In: C.-M. Karat, J. Blom and J. Karat, eds: Designing Personalized User Experiences for eCommerce.
Dordrecht, Netherlands: Kluwer Academic Publishers, 2004

MAXIMILIAN TELTZROW, ALFRED KOBSA

IMPACTS OF USER PRIVACY PREFERENCES ON PERSONALIZED SYSTEMS

A Comparative Study [1]

1. INTRODUCTION

Personalized (or "user-adaptive") systems have gained substantial momentum with the rise of the World Wide Web. The market research firm Jupiter (Foster, 2000) defines personalization as predictive analysis of consumer data used to adapt targeted media, advertising and merchandising to consumer needs. According to Jupiter, personalization can be viewed as a cycle of recurring processes consisting of 'data collection', 'profiling' and 'matching': from collected data, user profiles can be created that are the basis for adapting user interfaces to individuals or groups of individuals. A more web-oriented definition was proposed by Kobsa et al. (2001) who regard a personalized hypermedia application as a hypermedia system that adapts the content, structure and/or presentation of the networked hypermedia objects to each individual user's characteristics, usage behavior and/or usage environment. In contrast to user-adaptable systems where the user is in control of the initiation, proposal, selection and production of the adaptation, user-adaptive systems perform all steps autonomously.

The advantages of personalization can be manifold. Website visitors see the major benefits in sites being able to offer more relevant content and to recall user preferences and interests (Cyber Dialogue, 2000). The personalization of hypermedia is beneficial for several other purposes as well, most notably for improving the learning progress in educational software (Brusilovsky *et al.*, 1998; Specht, 1998). Given the increasing amount of information offered on the Internet, the development of advanced personalized services seems to become inevitable.

Personalization systems need to acquire a certain amount of data about users' interests, behavior, demographics and actions before they can start adapting to them. Thus, they are often useful in domains only where users engage in extended (and

[1] This research has been supported by the National Science Foundation under Grant No. 0308277, the Deutsche Forschungsgemeinschaft Berlin-Brandenburg Graduate School in Distributed Information Systems (DFG grant no. GRK 316/2), and the Humboldt Foundation (TransCoop Program). We would like to thank Stefan Wrobel for his help in the analysis of survey results.

[Editor(s) here] (ed.), [Book Titlen here], 1—18.

most often repeated) system use. They may not be appropriate for infrequent users with typically short sessions. The extensive and repeated collection of detailed user data, however, may provoke consumer privacy concerns, which have been identified as a primary impediment to users' willingness to buy online (Pavlou, 2003). Consumer surveys show that the number of consumers refusing to shop online because of privacy concerns is as high as 64% (Culnan and Milne, 2001). Finding the right balance between privacy protection and personalization remains a challenging task. Therefore, our chapter discusses the impacts of consumers' privacy concerns on personalization systems. Furthermore, we will provide suggestions on how to increase consumer trust in such systems.

Section 2 categorizes personalization systems according to the input data they require. Section 3 summarizes privacy concerns from more than 30 consumer surveys and relates their effects on personalization systems to the categories of Section 2. Differences between consumers' privacy views and their actual behaviors, and between consumer and industry opinions on privacy, are also presented. Section 4 discusses the impact of the privacy concerns that were summarized in Section 2 on the personalization systems described in previous chapters of this book. Finally, Section 5 summarizes possible approaches to better address the tension between privacy and personalization.

2. INPUT DATA FOR PERSONALIZATION

Kobsa (2001) divides the data that are relevant for personalization purposes into 'user data', 'usage data', and 'environment data'. 'User data' denote information about personal characteristics of a user, while 'usage data' relate to a user's (interactive) behavior. A special kind of 'usage data' is 'usage regularities', which describe frequently re-occurring interactions of users. 'Environment data' refer to the user's software and hardware, and the characteristics of the user's current locale.

Table 1 lists the most frequently occurring subtypes of these data. The taxonomy allows one to refer to specific kinds of personalization systems more easily, and facilitates our analysis of privacy concerns and their impacts on certain system types.

3. RESULTS FROM PRIVACY SURVEYS

3.1 Impacts on user-adaptive systems

We categorized 30 recent consumer surveys on Internet privacy (or summaries of such surveys), and analyzed their potential impacts on the different types of personalization systems listed in Table 1. Questions from different surveys addressing the same privacy aspects were grouped together, to convey a more complete picture of user concerns. Eleven documents included all questions, six provided an extensive discussion of survey results, and ten contained factual executive summaries. For three studies, only press releases were available.

Input Data	**Examples of User-Adaptive Systems**
A) User data:	
Demographic data	Personalized web sites based on user profiles; software providers: Broadvision, Personify, NetPerceptions etc.
User knowledge	Expertise-dependent personalization; product and technical descriptions: Sales Assistant (Popp and Lödel, 1996), SETA (Ardissono and Goy, 2000); learning systems: KN-AHS (Kobsa *et al.,* 1994), Brusilovsky, 2001
User skills and capabilities	Help Systems: Unix Consultant (Chin, 1989), Küpper and Kobsa, 1999; disabilities: AVANTI (Fink *et al.,* 1998)
User interests and preferences	Recommender systems (Resnick and Varian, 1997); used car domain: Jameson *et al.,* 1995; domain of telephony devices: Ardissono and Goy, 1999
User goals and plans	Personalized support for users with targeted browsing behavior, plan recognition: (Lesh *et al.,* 1999), PUSH (Höök *et al.*, 1996), HYPERFLEX (Kaplan *et al.*, 1993)
B) Usage data:	
Selective actions	Adaptation based on link-selection: WebWatcher (Joachims *et al.,* 1997), Letizia (Lieberman, 1995); image-selection: Adaptive Graphics Analyser (Holynski, 1988)
Temporal viewing behavior	Adaptation based on viewing time; streaming objects: Joerding, 1999; temporal navigation behavior: Chittaro and Ronan, 2000; micro-interaction: Sakagami *et al.,* 1998
Ratings	Adaptation based on object ratings; product suggestions: Firefly (Shardanand and Maes, 1995), GroupLens (Konstan *et al.,* 1997); web pages: Pazzani and Billsus, 1997
Purchases and purchase-related actions	Suggestions of similar goods after product selection: Amazon.com; other purchase-related actions: registering, transferring products into virtual shopping cart, quizzes
Other (dis-) confirmatory actions	Adaptation based on other user actions, e.g. saving, printing documents, bookmarking a web page: Konstan *et al.,* 1997
C) Usage regularities:	
Usage frequency	Adaptation based on usage frequency; icon toolbar: Debevc *et al.,* 1996, Flexcel (Krogsæter *et al.*, 1994); web page visits: AVANTI (Fink *et al.*, 1998)
Situation-action correlations	Interface agents; routing mails: Mitchell *et al.,* 1994, Maes, 1994, meeting requests: Kozierok and Maes, 1993
Action Sequences	Recommendations based on frequently used action sequences, e.g. past actions, action sequences of other users
D) Environment data:	
Software environment	Adaptation based on users' browser versions and platforms, availability of plug-ins, Java and JavaScript versions
Hardware environment	Adaptation based on users' bandwidth, processor speed, display devices (e.g. resolution), input devices
Locale	Adaptation based on users' current location (e.g. country code), characteristics of usage locale

