

PERSONALIZED HYPERMEDIA AND INTERNATIONAL PRIVACY [ALFRED KOBSA]

Personalized hypermedia systems may be in conflict with privacy concerns of computer users, and with privacy laws that are in effect in many countries.

U ser-adaptive (or “personalized”) hypermedia systems are able to cater to users more effectively given the more information they possess about them. The adaptations that will be necessary are usually not known at the time when different pieces of information about users become available. Personalized hypermedia systems therefore have to lay this collected data “in stock” and maintain it for possible future usage. Moreover, data gathering is moreover mostly performed in an unobtrusive manner and often without users’ awareness. This is done to avoid distracting users from their tasks and with consideration that users are very reluctant to perform actions not directed toward their immediate goals (like providing data about themselves) if they do not receive immediate benefits, even when they would benefit in the long run [3].

This current practice of data collection and processing in personalized hypermedia systems, specifically in personalized Web sites [7], seems to be in conflict with privacy concerns of Internet users. These became manifest in numerous recent consumer polls,¹ where respondents reported being concerned about threats to their privacy when using the Internet; being concerned about divulging personal information online; and being concerned about being tracked online. Web users are not only concerned but already counteract. They reported leaving Web sites that required registration information; having entered false registration information; and having refrained from shopping online due to pri-

vacancy concerns, or having bought less. Internet users who are concerned about privacy are thereby not naive isolationists, but have very pragmatic demands. They want Internet sites to ask for permission to use personal data and are willing to give out personal data for getting something valuable in return.

The current practice of personal data processing in personalized hypermedia also appears to be in conflict with privacy laws and guidelines that usually call for parsimony, purpose-specificity, and user awareness or even user consent in the collection and processing of personal data. Privacy laws protect the data of identified or *identifiable*² individ-

¹See www.privacyexchange.org/iss/surveys/surveys.html for a very comprehensive overview of the more than 100 surveys that have been conducted to date.

²It is thus not required that the system actually identifies the user. Privacy laws already apply when it is possible to identify the user with reasonable efforts based on the data that the system collects. Static IP numbers are generally regarded as identifying characteristics of Web users.

uals in more than 30 countries, and a few states and provinces. Several countries, including the U.S., have some sector-specific regulations. In another 10–15 countries, privacy regulations are currently at various stages in the legislative process.

While at a surface level privacy laws are often widely different, they are usually grounded in a comparatively small number of basic privacy principles. As an example, the sidebar “OECD Basic Privacy Principles of National Application” lists recommended privacy guidelines adopted by the Organization for Economic Co-operation and Development (OECD) in 1980 (these are not binding for its current 30 member countries, which include the U.S.). The privacy laws of many countries also restrict the transborder flow of personal data, or even extend their coverage beyond the national boundaries. Such laws then also affect personalized hypermedia sites abroad that serve users in these regulated countries, even if there is no privacy law in place in the country where the site is located (see the sidebar for a brief overview).³ For comprehensive privacy resources, see for example, www.epic.org, www.privacy.org, www.privacyinternational.org, www.privacyexchange.org, [1], [8], and [11].

If privacy laws apply to a personalized hypermedia application, they often not only affect the conditions under which personal data may be collected and the rights that data subjects have with respect to their data, but also the methods that may be used for processing the data. Following is a sample of restrictions from the German Teleservices Data Protection Act and the European Data Protection Directive that substantially affect the internal operation of personalized hypermedia applications:

- Usage logs must be deleted after each session. This provision affects, for example, the machine-learning methods that can be employed in a personalized hypermedia system. If learning takes place over several sessions, only incremental methods can be employed since the raw usage data from previous sessions has all been discarded.
- Usage logs of different services may not be combined, except for accounting purposes. This is a severe restriction for so-called central user modeling servers that collect user data from, and make it

³When users in country A interact with a Web site that is located in country B, data collection is generally deemed to take place in country A, and a transfer of data takes place from A to B.

available to, different user-adaptive applications [5].

- User profiles are permissible only if pseudonyms are used. Profiles retrievable under pseudonyms shall not be combined with data relating to the bearer of the pseudonym. This clause mandates a Chinese wall between the component that receives data from identifiable users, and the user-modeling component that makes generalizations about pseudonymous users and adapts hypermedia pages accordingly. Communication between these components may only take place through a trusted third component that manages the directory of pseudonyms, or through more complex pseudonymization procedures.
- No fully automated individual decisions are allowed that produce legal effects concerning the data subject or significantly affect him and that are based solely on automated processing of data intended to evaluate certain personal aspects relating to him, such as his performance at work, creditworthiness, reliability, conduct, and so forth. This prohibition has impacts on learner-adaptive hypermedia systems for tutoring [2]. If such systems assign formal grades, there has to be a human somewhere in the loop.
- Technical and organizational security measures must be implemented to protect personal data against accidental or unlawful destruction or accidental loss, alteration, unauthorized disclosure, or access.
- Anonymous or pseudonymous access and payment must be offered possible and reasonable.

