to the user (such as informatics.uci.edu or even more alfredkobsa.name) may compromise the user's anonymity. To safeguard it, Kobsa and Schreck [110, 166] extend their pseudonymity infrastructure to also protect the anonymity of user modeling servers.

Some authors expect that Internet users are more likely to provide information when they are not identified [39, 110], which may improve the quality of personaliz-

ation and the benefits that users receive from it. To date, this claim has however not found much empirical substantiation. In an online survey from 1998 [76], 66.3% of respondents strongly agreed and 21.8% somewhat agreed with the statement "I value being able to visit sites on the Internet in an anonymous manner." 30.5% also strongly agreed and 22.1% somewhat agreed with the statement "I would prefer Internet payment systems that are anonymous to those that are user identified". The demographics of the survey respondents was however considerably skewed towards higher education (nearly 80% had at least some college-level education) and towards fairly advanced web skills. Ordinary consumers tend to be unfamiliar with many basic security features, and base their perception of security rather on the company's reputation, their experience with the site, and recommendations from independent third parties [181].

The implications of these limited findings for the design of personalized web-based systems seem a bit unclear. Designers should definitely allow for pseudonymous access and pseudonymous user models (and even allow for anonymization architectures with the above properties if one is readily available). This follows from the data minimization and security requirements of the Principles of Fair Information Practices that were discussed in Section 21.4.3. Some privacy laws also mandate [51] or recommend [56] the provision of pseudonymous access if it is technically possible and not unreasonable (an interesting side effect of pseudonymous access is that in most cases privacy laws do not apply any more when users cannot be identified with reasonable means).

Due to a lack of relevant studies, it is unclear though whether increased anonymity will lead to more disclosure and better personalization. Anonymity is currently also difficult and/or tedious to preserve when payments, physical goods and non-electronic services are being exchanged. It also harbors the risk of misuse and hinders vendors from cross-channel marketing (e.g. sending a products catalog to a web customer by postal mail). Finally, research shows that the anonymity of database entries [176], web trails [123], query terms [140], ratings [64] and textual data [157] can be surprisingly well defeated by a resourceful attacker who has identified data available that can be partly matched with the "anonymous" data.

21.5.2 Client-Side Personalization

A number of authors [28, 29, 36, 138] have worked on personalized systems in which users' data are located at the client rather than the server side. Likewise, all personalization processes that rely on this data are also carried out at the client side only. From a privacy perspective, this approach has two major advantages:

1. The privacy problem becomes smaller since very few, if any, personal data of users will be stored on the server. In fact, if a website with client-side personalization does not have control over any data that would allow for the identification of users with reasonable means, it will generally not be subject to privacy laws.
2. Users may possibly be more inclined to disclose their personal data if personalization is performed locally upon locally stored data rather than remotely on remotely

stored data, since they may feel more in control of their local physical environment.[17]

Client-side personalization also poses a number of challenges though:

1. Popular user modeling and personalization methods that rely on an analysis of data from the whole user population, such as collaborative filtering and stereotype learning (see [109]), cannot be applied any more or will have to be radically redesigned (see the next section).
2. Personalization processes will also have to operate at the client side since even only a temporary or partial transmission of personal data to the server is likely to annul the abovementioned advantages of client-side personalization. However, program code that is used for personalization often incorporates confidential business rules or methods, and must be protected from disclosure through reverse engineering. Trusted computing platforms will therefore have to be developed for this purpose, similar to the one that Coroama and Langheinrich [35, 36] envisage to ensure the integrity of their client-side collection of personal data.

If these drawbacks pose no problems in a specific application domain, then developers of personalized web-based systems should definitely adopt client-side personalization as soon as suitable tools become available. Doing so would constitute a great step forward in terms of the data minimization principle (see Section 21.4.3) and is also likely to increase users' trust.

21.5.3 Distribution, Encrypted Aggregation, Perturbation and Obfuscation

A number of techniques have been proposed and partially also technically evaluated that can help protect the privacy of users of recommender systems that employ collaborative filtering (see Section 9 of this book [164]). Traditional collaborative filtering systems collect large amounts of information about their users in a central repository (e.g., users' product ratings, purchased products or visited web pages), to find regularities that allow for future recommendations. Such central repositories may not always be trustworthy though, and they are also likely to constitute an attractive target for unauthorized access. To some extent, central repositories may also be mined for individual user data by requesting recommendations using cleverly constructed profiles [27]. For instance, personal websites tend to be visited by their owners more frequently than by anyone else. In a recommender system that tracks users' website visits, websites that are highly correlated with personal websites are hence likely to

[17] No empirical verification for this assumption seems to exist as yet. In times of global network connectivity, this purported feeling of local control may be illusionary though. For instance, probably not many Skype users are aware that if they are not sitting behind a firewall or broadband gateway, but have good connectivity to the network, then they are pretty likely to have other people's traffic flowing through their computers (and using their network bandwidth). The pervasiveness of malware on people's computers also does not speak for a higher safety of locally stored personal data.

have been visited by those owners as well. Requesting a recommendation for pages to visit using a profile that contains this home page only may therefore reveal frequently visited web pages of its owner. Another statistical vulnerability is that correlations between an item and others will disclose much information about the choices of its raters if this item has very few raters only.

Client-side personalization (see Section 21.5.2) alone is not a remedy against such privacy attacks in collaborative filtering systems. Even when all user profiles are stored at the clients' sides, a considerable number of them (or even all) must still be merged and compiled in order that recommendations can be generated. Below we describe several strategies that are currently investigated to thwart such risks.

21.5.3.1 Distribution

One possible strategy to better safeguard individuals' data is to abandon central repositories that contain the data of all users, in favor of distributed clusters that contain information about some users only. Distribution may also improve performance and availability of the recommender system.

For instance, in the distributed match-making system Yenta [62], agents representing a user continuously form clusters of like-minded agents by exchanging information about their users and referring agents to potentially similar other agents. While this work is not explicitly aimed at protecting privacy, it does so to some extent by virtue of the fact that at any given time, agents only maintain the data of a limited number of like-minded agents and that a pseudonymity scheme can by added to protect users' identity.

The distributed PocketLens collaborative filtering algorithm [131] goes even further in terms of data avoidance. For each user, PocketLens first searches for neighbors in a P2P network and then incrementally updates the user's individual item-item similarity model by incorporating one neighbor's ratings at a time (ratings are immediately discarded thereafter). The recommendations produced by PocketLens were shown to be as good as those of the best "centralized" collaborative filtering algorithms published to date.

21.5.3.2 Aggregation of encrypted data

Canny [26, 27] proposed the usage of a secure multi-party computation scheme that allows users to privately maintain their own individual ratings, and a community of such users to compute an aggregate of their private data without disclosing them by using homomorphic encryption and peer-to-peer communication. The aggregate (a single-value decomposition of a user-item matrix) then allows personalized recommendations to be generated at the client side using one's own ratings. The scheme is however still prone to the above-mentioned statistical vulnerabilities. The PocketLens system [131] was also connected to a blackboard based on the same security schemes as those used by Canny, to allow a community of users to compute a similarity model without having to reveal their individual rankings.

21.5.3.3 Perturbation

In the perturbation approach, users' ratings are submitted to a central server which performs all collaborative filtering. These ratings become systematically altered before submission though, to hide users' true values from the server. Polat and Du [152, 153] show that adding random numbers to user ratings may still yield acceptable recommendations. The quality of recommendation based on perturbed data improves when the number of items and users increases and when the standard deviation of the perturbation function decreases (the latter obviously reduces privacy). The authors conducted a series of experiments with two databases of user rankings, namely Jester [75] and MovieLens [74], using a privacy measure proposed by Agrawal and Agrawal [4] that is based on differential entropy between the unperturbed and the perturbed data. For the Jester database, the authors find that privacy levels of about 97% and 90% will introduce average errors of about 13% and 5%, respectively, compared with predictions based on unperturbed data. For MovieLens, the average relative errors due to perturbation at these privacy levels were 10% and 5%, respectively.