Table 1: *Types of personalization-relevant data and examined systems (summary of the taxonomy in Kobsa et al., 2001)*

We distinguished several categories of privacy aspects. The category ‘privacy of user data in general’ has a direct impact on any personalization systems that requires personal data (such as the user’s name, address, income etc.). The category ‘privacy in a commercial context’ primarily affects personalized systems in e-commerce. ‘Tracking of user sessions’ and ‘use of cookies’ influence user-adaptive systems requiring usage data. A few studies focus on ‘e-mail privacy’. This category might have an impact on user-adaptive systems that generate targeted e-mails. Two studies directly address the topic of privacy and personalization (Mabley, 2000; Personalization Consortium, 2000). They are highly interesting because they directly affect most personalization systems.

Results regarding user data in general

1. Internet Users who are concerned about the security of personal information: 83% (Cyber Dialogue, 2001), 70% (Behrens, 2001), 72% (UMR, 2001), 84% (Fox *et al*., 2000)
2. People who have refused to give (personal) information to a web site: 82% (Culnan and Milne, 2001)
3. Internet users who would never provide personal information to a web site: 27% (Fox *et al*., 2000)
4. Internet users who supplied false or fictitious information to a web site when asked to register: 34% (Culnan and Milne, 2001), 24% (Fox *et al*., 2000)
5. Online users who think that sites who share personal information with other sites invade privacy: 49% (Cyber Dialogue, 2001)

A significant concern about the use of personal information can be seen in these results, which is a problem for those personalization systems in Table 1 that require ‘user data’ (such as demographic data’, data about ‘user knowledge’, etc.). Systems that record ‘purchases and purchase-related actions’ may also be affected. More than a quarter of the respondents even indicated 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 user linking across sessions and thereby accurate recommendations based on ‘user interests and preferences’ very difficult.

Results regarding user data in a commercial context

1. People wanting businesses to seek permission before using their personal information for marketing: 90% (Roy Morgan Research, 2001)
2. Non-online shoppers who did not purchase online because of privacy concerns: 66% (Ipsos Reid, 2001), 68% (Interactive Policy, 2002), 64% (Culnan and Milne, 2001)
3. Online shoppers who would buy more if they were not worried about privacy/security issues: 37% (Forrester, 2001), 20% (Department for Trade and Industry, 2001)
4. Shoppers who abandoned online shopping carts because of privacy reasons: 27% (Cyber Dialogue, 2001)

5. People who are concerned if a business shares their data for a different than the original purpose: 91% (UMR, 2001), 90% (Roy Morgan Research, 2001)

These results suggest that in a commercial context, privacy concerns may play an even more important role than for general personalized systems. Most people want to be asked before their personal information is used, and many regard privacy as a must for Internet shopping. Thus, commercial personalization systems need to include privacy features. In particular, those systems in Table 1 that require ‘demographic data’, ‘user knowledge’, ‘user interests and preferences’, ‘user goals and plans’ and ‘purchase-related actions’ are affected.

Furthermore, more than 90% of respondents are concerned if a business shares their information for a different than the original purpose. This has a severe impact on central user modeling servers that collect data from, and share them with, different user-adaptive applications, unless sharing can be controlled by the user (Kobsa, 2001; Schreck and Kobsa, 2003).

Results regarding user tracking and cookies

1. People who are concerned about being tracked on the Internet: 60% (Cyber Dialogue, 2001), 54% (Fox *et al*., 2000), 63% (Harris, 2000)
2. People who are concerned that someone might know what web sites they visited: 31% (Fox *et al*., 2000)
3. Internet users who generally accept cookies: 62% (Personalization Consortium, 2000)
4. Internet users who set their computers to reject cookies: 25% (Culnan and Milne, 2001), 3% (Cyber Dialogue, 2001), 31% in warning modus (Cyber Dialogue, 2001), 10% (Fox *et al*., 2000)
5. Internet users who delete cookies periodically: 52% (Personalization Consortium, 2000)
6. User who feel uncomfortable being tracked across multiple web sites: 91% (Harris, 2000)

Users’ privacy concerns about tracking and cookies that became manifest in the first result affect the acceptance of personalization systems based on ‘usage data’ and ‘usage regularities’ (see Table 1). In particular, systems using ‘selective actions’, ‘temporal viewing behavior’ and ‘action sequences’ conflict with users’ privacy preferences. Results 2 – 5 directly affect machine-learning methods that operate on user log data since without cookies, sessions of the same user cannot be linked any more. Result 6 affects personalization systems that combine information from several sources, in particular those systems that use data from ‘action sequences’, ‘demographics’, ‘purchase-related actions’ and the user’s ‘locale’.

