On the Relative Importance of Privacy Guidelines for Ambient Health Care

Evelien van de Garde–Perik, Panos Markopoulos
Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
{e.m.v.d.garde; p.markopoulos}@tue.nl

Boris de Ruyter
Philips Research, Media Interaction Group, High Tech Campus 34, 5656 AE Eindhoven, The Netherlands
boris.de.ruyter@philips.com

ABSTRACT
We present an empirical study regarding the relative importance of complying with privacy related guidelines in the context of a Health Monitoring System. Participants were confronted with text scenarios describing privacy related aspects of a health monitoring service for daily use at home. Participants assessed the relative importance to them of simplified variants of the OECD (Organization of Economic Cooperation and Development) guidelines for the protection of personal data. The guidelines that relate to Insight and Openness were most valued. The guidelines relating to Modification and Data Quality were valued least by most participants in this context. Methodological challenges were encountered on the way, which reveal the complexity of conducting empirical investigations of privacy aspects of human-computer interaction.

Author Keywords
Fair Information Practices, privacy guidelines, ambient intelligence.

ACM Classification Keywords
H5.m. Information interfaces and presentation: Miscellaneous.

INTRODUCTION
Indicating privacy as a major concern appears to be inevitable in most discussions regarding Ambient Intelligence and its acceptance by the wider public. Casual discussions often bring up 'Big Brother' referring to Orwellian visions of state-control over individuals. In such informal contexts but also in empirical surveys, people will often declare emphatically their resentment to being tracked, monitored or recorded demonstrating a clear dissent to a technological landscape that seems to be approaching with obvious threats to personal freedom.

Privacy researchers have for some time realized this issue, sometimes proposing systematic analyses of privacy risks [2], structured methods to guide the design of context aware and adaptive systems with respect to personal privacy [3]. One commonly traveled avenue for addressing privacy concerns is to rely on Fair Information Practices [as suggested in 1] or other legal guidelines prescribing how to deal with collection, storage and use of personal data. Despite the fact that many scholars refer to the solution that Fair Information Practices may offer in minimizing privacy concerns, hardly any research to date provides evidence regarding their relevance and importance for users of systems and services they apply to.

There are many variations of Fair Information Practices; each country seems to have its own rules and regulations. However, due to the high similarity between the various principles, we focus on only one set of guidelines, namely those by the Organization of Economic Cooperation and Development (OECD). The OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data [6], were adopted on 23 September 1980, represent an international consensus on general guidance concerning the collection and management of personal information. The guidelines, encapsulated in eight core principles, are used by governments, business and consumer representatives in their efforts to protect privacy and personal data, and in preventing unnecessary restrictions to data flows across borders, both on and off line. This set of guidelines has provided the basis of the extensive and influential treatment of privacy in ubiquitous computing by M. Langheinrich [5]. His work, while providing a useful analysis and set of concepts does not provide any empirical evidence to demonstrate that these principles do indeed lead to higher acceptance of such systems by end-users. There is a need to investigate whether the incorporated functionality of systems to guarantee privacy is usable and understandable to the people who interact with it. Otherwise it is as if the functionality does not exist [4].

In trying to provide empirically based advice to the designers of Ambient Intelligence systems we study the relevant importance of these guidelines for end-users. This study is the first stage where participants are asked to

NordiCHI 2006: Changing Roles, 14-18 October 2006, Oslo, Norway

explicitly rate the importance of guidelines based on system descriptions. In a later stage these findings will be verified by having participants rate the importance of the guidelines implicitly. This investigation is conducted in the context of a health support system that allows the daily monitoring of health parameters of individuals from the comfort of their homes.

SET UP OF THE STUDY

Design of the study

Participants were shown a text scenario describing the health care system (see figure 1). The description provided general context information about the system and then detailed its privacy related features; the described system does not adhere to any of the OECD guidelines. Participants were then presented with potential 'fixes' to the system, each of which would make it comply with one specific OECD guideline. These were presented in pairs, and participants were asked to choose which of the pair they thought was most important to them.

