PERSONAL PRIVACY THREATS: A TAXONOMY FOR RISK ASSESSMENT

Shirley Atkinson[1], Christopher Johnson[2] and Andrew Phippen[1]

[1]Network Research Group, University of Plymouth, UK
shirley.atkinson@plymouth.ac.uk
[2]University of Plymouth, UK
c.johnson@plymouth.ac.uk

ABSTRACT

The explosion in the use of the Internet and the growth of the volume of available data has made collecting personal information about an individual easier than ever before. Problems faced by vulnerable individuals which stem from the abuse of gathered information are exacerbated. Abuse and harm of individuals, through grooming, harassment and bullying, coexist with identity theft as examples of criminal behaviour that are, aggravated by the ready availability of personal information. This paper presents a Taxonomy of threat to be utilised when assessing risks to vulnerable individuals and concludes with a description of the application of our taxonomy to five social networking websites.

KEYWORDS

Internet, Personal Privacy, Teenagers, Domestic Abuse Survivors, Semantic Web

1. INTRODUCTION

Modern privacy problems have been identified as a result of unchecked information flows between a variety of different entities [1]. The ubiquitous nature of the Internet facilitates the gathering, storage and onward transmission of personal data - something which business enterprises turn into a commodity [2]. One impediment to the free flow of information between entities has been the format of data, in that not all formats have been recognised or accepted. The Semantic Web [3] concept aims to address this issue and provide the standards required to allow data to be shared in a more seamless fashion.

Personal information is a difficult entity to control, once it has been divulged it is difficult to ascertain exactly where else it may be divulged. As Tavani [4] highlights the three types of threat to personal privacy add to the complexity of protecting personal data: implicit or explicit data gathering techniques; data exchange techniques; and data mining techniques add to the pressure upon personal data. Social networking web applications, where individuals link to each other give a good example of uncontrolled data exchange. One thing divulged to a friend with a direct link can then be observed by somebody else who does not have a direct link.

Divulging information is observed in two distinct areas: personal websites, on-line diaries and other internet-mediated communications encourage individuals to divulge their personal information; and public personal information [4], information about an individual held by third parties. Public records have always been available to those who take the time to enter public buildings and search records, however, government services now make many of those public records available through the Internet. Examples of these are planning application details, and access to the general record office indexes of births, marriage and deaths.

Divulging personal information does not in itself pose a problem, however problems arise when the information divulged is abused. In this respect some individuals are considered more prone to harm than others. Abuse and harm of individuals, through grooming, harassment and bullying, coexist with identity theft as examples of criminal behaviours, all aggravated by the ready availability of personal information. Posting information on social networking websites has been linked to murder [5]; Bocij [6] identifies the Internet as a tool for stalking behaviour; Southworth et al [7] illustrates how modern technology is being used in situations of domestic abuse; and Mitchell et al [8] and Hughes [9] observe how the Internet has facilitated sexual exploitation of women and children.

In this paper we present a taxonomy of threat designed for use within risk assessment. This taxonomy has been designed by examining the problems faced by vulnerable groups with regard to the impact of the Internet upon their personal privacy. In section two we introduced the methodology, describe how the vulnerable groups were selected and the research methods used. In section three the findings are presented and section four presents the taxonomy and applies it to a selection of social networking websites. Section five concludes with the direction for future research.

2. METHODOLOGY

Raab and Bennett [10] propose that more effective privacy solutions are created by focussing on the privacy issues for vulnerable groups. With this in mind, two groups were selected. Qualitative research methods were utilised to explore the social context [11] within which those groups found themselves. The data collected was considered to be responsive to their thoughts, feelings, experiences and behaviours [12] which would allow a richer understanding of their personal privacy situation. Taking note of the complex interplay of issues would therefore lead to an enhanced understanding of the social context, which in turn would lead to more meaningful action to be taken. For the mitigation of risks to be effective, a good understanding of the factors involved is required and the qualitative approach lends itself to gaining that understanding.

2.1. Selection of Vulnerable groups

Two groups were considered to represent vulnerable groups for the purposes of this research, domestic abuse survivors and teenagers.

Domestic abuse survivors, hereafter referred to as Survivors, exhibit vulnerability in different situations. As a group of individuals they are most likely to experience "dataveillance" [13] technologies being used against them. Whilst Survivors will face many violent and controlling episodes before they seek help [14], the time when they are at their most vulnerable is when the decision to leave the abusive relationship and seek refuge is made [15]. Any release of personal information at this time can lead to serious harm or death.