21.5.3.4 Obfuscation

In the obfuscation approach of Berkovsky et al. [19], a certain percentage of users' ratings become replaced by different values before the ratings are submitted to a central server for collaborative filtering. Users are supposed to be able to freely choose which of their data should be obfuscated, and to "plausibly deny" the accuracy of any of their data should they become compromised. In subsequent work, Berkovsky et al. [20] combined obfuscation with distributed recommendation generation by ad-hoc peers, which adds an additional layer of privacy protection through distribution (see Section 21.5.3.1).

The authors performed experiments on the user ratings of the Jester [75], MovieLens [74] and EachMovie [125] recommender systems. They varied the ratio of obfuscated data in users' submitted rankings and compared the ensuing loss of prediction accuracy. They found that obfuscation of the true rating through replacement by the following values had the smallest impact on the prediction error (in the range of 5-7% at an obfuscation rate of 90%): the means of the ratings scale, a random value from the scale, and a random value from the scale taking the means and variance of the ratings in the data set into account. In contrast, uniform replacement by the highest or lowest scale value resulted in an about 300% increased prediction error at a 90% obfuscation rate.

In all these experiments, the data to be obfuscated were randomly selected for each individual user. This strategy does not take into account that users are likely to prefer obfuscation for certain kinds of data rather than random data (see Section 21.2.3). Such a tendency is likely to further increase the prediction error. Recent experiments by the authors showed that obfuscating 10% of the ratings at the high end of the scale affected the prediction error more than obfuscating 10% of mid-scale ratings [18].

21.5.3.5 Consequences for the design of personalized systems

Distribution, aggregation of encrypted user data, perturbation and obfuscation constitute promising privacy-protecting techniques. They can be supplemented by pseudonymity in applications where anonymity of users or their user models is additionally desired (see Section 21.5.1). While aggregation of encrypted user data cannot defeat

attacks on statistical vulnerabilities that were discussed at the beginning of Section 21.5.3, perturbation and obfuscation may be able to thwart them (specifically if users are aware of their "weak statistical spots" and elect to obfuscate them). Experiments will need to determine the required level of perturbation or obfuscation that guarantees a high degree of protection.

While these techniques have so far only been investigated in the area of recommender systems, it is likely that distribution, perturbation and obfuscation can in principle be applied to virtually any machine learning technique that computes aggregate data based on individual user data (learning of encrypted user data will only be possible if a suitable homomorphic encryption can be found). The effects on the quality of the learning results still remain to be seen, however.

21.5.4 Personalizing Privacy

Individual privacy preferences may differ between users (see Section 21.2), and applicable privacy laws may also be different for users from different states and countries (see Section 21.4). Different privacy preferences and laws impose different requirements on admissible personalization methods for each user. Personalized systems should therefore cater to the different privacy needs of individual users, i.e. they should "personalize privacy" [37, 106].

So far, there only exist two simplistic "solutions" to this problem:

1. *Largest permissible common subset approach*. In this approach, only those personalization methods are used that satisfy the privacy laws and regulations of all jurisdictions of all users of a website. The Disney website, for instance, observes both the U.S. Children's Online Privacy Protection Act [184], and the European Union Directive [51]. This solution is likely to run into problems if more than a very few jurisdictions are involved, since the largest common subset of permissible personalization methods may then become very small. The approach also does not take users' individual privacy preferences into account.
2. *Different country/region versions*. In this approach, personalized systems have different country versions, each of which uses only those personalization methods that are permitted in the respective country. If countries have similar privacy laws, these countries can be pooled using the above-described largest permissible common subset approach. For example, IBM's German-language pages comply with the privacy laws of Germany, Austria and Switzerland [91], while IBM's U.S. site meets the legal constraints of the U.S. only. As with the largest permissible common subset method, this approach also does not scale well when the number of countries/regions, and hence the number of different versions of the personalized system, increases. It also does not take users' individual privacy preferences into account.

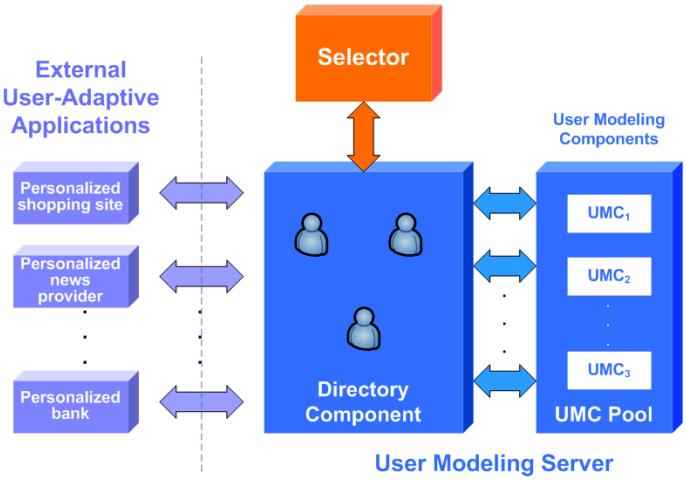

Fig. 21.1. Dynamic privacy-enabling personalization infrastructure (from [192])

Wang and Kobsa [191, 192] developed an architecture that allows personalized systems to provide optimal personalization benefits for each user, while at the same time satisfying the privacy constraints that apply to each individual user (e.g., their privacy preferences, and applicable laws and regulations). Figure 21.1 gives an overview of this architecture. The Directory Component is a repository of user models, each of which also includes the user's privacy constraints stemming from personal preferences and applicable laws and regulations. The UMC Pool contains a set of User Modeling Components, each of which encapsulates a user modeling method that operates upon the user models in the Directory Component, such as collaborative filtering (see Chapter 9 of this book [164]) or case-based recommendation (see Chapter 11 of this book [171]). On the left-hand side we see user-adaptive clients that access models of their current users in order to personalize their interaction with them.

As described so far, this architecture is similar to the one presented by Fink and Kobsa [57, 130], which was also used in a commercial user modeling server. The novel privacy enhancement consists in each user having his or her own instance of the UMC Pool, each containing only those user modeling components that meet the privacy requirements for the respective user (users with identical UMC Pool instances share the same instance). To realize this, the above architecture has been implemented as a Software Product Line (SPL) architecture [22, 33], with the UMCs as optional elements. At the beginning of the interaction with a user, a Selector verifies for every UMC whether it is allowed to operate under the privacy constraints that apply to the specific user, and creates an architectural instance with those permissible UMCs (or lets the user share this instance if one already exists). The special SPL management environment that we employ [7, 67, 189] even supports dynamic runtime (re-) configuration, which allows the Selector to react immediately should, e.g., users change their privacy preferences during the current session. The architecture therefore fully supports compliance with the consent principles discussed in Section 21.4.1 and

21.4.3, allowing a website to adjust its data practices to the user's preferences in a nuanced and highly dynamic manner.

21.6 Conclusion

A tension exists between personalization and privacy in web-based systems. On the one hand, personalization provides benefits to both users and operators of personalized websites. On the other hand, Internet users have high concerns regarding their privacy online, which may make them reluctant to disclose data about themselves to personalized systems. This poses a threat to personalization, whose quality hinges strongly on the amount of personal data supplied. The problem is exacerbated by the fact that many countries and states have privacy laws enacted that affect the permissibility of personalization methods, and that some company and industry regulations as well as principles of fair information practices have the same effect.

This chapter described a number of approaches that can be taken to render personalization more compatible with privacy. It first discussed measures that have proven to increase users' willingness to disclose data about themselves, mostly through increased trust (one of these measures, pointedly, consists in increasing the value of personalization as perceived by the user). It then analyzed how privacy legislation, self-regulation and principles of fair information practices impact the usage of personalization methods. Finally, it presented a number of technical solutions specifically intended for personalized systems that may either lessen the privacy problem in the first place (albeit no verification through user studies seems to have taken place as yet), or help developers of personalized systems adjust personalization individually to users' privacy preferences and to normative demands stemming from privacy laws, regulations and principles.