Most users do not consider current forms of tracking as helpful methods to collect data for personalization. Users’ participation in deciding when and what usage information should be tracked might decrease such privacy concerns and will be discussed in Section 4 of this chapter.

Results regarding e-mail privacy

1. People who have asked for removal from e-mail lists: 78% (Cyber Dialogue, 2001), 80% (Culnan and Milne, 2001)
2. People who complain about irrelevant e-mail: 62% (Ipsos Reid, 2001)
3. People who have received unsolicited e-mail: 95% (Cyber Dialogue, 2001)
4. People who have received offensive e-mail: 28% (Fox *et al*., 2000)

Results 2 and 3 constitute a problem for the acceptance of personalized e-mail. The problem affects primarily those systems in Table 1 that use 'situation-action correlation'. The findings indicate that many deployed e-mail personalization systems, such as software for the management of targeted marketing campaigns, are not yet able to address user needs specifically enough to evoke positive reactions among the recipients.

Results regarding privacy and personalization

1. Online users who see personalization as a good thing: 59% (Harris, 2000)
2. Online users who do not see personalization as a good thing: 37% (Harris, 2000)
3. Types of information users are willing to provide in return for personalized content: name: 88%, education: 88%, age: 86%, hobbies: 83%, salary 59%, credit card number: 13% (Cyber Dialogue, 2001)
4. Internet users who think tracking allows the site to provide information tailored to specific users: 27% (Fox *et al*., 2000)
5. Online users who think that sites who share information with other sites try to better interact: 28% (Cyber Dialogue, 2001)
6. Online users who find it useful if site remembers basic information (name, address): 73% (Personalization Consortium, 2000)
7. Online users who find it useful if a site remembers information (preferred colors, delivery options etc.): 50% (Personalization Consortium, 2000)
8. People who are bothered if a web site asks for information one has already provided (e.g., mailing address): 62% (Personalization Consortium, 2000)
9. People who are willing to give information to receive a personalized online experience: 51% (Personalization Consortium, 2000), 40% (Roy Morgan Research, 2001), 51% (Privacy & American Business, 1999)

The results of the last category directly reflect users' attitudes towards personalization, and their willingness to share personal information in return for personalized content. Results 1 and 2 affect all systems in Table 1: a significant portion of the respondents does not seem to see enough value in personalization that they would be willing to give out personal data. If any possible, personalization should therefore be designed as an option that can be switched off. Results 3 and 9 affect all systems that rely on 'user data', and results 6 - 7 additionally those that rely on 'purchases and purchase-related actions'. Result 4 applies to all systems that utilize 'usage data' and information about the 'locale' of the user. Result 5, finally, applies to all personalized systems that share information via a central user modeling server (Kobsa, 2001).

3.2 Differences in consumer statements and actual privacy practices

This meta-analysis demonstrates that consumers are highly concerned about the privacy implications of various data collection methods, but many would share some data in return for personalization.[2] Users however do not seem to always have a good understanding of their privacy needs in a personalization context. Stated privacy preferences and actual behavior often diverge:

– User tracking evokes significant privacy concerns, but only 10% (27%) of American Internet users have set their browsers to reject cookies (Fox *et al.*, 2000; Roy Morgan Research, 2001).
– 76% of survey respondents claimed that privacy policies on web sites were very important to them (Behrens, 2001), but in fact users barely view such pages when visiting web sites.[3]
– In an experiment, Berendt *et al.* (2004) found that users often do not live up to their self-reported privacy preferences: subjects claimed to be highly concerned about their privacy, but shared very personal and sensitive information with a personalized web site.

3.3 Differences in the privacy views of consumers and industry

Besides differences in consumers' self-perception and actual behavior, our analysis of survey results also uncovered a few major discrepancies in the privacy views of consumers and industry. Consumer expectations and actual industry practices should however be in line with each other, so that consumers can build trust which is the basis for the acceptance of personalization. For instance, 54% do not believe that most businesses handle the personal information they collect in a proper and confidential way (Responsys.com, 2000; Harris Interactive, 2003). In contrast, 90% of industry respondents believe that this is the case for their own business, and 46% that this is the case for industry in general.[4]

Consumer demands and current practice in companies also diverge significantly on the issue of data control. Most Internet users (86%) believe that they should be allowed control over what information is stored by a business (Fox *et al*., 2000), but only 17% of businesses allow users to delete at least some personal information (Andersen, 2001). Furthermore, 40% of businesses do not provide access to personal data for verification, correction and updates (Deloitte, 2001).

[2] Users' willingness to share information with a web site may also depend on other factors that are not considered here such as the usability of a site, users' general level of trust towards a site, and the company or industry to which the site belongs (Princeton Survey Research Associates 2002). For example, good company reputation makes 74% of the surveyed Internet users more comfortable disclosing personal information (Ipsos Reid 2001).

[3] Web site operators report a fairly low attention to privacy policies. For example, on the day after the company Excite@home was featured in a 60 Minutes segment about Internet privacy, only 100 out of 20 million unique visitors accessed that company's privacy pages (Wham 2001).

[4] However, only 40% of businesses say steps have been taken to secure personal information held by a site (Internet Privacy Survey 2001), and 55% do not store personal data in encrypted form. 15% share user data with third parties without having obtained users' permission (Deloitte 2001).

Industry and consumers also disagree significantly on the value of privacy laws. Nine of ten marketers claim that the current regime of self-regulation works for their companies, and 64% think that government involvement will ultimately hurt the growth of e-commerce (Responsys.com, 2002). In contrast, two-thirds of e-mail users think that the federal government should pass more laws to ensure citizens' privacy online (Gallup Organization, 2001), while only 15% supported self-regulation (Harris, 2000). However, Harris (2001a) found that trust in the effectiveness of privacy legislation has meanwhile decreased among consumers.

Although both governments and private organizations have made serious efforts to ease users' privacy concerns, much remains to be done to build and maintain customer confidence, which is a prerequisite for successful personalization.