Impacts on Personalized Hypermedia Systems

If users can interact with a personalized system in an anonymous or pseudonymous manner, privacy laws generally do not apply and users’ privacy concerns seem to abate. Internet users that remain anonymous are more likely to provide data about themselves, which will improve the quality of the personalization. To facilitate the development of such systems, [12] defines a reference model that preserves the pseudonymity of both users and user modeling servers while at the same time permitting full personalization.

If anonymous or pseudonymous access is not possible, users’ privacy preferences need to be taken into account as well as national privacy laws that possibly apply. The following are suggested guidelines for the design of personalized hypermedia systems in an

OECD Basic Privacy Principles of National Application

Collection Limitation Principle: There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject.

Data Quality Principle: Personal data should be relevant to the purposes for which it is to be used, and, to the extent necessary for those purposes, should be accurate, complete and kept up-to-date.

Purpose Specification Principle: The purposes for which personal data is collected should be specified not later than at the time of data collection and the subsequent use limited to the fulfillment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose.

Use Limitation Principle: Personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with [the Purpose Specification Principle] except: a) with the consent of the data subject; or b) by the authority of law.

Security Safeguards Principle: Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorized access, destruction, use, modification, or disclosure of data.

Openness Principle: There should be a general policy of openness about developments, practices, and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purposes of its use, as well as the identity and usual residence of the data controller.

Individual Participation Principle: An individual should have the right: a) to obtain from a data controller, or otherwise, confirmation of whether or not the data controller has data relating to him; b) to have communicated to him, data relating to him within a reasonable time; at a charge, if any, that is not excessive; in a reasonable manner; and in a form that is readily intelligible to him; c) to be given reasons if a request made under subparagraphs (a) and (b) is denied, and to be able to challenge such denial; and d) to challenge data relating to him and, if the challenge is successful, to have the data erased, rectified, completed, or amended.

Accountability Principle: A data controller should be accountable for complying with measures that give effect to the principles stated here. **C**

identifiable usage context. The guidelines constitute keystones in most privacy laws and probably also meet the current privacy preferences of many users:

- Make (long-term) personalization a clear purpose of your service. This may improve the accuracy of the data needed for personalization purposes, and will also facilitate the next guideline.
- Provide comprehensive and intelligible advance notice to users about all data that is to be collected, processed and transferred, and indicate the purposes for which this is being done. This is likely to increase users' trust in the application and is mandated by virtually all privacy laws.
- Obtain users' informed and voluntary consent to processing their data for personalization purposes.
- Provide organizational and technical means to allow users to inspect, block, rectify and erase both the data they provided, and specifically the assumptions the system inferred about them.
- Provide security mechanisms that are commensurate with the technical state of the art and the sensitivity of the stored user data.

These guidelines can be complemented or modified by users' individual privacy preferences (as for example, expressed in the APPEL language [4] in combination with the PGP protocol [8]), yielding a privacy policy adapted to each individual user. When personalized hypermedia systems are aimed at users from foreign countries, care must be taken whether privacy laws are in place that affect the trans-border flow or the extraterritorial processing of personal data gathered in these countries. If so, the regulations that must be observed are in most cases different between these countries, and different from those of the country where the personalized system is located. It is therefore proposed in [6] that a flexible architecture for personalized systems should tailor privacy dynamically to each individual user, taking his or her preferences and the privacy laws at both the system's and the user's location into account.

Future Directions

New architectures for personalized systems (which are currently too resource-intensive to build) would considerably reduce the amount of personal data that becomes available to others, and would thus mitigate the impact of privacy laws and privacy concerns. Client-side instead of server-side personalization would give users exclusive control of all purposely-collected personal data as well as all processes that operate on the data. Mobile personalized hypermedia services that are

executed on the client side only would additionally leave all information about a user's interaction with these services exclusively in the user's domain.

On the legal front, an international treaty on information privacy [10] is expedient that would allow personalized systems to follow a single privacy standard instead of catering to different requirements of different countries. Such an agreement could build on a revised version of the OECD guidelines mentioned previously, or other privacy principles that are currently being discussed. **C**

An electronic version of this article that includes additional references to all discussed privacy surveys and privacy laws is available at www.ics.uci.edu/~kobsa/papers/2002-kobsa-CACM.pdf. A collection of international privacy laws with a comparison of the stipulations that specifically affect personalized systems can be found at www.ics.uci.edu/~kobsa/privacy.