Teenagers have been identified by Magid [16] as those most at risk from predatory behaviour. These young people increasingly explore the boundaries of the technology that surrounds them, often in such ways that their parents do not understand and therefore find difficult to monitor. The Internet has become more of a social space with many of them creating and uploading content [17] through the many and varied social networking websites. These teens utilise the different web applications as a way of keeping up with their peers and often do not consider the consequences of their actions. Advice given by the government education campaigns [18], and by researchers [7; 19] centre around keeping personal information private. However, here lies the dichotomy: young people should keep their information safe, but they want to share it with their friends using the technology that is part of their social world.

2.2 Research Methods

The opinions from members of statutory and voluntary sector bodies in the field of domestic abuse were sought through interviews and workshops. Focus groups were carried out where young people could discuss their views on personal privacy and interaction with the Internet. These were backed up through an on-line survey.

Semi-structured interviews were held with two managers of refuges and two outreach workers; two conference workshops were held at the Women's Aid National Conference. The participants of the workshops were workers from both refuges and outreach services, all working with domestic abuse survivors. The workshops were aimed at exploring the uses and abuses of technology.

Seven focus groups were conducted involving teenagers; two were held with year twelve teenagers, one with year ten and another with year eight. Four more focus group transcripts were made available from the Trustguide[1] project which utilised a similar methodology and questions. Prior to conducting the discussions the young people within the focus groups were issued with a questionnaire designed to elicit some broad demographic data around their Internet usage. The discussions were generated by describing a set of scenarios based upon an understanding of the problem situation, and encouraging the respondents to discuss their views, thoughts and feelings that the scenarios generated. The discussions were recorded onto tape for later transcription which were analysed using N6 software to extract the concepts. The concepts emerging from the data were refined and structured into the taxonomy that is presented in the next section.

3. FINDINGS

Those who had responsibility for the safety and well being of others, managers of refuges, and teachers, voiced their concerns about the risk potential that the Internet and related technologies posed. Primarily their concerns were connected with the ease with which personal information was divulged through such things as mobile phones, emails, social networking websites, public records and third party databases.

Examples were given where personal information made available through the Internet had compromised women's safety; 83% of the teenagers interviewed divulged personal information with 27% expressing concern about having done so.

Teenagers employed a variety of coping mechanisms: they made good use of any blocking techniques made available by the different software used; where requests for personal information were considered to be excessive, these were ignored or if the request was mandatory, false information supplied. However, the descriptions given by the teenagers of the information they divulged did not entirely match the public information given by themselves in social networks. For example, some chose to claim they were older than they really were, others posted photographs of themselves wearing revealing clothing. On examining the top three social networking websites listed by the teenagers, each of the schools taking part in the focus groups had a substantial presence on them.

4. TAXONOMY OF THREAT

Focussing upon the concerns raised by teachers and managers of refuges, the data collected was evaluated to ascertain where personal privacy risks to teenagers and survivors lay. Risk categories were identified in terms of the potential impact where damage to personal privacy

[1] http://trustguide.org.uk

could take place; where threats to giving out personal information might lie; and where there was a potential for unwanted intervention.

Within these three areas, the manner in which the risks to individuals manifested themselves were considered within four different categories:

- e-Sociability
  - This considered the act of being sociable within the electronically connected context and examined the methods employed for keeping in touch with peers.
- Data boundaries
  - How individuals determined what elements of personal data needed protecting and how boundaries were set around personal data.
- Access control
  - Once the boundaries were determined, consideration was given to how they were enforced, along with who provided the tools to enforce those boundaries.
- Technological impact
  - Consideration was given to the effects the technology had on an individual's behaviour.

Our taxonomy of threat has the form illustrated below in table 1.

Table 1: Taxonomy of Threat

	Propensity for Harm	***Divulging Personal Information***	***Unauthorised Intrusion***
e-Sociability	Manifestation of risk		
Data Boundaries			
Access Control			
Technological Impact			

Risk assessments are usually carried out by experts within a field to consider specific hazards and give proposals on how to minimise or remove the identified risks. However, problems arise in a number of ways. Internet devices, their uses and impacts pose many different hazards in many different situations. Problems also arise when different experts attribute different meanings and weightings to risks. Utilising a framework provides consistency and by basing the framework upon a taxonomy a less restrictive approach structure will allow risks to be evaluated. Different experts can utilise the same taxonomy structure but have the ability to adapt the structure to different local contexts [20].