3.4 Discussion of the methodology

The cited studies were mostly conducted by well-known research institutions and market research firms between 2000 and 2003. The number of respondents in the studies varied between 500 and 4500, with an average of about 2000. The answers were collected by telephone interviews and online questionnaires. From the 30 surveys analyzed, 21 were conducted in the U.S., three in Canada, two in Australia and New Zealand, two in Britain and one in the European Union. One survey was based on an international respondent sample.

Though this meta-analysis provides a more comprehensive and objective overview of privacy concerns and their impacts on personalization than can be expected from a single study, some caution should be exercised. A general problem is the lack of comparability of the studies: small differences in the wording of the questions, their context in the questionnaires, the recruitment method and the sample population make user statements difficult to compare. Harper and Singleton (2001) criticized the use of manipulative questions in many privacy studies, a lack of trade-offs between privacy and other desires, and imprecise terminology (e.g. the term "privacy" is often understood as a synonym for security, or a panacea against identity fraud and spam). Finally, as mentioned above, disparities seem to exist between people's responses to general, context-less privacy questions, and their behavior when working with concrete websites having specific goals in mind.

4. PRIVACY IMPACTS IN PERSONALIZATION DOMAINS ADDRESSED BY OTHER CHAPTERS OF THIS BOOK

The previous sections discussed impacts of users' privacy concerns on different types of personalization systems. To round out our analysis, we looked at the personalization approaches described in other chapters of this book and analyzed three privacy-related aspects: (1) which input data from Table 1 is required by these personalization systems, (2) which user privacy concerns might therefore affect these systems following the results of Section 3.1 and additional findings from privacy surveys, and (3) how users' privacy concerns could be addressed to increase trust in these systems (this third aspect will be discussed in more detail in Section 5). Most chapters focus on personalization in e-commerce, and a few on e-finance, e-travel and e-government. We will group the chapters of this book accordingly.

4.1 E-Commerce, retail

In Brodie *et al.*'s chapter on "How Personalization of an E-Commerce Website Affects Consumer Trust", the authors address the issue of user trust in personalized e-commerce applications. The chapter explores how a company's choice of personalization policies and features might affect visitors' willingness to share personal information. The authors' findings and recommendations are relevant for most types of personalization systems listed in Table 1.

Their study with a large B2B website identified several requirements for personalization. Most importantly, they found that users should be given more control of their data, since users were more willing to share personal information with the company when they were allowed to view, edit and delete their data. Brodie *et al.* suggest letting users specify when they think data collection could be useful. Their findings confirm results from several consumer surveys: in one study, 69% of consumers found it important to have control of their data (Harris, 2003). Consumers also react more positively to organizations if they have a higher perceived level control (Hine and Eve, 1998). Giving users control of when to collect data might also increase the acceptance of online user tracking, which is often considered as a privacy threat (Harris, 2000).

Second, Brodie *et al.* investigate if asking for the information only that is needed to provide an immediate service would increase the willingness of website users to share data. One survey had already revealed that most users would be willing to provide personal information in return for personalization (Personalization Consortium, 2000). The results of Brodie *et al.* enhance our understanding of this willingness in that indeed only information needed to provide an immediate service should be requested, and not all information for all personalized services that the website can potentially provide at some point.

Third, the authors suggest to let users select among different identities when interacting with a site. Users should be able to specify different interaction roles for personalization, e.g. as an employee at work or a private individual at home. Giving users more control over the persona they disclose could increase their trust in a company's ability to provide useful personalization.

Similar methods to ease privacy concerns could be applied in the system described by Hoelscher and Dietrich in their chapter "E-Commerce Personalization and Real-Time Site Monitoring". Their 7d system is based on 'demographic data', 'usage data' and 'usage regularities', aggregating the entirety of a user's interactions with a site and attempting to draw useful conclusions based thereon.

This combination of data sources may raise severe privacy concerns however: more than half of all Internet users are concerned about being tracked on the Internet (Cyber Dialogue, 2001). Thus, user concerns about merging different data need to be addressed adequately to increase acceptance and usage of such personalization systems. Asking users for explicit consent – as described by Brodie *et al.* – might be a reasonable way to lessen consumer concerns. Furthermore, giving users more control over the persona they disclose could be a protection against privacy threats. Personal information could be saved under pseudonyms and not users' real names. Communicating such a policy could significantly ease users' privacy concerns. Furthermore, privacy legislation needs to be considered. According to the EU Directive on Privacy and Electronic Communications (EU, 2002), purchase data must be deleted if they are not needed any more for the original purpose, unless the user explicitly permitted a longer retention or their use for secondary purposes.

In their chapter "Recommending as Personalized Teaching", Stolze and Ströbel focus on needs-based recommendation systems for web sites. The authors developed a recommender prototype that enhances customers' understanding of the mapping between stated needs and personalized product recommendations.

The described system requires data from the categories 'user interests and preferences' and 'user goals and plans' (see Table 1). Privacy concerns about personal information may hinder the acceptance and use of such systems, however. Only about 50% of Internet users indicate to be willing to divulge personal information to receive a personalized online experience (Personalization Consortium, 2000; Roy Morgan Research, 2001; Privacy & American Business, 1999). The authors suggest explaining the underlying personalization algorithms to users, who thereby may feel they can better evaluate the appropriateness and credibility of such a system. This approach is useful from a privacy perspective, as users may develop more trust in personalization systems if they better understand their potential benefits, and specifically, how the personalization relates to their own needs (Hine and Eve, 1998).

Häubl *et al.*'s work on personalization systems deals with certain 'user data' and 'usage data'. Specifically, it employs information about 'user interests and preferences', 'selective actions' and 'ratings' to create recommendations for users. The chapter describes the benefits that personalized electronic shopping environments may provide. For example, one user experiment showed that product recommendations might allow consumers to reduce search costs and improve decision quality. This finding confirms results from user surveys: most people would welcome personalization services that simplify their browsing experience (Personalization Consortium, 2000; Harris, 2000). Besides the various personalization benefits described in this chapter, privacy concerns need to be considered as well. 82% of users have refused to give personal information to a web site due to privacy concerns (Culnan and Milne, 2001), and 27% would never share personal information with a web site (Fox *et al.* 2000). Thus, for a decision aid to be effective, users must be convinced to trust it. The approaches suggested in the chapters by Brodie *et al.* and Stolze and Ströbel to increase users' confidence and trust in personalization systems may apply here as well.

