Who knows about me? An analysis of age-related disclosure preferences

Linda Little
PaCT Lab
Northumbria University
Newcastle, UK

Pam Briggs
PaCT Lab
Northumbria University
Newcastle, UK

Lynne Coventry
PaCT Lab
Northumbria University
Newcastle, UK

Users are increasingly willing to disclose sensitive personal information online, seemingly without great regard for privacy protection. We surveyed over 1200 people to measure user attitudes and behaviours in terms of: (i) the type and perceived sensitivity of information they regularly disclosed, and (ii) who the recipients of different types of information were. In our initial analysis of the data we have observed an interesting age-related trend: a U shaped curve whereby the youngest and oldest members of society are less protective of their privacy than the middle-aged cohort.

Information disclosure, privacy, age

INTRODUCTION

The disclosure of highly sensitive personal information, is happening on an unprecedented scale, raising important questions about user-preferences in respect of privacy, defined here as 'an individual right to determine how, when and to what extent information about the self will be released to another person' (USACM, 2006). New privacy challenges are inevitable with the growth of online interaction (Olson, Grudin & Horvitz, 2005). For example, users wishing to sign-up for an online product or service or wishing to simply join a group are typically asked to provide a significant profile of personal data on enrolment and are, thereafter, subject to a subtle data-collection process in respect of personal choices and preferences.

Unsurprisingly, then, a research agenda has grown up around understanding disclosure patterns and privacy preferences (Guha, Tang & Francis, 2008), in part driven by the perceived need to automate such disclosure processes and preferences and in part driven by the rather worrying observation that, in the face of known privacy risks, users seem willing to give away personal identity information very cheaply (Leathern, 2002).

Internet users who hold privacy in high regard can recognise not only the costs, but also the potential benefits to information disclosure (Guha, et al, 2008). Researchers have recently begun to understand privacy management in terms of the costs and benefits of informational trade-offs, and a new study of 'privacy economics' has emerged, where privacy benefits can include better access to relevant information and targeted sales advice, while privacy costs can include identity-threats and physical vulnerabilities dependent upon location-disclosures e.g. (Ahern, et al., 2007)). As information exchange becomes more ubiquitous, responsibility for the calculation of costs and benefits for each single exchange can become too much for any individual, with the result that a significant research effort is being targeted at the construction of automated 'trust agents' or 'privacy wizards, capable of managing disclosure decisions seamlessly (Gross & Acquisti, 2005).

Effective risk management is conceptually possible in the privacy domain, however many users act as if they are either unaware of the risks and consequences of revealing too much personal information in online environments (Young & Quan-Haase, 2009) or as if they simply don't care (Guha, et al., 2008). Indeed, the privacy literature is peppered with examples of users who vouchsafe the importance of privacy protection while, at the same time, eschew privacy protection behaviours (Lederer, Mankoff & Dey, 2003).

Nowhere is this more obvious than on social networking sites. For example, Facebook users seem very willing to disclose highly sensitive information in relation to their personal identity and lifestyle (Leathern, 2002; Schrammel, Koffel & Tscheligi, 2009). Real names, pictures, date of birth, telephone numbers, employment status and even physical addresses have all been recorded as being openly accessible by strangers, despite the

fact that such personal disclosure can lead to identity theft and more disturbing outcomes such as stalking (Stater & Lipford, 2008). Subsequent studies of different student communities (e.g.Schrammel, et al., 2009) have showed very similar patterns of high levels of personal disclosure. In the Tufekci (2008) study, for example, two thirds of students from a sample of 601 disclosed their sexual orientation and relationship status, half disclosed their political and religious orientation and almost a third gave out their phone number.

This study is the first in a series of investigations in which we seek to move beyond a generic predictive model of disclosure and in which we are beginning to explore tools that can capture individual differences in stated disclosure preferences. Our motivation is to provide new data and techniques in support of the development of individual privacy management tools while, at the same time, exploring emergent patterns in disclosure data. In this paper, we present a preliminary analysis of a dataset (N=1218), focusing on age-related trends in disclosure preferences. Our rationale for this focus is that, while the willingness of young people to disclose personal data about themselves is well established (Tufekci, 2008), other possible age-related patterns have been left unexplored in the literature.