Personalization has already made some inroads into current commercial websites (see Chapter 16 of this book [72]). Given the high privacy concerns of today's Internet users, further advances are likely to only take place if privacy plays a much more important role in the future. Research on Privacy-Enhanced Personalization aims at reconciling the goals and methods of user modeling and personalization with privacy considerations, and at achieving the best possible personalization within the boundaries set by privacy. Many of the approaches described in this chapter are ready to be deployed to practical systems, and feedback from such deployments will in turn be very informative for research. Other approaches still need further technical development or evaluation in user experiments and may yield fruitful solutions in the future.

Acknowledgments

The preparation of this article has been supported by grant IIS 0308277 of the National Science Foundation, by a Trans-Coop grant, and by an Alexander von Humboldt Research Award. The author would like to thank Peter Brusilovsky, Lorrie Cranor, Miriam Metzger, Sameer Patil, Yang Wang and the anonymous reviewers for valuable comments on an earlier draft.

References

1. Ackerman, M. S., Cranor, L. F., and Reagle, J.: Privacy in E-commerce: Examining User Scenarios and Privacy Preferences. First ACM Conference on Electronic Commerce, Denver, CO (1999) 1-8, DOI 10.1145/336992.336995.
2. Acquisti, A.: Privacy in Electronic Commerce and the Economics of Immediate Gratification. EC'04 ACM Conference on Electronic Commerce, New York, NY (2004) 21-29, DOI 10.1145/988772.988777.
3. Acquisti, A. and Grossklags, J.: Privacy and Rationality in Individual Decision Making. IEEE Security & Privacy 3, (2005) 26-33, DOI 10.1109/MSP.2005.22.
4. Agrawal, D. and Aggarwal, C. C.: On the Design and Quantification of Privacy Preserving Data Mining Algorithms. 20th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database System, Santa Barbara, CA (2001) 247–255.
5. Andrade, E. B., Kaltcheva, V., and Weitz, B.: Self-Disclosure on the Web: The Impact of Privacy Policy, Reward, and Brand Reputation. In: Advances in Consumer Research, Broniarczyk, S. M. and Nakamoto, K., Eds. Valdosta, GA: Assoc for Consumer Research (2002) 350-353, http://bear.cba.ufl.edu/weitz/papers/Andrade, Kaltcheva, Weitz.pdf.
6. APEC Privacy Framework. Asia-Pacific Economic Cooperation, Singapore, Report APEC#205-SO-01.2, (2005).
7. Archstudio: ArchStudio 3.0. (2006), http://www.isr.uci.edu/projects/archstudio/
8. Arlein, R. M., Jai, B., Jakobsson, M., Monrose, F., and Reiter, M. K.: Privacy-Preserving Global Customization. 2nd ACM Conference on Electronic Commerce, Minneapolis, MN (2000) 176-184, DOI 10.1145/352871.352891.
9. AT&T: AT&T Privacybird. (2006), http://www.privacybird.com/
10. AU-NPP: National Privacy Principles. The Office of the Privacy Commissioner, Australia (2001), http://www.privacy.gov.au/publications/npps01.html
11. Bachem, C.: Profilgestütztes Online Marketing. Personalisierung im E-Commerce, Hamburg, Germany (1999).
12. Bart, Y., Shankar, V., Sultan3, F., and Urban, G. L.: Are the Drivers and Role of Online Trust the Same for All Web Sites and Consumers? A Large-Scale Exploratory Empirical Study. Journal of Marketing 69, (2005) 133-152, DOI 10.1509/jmkg.2005.69.4.133.
13. Basso, A., Goldberg, D., Greenspan, S., and Weimer, D.: First Impressions: Emotional and Cognitive Factors Underlying Judgments of Trust E-Commerce. 3rd ACM Conference on Electronic Commerce (2001) 137-143, DOI 10.1145/501158.501173.
14. Baumeister, R. F. and Leary, M. R.: The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation. Psychological Bulletin 117, (1995) 497-529.
15. Behrens, L., Ed.: Privacy and Security: The Hidden Growth Strategy (2001), http://www.gartner.com/5_about/press_releases/2001/pr20010807d.html.
16. Bellman, S., Lohse, G. L., and Johnson, E. J.: Predictors of Online Buying Behavior. Communications of the ACM 42, (1999) 32-38, DOI 10.1145/322796.322805.
17. Berendt, B., Günther, O., and Spiekermann, S.: Privacy in E-Commerce: Stated Preferences vs. Actual Behavior. Communications of the ACM 48, (2005) 101-106, DOI 10.1145/1053291.1053295.
18. Berkovsky, S.: Personal Communication. (2006).
19. Berkovsky, S., Eytani, Y., Kuflik, T., and Ricci, F.: Privacy-Enhanced Collaborative Filtering. In: PEP05, UM05 Workshop on Privacy-Enhanced Personalization, Kobsa, A. and Cranor, L., Eds. Edinburgh, Scotland (2005) 75-83, http://www.isr.uci.edu/pep05/papers/PEPfinal.pdf.
20. Berkovsky, S., Eytani, Y., Kuflik, T., and Ricci, F.: Hierarchical Neighborhood Topology for Privacy Enhanced Collaborative Filtering. Proceedings of PEP06, CHI 2006

Workshop on Privacy-Enhanced Personalization, Montreal, Canada (2006) 6-13, http://www.isr.uci.edu/pep06/papers/PEP06_BerkovskyEtAl.pdf.
21. Borking, J. J. and Raab, C. D.: Laws, PETs and other Technologies for Privacy Protection. Journal of Information, Law and Technology, (2001), http://elj.warwick.ac.uk/jilt/01-1/borking.html.
22. Bosch, J.: Design and Use of Software Architectures: Adopting and Evolving a Product-Line Approach. New York: Addison-Wesley (2000).
23. Brusilovsky, P. and Millán, E.: User Models for Adaptive Hypermedia and Adaptive Educational Systems. In: The Adaptive Web: Methods and Strategies of Web Personalization, Brusilovsky, P., Kobsa, A., and Nejdl, W., Eds. Berlin Heidelberg New York: Springer Verlag (2007) this volume.
24. Burke, R.: Hybrid Web Recommender Systems. In: The Adaptive Web: Methods and Strategies of Web Personalization, Nejdl, W., Ed. Heidelberg, Germany: Springer Verlag (2006) In this volume.
25. Burkert, H.: Privacy-Enhancing Technologies: Typology, Critique, Vision. In: Technology and Privacy: The New Landscape, Agre, P. E. and Rotenberg, M., Eds. Boston, MA: MIT Press (1997) 126-143.
26. Canny, J.: Collaborative Filtering with Privacy. IEEE Symposium on Security and Privacy, Oakland, CA (2002) 45-57, DOI 10.1109/SECPRI.2002.1004361.
27. Canny, J.: Collaborative Filtering with Privacy via Factor Analysis. 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2002), Tampere, Finland (2002) 238-245, DOI 10.1145/564376.564419.
28. Cassel, L. and Wolz, U.: Client Side Personalization. DELOS Workshop: Personalisation and Recommender Systems in Digital Libraries, Dublin, Ireland (2001), http://www.ercim.org/publication/ws-proceedings/DelNoe02/CasselWolz.pdf.
29. Ceri, S., Dolog, P., Matera, M., and Nejdl, W.: Model-Driven Design of Web Applications with Client-Side Adaptation. In: Web Engineering: 4th International Conference, ICWE 2004, Koch, N., Fraternali, P., and MartinWirsing, Eds. Berlin – Heidelberg: Springer Verlag (2004) 201-214, DOI 10.1007/b99180.
30. Privacy Policies Critical to Online Consumer Trust. Columbus Group and Ipsos-Reid, Canadian Inter@ctive Reid Report (2001), http://www.ipsos-na.com/news/pressrelease.cfm?id=1171.
31. Chellappa, R. K. and Sin, R.: Personalization versus Privacy: An Empirical Examination of the Online Consumer's Dilemma. Information Technology and Management 6, (2005) 181-202, DOI 10.1007/s10799-005-5879-y.
32. ChoiceStream Personalization Survey: Consumer Trends and Perceptions. (2005), http://www.choicestream.com/pdf/ChoiceStream_PersonalizationSurveyResults2005.pdf
33. Clements, P. and Northrop, L.: Software Product Lines: Practices and Patterns. New York: Addison-Wesley (2002).
34. Cooperstein, D., Delhagen, K., Aber, A., and Levin, K.: Making Net Shoppers Loyal. Forrester Research, Cambridge, MA June 1999 (1999).
35. Coroama, V.: The Smart Tachograph: Individual Accounting of Traffic Costs and Its Implications. In: Pervasive Computing: 4th International Conference, PERVASIVE 2006,, Fishkin, K. P., Schiele, B., Nixon, P., and Quigley, A., Eds. Berlin – Heidelberg: Springer Verlag (2006) 135-152, DOI 10.1007/11748625.
36. Coroama, V. and Langheinrich, M.: Personalized Vehicle Insurance Rates: A Case for Client-Side Personalization in Ubiquitous Computing. Proceedings of PEP06, CHI 2006 Workshop on Privacy-Enhanced Personalization, Montreal, Canada (2006) 56-59, http://www.isr.uci.edu/pep06/papers/PEP06_CoroamaLangheinrich.pdf.
37. Cranor, L. F.: 'I Didn't Buy it for Myself': Privacy and Ecommerce Personalization. 2003 ACM Workshop on Privacy in the Electronic Society, Washington, DC (2003), DOI 10.1145/1005140.1005158.