Degemmis *et al.*'s chapter on "Improving Collaborative Recommender Systems by Means of User Profiles" presents a hybrid approach combining collaborative and feature-based filtering to construct individual user models. Their system exploits 'usage data' and may raise similar privacy concerns as the one by Hoelscher and Dietrich. Additionally, people using recommender systems may not want their habits or views to be widely known (Resnick, 1997). Actions need to be taken to decrease user concerns about merging different data sources, which is disliked by most people (Harris, 2000). Thus, the described system needs privacy preservation mechanisms such as explicit user consent and data access control, to reduce consumer concerns and to comply with international privacy regulations.

4.2 Travel

In their chapter "Supporting Travel Decision Making Through Personalized Recommendation", Ricci and Del Missier propose a combination of collaborative and content-based recommendation to personalize a travel site. Users can provide both content features (characteristics of their planned travel) and collaborative features (characteristics of themselves, particularly their travel preferences). The system first uses content features to retrieve travel options, where the user may relax or tighten the result set. Then the collaborative component searches for similar past travel plans of other people, to produce a ranked list of recommendations.

The system collects a variety of personal information related to a user's travel plan. Personal information is stored in a recommendation session and includes travel preferences, travel products chosen and – as soon as the user registers – also personally identifiable information (PII) such as age, address, nationality and gender. According to the terminology in Table 1, the system requires the data types 'user interests and preferences', 'user goals and plans' and 'demographic data'.

From a privacy point of view, the proposed collaborative filtering may raise few concerns as long as users can remain anonymous. The content-based filtering technique tends to be more privacy-critical since it is based on explicit user needs, which may entail that users become uniquely identified if considerable information is available about them (cf. Sweeney, 2001). The privacy criticality increases when users must register with the travel recommender system, since then travel preferences can be linked with PII, which evokes severe privacy concerns and may also conflict with existing laws (Harris, 2000; EU, 2002). Thus, privacy approaches as discussed before and in section 5 of this chapter are necessary to balance privacy concerns and personalization benefits.

4.3 E-Government

One of the main issues in the chapter by Halstead-Nussloch is the relationship between personalization, privacy and trust in e-government. It draws comparisons between e-government and e-commerce to identify characteristics of personalized user experiences. A central argument is that e-government is different from e-commerce because e-government has a wider range of requirements for accountability and openness. For example, governments issue basic identity records such as birth certificates, driver's licenses and passports, which are often shared with public and private organizations. In the future, biometrics and other privacy-sensitive information may be added to such records. On the one hand, personalization in e-government could decrease paperwork and increase the efficiency of the administration. On the other hand, however, privacy concerns may limit the acceptance of personalization in e-government: 60% believe that the government possesses too much personal information about individuals (First Amendment Center, 2002). Thus, clear privacy commitments are necessary to increase trust in the government's fair use of personal information to provide a basis for accepted personalization in e-government.

Halstead-Nussloch discusses several approaches to achieve these goals: Legislation – as promoted in the EU – is seen as one solution. Numerous privacy regulations have been enacted in the US as well (cf. section 5.1 in Halstead-Nussloch's chapter). The author recommends to further evaluate regulation with respect to usability, to keep personalization intuitive and natural. Furthermore, opt-in processes should be used to obtain users' explicit consent. Most Internet users prefer this solution: 86% are in favor of "opt-in" privacy policies that require companies to ask people for permission to use their personal information (Fox *et al.*, 2000). Halstead-Nussloch also suggests the use of group identities, which entail fewer privacy risks than individual e-identities. Aggregation methods as suggested in the privacy and security community (e.g. Agrawal, 2002) could be a helpful addendum to this endeavor. Finally, the author argues for awareness-raising about privacy and personalization, and for reducing the mismatch between user concerns and actual privacy impacts (cf. section 3.1 of this chapter).

4.4 Banking and insurance

Hiltunen *et al.*'s chapter on "Personalized Electronic Banking Services: Case Nordea" deals with personalization in Internet banking. Banks dispose of a wealth of customer data. Multi-channel banks have the potential to combine data from numerous sources, including physical branches, call centers and Internet sites. The chapter introduces the business case of the Nordea bank, where a variety of data is collected including demographics, market research data, usage patterns, customization choices, customer behavior in traditional businesses, attitudes, interests and life events, customer programs, and customer feedback. The authors' personalization approach includes virtually all data types in the taxonomy in Table 1. Screen size adaptation, browser optimization, personal greetings and individual product recommendations are just a few examples of how this data can be used for personalization.

Since data collection is so extensive, virtually all privacy concerns apply that were discussed in Section 3.1. Consumers may however harbor additional privacy concerns specifically about e-banking. The market research firm Jupiter found that security and privacy fears prevent more than 40% of online banking customers in

Europe from managing their finances online (Jones *et al.*, 1999). Since most countries do not permit anonymous bank accounts, anonymous access cannot be used to alleviate users' privacy concerns. The trade-off between privacy threats and personalization benefits can therefore only be addressed by laws, user participation and control. From a regulatory point of view, most countries do not have specific provisions for financial service institutions beyond their national privacy laws. In the US, the "Gramm-Leach-Bliley Act" of 1999 regulates how financial institutions may handle personal information. The Act requires financial institutions to give consumers an opportunity to "opt out" before PII is shared with nonaffiliated third parties. Furthermore, a privacy policy must be adopted describing the categories of data collected and recipients of the information.

Communicating the bank's compliance with privacy laws and corporate privacy policies at its website may decrease privacy concerns. Giving users the choice to "opt-in" – as discussed in Halstead-Nussloch's chapter – might be an additional means to lessen privacy concerns. Finally, giving users access to their data and allowing them to specify whether or not they desire personalization services also seems recommendable to increase the acceptance of personalization in e-banking.