METHOD

A new questionnaire, grounded in earlier work (Little & Briggs, 2009), was developed to measure beliefs about information sensitivity and availability and to capture disclosure preferences. The questionnaire was promoted on the Zoomerang.com website. The first section of the questionnaire addressed patterns of disclosure across different domains. Demographic variables were recorded related to age, gender, level of education, employment status and country of origin.

From an initial 1687 responses, the data set was cleaned up and any incomplete questionnaires were removed. A total of 469 replies were removed from the set (mainly through incomplete answers) leaving a total of 1218 respondents. Of the respondents, 657 were males and 561 females. Respondents reported locations from all over the world. As might be expected, the vast majority (1037) were from the United States, 165 from the UK and 16 were from other locations. This reflects the online population but also the bias as the survey was placed on Zoomerang.com a US site.
Respondents reflected a wide range of ages. The majority falling in the 36-45 age group, though with a strong representation in all age groups from 18 to 65. Only the under 18 and over 75 groups showed any tailing off.

A ubiquitous computing scenario was used at the beginning of the questionnaire to communicate the concept of seamless disclosure to the average citizen. In this scenario they were asked to imagine a future where devices embedded in the environment would be in constant communication about their health, lifestyle, finances etc. but that they would be charged with setting up their privacy preferences to ensure appropriate disclosures.
Participants were then presented with a 'disclosure grid' that asked them to consider a range of different information types set against a range of different information recipients. In each cell of the grid they were asked to rate on a scale of 1-5 how happy they would be to reveal this information to that individual or group. For the purpose of this, preliminary analysis, information type was collapsed into four main areas:

(i) Personal Identity- name, date of birth;
(ii) Health- medical history, family history, allergies, medication, genetic disorders, GP visits;
(iii) Lifestyle- lifestyle, shopping habits, employment status;
(iv) Financial- financial details, current bank balance.

Information recipient was also collapsed into four types

(v) Health professional-doctor, hospital consultant.
(vi) Family/friend- partner, family member, friend.
(vii) Work acquaintance-employer, work colleague;
(viii) External companies-private company, government agency, financial institution.

RESULTS

All data was entered into PASW and screened for normality prior to further analyses. In a preliminary analysis of age-related trends, participant responses were divided into three age groups (>35, 36-55, 56< years) multivariate ANOVA's were conducted to find any differences between the groups in terms of how happy they were to disclose different types of personal information. Findings are given below, categorized by information type, with patterns in age-related data shown in the series of figures that follow.

Personal Identity information

The results show, firstly, an extreme willingness to disclose personal identity data. There were no significant differences between age groups for happiness to disclose personal identity information to either health professionals or family/friends (perhaps related to a ceiling effect). Significant

differences were found between groups for happiness to disclose information to work acquaintances F (2, 749) = 8.08, p<.001. Post hoc comparisons revealed participants aged 36-55 feel significantly less comfortable disclosing personal identity information to external companies in comparison to both the under 35 (p<.01) and over 56 (p<.01) age groups. Significant differences were found between groups for happiness to disclose information to external companies F (2, 749) = 6.49, p<.01. Post hoc comparisons revealed participants aged 36-55 feel significantly less comfortable disclosing personal identity information to external companies in comparison to both the under 35 (p<.01) and over 56 (p<.05) age groups. See figure 1.

Health information

The results show there were no significant differences between age groups for happiness to disclose health information to either health professionals, family/friends, work acquaintances or external companies. See figure 2.

Lifestyle information

The results show there were no significant differences between groups for happiness to disclose lifestyle information to either health professionals, work acquaintances or external companies. Significant differences were found between groups for happiness to disclose information to family/friends F (2, 748) = 3.49, p<.05. Post hoc comparisons revealed participants aged 36-55 feel significantly less comfortable disclosing lifestyle information to family/friends in comparison to both the under 35 (p<.05) and over 56 (p<.01) age groups. See figure 3.