38. Cranor, L. F., Guduru, P., and Arjula, M.: User Interfaces for Privacy Agents. ACM Transactions on Human-Computer Interactions, (2006), DOI 10.1145/1165734.1165735.
39. Cranor, L. F., Reagle, J., and Ackerman, M. S.: Beyond Concern: Understanding Net Users' Attitudes About Online Privacy. AT&T Labs - Research, Technical Report TR 99.4.3, (1999), http://www.research.att.com/resources/trs/TRs/99/99.4/99.4.3/report.htm.
40. Largest 100 US Firms Rated on Customer Online Experience They Provide in Third Annual Customer Respect Group Study. The Customer Respect Group (2004), http://www.customerrespect.com
41. Model Code for the Protection of Personal Information. Canadian Standards Association (1996), http://www.csa.ca/standards/privacy/code/Default.asp?language=English
42. Culnan, M. J.: Managing Privacy Concerns Through Fairness and Trust: Implications for Marketing. Visions for Privacy in the 21st Century: A Search for Solutions. Conference Papers., Victoria, BC (1996), http://www.privacyexchange.org/iss/confpro/bcculnan.html.
43. Culnan, M. J. and Armstrong, P. K.: Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation. Organization Science 10, (1999) 104-115, http://links.jstor.org/sici?sici=1047-7039%28199901%2F02%2910%3A1%3C104%3AIPCPFA%3E2.0.CO%3B2-H.
44. Culnan, M. J. and Milne, G. R.: The Culnan-Milne Survey on Consumers & Online Privacy Notices: Summary of Responses. Interagency Public Workshop: Get Noticed: Effective Financial Privacy Notices, Washington, D.C. (2001), http://www.ftc.gov/bcp/workshops/glb/supporting/culnan-milne.pdf.
45. UCO Software To Address Retailers $6.2 Billion Privacy Problem: Cyber Dialogue Survey Data Reveals Lost Revenue for Retailers Due to Widespread Consumer Privacy Concerns. Cyber Dialogue, (2001), http://www.cyberdialogue.com/news/releases/2001/11-07-uco-retail.pdf
46. Act 101 of April 4, 2000 on the Protection of Personal Data and on Amendment to Some Acts. Czech Republic (2000), http://www.uoou.cz/index.php?l=en&m=left&mid=01:105
47. Czarkowski, M. and Kay, J.: Bringing Scrutability to Adaptive Hypertext Teaching. In: Intelligent Tutoring Systems 2000, Gauthier, G., Frasson, C., and VanLehn, K., Eds. Berlin: Springer (2000) 423-433, www.springerlink.com/link.asp?id=q0qvb6qu7nbtxkxp.
48. Czarkowski, M. and Kay, J.: How to Give the User a Sense of Control Over the Personalization of AH? AH2003: Workshop on Adaptive Hypermedia and Adaptive Web-Based Systems, Budapest, Hungary; Johnstown, PA; Nottingham, England (2003), http://wwwis.win.tue.nl/ah2003/proceedings/paper11.pdf.
49. D'Hertefelt, S.: Trust and the Perception of Security. 3 January 2000 (2000), http://www.interactionarchitect.com/research/report20000103shd.htm.
50. Teleservices Data Protection Act 1997, as amended on 14 Dec. 2001. (2001), http://bundesrecht.juris.de/tddsg/BJNR187100997.html
51. Personal Communication, Chief Privacy Officer, Disney Corporation. (2002).
52. Dommeyer, C. J. and Gross, B. L.: What Consumers Know and What They Do: An Investigation of Consumer Knowledge, Awareness, and Use of Privacy Protection Strategies. Journal of Interactive Marketing 17, (2003) 34-51, DOI 10.1002/dir.10053.
53. Earp, J. B. and Baumer, D.: Innovative Web Use to Learn About Consumer Behavior and Online Privacy. Communications of the ACM Archive 46, (2003) 81 - 83, DOI 10.1145/641205.641209.
54. Edelman, B.: Adverse Selection in Online "Trust" Certifications. Harvard University, Working Draft 15 Oct. 2006 (2006), http://www.benedelman.org/publications/advsel-trust-draft.pdf.
55. 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. Official Journal of the European Communities, (1995) 31ff, http://158.169.50.95:10080/legal/en/dataprot/directiv/directiv.html.

56. Directive 2002/58/EC of the European Parliament and of the Council Concerning the Processing of Personal Data and the Protection of Privacy in the Electronic Communications Sector. (2002), http://register.consilium.eu.int/pdf/en/02/st03/03636en2.pdf
57. Fink, J.: User Modeling Servers - Requirements, Design, and Evaluation. Amsterdam, Netherlands: IOS Press (2004), http://books.google.com/books?q=isbn:1586034057.
58. Fink, J., Kobsa, A., and Nill, A.: Adaptable and Adaptive Information Provision for All Users, Including Disabled and Elderly People. The New Review of Hypermedia and Multimedia 4, (1998) 163-188, http://www.ics.uci.edu/~kobsa/papers/1998-NRHM-kobsa.pdf.
59. Fogg, B. J.: Persuasive Technology: Using Computers to Change What We Think and Do. San Francisco: Morgan Kaufmann Publishers (2003).
60. Fogg, B. J., Kameda, T., Boyd, J., Marshall, J., Sethi, R., Sockol, M., and Trowbridge, T.: Stanford-Makovsky Web Credibility Study 2002: Investigating What Makes Web Sites Credible Today. A Research Report by the Stanford Persuasive Technology Lab & Makovsky & Company, Stanford, CA (2002), http://www.webcredibility.org/pdf/Stanford-MakovskyWebCredStudy2002-prelim.pdf.
61. Fogg, B. J., Soohoo, C., Danielson, D. R., Marable, L., Stanford, J., and Tauber, E. R.: How Do Users Evaluate the Credibility of Web Sites?: a Study with over 2,500 Participants. Conference on Designing for User Experiences, San Francisco, CA (2003) 1-15, DOI 10.1145/997078.997097.
62. Foner, L. N.: Yenta: A Multi-Agent Referral-Based Matchmaking System. International Conference on Autonomous Agents, Marina del Rey, CA (1997) 301-307, DOI 10.1145/267658.267732.
63. Fox, S., Rainie, L., Horrigan, J., Lenhart, A., Spooner, T., and Carter, C.: Trust and Privacy Online: Why Americans Want to Rewrite the Rules. The Pew Internet & American Life Project, Washington, DC (2000), http://www.pewinternet.org/report_display.asp?r=19.
64. Frankowski, D., Cosley, D., Sen, S., Terveen, L., and Riedl, J.: You Are What You Say: Privacy Risks of Public Mentions. 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Seattle, WA (2006) 565-572, DOI 10.1145/1148170.1148267.
65. Privacy Online: Fair Information Practices in the Electronic Marketplace. A Report to Congress. Federal Trade Commission (2000), http://www.ftc.gov/reports/privacy2000/privacy2000.pdf
66. Gabber, E., Gibbons, P. B., Matias, Y., and Mayer, A.: How to Make Personalized Web Browsing Simple, Secure, and Anonymous. In: Financial Cryptography'97. Berlin - Heidelberg - New York: Springer Verlag (1997), DOI 10.1007/3-540-63594-7_64.
67. Garg, A., Critchlow, M., Chen, P., van derWesthuizen, C., and van der Hoek, A.: An Environment for Managing Evolving Product Line Architectures. 19th IEEE International Conference on Software Maintenance, Amsterdam, Netherlands (2003) 358-367, http://www.ics.uci.edu/~andre/papers/C31.pdf.
68. Gefen, D., Karahanna, E., and Straub, D. W.: Trust and TAM in Online Shopping: An Integrated Model. MIS Quarterly 27, (2003) 51-90, http://search.epnet.com/login.aspx?direct=true&db=buh&an=9284295.
69. Gefen, D., Srinivasan Rao, V., and Tractinsky, N.: The Conceptualization of Trust, Risk and Their Electronic Commerce: the Need for Clarifications. 36th Annual Hawaii International Conference on Systems Sciences, Big Island, HI (2003), DOI 10.1109/HICSS.2003.1174442.
70. Gideon, J., Cranor, L., Egelman, S., and Acquisti, A.: Power Strips, Prophylactics, and Privacy, Oh My! Second Symposium on Usable Privacy and Security, Pittsburgh, Pennsylvania (2006) 133-144, DOI 10.1145/1143120.1143137.