Participants

Participants were recruited by placing adverts on local mailing lists in the respective organizations of the authors (a University and an Industrial Research lab). Recruitment was aimed at obtaining two groups of individuals depending on their need for medical attention. The first group consisted of people with a chronic health condition and people aged over 65. The second group consisted of individuals with no specific need for medical attention. A total of 50 persons participated.

Materials

The original OECD guidelines are written in quite terse and lengthy language intended for legal purposes. For this reason simplified expressions relevant to the scenario of our study were created. These are shown in table 1. Every row of table 1 represents one guideline. It provides a title, shorthand and a single-sentence description that captures its essence. The list we developed in this way does not include the principle of accountability which is relevant in a legal context (it makes explicit the responsibility of organizations handling personal data to adhere to the other guidelines). Further the guideline pertaining to personal involvement which is a conjunction of several clauses was broken up to two constituents, Insight (IN) and Modification (MO).

> The system does not inform John that it will collect data regarding his health, blood pressure, pulse and glucose level. The system does not inform John that data is collected in order to monitor his diabetes condition. The system informs John that it also collects data that is useful for things other than tracking his diabetes condition. The system informs John that it uses his data for other reasons than the main purpose of the system as well. The system does not inform John regarding all the organizations or individuals who can access his data. The system informs John that his data is not protected by any security safeguards. The system does not provide facilities for John to inspect all data collected about him. The system does not provide facilities to allow John to modify or erase any data about him.

Figure 1. Extract of the scenario describing the deviant behavior of the health monitoring system.

CL: Collection Limitation	The user is informed about the type of data that will be collected.
PS: Purpose Specification	The user is informed about the main purpose for which the data will be used.
DQ: Data Quality	The system only collects data that is relevant to the main purpose of the system.
UL: Use Limitation	The data will be used solely to serve the main purpose of the system.
OP: Openness	The user is informed about which other parties have access to the collected data.
SS: Security Safeguards	The data is securely stored.
IN: Insight	The user can inspect the stored personal data.
MO: Modification	The user has the possibility to make changes in the stored data

Table 1. Simplified expressions of OECD guidelines to address.

Participants were asked to make their judgment regarding the privacy guidelines through a web-based questionnaire. The first page contained information about the study and the kind of participants needed. The next pages consecutively described a context of a diabetic person that may benefit from a health monitoring system, the purpose of the questionnaire, the base scenario (see figure 1) and the explanation of the pairs of adaptations that would be offered to them.

To ensure that the text and the guidelines were understood properly two preparatory studies were conducted with eight participants each. These studies revealed the difficulties of comprehension of privacy related statements; to an extent these are due to participants inferring privacy related functionality beyond the text. The preparatory studies lead to a rephrasing of the texts as shown in figure 1 and table 1; with a marked improvement in comprehension.

Measures

The relative importance of complying with the eight privacy guidelines was measured by pairwise comparison. Participants were offered all combinations of complying with the guidelines in pairs of two (a total of 28 pairs). Participants were asked to choose their most preferred adaptation for each pair. We obtained from each participant

a total of 7 judgments per guideline indicating the importance of complying with that guideline compared to the other guidelines. This total score per guideline was divided by 7 to obtain a score for importance between 0 (never regarded more important than other guidelines) and 1 (always regarded more important than all other guidelines). Besides the preferences for each of the guideline, participants were also inquired about chronic conditions, age and land of residence.

Procedure

After accepting to take part in the study, participants entered the website, read the context description and were informed about the purpose of the study. Then they were offered the base deviant scenario. Subsequently, participants were offered a pair of possible adaptations. They were asked to indicate which of the two adaptations they would prefer. This process was continued until the participant judged all 28 possible pairs of guidelines. Finally, some additional questions were asked about chronic conditions, age and country of residence.

All pairs of possible adaptations were offered to participants in random order. Besides, the position of each adaptation was alternated so that each adaptation would be offered a similar amount of times as the first or the second option within a pair. (For eight participants that were recruited for a paper-based pilot study a different procedure was followed. The position of each adaptation was still alternated, but the pairs of possible adaptations were not presented in random order to prevent administrative problems).