The chapter "Personalization and Trust: a Reciprocal Relationship?" by Briggs *et al.*, finally, discusses four fictitious online insurance websites that differed in the factors "personalized vs. non-personalized" and "established vs. new". In the categorization of Table 1, the required input was 'user data'. The authors found no significant differences in the four conditions regarding subjects' willingness to disclose personal data, nor in their trust in the quality of the advice.

Since insurance companies are similar to banks in many ways, the same privacy considerations as spelled out for the chapter of Hiltunen *et al.* seem to apply.

5. FUTURE RESEARCH DIRECTIONS FOR PRIVACY-PRESERVING PERSONALIZATION

Our meta-analysis of consumer surveys demonstrated that users' privacy concerns are major and have a direct impact on personalization systems. Two different directions can be pursued to alleviate these concerns. In one approach, users receive commitments that their personal data will be used for specific purposes only, including personalization. Such commitments can be given in, e.g., individual negotiations or publicly displayed privacy promises ("privacy policies"), or they can be mandated in privacy laws. It is necessary though that these privacy commitments be guaranteed. Ideally, they ought to be enforced through technical means (Agrawal *et al.*, 2002; Karjoth *et al.*, 2003; Fischer-Hübner, 2001), or otherwise through audits and legal recourse. Since individual privacy preferences may considerably vary between users, Kobsa (2003) proposes a meta-architecture for personalized systems that allows them to cater to individual privacy preferences and to the privacy laws that apply to the current usage situation. The personalized system would then exhibit the maximum degree of personalization that is permissible under these constraints.

From the analysis of personalization systems in previous chapters, additional privacy requirements can be inferred. First, users of personalized systems should be given ample control of their data (Karat *et al*., 2003; Harris, 2000; Fox *et al*., 2000). The chapters by Brodie *et al*. and Stolze and Ströbel elaborate on ways to do this. Second, personalization systems should better communicate potential privacy impacts. The AT&T Privacy Bird[5], IE6 and Stolze and Ströbel's approach to explain recommendation algorithms are first steps towards helping users better understand personalization and the potential privacy impact of their interactions. The negotiation and contextualized explanation of privacy and personalization needs to be further researched (Teltzrow and Kobsa, 2004).

The other approach is to allow users to remain anonymous with regard to the personalized system and the whole network infrastructure, whilst enabling the system to still recognize the same user in different sessions so that it can cater to her individually (Kobsa and Schreck, 2003). Karat *et al*. (2003) also address this requirement through different levels of identity. Anonymous interaction seems to be desired by users (however, only a single user poll addressed this question explicitly so far (GVU, 1998)). One can expect that anonymity will encourage users to be more open when interacting with a personalized system, thus facilitating and improving the adaptation to the respective user. The fact that privacy laws do not apply any more when the interaction is anonymous also relieves the providers of personalized systems from restrictions and duties imposed by such laws (they may however choose to observe these laws nevertheless, or to provide other privacy guarantees on top of anonymous access). Finally, anonymous interaction is even legally mandated in some countries if it can be realized with reasonable efforts (TSDPL, 2001).

It is currently unclear which of these two directions should be preferably pursued. Each alternative has several advantages and disadvantages. Neither is a full substitute for the other, and neither is guaranteed to alleviate users' privacy concerns, which ultimately result from a lack of trust. For the time being, both directions need to be pursued.

6. REFERENCES

Agrawal, R., Kiernan, J., Srikant, R., and Xu, Y. (2002) *Hippocratic Databases. 28th International Conference on Very Large Databases*, Hong Kong, China.

Andersen Legal, Internet Privacy Survey (2001) *A Re-Survey of the Privacy Practices of Australia's Most Popular Websites*. Sydney: Andersen Legal, 12 April 2001.

Ardissono, L. and Goy, A. (2000) *Tailoring the interaction with users in web stores*. User Modeling and User-Adapted Interaction **10**(4) 251–303.

Behrens, L. (2001) *Privacy and Security: The Hidden Growth Strategy*. In Gartner G2, 31 May 2001.

Berendt, B., Günther, O., and Spiekermann, S. (to appear). *Privacy in E-Commerce: Stated preferences vs. actual behavior*. Communications of the ACM, 2004.

[5] http://www.privacybird.com

Brusilovsky, P., Kobsa, A. and Vassileva, J. (Eds), 1998, *Adaptive Hypertext and Hypermedia*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Brusilovsky, P. (2001) *Adaptive hypermedia.* User Modeling and User-Adapted Interaction, 11(1–2), 87–110.

Center for Democracy Technology (2001) *Online Banking Privacy: A Slow, Confusing Start to Giving Customers Control Over Their Information*. Washington DC: CDT.

Chin, D.N. (1989) *KNOME: modeling what the user knows in UC.* In Kobsa, A. and Wahlster, W. (Eds.) *User Models in Dialog Systems* Springer Verlag 74–107.

Chittaro, L. and Ranon, R. (2000) *Adding adaptive features to virtual reality interfaces for e-commerce*. In Brusilivsky, P., Stock O. and Strappavara, C. (Eds.) *Adaptive Hypermedia and Adaptive Web-Based Systems*. Springer, 86–91.

Culnan, M J. and Milne, G. R. (2001) *The Culnan-Milne Survey on Consumers & Online Privacy Notices: Summary of Responses*. In Interagency Public Workshop (Ed.) *Get Noticed: Effective Financial Privacy Notices*, Washington, D.C.

Cyber Dialogue (2001) *UCO Software To Address Retailers $6.2 Billion Privacy Problem*, Press Release, http://www.cyberdialogue.com/news/releases/2001/11-07-uco-retail.pdf

Debevc, M., Meyer, B., Donlagic, D., and Svecko, R (1996) *Design and evaluation of an adaptive icon toolbar.* User Modeling and User-Adapted Interaction **6**(1) 1–21.