71. Goldberg, I. A.: Privacy-Enhancing Technologies for the Internet, II: Five Years Later. In: Privacy Enhancing Technologies – Second International Workshop, PET 2002, Dingledine, R. and Syverson, P., Eds. Berlin - Heidelberg: Springer Verlag (2003) 1-12.
72. Goy, A., Ardissono, L., and Petrone, G.: Personalization in E-Commerce Applications. In: The Adaptive Web: Methods and Strategies of Web Personalization, Brusilovsky, P., Kobsa, A., and Nejdl, W., Eds. Berlin Heidelberg New York: Springer Verlag (2007) this volume.
73. Grabner-Kräuter, S. and Kaluscha, E. A.: Empirical Research in On-line Trust: a Review and Critical Assessment. International Journal of Human-Computer Studies 58, (2003) 783-812, DOI 10.1016/S1071-5819(03)00043-0.
74. GroupLens Research: movielens: Helping You Find the Right Movies. (2007), http://movielens.umn.edu/login
75. Gupta, D., Digiovanni, M., Narita, H., and Goldberg, K.: Jester 2.0: A New Linear-Time Collaborative Filtering Algorithm Applied to Jokes. Workshop on Recommender Systems Algorithms and Evaluation, 22nd International Conference on Research and Development in Information Retrieval, Berkeley, CA (2000).
76. GVU: GVU's 10th WWW User Survey. Graphics, Visualization and Usability Lab, Georgia Tech (1998),
77. Hagen, P. R., Manning, H., and Souza, R.: Smart Personalization. Forrester Research, Cambridge, MA (1999).
78. Hann, I.-H., Hui, K.-L., Lee, T. S., and Png, I. P. L.: Online Information Privacy: Measuring the Cost-Benefit Tradeoff. Proceedings of the Twenty-Third Annual International Conference on Information Systems, Barcelona, Spain (2002) 1-10, http://aisel.isworld.org/pdf.asp?Vpath=ICIS/2002&PDFpath=02CRP01.pdf.
79. Hansen, M.: Personal communication. (2002).
80. Harper, J. and Singleton, S.: With a Grain of Salt: What Consumer Privacy Surveys Don't Tell Us. Competative Enterprises Institute (2001), http://www.cei.org/PDFs/with_a_grain_of_salt.pdf.
81. Harris, Louis and Associates and Alan F. Westin: Harris-Equifax Consumer Privacy Survey 1991. Atlanta, GA: Equifax Inc. (1991).
82. A Survey of Consumer Privacy Attitudes and Behaviors. Harris Interactive. (2000), http://www.bbbonline.org/UnderstandingPrivacy/library/harrissummary.pdf
83. Privacy Notices Research: Final Results. Harris Interactive, Inc. Study No. 15338, December 2001, http://www.bbbonline.org/UnderstandingPrivacy/library/datasum.pdf.
84. Henze, N. and Brusilivsky, P.: Open Corpus Adaptive Educational Hypermedia. In: The Adaptive Web: Methods and Strategies of Web Personalization, Brusilovsky, P., Kobsa, A., and Nejdl, W., Eds. Berlin Heidelberg New York: Springer Verlag (2007) this volume.
85. Hitchens, M., Kay, J., Kummerfeld, B., and Brar, A.: Secure Identity Management for Pseudo-Anonymous Service Access. In: Security in Pervasive Computing: Second International Conference, SPC 2005, Boppard, Germany, April 6-8, 2005. Proceedings, Hutter, D. and Ullmann, M., Eds. Berlin - Heidelberg: Springer Verlag (2005) 48-55, DOI 10.1007/b135497.
86. Hof, R., Green, H., and Himmelstein, L.: Now it's YOUR WEB. Business Week October 5, (1998) 68-75.
87. Hoffman, D. L., Novak, T. P., and Peralta, M.: Building Consumer Trust Online. Communications of the ACM 42, (1999) 80-85, DOI 10.1145/299157.299175.
88. Huberman, B. A., Adar, E., and Fine, L. R.: Valuating Privacy. Fourth Workshop on the Economics of Information Security (WEIS05), Cambridge, MA (2005), http://infosecon.net/workshop/pdf/58.pdf.