Financial Information

The results show there were no significant differences between groups for happiness to disclose financial information to either health professionals or family/friends. Significant differences were found between groups for happiness to disclose information to work acquaintances F (2, 573) = 3.09, p<.05, however, post hoc comparisons failed to support this difference. Significant differences were found between groups for happiness to disclose information to external companies F (2, 573) = 3.12, p<.05 participants aged over 56 feel significantly less comfortable disclosing financial information to external companies in comparison to the under 35 (p<.05). See figure 4.

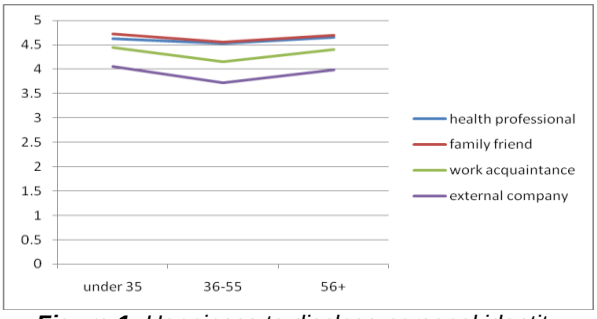

Figure 1: *Happiness to disclose personal identity information as a function of age*

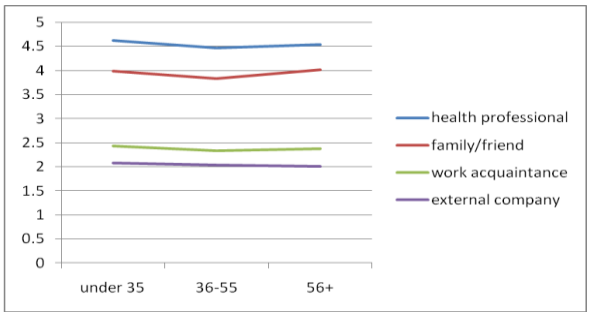

Figure 2: *Happiness to disclose health information as a function of age*

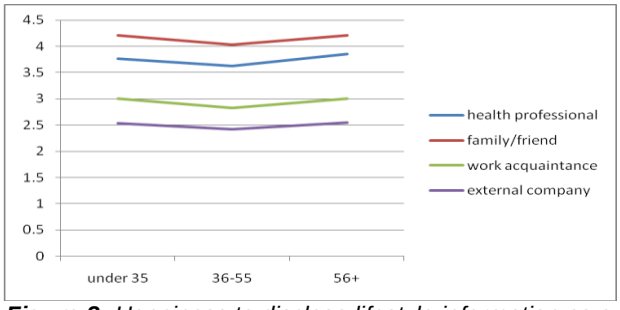

Figure 3: *Happiness to disclose lifestyle information as a function of age*

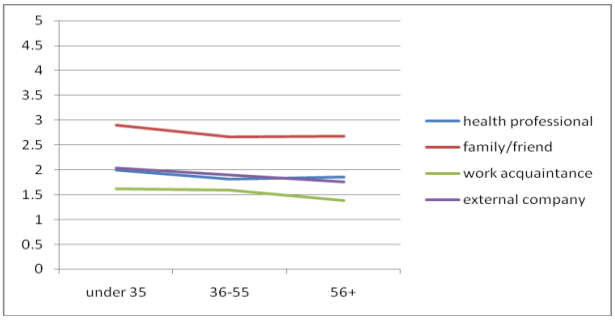

Figure 4: *Happiness to disclose financial information as a function of age*

DISCUSSION

Overwhelmingly, our first observation should be the remarkable willingness of our participants to disclose identity information in comparison with other information types. The data suggests we have become rather blasé in the disclosure of names and dates of birth and have reconciled ourselves to giving up such information as the first bargaining chip in an information exchange. Our

second point concerns age. Our initial justification for exploring age-related disclosure preferences derives from (i) the existing literature concerning young people's relative insensitivity to privacy issues and (ii) our own interviews with some of our oldest old, where a certain carefree attitude also seemed to reign (in part derived from a sense that they weren't doing anything to be ashamed of). Looking across the first three data sets in the figures above, this intuition is born out in the U-shaped curve that suggests a dip in willingness to disclose information during middle age, the notable exception to this pattern being the disclosure of financial information. We believe that this is the first time that such a U-shaped age related pattern has been reported.