Deloitte Touche Tohmatsu and Dimension Data (2001) *Survey Reveals Major Corporations are Getting Ready for New Privacy Law*. Canberra: Deloitte, September 17, 2001.

Department for Trade and Industry (2001) Informing Consumers about E-Commerce conducted by MORI, London: DTI, http://www.mori.com/polls/2001/pdf/dti-e-commerce.pdf

EU (2002) *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*. http://europa.eu.int/eur-lex/pri/en/oj/dat/2002/l_201/l_20120020731en00370047.pdf

Fink, J., Kobsa, A. and Nill, A. (1998) *Adaptable and adaptive information provision for all users, including disabled and elderly people.* The New Review of Hypermedia and Multimedia **4** 163–188.

First Amendment Center (2001) *Freedom of Information in the Digital Age*. Project of the ASNE Freedom of Information Committee and The First Amendment Center. http://www.freedomforum.org /publications/first/foi/foiinthedigitalage.pdf

Fischer-Hübner, S. (2001) *IT-Security and Privacy: Design and Use of Privacy-Enhancing Security Mechanisms*. LNCS 1958, Heidelberg, Germany: Springer.

Forrester Research (2001) *Privacy Issues Inhibit Online Spending* (survey summary). Cambridge, MA. http://www.nua.com/surveys/?f=VS&art_id=905357259&rel=true

Foster, C. (2000) *The Personalization Chain*. In Jupiter Communications, *Site Operations*, Vol. 3, 2000.

Fox, S., Rainie, L. (2000) *Trust and Privacy Online: Why Americans Want to Rewrite the Rules.* Pew Internet & American Life Project, Washington DC. http://www.pewinternet.org/reports/toc.asp?Report=19

Gallup Organization (2001) *Majority of E-mail Users Express Concern about Internet Privacy*. Washington DC: The Gallup Organization, June 28, 2001. http://www.gallup.com/subscription/?m=f&c_id=10732

GVU (1998) *GVU's 10th WWW User Survey*. Graphics, Visualization and Usability Lab, Georgia Tech.

Harper, J., Singleton, S. (2001) *With a grain of salt, What Consumer Privacy Survey don't tell us*. Competitive Enterprise Institute. http://www.cei.org/PDFs/with_a_grain_of_salt.pdf

Harris Interactive (2000) *A Survey of Consumer Privacy Attitudes and Behaviors*. Rochester, NY.

Harris Interactive (2001) *Privacy Notices Research, Final Results.* Privacy Leadership Initiative, Rochester, NY.

Harris Interactive (2003) *Most People Are Privacy Pragmatists.* Rochester NY. March 19, 2003.

Hine, C, Eve, J. (1998) *Privacy in the Marketplace*. Information Society 14 (4), 253-262.

Holynski, M. (1988) *User-adaptive computer graphics.* International Journal of Man-Machine Studies **29,** 539–548.

Höök, K., Karlgren, J., Waern, A., Dahlbäck, N., Jansson, C., Karlgren, K. and Lemaire, B. (1996) *A glass box approach to adaptive hypermedia*. User Modeling and User-Adapted Interaction **6**(2–3), 157–184.

Interactive Policy Making (2002) *Views on Data Protection, Questionnaire on the Implementation of the Data Protection Directive (95/46/EC).* Results of Online Consultation 20 June - 15 September 2002, Brussels.

Ipsos-Reid and Columbus Group, 2001, Privacy Policies Critical to Online Consumer Trust, Canadian Inter@ctive Reid Report.

Jameson, A., Schäfer, R., Simons, J. and Weis, T. (1995) *Adaptive provision of evaluation-oriented information: tasks and techniques.* Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, Montreal, Canada, Morgan Kaufmann, 1886–1893.

Joachims, T., Freitag, D., and Mitchell, T., (1997) *Webwatcher: a tour guide for the World Wide Web.* In Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence. Nagoya, Japan, Morgan Kaufmann Publishers, 770-777.

Joerding, T. (1999) *A temporary user modeling approach for adaptive shopping on the web*. Proceedings of the 2nd Workshop on Adaptive Systems and User Modeling on the WWW, WWW-8, Toronto, Canada and UM99, Banff, Canada.

Jones, N., Neufeld, E., Waagstein, L., Stemmer, A. (1999) *Online Financial Services - Integrated Service Is Key To Online Money Management.* Jupiter Communications, Strategic Planning Services, (10), 1-19.

Kaplan, C., Fenwick, J., and Chen, J. (1993) *Adaptive hypertext navigation based on user goals and context.* User Modeling and User-Adapted Interaction **3**(3) 193–220.

Karat, C., Brodie, C., Karat, J., Vergo, J., and Alpert, S. (2003) *Personalizing the User Experience on ibm.com*. In Vredenburg, K. (Ed.), *IBM Systems Journal*, 42, 2, 686-701.

Karjoth, G., Schunter, M., and Waidner, M. (2003) *Platform for Enterprise Privacy Practices: Privacy-enabled Management of Customer Data*. In *2nd Workshop on Privacy Enhancing Technologies*, LNCS, Volume 2482 / 2003, Berlin: Springer-Verlag, 69 – 84.

Kobsa, A., Müller, D., and Nill, A. (1994) *KN-AHS: an adaptive hypertext client of the user modeling system BGP-MS*. In *Proceedings of the Fourth International Conference on User Modeling*, Cape Cod, MA, 99–105. Reprinted in M. Marbury and W. Wahlster (eds), 1998, *Intelligent User Interfaces* Morgan Kaufman, 372–378.

Kobsa, A. and Schreck, J. (2003) *Privacy through Pseudonymity in User-Adaptive Systems*. ACM Transactions on Internet Technology 3 (2), 149-183

Kobsa, A. (2001) *Generic User Modeling Systems*. User Modeling and User-Adapted Interaction, *11*. 49-63.