A taxonomy is an organised structure that serves as a useful lens for classifying and understanding a body of knowledge [21]. Concepts are logically ordered into groups and categories as illustrated in Table 1 thus allowing preventative measures to be applied.

4.1 Existing Taxonomies

There are three taxonomies proposed that are connected with personal privacy: the Privacy Goals Taxonomy [22], Young people and risk on-line [23] and the Taxonomy of Privacy [24]. The first from Anton et al [22] primarily focuses upon business privacy data and the field of commerce. Existing threats to consumer privacy are categorised into seven classes of threat. The second taxonomy developed by the Cyberspace Research Unit at Lancaster is more relevant

than the first to the problems faced by teenagers [23]; behaviours are represented in terms of physical, psychological and social well being of children and young people. The third taxonomy proposed by Solove [24] identifies different privacy harms and problems that have already achieved a significant amount of social recognition. Four categories and many related sub-categories are identified: Information Collection, Information processing; Information dissemination and Intrusion.

Whilst these existing taxonomies assist in providing different viewpoints of the privacy field, they concentrate on different areas to that which has been examined in our research. The Taxonomy of Threat introduced in Table 1 is discussed further by use of an example below.

4.2 Social Networks

To demonstrate how the taxonomy presented in table 1 assists risk assessment, it is applied to an assessment of a selection of five social networking web applications: www.myspace.com, www.bebo.com, www.spaceslive.com (Windows Live Space), www.facebook.com and www.zorpia.com. These applications were sampled from those listed by teenagers in the questionnaires.

4.2.1 E-Sociability

The Internet provides different methods for young people to keep connected with their peers either through Internet-mediated communication such as emails and messenger, but also through web applications and social websites. "Blogging", creating on-line diaries, has become a popular past-time and is considered to be a growing phenomenon [25] but one which has been identified by CEOP [26] as an area of concern. McMillan and Morrison [27] observe how young people build their community around the interactive technologies.

New Nokia phones contain LifeBlog software offering the ability to create an on-line diary, a blog, whilst on the move. O2 encourages people to upload content in return for payment [29] and video social networking websites such as YouTube are launching the facility to use mobile phones to view the videos posted on the website [30].

Gross et al [31] suggest that the interface of social networks combined with peer pressure, herding behaviour, and short-sighted privacy attitudes contribute to the situation where young people reveal quantities of personal information.

Each of the applications considered allowed people to link and connect with each other. Common features include photographs and some form of comment whether in the form of blogs, journals or discussion boards. MySpace, Bebo and Facebook all link people together in groups; these can be based on school, university or the workplace. Bebo and Facebook provide differing levels of control over who joins the different groups. In the case of Bebo you cannot join a group uninvited, another member must enrol you. Facebook allows you access to school networks for two weeks before removing you from the group if you have not been accepted by another member of the group. Bebo, Facebook and MySpace offer a multimedia rich environment allowing music and videos to be shared and played. Zorpia is aiming at the over 16 year old market. Facebook provides the facility to upload photographs and place description tags of individual's within the photographs that link to the personal profiles of those people.

A summary of the potential risks is given below:

Table 2: e-Sociability Summary

Propensity for Harm	*Divulging Personal Information*	*Unauthorised Intrusion*
Personal data gleaned for use in situations of bullying, stalking and harassment.	Photographs, blogs, journals, discussions linked to personal profiles. Name and address used in search and display terms.	Photographs and video's uploaded by third parties. Access to profiles through friends rather than direct link.

4.2.2 Data Boundaries

Tavani [4] identifies personal data in two areas: public personal data and normative personal data. Normative personal data is where an individual would expect their data to be kept private. The public personal data has boundaries placed around it that are not under the control of the individual to whom the data belongs. Third party actions upon these boundaries can cause issues of concern. One area of concern is the release of public records, for example the electoral role, combined with the information released through social networks. This data boundary is infringed when websites contain personal data that has been posted by other individuals.

Each of the five websites collect and display a wide variety of personal information. As a minimum MySpace collects first name, last name, postcode, country and email address. The others following similar lines. Each allow personal photographs to be uploaded. Facebook and Zorpia do not make as much information mandatory.