89. Hui, K.-L., Tan, B. C. Y., and Goh, C.-Y.: Online Information Disclosure: Motivators and Measurements. ACM Transactions on Internet Technology 6, (2006) 415 - 441, DOI 10.1145/1183463.1183467.
90. Hui, K.-L., Teo, H. H., and Lee, S.-Y. T.: The Value of Privacy Assurance: An Exploratory Field Experiment. MIS Quarterly 31, (2007), www.comp.nus.edu.sg/~lung/PrivacyAssurance.pdf.
91. Personal Communication, Chief Privacy Officer, IBM Zurich. (2003).
92. Ishitani, L., Almeida, V., and Wagner, M., Jr.: Masks: Bringing Anonymity and Personalization Together. IEEE Security & Privacy Magazine 1, (2003) 18-23, DOI 10.1109/MSECP.2003.1203218.
93. ISO: ISO/IEC 15408-2: Information Technology — Security Techniques — Evaluation Criteria for IT Security: Part 2: Security Functional Requirements. (1999), http://csrc.nist.gov/cc/t4/wg3/15408-2.zip
94. Jarvenpaa, S. and Tractinsky, N.: Consumer Trust in an Internet Store: A Cross-Cultural Validation. Journal of Computer Mediated Communication 5, (1999) 1-36, http://jcmc.indiana.edu/vol5/issue2/jarvenpaa.html.
95. Jarvenpaa, S. L., Tractinsky, N., and Vitale, M.: Consumer Trust in an Internet Store. Information Technology and Management 1, (2000) 45-71, DOI 10.1023/A:1019104520776.
96. Jensen, C. and Potts, C.: Privacy Policies as Decision-Making Tools: An Evaluation of Online Privacy Notices. 2004 Conference on Human Factors in Computing Systems, Vienna, Austria (2004) 471-478.
97. Jensen, C., Potts, C., and Jensen, C.: Privacy Practices of Internet Users: Self-Reports versus Observed Behavior. International Journal of Human-Computer Studies 63, (2005) 203–227, DOI 10.1016/j.ijhcs.2005.04.019.
98. Kahneman, D. and Tversky, A.: Choices, Values, and Frames. Cambridge: Cambridge Univ. Press (2000).
99. Kay, J.: A Scrutable User Modelling Shell for User-Adapted Interaction. In Basser Department of Computer Science: University of Sydney, Australia (1999), http://www.cs.usyd.edu.au/~judy/Homec/Pubs/thesis.bz2.
100. Kay, J.: Accretion Representation for Scrutable User Modeling. In: Intelligent Tutoring Systems 2000, Gauthier, G., Frasson, C., and VanLehn, K., Eds. Berlin: Springer (2000).
101. Kay, J.: Stereotypes, Student Models and Scrutability. In: Intelligent Tutoring Systems 2000, Gauthier, G., Frasson, C., and VanLehn, K., Eds. Berlin: Springer (2000).
102. Kay, J., Kummerfeld, R. J., and Lauder, P.: Foundations for Personalized Documents: a Scrutable User Model Server. Proceedings of ADCS'2001, Australian Document Computing Symposium (2001) 43-50, http://www.cs.usyd.edu.au/~judy/Homec/Pubs/2001_adcs_personis.pdf.
103. Kilian, J. and Petrank, E.: Identity Escrow. In: Advances in Cryptology — CRYPTO '98. Heidelberg - Berlin: Springer Verlag (1998) 169-185, DOI 10.1007/BFb0055715.
104. Kobsa, A.: User Modeling in Dialog Systems: Potentials and Hazards. IFIP/GI Conference on Opportunities and Risks of Artificial Intelligence Systems, Hamburg, Germany (1989) 147-165.
105. Kobsa, A.: User Modeling in Dialog Systems: Potentials and Hazards. AI & Society 4, (1990) 214-240, DOI 10.1007/BF01889941.
106. Kobsa, A.: Tailoring Privacy to Users' Needs (Invited Keynote). In: User Modeling 2001: 8th International Conference, Bauer, M., Gmytrasiewicz, P. J., and Vassileva, J., Eds. Berlin - Heidelberg: Springer Verlag (2001) 303-313, http://www.ics.uci.edu/~kobsa/papers/2001-UM01-kobsa.pdf.
107. Kobsa, A.: Personalization and International Privacy. Communications of the ACM 45, (2002) 64-67, DOI 10.1145/767193.767196.

108. Kobsa, A. and Cranor, L., Eds.: Proceedings of the UM05 Workshop 'Privacy-Enhanced Personalization'. Edinburgh, Scotland (2005), http://www.isr.uci.edu/pep05/papers/w9-proceedings.pdf.
109. Kobsa, A., Koenemann, J., and Pohl, W.: Personalized Hypermedia Presentation Techniques for Improving Customer Relationships. The Knowledge Engineering Review 16, (2001) 111-155, DOI 10.1017/S0269888901000108.
110. Kobsa, A. and Schreck, J.: Privacy through Pseudonymity in User-Adaptive Systems. ACM Transactions on Internet Technology 3, (2003) 149–183, DOI 10.1145/767193.767196.
111. Kobsa, A. and Teltzrow, M.: Contextualized Communication of Privacy Practices and Personalization Benefits: Impacts on Users' Data Sharing Behavior. In: Privacy Enhancing Technologies: Fourth International Workshop, PET 2004, Toronto, Canada, Martin, D. and Serjantov, A., Eds. Heidelberg, Germany: Springer Verlag (2005) 329-343, DOI 10.1007/11423409_21.
112. Kohavi, R.: Mining E-Commerce Data: the Good, the Bad, and the Ugly. Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA (2001) 8-13, DOI 10.1145/502512.502518.
113. Krause, J., Hirschmann, A., and Mittermaier, E.: The Intelligent Help System COMFOHELP: Towards a Solution of the Practicability Problem for User Modeling and Adaptive Systems. User Modeling and User-Adapted Interaction: The Journal of Personalization Research 3, (1993) 249-282, DOI 10.1007/BF01257891.
114. Kumaraguru, P. and Cranor, L. F.: Privacy Indexes: A Survey of Westin's Studies. Institute for Software Research International, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, Technical Report CMU-ISRI-5-138, December 2005 (2005), http://reports-archive.adm.cs.cmu.edu/anon/isri2005/CMU-ISRI-05-138.pdf.
115. LaRose, R. and Rifon, N. J.: Your Privacy Is Assured—of Being Disturbed: Comparing Web Sites with and Without Privacy Seals. New Media and Society 8, (2006) 1009-1029, DOI 10.1177/1461444806069652.
116. LaRose, R., Rifon, N. J., and Lee, A.: Promoting I-Safety: Effects of Privacy Warning Boxes and Privacy Seals on Risk Assessment and Online Privacy Behaviors. Paper presented at AMA, Marketing and Public Policy Conference, Salt Lake City, UT (2004).
117. Laufer, R. S., Proshansky, H. M., and Wolfe, M.: Some Analytic Dimensions of Privacy. In: Architectural Psychology. Proceedings of the Lund Conference, Rikkard Kuller, Ed. Stroundsbourg, PA: Dowden, Hutchinson & Ross (1974) 353-372.
118. Laufer, R. S. and Wolfe, M.: Privacy as a Concept and a Social Issue: A Multidimensional Developmental Theory. Journal of Social Issues 33, (1977) 22-42.
119. Louis Harris and Associates and Westin, A. F.: Commerce, Communications, and Privacy Online: A National Survey of Computer Users. (1997), http://www.harrisinteractive.org.
120. Lu, Y., Tan, B. C. Y., and Hui, K.-L.: Inducing Customers to Disclose Personal Information to Internet Businesses with Social Adjustment Benefits. ICIS 2004: Twenty-Fifth International Conference on Information Systems, Washington, D.C. (2004) 272-281, http://aisel.isworld.org/Publications/ICIS/2004/2004RP45.pdf.
121. Lutz, W.: Statement of William Lutz. American Council of Life Insurers, et al. vs. Vermont Department of Banking, Securities, and Healthcare Administration, et al. (2004), http://www.epic.org/privacy/glba/vtlutz.pdf
122. Mabley, K.: Privacy vs. Personalization: Part III. Cyber Dialogue, Inc. (2000), http://www.cyberdialogue.com/library/pdfs/wp-cd-2000-privacy.pdf
123. Malin, B., Sweeney, L., and Newton, E.: Trail Re-Identification: Learning Who You Are From Where You Have Been. Carnegie Mellon University, Laboratory for International Data Privacy, Pittsburgh, PA, Technical Report LIDAP-WP12, March 2003 (2003), http://privacy.cs.cmu.edu/people/sweeney/trails1.pdf.