RESULTS

In total 50 participants completed the questionnaire – this number includes 8 participants from the paper-based study for whom health related information had not been obtained. Most participants (69%) were between 26 and 45 years old and reside in the Netherlands. Table 2 shows the occurrence of different chronic conditions among the remaining 42 participants in our sample. It turns out that 6 of the participants asked suffer from heart failure, 5 suffer from diabetes and also 5 from asthma. COPD and Depression were mentioned by only 1 participant as a current condition. On the other hand 6 out of 30 participants indicated to suffer from another condition than the ones specified (this question was added later, and hence only answered by a small group of participants).

Chronic condition	Participants
Heart Failure (N=42)	14%
Diabetes (N=42)	12%
COPD (N=42)	2%
Asthma (N=42)	12%
Depression (N=42)	2%
Other Chronic Condition (N=30)	20%

Table 2. Data on health condition of participants.

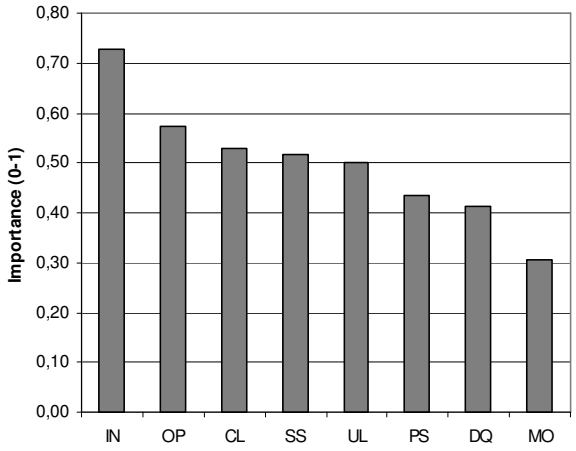

Figure 2. Relative importance of OECD guidelines for health monitoring scenario.

Figure 2 shows the importance of all eight guidelines compared to the other guidelines for the whole sample on average. Insight is the preferred adaptation in almost three quarters of the situations (0.73), and Openness is found more important than other guidelines in more than half of the situations (0.57). Least preferred were the guidelines of Data Quality and Modification. Data Quality was preferred to other guidelines only in 41% of the situations, and Modification in 31%. There is however, quite some difference in preference by participants for the guidelines Modification, Purpose Specification, Security Safeguards and Collection Limitation.

In order to check whether there is a difference in guideline importance between users depending on their need for medical monitoring (people with a chronic condition, people over 65 and those with no specific need for medical monitoring), we performed a one-way ANOVA analysis. The analysis revealed that there is no significant difference in mean guideline importance between the groups based on their need for medical monitoring except for Security Safeguards (p=.028) and Purpose Specification (p=.036). Since there were no significant differences for the other six guidelines, we decided it was feasible to add the data of all participants together and treat them as one single group. Apparently, there is no clear difference in guideline importance depending on people's need for medical monitoring. However, there is quite some variation in guideline importance for some guidelines. To explore the reasons for different judgments regarding the relative importance of the guidelines, it was decided to analyze the data for different clusters of users. Due to the small number of participants in this study the outcome should be treated with care. A K-means cluster analysis was performed. Different numbers of clusters were analyzed. Finally a cluster of 4 was chosen. The number of participants within each cluster ranges from 10 to 19. It is interesting to see

that the participants with a need for medical attention and those without such a need are spread over the different clusters.