Table 3: Data Boundaries Summary

Propensity for Harm	*Divulging Personal Information*	*Unauthorised Intrusion*
Linking postcode to mapping applications. Linking date of birth to GRO indexes to obtain mothers maiden name.	Profile made public. Third parties divulging information.	Third parties posting photographs, names, addresses and other personal details.

4.2.3 Access Control

The approaches and tools for profile protection differ between the five websites. Facebook is the only website to offer a fine-grained approach to controlling what is made publicly available. Many of the personal data elements can be toggled between public or friends only viewing. Photographs that are tagged with an individual's name are notified to that individual, thus allowing them the opportunity to request their removal if required. Profiles and photographs of individual's can only be seen once a link has been made and approved by the other party.

Bebo allows the whole profile to be made private only. Each of the websites assessed allow the individual to hide their date of birth, friends who have made connections are easily seen along with their friends. They allow you to browse freely the friends of friends who are connected to your profile.

Searching for individual's differs amongst the websites, some allow searching for individual's by location, age and gender, others are more restrictive only allowing searching to be carried out

on networks that the searcher has been invited into. Zorpia allows searching to be carried out by gender, age and location. MySpace uses first name, last name and location for the searching and state in their privacy policy that pictures and first names will be displayed to users who search for you. Windows Live Space and Bebo provide a free text search box and do not have a facility to refine the search any further.

Table 4: Access Control Summary

Propensity for Harm	***Divulging Personal Information***	***Unauthorised Intrusion***
Search on location, gender, age.	Control of profile vs finer grained control of individual data elements. First name and photograph returned in search.	Tagging and linking of photographs. Searching

4.2.4 Technological Impact

Each of the websites allows and encourages personal information to be shared, however each has a different approach to protecting the user's privacy.

MySpace is the only one to make the majority of important personal information mandatory to join the site. First name, last name, email address, postcode, country and gender are all essential for registration, date of birth however can be omitted. It does provide safety tips and the privacy policy from links at the bottom of the page and the registration page reminds potential users that their data will be stored and bound by US data rules.

Bebo and Windows Live Space make more of the interaction with CEOP and blog safety campaigns by placing the links to report abuse and safety guidelines in prominent positions. Bebo places reminders for those under 21 next to the text boxes so that the safety tips are more prominent. Zorpia, being aimed at those over 16, carries no such warnings.

Table 5: Technological Impact

Propensity for Harm	***Divulging Personal Information***	***Unauthorised Intrusion***
Lack of safety warnings on some sites	Important personal information mandatory for registration	Data control rules applied from different country.

5. CONCLUSIONS

By applying the taxonomy of threat as detailed in Table 1 to the field of social networking applications, issues that pose potential for risks to individuals can be highlighted and from there action can be taken by individuals or those who have responsibility for other individuals.

The results from the early stages of the research indicate that individuals use of Internet technology should be combined with empowered, informed consent. Technology should therefore be designed to facilitate an individuals control of the flow of personal information.

The next phase of the research is to evaluate the success or otherwise of a technological approach in providing privacy protection whilst using the Internet. Evaluation of the technological approach is to be carried out by the user groups themselves. The effectiveness will be assessed in terms of the understanding and perception of levels of control of personal information and understanding the potential consequences for actions taken.

REFERENCES

[1] Solove, D.J, (2004), *The Digital Person*, New York University Press, New York

[2] Tynan, D, (2005), *Computer Privacy Annoyances*, O'Reilly, USA

[3] Berners-Lee T, (2000), *Weaving the Web*, Texere, London

[4] Tavani, H.T, (2007), *Ethics and Technology 2nd Edition*, Wiley, USA

[5] Wired News, (2006), *Teens Reveal Too Much Online*, Associated Press, 5th February, 2006, http://www.wired.com/news/wireservice/1,70163-0.html, date accessed 31/01/07

[6] Bocij, P, (2004), *Cyberstalking*, Praeger, Conneticut

[7] Southworth, C., Dawson, S., Fraser, C., Tucker, S., (2005), "A High Tech Twist on Abuse: Technology, Intimate Partner Stalking and Advocacy", *Violence Against Women Online Resources*, Minesota, www.mincava.umn.edu/documents/commissioned/stalkingandtech/stalkingandtech.html date accessed 31/01/07

[8] Mitchell KJ, Finkelhor D, Wolak J, (2005), "The Internet and family and acquaintance sexual abuse", *Child Maltreatment*, Vol. 10, No. 1, pp49-60.