124. Mayer, R. C., Davis, J. H., and Schoorman, F. D.: An Integrative Model of Organizational Trust. Academy of Management Review 20, (1995) 709-734, links.jstor.org/sici?sici=0363-7425%28199507%2920%3A3%3C709%3AAIMOOT%3E2.0.CO%3B2-9.
125. McJones, P.: Eachmovie Collaborative Filtering Data Set. (1997), http://research.compaq.com/SRC/eachmovie/
126. Metzger, M.: Communication Privacy Management in Electronic Commerce. Journal of Computer-Mediated Communication 12, (2007), http://jcmc.indiana.edu/vol12/issue2/metzger.html.
127. Metzger, M. J.: Privacy, Trust, and Disclosure: Exploring Barriers to Electronic Commerce. Journal of Computer-Mediated Communication 9, (2004), http://jcmc.indiana.edu/vol9/issue4/metzger.html.
128. Metzger, M. J.: Effects of Site, Vendor, and Consumer Characteristics on Web Site Trust and Disclosure. Communication Research 33, (2006) 155-179, DOI 10.1177/0093650206287076.
129. Metzger, M. J.: Personal Communication. (2007).
130. Micarelli, A., Gasparetti, F., Sciarrone, F., and Gauch, S.: Personalized Search on the World Wide Web. In: The Adaptive Web: Methods and Strategies of Web Personalization, Brusilovsky, P., Kobsa, A., and Nejdl, W., Eds. Berlin Heidelberg New York: Springer Verlag (2007) this volume.
131. Miller, B. N., Konstan, J. A., and Riedl, J.: PocketLens: Toward a Personal Recommender System. ACM Transactions on Information Systems 22, (2004) 437-476, DOI 10.1145/1010614.1010618.
132. Milne, G. R. and Boza, M.-E.: Trust and Concern in Consumers' Perceptions of Marketing Information Management Practices. Journal of Interactive Marketing 13, (1999) 5-24, DOI 10.1002/(SICI)1520-6653(199924)13:1<5::AID-DIR2>3.0.CO;2-9.
133. Milne, G. R. and Culnan, M. J.: Strategies for Reducing Online Privacy Risks: Why Consumers Read (or Don't Read) Online Privacy Notices. Journal of Interactive Marketing 18, (2004) 15-29, DOI 10.1002/dir.20009.
134. Miyazaki, A. D. and Krishnamurthy, S.: Internet Seals of Approval: Effects on Online Privacy Policies and Consumer Perceptions. Journal of Consumer Affairs 36, (2002) 28, DOI 10.1111/j.1745-6606.2002.tb00419.x.
135. Mobasher, B.: Data Mining for Web Personalization. In: The Adaptive Web: Methods and Strategies of Web Personalization, Brusilovsky, P., Kobsa, A., and Nejdl, W., Eds. Berlin Heidelberg New York: Springer Verlag (2007) this volume.
136. Moores, T.: Do Consumers Understand the Role of Privacy Seals in E-Commerce? Communications of the ACM 48, (2005) 86-91, DOI 10.1145/1047671.1047674.
137. Moores, T. T. and Dhillon, G.: Do Privacy Seals in E-Commerce Really Work? Communications of the ACM 46, (2003) 265 - 271, DOI 10.1145/953460.953510.
138. Mulligan, D. and Schwartz, A.: Your Place or Mine?: Privacy Concerns and Solutions for Server and Client-Side Storage of Personal Information. Computers, Freedom & Privacy Conference (1999) 81-84, DOI 10.1145/332186.332255.
139. Self-Regulatory Principles for Online Preference Marketing by Network Advertisers. Network Advertising Initiative (2006), http://www.networkadvertising.org/pdfs/NAI_principles.pdf
140. Nakashima, E.: AOL Search Queries Open Window Onto Users' Worlds. washingtonpost.com (2006), http://www.washingtonpost.com/wp-dyn/content/article/2006/08/16/AR2006081601751_pf.html
141. Neale, J. M. and Liebert, R. M.: Science and Behavior: An Introduction to Methods of Research. Englewood Cliffs, NJ: Prentice-Hall (1973).
142. Recommendation of the Council Concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data. OECD (1980), www1.oecd.org/publications/e-book/9302011E.PDF

143. Orwant, J.: Heterogenous Learning in the Doppelänger User Modeling System. User Modeling and User-Adapted Interaction: The Journal of Personalization Research 4, (1995) 107-130, DOI: 10.1007/BF01099429.
144. Microsoft Passport Network. (2006), http://www.passport.net
145. Pavlou, P. A.: Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model. International Journal of Electronic Commerce 7, (2003) 101-134, http://mesharpe.metapress.com/link.asp?id=ymy1p2ngk06wt39f.
146. Pazzani, M. J. and Billsus, D.: Content-Based Recommendation Systems. In: The Adaptive Web: Methods and Strategies of Web Personalization, Brusilovsky, P., Kobsa, A., and Nejdl, W., Eds. Heidelberg, Germany: Springer Verlag (2006) In this volume.
147. Summary Report. Privacy Commissioner – Te Mana Matapono Matatapu, New Zealand (2006), http://www.privacy.org.nz/filestore/docfiles/24153322.pdf.
148. Personalization & Privacy Survey. Personalization Consortium (2000), http://www.personalization.org/SurveyResults.pdf
149. New Survey Shows Consumers Are More Likely to Purchase At Web Sites That Offer Personalization: Consumers Willing to Provide Personal Information in Exchange for Improved Service and Benefits. Personalization Consortium, Wakefield, MA, Press Release May 9, 2001, http://web.archive.org/web/20010526174824/http://www.personalization.org/pr050901.html.
150. Pfitzmann, A. and Köhntopp, M.: Anonymity, Unobservability, and Pseudonymity: A Proposal for Terminology. In: Anonymity 2000, Federrath, H., Ed. Berlin-Heidelberg, Germany: Springer-Verlag (2001) 1-9,
151. Phelps, J., Nowak, G., and Ferrell, E.: Privacy Concerns and Consumer Willingness to Provide Personal Information. Journal of Public Policy & Marketing 19, (2000) 27-41, http://search.epnet.com/login.aspx?direct=true&db=buh&an=3215141.
152. Polat, H. and Du, W.: Privacy-Preserving Collaborative Filtering. International Journal of Electronic Commerce 9, (2003) 9-35, http://ejournals.ebsco.com/direct.asp?ArticleID=1A72U87WVYJ61B9C.
153. Polat, H. and Du, W.: SVD-based Collaborative Filtering with Privacy. ACM Symposium on Applied Computing, Santa Fe, New Mexico (2005) 791-795, DOI 10.1145/1066677.1066860.
154. Portz, K., Strong, J. M., Busta, B., and Schneider, K.: Do Consumers Understand What Web-Trust Means? The CPA Journal 70, (2000) 46-52, http://www.nysscpa.org/cpajournal/2000/1000/features/f104600a.htm.
155. Preece, J., Rogers, Y., and Sharp, H.: Interaction Design: Beyond Human-Computer Interaction. New York, NY: Wiley (2002).
156. Opinion surveys. Privacy Exchange (2003), http://www.privacyexchange.org/iss/surveys/surveys.html
157. Rao, J. R. and Rohatgi, P.: Can Pseudonymity Really Guarantee Privacy? 9th USENIX Security Symposium (2000) 85–96, http://www.usenix.org/publications/library/proceedings/sec2000/full_papers/rao/rao_html.
158. Regan, K.: Does Anyone Read Online Privacy Policies? E-Commerce Times, (2001), http://www.ecommercetimes.com/story/11303.html.
159. Riegelsberger, J., Sasse, M. A., and McCarthy, J. D.: Shiny Happy People Building Trust?: Photos on E-commerce Websites and Consumer Trust. SIGCHI Conference on Human Factors in Computing Systems, Ft. Lauderdale, FL (2003) 121-128, DOI 10.1145/642611.642634.
160. Rifon, N. J., LaRose, R., and Choi, S. M.: Your Privacy Is Sealed: Effects of Web Privacy Seals on Trust and Personal Disclosures. Journal of Consumer Affairs 39, (2005) 339-360, DOI 10.1111/j.1745-6606.2005.00018.x.
161. Rotenberg, M.: The Privacy Law Sourcebook 2004: United States Law, International Law, and Recent Developments. Washington, DC: EPIC (2005).