Our youngest and oldest adults have told us they're happier to share personal information than those in middle-age (aged 35-55). Why might this be? One interpretation is that, as we enter middle-age, we become more aware of the implications of data sharing – we understand the ways in which personal data has value. Earlier we discussed the kinds of cost-benefit analyses that any individual may undertake in weighing up a disclosure decision. As we engage more fully with society, we come to realize both costs and benefits more fully and also realize that we may have more to lose – in terms of our status in society and specific monetary issues such as salary and insurance privileges. Later, as we move into our sixties and beyond, we regain a willingness to disclose information in all but the financial domain. Why might this be? The answer we seek may come from other studies we've conducted in relation to ubiquitous technologies (Little & Briggs, 2009) and where, in response to questions about privacy preferences, we've often been told by our older adults '*I don't care who knows where I am, what have I got to hide?*'. In other words, a sense that our oldest adults fear little in terms of others knowing their business.

SUMMARY AND FUTURE WORK

In this paper we have identified some initial age-related patterns in the data showing, for the first time, a U shaped curve in relation to privacy management. An interesting question for the future is whether or not young adults living in the internet age continue to be open with regard to revealing personal information or if experience will change their behaviour? We are aware that, in this work-in-progress, we are limited to self-report data, however this study forms part of a series of studies in which we seek to combine the results from such surveys with observational studies of real data sharing in order to better understand both stated preferences and actual behaviours. What we would argue here is that our understanding of the way privacy concerns are developed and shaped over time, and the way that such concerns come to shape behavior is as yet very poorly understood. Furthermore, this paucity of understanding is reflected in the limited range of technological solutions available. Better tools are needed to capture and understand privacy preferences and it is these that will be the focus of our future work in this area.

REFERENCES

Ahern, S., Eckles, D., Good, N.S., King, S., Naaman, M., & Nair, R. 2007 Over-exposed?: privacy patterns and considerations in online and mobile photo sharing, *SIGCHI conference on Human Factors in Computing Systems*, San Jose, California, USA

Gross, R. and Acquisti, A.2005 Information revelation and privacy in online social networks. *Proceedings of the ACM Workshop on Privacy in the Electronic Society.*

Guha, S., Tang, K., & Francis, P.2008 NOYB: Privacy in Online Social Networks. *WOSN'08, August 18,* Seattle, Washington, USA. 49-54

Leathern, R. 2002. *Online Privacy: Managing Complexity to Realize Marketing Benefits.* May, 2002. www.forrester.com.

Lederer, S., Mankoff, J., and Dey, A. K. 2003 Who wants to know what when? Privacy preference determinants in ubiquitous computing. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA, April 05 - 10

Little, L., & Briggs, P. Privacy factors for successful ubiquitous computing. *International Journal of E-Business Research, (2009), 5 (2),* 1-20

Olson, J., Grudin, J., Horvitz, E. 2005 A study of preferences for sharing and privacy'. *CHI '05 Extended Abstracts on Human Factors in Computing Systems*.

Schrammel, J., Koffel, C., & Tscheligi, M. 2009 How much do you tell? Information disclosure behaviour in different types of online communities. *C&T'* University of Pennsylvania, USA, 275-284

Strater, K., & Lipford, H.R. 2008 Strategies and struggles with privacy in an online social networking community, Proceedings of the 22nd *British HCI Group Annual Conference on HCI 2008: People and Computers XX II: Culture, Creativity, Interaction,* September 01-05, Liverpool, United Kingdom

Tufekci, Z. 2008 Can you see me now? Audience and disclosure regulation in online social network sites. *Bulletin of Science, Technology and Society*. 28, 20--36.

USACM 2006 USACM Policy Recommendations on Privacy, June 2006
. http://usacm.acm.org/usacm/Issues/Privacy.htm

Young, A.L. and Quan-Haase, 2009 A. Information revelation and Internet privacy concerns on social network sites: A case study of Facebook. *Proceedings of the fourth international conference on Communities and technologies,* June 25-27, University Park, Pennsylvania, USA. ACM Press.