Convincing Users to Disclose Personal Data

Alfred Kobsa
Department of Informatics
University of California, Irvine
+1 949 485-5020
kobsa@uci.edu

Max Teltzrow
International Computer Science Institute
Berkeley, CA
+1 510 666-2881
teltzrow@wiwi.hu-berlin.de

ABSTRACT[*]
To provide users with tailored services, personalized systems need considerable amounts of personal data. Users are however often reluctant to divulge those data due to privacy concerns. We conducted an experiment with a traditionally designed web store and one in which the privacy practices of the website and users' benefits from providing their personal data are clearly explained in a contextualized manner. We found that subjects in the second condition were significantly more willing to share personal data with the website, rated the perceived benefit resulting from data disclosure significantly higher, and also made considerably more purchases.

Keywords
Privacy disclosure, personalization, user benefit, trust, recommendation, perceived quality, adoption, purchases

A DESIGN PATTERN FOR WEBSITES THAT COLLECT PERSONAL DATA
To adequately address privacy concerns of users of personalized websites, we investigate web design patterns that communicate a site's privacy practices both at a global and a local (contextualized) level. Design patterns constitute descriptions of best practices within a given design domain based on research and application experience [8]. They give designers guidelines for the efficient and effective design of user interfaces.

Fig. 1 shows the application of the proposed interface design pattern to a web bookstore that offers personalized services. The top three links in the left-hand frame lead to the global disclosures: privacy, personalization benefits, and security. The main frame contains input fields and checkboxes for entering personal data. Each of them is accompanied by a contextualized explanation of the site's privacy practices regarding the respective personal data

(which focuses specifically on usage purposes), and the personalized services that these data afford.

A COMPARATIVE EXPERIMENT
We developed a mock book recommendation and sales website whose interface suggests a future version of a well-known online bookstore. The web forms ask a broad range of questions relating to users' interests (32 questions with 66 meaningful answer options). A few sensitive questions on political interests, religious interests and adherence, literary sexual preferences, and interest in certain medical subareas (including venereal diseases) are also present.

Two variants of this system were created, one without contextual explanations of privacy practices and personalization benefits ("-expl"), and one with such explanations present ("+expl"). Figure 1 shows an excerpt of the second variant, translated from German into English. A between-subjects experiment was carried out to verify the hypothesis that condition +expl would prompt users to answer more questions. Subjects were told that they would test an experimental new version of the online bookstore with an intelligent book recommendation engine inside. They were instructed that they could buy one of the recommended books at the end of the experiment at a high discount, and advised that the more and the better data they provided, the better would be the personalized book selection. They were also warned that their data would be given to the book retailer after the experiment, but that they were not required to answer any question.

After nine pages of data entry, fifty predetermined and invariant books were displayed that had been selected based on their low price and their presumable attractiveness for students (book topics include popular fiction, politics, tourism, and sex and health advisories). The prices of all books are visibly marked down by 70%. Users were free to choose whether or not to buy one single book. Those who did were asked for their address and payment data, whose veracity was verified after the experiment.

RESULTS AND THEIR INTERPRETATION
Subjects in condition +expl answered 8% more questions ($p<.001$), gave 20% more answers ($p<.001$), and decided 33% more often to buy a book ($p<0.07$). They also showed a 19% higher agreement to the statement that "The data allowed the store to select better books for me" ($p<.035$).
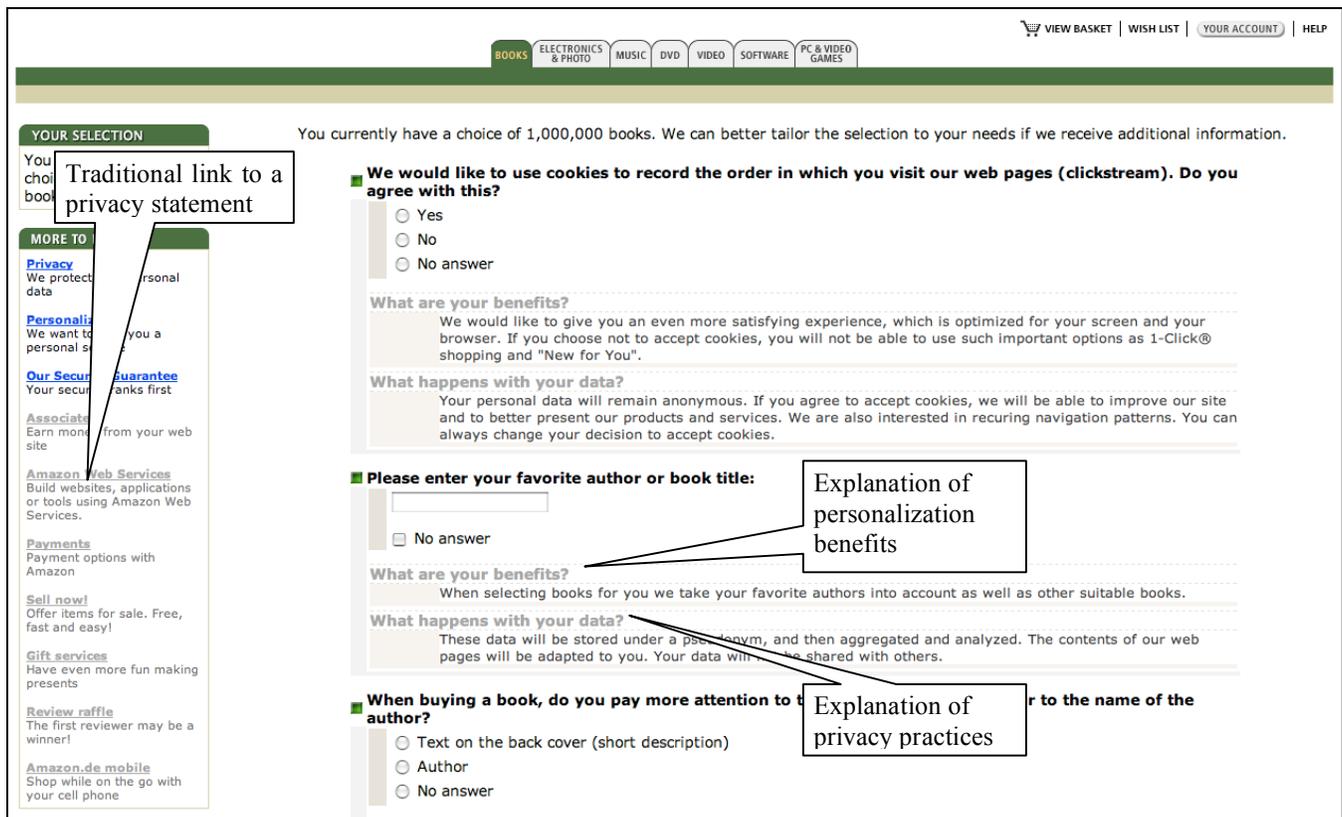
Figure 1: Global and contextual communication of privacy practices and personalization benefits