A closer look at the four clusters shows that the first cluster with 10 participants especially values Purpose Specification, Data Quality and Use Limitation (see Figure 3). In other words they find it important to know the purpose for which data is collected, they value the relevance of the collected data to that purpose, and want the data to be used solely for that purpose as well. They could be considered as the cluster that is mainly concerned about Purpose of Use. The second cluster, consisting of 11 participants, values the guidelines Openness, Security Safeguards and Insight. So this cluster appreciates to know which other parties have access to the data, and that data is protected by security safeguards, and to have access to the data themselves. They can be regarded as having a desire for guarantees. The third cluster of 10 participants finds the guidelines Insight and Modification relatively important. This means that they particularly value to have access to and control over their data. They could be considered as the cluster requiring User Control. The last cluster of 19 participants, scores high on Collection Limitation and Insight. So they especially care about the type of data that is collected, and want to be able to inspect that data. This cluster is mainly concerned about the type of data. Since this study is only the first stage of a larger research, we don't have enough data to explain why these differences occur. Neither age nor chronic conditions seem to explain these results.

DISCUSSION

We had anticipated that Collection Limitation and Purpose Specification would be more important than Data Quality. As knowledge of what data is collected and for what purpose may help infer the relevance of the data collected. Rather, our results show that there is not much difference in importance between Purpose Specification and Data Quality and that Collection Limitation is valued somewhat more. We also expected that Security Safeguards would be regarded as important, however in this study Security Safeguards turned out to be somewhat neutral compared to the other guidelines.

We had also anticipated that the ability to modify data would be valued more than having insight in the data, since being able to make modifications would imply having some insight in the data already. However, in this study Insight was valued much more (0.73) than being able to make modifications to data (0.31). This means that Insight was preferred over other guidelines in almost three quarters of the situations, whereas the ability to modify was only preferred over other guidelines in less than one third of the situations. From comments participants made we know that people feel that modifying health related data is not regarded useful.

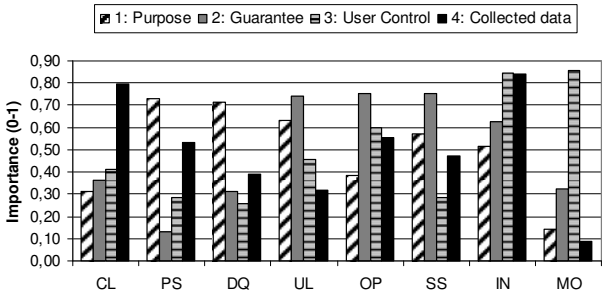

Figure 3. Cluster analysis based on guideline importance.

This study discovered different clusters of people with regard to the importance of certain privacy guidelines. Cluster membership is not determined by age or presence of chronic conditions. Rather, four different clusters emerge which can be described as people concerned respectively about the purpose of use, guarantees, user control, or type of data collected.

Further research is pursued to find a way to profile these groups of people. Our results do not support any explanation for this clustering. As a follow up to our experiment two focus groups were conducted, one with young diabetics and one with aging heart patients. The details of this study are not presented here; we note however that the qualitative data obtained reveals radically different perceptions regarding privacy between these two groups, especially with regards to the dimension of control, with the young diabetics being more concerned, more inclined to control when and whether their doctor could obtain access to their health data. For both groups the disinterest in data modification found in our current experiment was confirmed.

REFERENCES

1. Culnan, M.J. and Armstrong, P.K. Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation. *Organization Science*, 10, 1, (1999), 104-115
2. Hong, J.I., Ng, J.D., Lederer, S. and Landay, J.A. Privacy risk models for designing privacy-sensitive ubiquitous computing systems. In *Proc. DIS 2004*, ACM Press (2004), 91-100.
3. Iacchello, G. and Abowd, G. Privacy and Proportionality: Adapting Legal Evaluation Techniques to Inform Design in Ubiquitous Computing. In *Proc. CHI 2005*, ACM Press. (2005), 91-100.
4. Karat, C., Karat, J. and Brodie, C. Why HCI research in privacy and security is critical now. *Internai. Journal of Human-Computer Studies*, 63, 1-2, (2005), 1-4.
5. Langheinrich, M. Privacy by Design – Principles of Privacy-Aware Ubiquitous Systems. In *Proc. of UbiComp 2001*, Springer-Verlag, LNCS 2201, (2001), 273—291.
6. OECD. Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. 23-Sep-1980.