[9] Hughes, D.M., (2003), "Prostitution online", *Journal of Trauma Practice*, Vol 2. No 3/4, pp115-132, www.uri.edu/artsci/wms/hughes/internet.pdf, date accessed 31/01/07

[10] Raab, C.D and Bennett, C.J, (1998), "Distribution of Privacy Risks: Who Needs Protection", *Information Society*, Vol 14, Issue 4, p 263-274

[11] Dahlberg, L., (2004), "Internet Research Tracings: Towards Non-Reductionist Methodology", *JCMC*, 9 (3) April 2004.

[12] Strauss, A., Corbin, J., (1998), *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, Sage, London

[13] Clarke, R, (1999), *Introduction to Dataveillance and Information Privacy*, Xmamx Consultancy Pty Ltd, www.anu.edu.au/people/Roger.Clarke/DV/Intro.html#Intro date accessed 31/01/07

[14] Yearnshire, S. (1997) "Analysis of Cohort" in: Bewley, S., Friend, J. and Mezey, G. (Eds). *Violence Against Women* London: RCOG

[15] Womens Aid Federation of England, (2002), Domestic Violence Statistical Factsheet 2002, http://www.womensaid.org.uk/dv/dvfactsh2002.htm date accessed 31/01/07

[16] Magid, L, (2004), *Teen Safety on the Information Highway*, National Center for Missing and Exploited Children, www.safeteens.com/safeteens.htm#Guidelines_for_Parents_0, date accessed 31/01/07

[17] Lenhart, A., Madden, M, (2005), *Teen Content Creators and Consumers*, Pew Internet & American Life Project, Washington www.pewinternet.org/pdfs/PIP_Teens_Content_Creation.pdf, date accessed 31/01/07

[18] Fiveash, K, (2006), *Internet safety talks for UK kids*, The Register, www.theregister.co.uk/2006/09/20/internet_children_safety/ date accessed 31/01/07

[19] CRU, (2006), *Internet Safety Zone*, Cyberspace Research Unit, University of Lancaster, www.internetsafetyzone.co.uk/root/default.htm date accessed 31/01/07

[20] Kemshall, K., McIvor, G, (Eds) (2004), "Managing Sex Offender Risk", *Research Highlights in Social Work 46,* Aberdeen

[21] Carr, M.J., Konda, S.L., Monarch, I., Ulrich, F.C., Walker, C.F., (1993), *Taxonomy-based Risk Identification*, sei.cmu.edu, www2.cs.uh.edu/~zhibinma/tx.pdf, date accessed 31/01/07

[22] Anton, A.I., Earp, J.B., Reese, A., (2002), "Analysing Web Site Privacy Requirements Using a Privacy Goal Taxonomy", I*n Proceedings of 10th Anniversary IEEE Joint Requirements Engineering Conference (RE'02)*, February 2, 2002

[23] O'Connell, R., Bryce, J., (2006), *Young People, Well Being and Risk Online*, Council Of Europe, http://www.coe.int/t/e/human_rights/media/1_Intergovernmental_Co-operation/MC-S-IS/H-inf(2006)005_en.pdf date accessed 31/01/07

[24] Solove, D, (2006), "A Taxonomy of Privacy", *University of Pennsylvania Law Review*, Vol 154, No 3, p477

[25] BBC, (2006a), *Blogging set to peak next year*, BBC, 14 December 2006 http://news.bbc.co.uk/1/hi/technology/6178611.stm, date accessed 31/01/07

[26] CEOP, (2006), *Blogs, Think You Know*, www.thinkuknow.co.uk/control/blogs.aspx, date accessed 31/01/07

[27] McMillan, S.J. and Morrison, M, (2006), "Coming of age with the Internet", *In New Media & Society*, Vol 8 (1), PP 73-95, Sage, London

[28] Briscoe, K, (2006), "The schoolchildren bullied by email and text", *Evening News*, 17 November 2006

[29] O2, (2006), *LookAtMe*, O2, www.o2.co.uk/fungames/lookatme date accessed 31/01/07

[30] BBC, (2006), *YouTube moves to the small screen*, BBC, 28 November 2006 news.bbc.co.uk/1/hi/technology/6190984.stm date accessed 31/01/07

[31] Gross, R., Acquisti, A, (2005), "Information Revelation and Privacy in Online Social", *Proceedings of the 2005 ACM: Workshop on Privacy in the Electronic Society*, pp71 - 80