These results can be best explained by assuming that the treatment affected factors that are known to increase users' trust in web-based systems. In condition "+expl", users better understanding of the website's privacy practices and of the contribution of disclosed data to resulting personalization benefits is likely to have increased users' trust and alleviated their privacy concerns. This in turn lead to more disclosure of personal data (for related survey-based results, see e.g. [1, 4, 6, 9]).

The decision to buy a book was a significant step in our experiment since at this point users revealed personally identifiable information and accepted that earlier pseudonymous information can be linked to their identities.

It seems that the increased trust of users in condition "+expl" due contextualized privacy disclosure may have contributed to more users opting to buy this book.

Other factors besides the one that was systematically varied in this experiment may also play a role in users' data disclosure behavior but were kept constant. These include the reputation of the website, the stringency of the website's data handling practices, the permanent visibility of contextual explanations, and the presence of a reference to the website's full privacy policy that may be required for legal reasons.

As far as reputation is concerned, we chose a web store that enjoys a relatively high reputation in Germany (we conducted a pre-test that confirmed this). Survey based studies (e.g. [2, 3, 7]) have shown that higher reputation increases users' willingness to share personal data with a website. In a more recent version of the experiment we therefore changed the name and logo of the website to ones that had received a medium reputation rating in the pre-test. We found similar effects of condition "+expl" for the medium-reputation website also, but with lower figures for data disclosure and purchases in both conditions. There was no interaction between reputation and form of disclosure.

REFERENCES

1. CG&I-R. Privacy Policies Critical to Online Consumer Trust, Columbus Group and Ipsos-Reid, 2001, www.ipsos-na.com/news/pressrelease.cfm?id=1171.

2. Chang, M.K. and Cheung, W., Online Trust Production: Interactions among Trust Building Mechanisms. In 38th Hawaii International Conference on System Sciences, (Kona, HI, 2005), 181c, DOI 10.1109/HICSS.2005.471.

3. Earp, J.B. and Baumer, D. Innovative Web Use to Learn About Consumer Behavior and Online Privacy. Communications of the ACM Archive, 46 (4), 2003, 81-83, DOI 10.1145/641205.641209.

4. Hine, C. and Eve, J. Privacy in the Marketplace. The Information Society, 14 (4), 1998, 253-262, DOI 10.1080/019722498128700.

5. Kobsa, A. and Teltzrow, M. Contextualized Communication of Privacy Practices and Personalization

Benefits: Impacts on Users' Data Sharing Behavior. In Martin, D. and Serjantov, A. eds. Privacy Enhancing Technologies, Springer Verlag, Heidelberg, Germany, 2005, 329-343, DOI 10.1007/11423409_21.

6. Malhotra, N.K., Kim, S.S. and Agarwal, J. Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Nomological Framework. Information Systems Research, 15 (4), 2004, 336–355, DOI 10.1287/isre.1040.0032.

7. Teo, H.H., Wan, W. and Li, L., Volunteering Personal Information on the Internet: Effects of Reputation, Privacy Initiatives, and Reward on Online Consumer Behavior. In 37th Hawaii International Conference on System Sciences, (Big Island, HI, 2004), 181-190, http://csdl.computer.org/comp/proceedings/hicss/2004/2056/07/205670181c.pdf.

8. van Duyne, D.K., Landay, J.A. and Hong, J.I. The Design of Sites: Patterns, Principles, and Processes for Crafting a Customer-Centered Web Experience. Addison-Wesley, Boston, 2002.

9. Xu, Y., Tan, B.C.Y., Hui, K.-L. and Tang, W.-K., Consumer Trust and Online Information Privacy. In International Conference on Information Systems 2003, (Seattle, WA, 2003), 538-548, http://aisel.isworld.org/Publications/ICIS/2003/03CRP45.pdf