Kobsa, A. (2003) *A Component Architecture for Dynamically Managing Privacy Constraints in Personalized Web-Based Systems*. In Dingledine, R. (Ed.) Privacy Enhancing Technologies: Third International Workshop, PET 2003, Dresden, Germany, Springer-Verlag, LNCS 2760, 177-188.

Kobsa, A., Koenemann, J., Pohl, W. (2001) *Personalised hypermedia presentation techniques for improving online customer relationships*. *The Knowledge Engineering Review*, Vol. 16(2), 111–155. Cambridge University Press

Kobsa, A. (2001) *Tailoring Privacy to Users' Needs*. Invited Keynote, *8th International Conference on User Modeling*, Sonthofen, Germany, Springer Verlag, 303-313.

Konstan, J.A., Miller, B.N., Maltz, D., Herlocker, J.L., Gordon, L.R. and Riedl, J. (1997) *GroupLens: applying collaborative filtering to Usenet news*. Communications of the ACM **40**(3) 77–87.

Kozierok, R. and Maes, P. (1993) *A learning interface agent for scheduling meetings*. In Gray, W.D., Hefley, W.E. and Murray, D. (Eds.) *Proceedings of the International Workshop on Intelligent User Interfaces.* ACM Press, 81–88.

Krogsæter, M., Oppermann, R. and Thomas, C.G. (1994) *A user interface integrating adaptability and adaptivity*. In R Oppermann (Ed.) *Adaptive User Support: Ergonomic Design of Manually and Automatically Adaptable Software.* Lawrence Erlbaum, 97-125.

Küpper, D. and Kobsa, A. (1999) *User-tailored plan generation.* In Kay, J. (Ed.) *UM99 User Modeling: Proceedings of the Seventh International Conference,* Banff, Canada, Springer-Verlag 45–54.

Lesh, N., Rich, C. and Sidner, C.L. (1999) *Using plan recognition in human-computer collaboration.* In Kay, J. (Ed.) *UM99 User Modeling: Proceedings of the Seventh International Conference,* Banff, Canada, Springer-Verlag, 23–32.

Lieberman, H. (1995) *Letizia: An agent that assists web browsing*. In *Proceedings of the International Joint Conference on Artificial Intelligence*, Montreal, Canada Morgan Kaufmann, 924-929.

Mabley, K. (2000) *Privacy vs. personalization, part three: a delicate balance.* New York, NY: Cyber Dialogue Inc. http://www.cyberdialogue.com/library/pdfs/wp-cd-2000-privacy.pdf

Maes, P. (1994) *Agents that reduce work and information overload.* Communications of the ACM **37**(7) 31–40.

McAteer, S., Graves, L., Gluck, M., May, M., Allard, K. (1999) *Proactive personalization – learning to swim, not drown in consumer data*. Jupiter Study, New York City, 4-12.

Mitchell, T., Caruana, R., Freitag, D., McDermott, J. and Zabowski, D. (1994) *Experience with a learning personal assistant*. Communications of the ACM **37**(7) 81-91.

Pavlou, P. A. (2003) *Consumer Acceptance of Electronic Commerce – Integrating Trust and Risk with the Technology Acceptance Model*. International Journal of Electronic Commerce **7**(3) 69-103.

Pazzani, M. and Billsus, D. (1997) *Learning and revising user profiles: the identification of interesting web sites*. Machine Learning **27** 313–331.

Personalization Consortium (2000) *Personalization & Privacy Survey*. Edgewater Place, MA. http://www.personalization.org/SurveyResults.pdf

Popp, H. and Lödel, D. (1996) *Fuzzy techniques and user modeling in sales assistants*. User Modeling and User-Adapted Interaction **5**(3–4) 349–370.

Privacy & American Business (1999) *Personalized Marketing and Privacy on The Net: What Consumers Want*, November 1999. http://www.pandab.org/doubleclicksummary.html

Resnick, P. and Varian, H.R. (1997) *Recommender systems.* Communications of the ACM, **40**(3) 56–58.

Responsys.com (2000) *Online Marketers Have Little Confidence in Self-Regulation of Internet Privacy*. Sponsored by Millard Brown IntelliQuest. Palo Alto, CA, September 26, 2000.

Roy Morgan Research (2001) *Privacy and the Community.* Prepared for the Office of the Federal Privacy Commissioner, Sydney, July 31, 2001. http://privacy.gov.au/publications/rcommunity.html

Sakagami, H., Kamba, T., Sugiura, A. and Koseki, Y. (1998) *Effective personalization of push-type systems: visualizing information freshness.* In *Proceedings of the 7th World Wide Web Conference, Brisbane, Australia.*

Shardanand, U. and Maes, P. (1995) *Social information filtering: algorithms for automating word of mouth.* In *Proceedings of the Human Factors in Computing Systems Conference (CHI-95), Denver, CO.* ACM Press 210–217.

Specht, M. (1998) *Empirical evaluation of adaptive annotation in hypermedia*. In *Proceedings of the ED-MEDIA98*, Freiburg, Germany, p.1327–1332.

Sweeney, L. (2001) *Computational Disclosure Control: A Primer on Data Privacy Protection*, Ph.D. Thesis, MIT, Cambridge, MA.

Teltzrow, M. and Kobsa, A. (2004) *Communication of Privacy and Personalization in E-Business*. In: Proceedings of the Workshop WHOLES: A Multiple View of Individual Privacy in a Networked World, Stockholm, Sweden.

TSDPL (2001) *Teleservices Data Protection Law (Article 3 of the Law on the Legal Requirements for Electronic Business Dealings of 14 Dec. 2001)*. German Federal Law Gazette 1, 3721.

UMR (2001) *Privacy Concerns Loom Large*. Conducted for the Privacy Commissioner of New Zealand. Survey summary, Auckland: PC of New Zealand. http://www.privacy.org.nz/privword/42pr.html

Wham, T. (2001) *Workshop on the Information Marketplace: Merging and Exchanging Consumer Data.* Interview Transcript, Federal Trade Commission. http://www.ftc.gov/bcp/workshops/infomktplace/transcript.htm.