REFERENCES

1. Agre, P. and Rotenberg, M., Eds. *Technology and Privacy: The New Landscape*. MIT Press, Cambridge, MA, 1997.
2. Brusilovsky, P. Adaptive and intelligent technologies for Web-based education. *Künstliche Intelligenz* (Apr. 2000), 19–25; www2.sis.pitt.edu/~peterb/papers/KI-review.html.
3. Carrol, J. and Rosson, M.B. The paradox of the active user. In J. Carrol, J., Ed., *Interfacing Thought: Cognitive Aspects of Human-Computer Interaction*. MIT Press, Cambridge, MA, 1989.
4. Cranor, L., Langheinrich, M., and Marchiori, M. A P3P Preferences Exchange Language 1.0 (APPEL 1.0), 2001; www.w3.org/TR/P3P-preferences.
5. Kobsa, A. Generic user modeling systems. *User Modeling and User-Adapted Interaction* 11, 1–2 (2001), 49–63; www.ics.uci.edu/~kobsa/papers/2001-UMUAI-kobsa.pdf.
6. Kobsa, A. Tailoring privacy to users' needs. In M. Bauer, P.J. Gmytrasiewicz, and J. Vassileva, Eds., *User Modeling 2001: 8th International Conference*, Springer Verlag, Berlin-Heidelberg, 2001, 303–313; www.ics.uci.edu/~kobsa/papers/2001-UM01-kobsa.pdf.
7. Kobsa, A., Koenemann, J. and Pohl, W. Personalized hypermedia presentation techniques for improving customer relationships. *The Knowledge Engineering Review* 16, 2 (2001), 111–155; www.ics.uci.edu/~kobsa/papers/2001-KER-kobsa.pdf.
8. Peters, T.A. *Computerized Monitoring and Online Privacy*. McFarland, Jefferson, NC, 1999.
9. Reagle, J. and Cranor, L. The platform for privacy preferences. *Commun. ACM* 42, 2 (Feb. 1999), 48–55.
10. Reidenberg, J.R. Resolving conflicting international data privacy rules in cyberspace. *Stanford Law Review* 52 (2000), 1315–1376; reidenberg.home.sprynet.com/international_rules.pdf.
11. Rotenberg, M. *The Privacy Law Sourcebook 2001: United States Law, International Law, and Recent Developments*. EPIC, Washington, DC, 2001.
12. Schreck, J. *Security and Privacy in User Modeling*. Kluwer, Dordrecht, Netherlands; www.security-and-privacy-in-user-modeling.info.

ALFRED KOBSA (kobsa@uci.edu) is a professor at the University of California, Irvine, CA.

© 2002 ACM 0002-0782/02/0500 \$5.00

Some Countries With Restrictions on the Transborder Flow of Personal Data

Argentina: With some exceptions, the transfer of any type of personal information is prohibited to countries or international or supranational entities that do not provide adequate levels of protection.

Australia: Australian privacy law extends to overseas sites if the user involved is an Australian citizen or permanent resident, and ... the site conducts business with users in Australia and collects the personal data in Australia. Conflicting foreign law overrides Australian law.

European Union: The European Data Protection Directive requires European Union member states to prohibit in their national privacy laws the transfer of personal data to non-EU countries that do not ensure an adequate level of protection, as determined by the European Commission.* Member states may nevertheless permit a transfer with the consent of the user, for entering or performing a contract, on public interest grounds, and for some additional reasons.

Hong Kong: With some exceptions and additional restrictions, the transborder transfer of personal data collected in Hong Kong is only permitted if the user

has been notified about the destination country and consented in writing, and if the data processor has reason to believe that comparable privacy legislation exists in the destination country and that processing there will not be in violation of the Hong Kong law.

Hungary: Personal data may only be transferred abroad with the consent of the data subject or if permitted by law, provided that the processing abroad is in compliance with the Hungarian privacy law.

Lithuania: The transfer of personal data abroad may only take place with authorization through the government, with user consent, to conclude or fulfill a contract, on the ground of state interest, for court hearings, to protect the data subject's life, and under international agreements.

New Zealand: Certain provisions of the New Zealand Privacy Act also apply to personal data that is processed abroad or was transferred out of New Zealand. Conflicting foreign law overrides New Zealand law.

Taiwan: Government registration is required for data collection. Government authorities may restrict international transmission if, for example, the receiving country lacks adequate data protection laws and/or ordinances and where the rights and interests of concerned parties seem injured. **C**

*U.S. organizations that subject themselves to the so-called Safe Harbor Principles that were negotiated between the European Union and the U.S. meet this adequacy standard.