162. Roy, M. C., Dewit, O., and Aubert, B. A.: The Impact of Interface Usability on Trust in Web Retailers. Internet Research 11, (2001) 388-398, DOI 10.1108/10662240110410165.
163. Roy Morgan Research: Privacy and the Community. Prepared for the Office of the Federal Privacy Commissioner, Sydney (2001), http://www.privacy.gov.au/publications/rcommunity.pdf.
164. Schafer, J. B., Frankowski, D., Herlocker, J., and Sen, S.: Collaborative Filtering Recommender Systems. In: The Adaptive Web: Methods and Strategies of Web Personalization, Brusilovsky, P., Kobsa, A., and Nejdl, W., Eds. Berlin Heidelberg New York: Springer Verlag (2007) this volume.
165. Schoenbachler, D. D. and Gordon, G. L.: Trust and Customer Willingness to Provide Information in Database-Driven Relationship Marketing. Journal of Interactive Marketing 16, (2002) 2-16, DOI 10.1002/dir.10033.
166. Schreck, J.: Security and Privacy in User Modeling. Dordrecht, Netherlands: Kluwer Academic Publishers (2003), http://www.security-and-privacy-in-user-modeling.info.
167. Sheehan, K. B. and Hoy, M. G.: Flaming, Complaining, Abstaining: How Online Users Respond to Privacy Concerns. Journal of Advertising 28, (1999) 37-51, http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=2791549&site=ehost-live.
168. Sheehan, K. B. and Hoy, M. G.: Dimensions of Privacy Concern Among Online Consumers. Journal of Public Policy & Marketing 19, (2000) 62-73, http://search.epnet.com/login.aspx?direct=true&db=buh&an=3215144.
169. Singel, R.: 'Free IPod' Takes Privacy Toll. Wired, No. Issue, March 16, 2006, http://www.wired.com/news/technology/0,70420-0.html
170. Smith, H. J., Milberg, S. J., and Burke, S. J.: Information Privacy: Measuring Individuals' Concerns about Organizational Practices. MIS Quarterly 20, (1996) 167-196, http://links.jstor.org/sici?sici=0276-7783%28199606%2920%3A2%3C167%3AIPMICA%3E2.0.CO%3B2-W.
171. Smyth, B.: Case-Based Recommendation. In: The Adaptive Web: Methods and Strategies of Web Personalization, Brusilovsky, P., Kobsa, A., and Nejdl, W., Eds. Berlin Heidelberg New York: Springer Verlag (2007) this volume.
172. Spiekermann, S., Grossklags, J., and Berendt, B.: E-privacy in 2nd Generation E-Commerce: Privacy Preferences versus Actual Behavior. EC'01: Third ACM Conference on Electronic Commerce, Tampa, FL (2001) 38-47, DOI 10.1145/501158.501163.
173. Spiekermann, S., Grossklags, J., and Berendt, B.: Stated Privacy Preferences versus Actual Behaviour in EC Environments: a Reality Check. WI-IF 2001: the 5th International Conference Wirtschaftsinformatik - 3rd Conference Information Systems in Finance, Augsburg, Germany (2001) 129-148.
174. Stephanidis, C.: Adaptive Techniques for Universal Access. User Modeling and User-Adapted Interaction: The Journal of Personalization Research 11, (2001) 159-179, DOI 10.1023/A:1011144232235.
175. Stone, D., Jarrett, C., Woodroffe, M., and Minocha, S.: User Interface Design and Evaluation. San Francisco, CA: Morgan Kaufmann (2005).
176. Sweeney, L.: k-Anonymity: A Model for Protecting Privacy. International Journal on Uncertainty, Fuzziness, and Knowledge-based Systems 10, (2002) 557-570, DOI 10.1142/S0218488502001648.
177. Tam, E.-C., Hui, K.-L., and Tan, B. C. Y.: What Do They Want? Motivating Consumers to Disclose Personal Information to Internet Businesses. Proceedings of the Twenty-Third Annual International Conference on Information Systems, Barcelona, Spain (2002) 11-21, http://aisel.isworld.org/pdf.asp?Vpath=ICIS/2002&PDFpath=02CRP02.pdf.
178. Tam, K. Y. and Ho, S. Y.: Web Personalization: is it Effective? IT Professional 5, (2003) 53-57, DOI 10.1109/MITP.2003.1235611.

179. Taylor, H.: Most People Are "Privacy Pragmatists" Who, While Concerned about Privacy, Will Sometimes Trade It Off for Other Benefits. The Harris Poll #17, March 19, 2003 (2003), http://www.harrisinteractive.com/harris_poll/index.asp?PID=365.
180. Teltzrow, M. and Kobsa, A.: Impacts of User Privacy Preferences on Personalized Systems: a Comparative Study. In: Designing Personalized User Experiences for eCommerce, Karat, C.-M., Blom, J., and Karat, J., Eds. Dordrecht, Netherlands: Kluwer Academic Publishers (2004) 315-332, DOI 10.1007/1-4020-2148-8_17.
181. Turner, C. W., Zavod, M., and Yurcik, W.: Factors that Affect the Perception of Security and Privacy of Ecommerce Web Sites. Fourth International Conference on Electronic Commerce Research, Dallas TX (2001) 628-636, http://www.sosresearch.org/publications/icecr01.pdf.
182. Turow, J.: Americans and Online Privacy: The System is Broken. Annenberg Public Policy Center, University of Pennsylvania (2003), http://www.asc.upenn.edu/usr/jturow/internet-privacy-report/36-page-turow-version-9.pdf
183. Health Insurance Portability and Accountability Act of 1996. 104th Congress Aug. 21, 1996, http://aspe.hhs.gov/admnsimp/pl104191.htm.
184. Children's Online Privacy Protection Act of 1998, http://www.ftc.gov/ogc/coppa1.htm.
185. Gramm-Leach-Bliley Act of 1999. Public Law No 106-102 (1999), http://thomas.loc.gov/cgi-bin/bdquery/z?d106:SN00900:l.
186. USACM Policy Recommendations on Privacy. U.S. Public Policy Committee of the Association for Computing Machinery, New York, NY June 2006, http://www.acm.org/usacm/Issues/Privacy.htm.
187. Fifth Study of the Internet by the Digital Future Project Finds Major New Trends in Online Use for Political Campaigns. Center for the Digital Future, Annenberg School, University of Southern California (2005), http://www.digitalcenter.org/pdf/Center-for-the-Digital-Future-2005-Highlights.pdf
188. van Blarkom, G. W., Borking, J. J., and Olk, J. G. E., Eds.: Handbook of Privacy and Privacy-Enhancing Technologies: The Case of Intelligent Software Agents. The Hague, The Netherlands: TNO-FEL (2003), http://www.andrewpatrick.ca/pisa/handbook/Handbook_Privacy_and_PET_final.pdf.
189. van der Hoek, A., Rakic, M., Roshandel, R., and Medvidovic, N.: Taming Architectural Evolution. Sixth European Software Engineering Conference (ESEC) and the Ninth ACM SIGSOFT Symposium on the Foundations of Software Engineering (FSE-9), Irvine, CA (2001) 1-10, DOI 10.1145/503209.503211.
190. Wang, Y., Chen, Z., and Kobsa, A.: A Collection and Systematization of International Privacy Laws, with Special Consideration of Internationally Operating Personalized Websites. (2006), http://www.ics.uci.edu/~kobsa/privacy
191. Wang, Y. and Kobsa, A.: A Software Product Line Approach for Handling Privacy Constraints in Web Personalization. In: PEP05, UM05 Workshop on Privacy-Enhanced Personalization, Kobsa, A. and Cranor, L., Eds. Edinburgh, Scotland (2005) 35-45, http://www.ics.uci.edu/~kobsa/papers/2005-PEP-kobsa.pdf.
192. Wang, Y., Kobsa, A., van der Hoek, A., and White, J.: PLA-based Runtime Dynamism in Support of Privacy-Enhanced Web Personalization. 10th International Software Product Line Conference, Baltimore, MD (2006) 151-162, DOI 10.1109/SPLINE.2006.1691587.
193. Wenning, R., Ed.: The Platform for Privacy Preferences 1.1 (P3P1.1) Specification: W3C Working Draft (2006), http://www.w3.org/TR/P3P11.
194. White, T. B.: Consumer Disclosure and Disclosure Avoidance: A Motivational Framework. Journal of Consumer Psychology 14, (2004) 41-51, DOI 10.1207/s15327663jcp1401& 2_6.
195. Xu, Y., Tan, B. C. Y., Hui, K.-L., and Tang, W.-K.: Consumer Trust and Online Information Privacy. International Conference on Information Systems 2003, Seattle, WA (2003) 538-548, http://aisel.isworld.org/Publications/ICIS/2003/03CRP45.pdf.

196. Yankee: Interactive Consumers in the Twenty-First Century: Emerging Online Consumer Profiles, Access Strategies and Application Usage". Yankee Group 23 Oct. 2001 (2001), http://www.yankeegroup.com.
197. Zaslow, J.: If TiVo Thinks You Are Gay, Here's How to Set It Straight. Wall Street Journal (Eastern Edition), (2002) A.1, online.wsj.com/article_email/0,,SB1038261936872356908,00.html.