Personal Privacy: Exploitation or Control through Technology

S.Atkinson[1], P.Jagodzinski[2], C.Johnson[2], and A.Phippen[1]

[1]Network Research Group, University of Plymouth, Plymouth, United Kingdom
[2]University of Plymouth, Plymouth, United Kingdom
e-mail: shirley.atkinson@plymouth.ac.uk

Abstract

This paper presents the current stage of work exploring the potential impact the Semantic Web will have on the personal privacy of individuals. The argument presented is that personal privacy should become part of the underlying architecture for the Semantic Web in order to limit the vulnerability of individuals.

Issues of vulnerability are presented along with the current body of thought from the legal, social and technical perspectives. The paper concludes with the direction for further research where the design of a semantic web tool will address the issues faced by vulnerable groups and individuals.

Keywords

Semantic Web, Privacy, Vulnerability, Domestic Violence.

1. Introduction

Increasingly technology is being used by abusers either for tracking or in terms of power and control (Southworth, 2005). Monitoring people is facilitated by the use of caller ID, GPS devices and high resolution web cameras. A quick search with Google using the term “track spouse” reveals a multitude of spyware, email scanners and other products for the American market. Other tools range from mobile phone tracking, through GPS devices along with voice detection software (Power, 2006).

This paper firstly considers the role of technology with consideration to vulnerability. A brief overview of the current legal, social and technical approaches being taken are presented followed by an outline of the research. The paper concludes with an overview of the current findings and an outline of how the research should progress from there.

2. Vulnerability

Vulnerability, the perceived risk of mental or physical harm, has been strongly linked to disclosure of personal information (Dinev and Hart, 2004). Margulis (1977) linked the control of personal information to vulnerability, suggesting that the more information is disclosed, the more vulnerable a person becomes.

The “architecture of vulnerability” considers how legal, social and technical elements of life all combine to create the context where privacy problems arise (Solove, 2003). These problems are exacerbated because information flows easily, is highly desirable by business and few people take responsibility for where it goes and what is done with it. One example is over the increasing number of public records now available through the Internet allowing many pieces of information to be gathered about people.

The gathering of personal information in large quantities in an unobtrusive fashion is facilitated by ubiquitous devices and commonplace technology interlinked through Internet protocols (Cranor and Garfinkel, 2005). The drive from commercial activity has eroded the boundary between private and public information. Personal information makes up a valuable commodity that can be traded and raises revenues (Tynan, 2004) especially when used in recommender systems.

Location based tracking on ordinary mobile phones provides one example of vulnerability. Four mobile phone network providers, Orange, T-Mobile, O2 and Vodafone were asked about the following scenario:

> *If a mobile phone was given as a present to a child, that had a location tracking service set up on it without the knowledge of that child, could the network provider tell the user of the phone if there was such a service, and who it was so that it could be stopped?*

None could inform a user if their network was forwarding location information to a tracking service. The only indication was a text from the location service provider to remind the user that the service was active – a potential wait from between 14 to 30 days. Once the locating service had been identified, stopping it was easy.

Viruses and identity theft are examples of exploitation of vulnerabilities in different contexts. Virus protection is more advanced with regularly issued operating system patches as the vulnerabilities are exposed. Anti-virus software, firewalls, encryption all combine to give a level of protection. Identity theft is a new growth area.

3. Legal, Social and Technical

3.1 Legal

Advances in technology create dilemmas for the legal profession from the privacy perspective. Law is used to redress harm done to the individual by the infringement of their privacy. However, this relies on being able to identify somebody to be sued, and the other assumes that the individual is empowered enough to know about the legal recourse they can take.

One dilemma emerges between the protection of society balanced against the privacy of the individual. European law attempts to address this in Articles 8 and 10 of the Convention of Human Rights which enshrine the right to privacy of the individual and set out the principles for freedom of speech. The state is allowed to intervene should there be a threat to the "economic well-being" of the country (Colvin, 2000).

Problems arise when privacy provides cover for those who would harm (Schoeman, 1984). Increased surveillance of public spaces (Garfinkel, 2000), and tracking of individuals through DNA testing (Davies, 2000) or National Identity Register (Blunkett, 2004) seek to allay the fear of crime or terrorism.

Copyright protection raises interesting dilemmas from the legal perspective between the protection of copyright and an individual's basic right to privacy (Katyal, 2004). Secretly installing software that cannot be removed (Security Focus, 2005) or digital rights management software building profiles of users without their knowledge (Solve, 2004) are being justified as a weapon in the fight against fraud (Oppliger, 2005).

3.2 Social

Privacy is a key element in human dignity for the maintaining social relationships and the sense of self (Schoeman, 1984). Concerns are raised about the risk of discrimination in terms of obtaining credit or work (Carins, 2005), participation in civic society along with the resulting detrimental impact on democracy (Schwartz, 1999), and the increase in gender divisions from the feminist perspectives (Phillips, 2004).

Education about technology has become important. Basic computer literacy courses (CLAIT) are run for free at many local colleges. These courses aim to build self-confidence in using computers for word processing, spreadsheets and using the Internet. America has an ongoing project teaching survivors about safe use of technology (Atkinson, 2004).

3.3 Technological

Privacy Enhancing Technologies seek to redress privacy concerns by making use of solutions such as anonymisation (HISPEC, 2002) to hide activity or location; control based access where roles or detection of intrusions (Lecomte et al, 2005). Animated, user-friendly visual aids like the Privacy Bird (Byers et al, 2004) or Hector the Protector (Microsoft, 2005) whilst visually appealing, have their limitations.

Finally, digital tokens provide solutions to e-commerce activity. These technologies can be considered as trust-enhancing or information exploitation, the difference being in how much information is gathered and how it is used (Guerra et al, 2003).

4. Research

A thorough understanding of privacy issues requires both an understanding of the privacy context within which people find themselves (Solove, 2003) and an understanding of the reality of privacy risks to vulnerable people (Raab and Bennet, 1998). The purpose of this phase of the research is to gain an understanding to be used in the creation of a prototype Semantic Web application.

Three groups of people have been selected. Group one considers the issues for those for whom privacy is essential, those who are at risk from violence from an intimate partner (domestic violence). Women who flee abusive relationships are at most risk when they leave (Women's Aid, 2002). Group two is a cross section of individuals who are not IT professionals. Group three are teenagers, young people who wholeheartedly embrace technology and who may not be completely aware of some of the risks. Magid (2004) proposes that teenagers are most at risk from predatory behaviour.

4.1 Methodology

A qualitative approach of semi-structured interviewing has been used so far in the gathering information. The qualitative approach was chosen as the best method for exploring the context of privacy and it was felt that quantitative methods would not address fully the complexity of that context.

In the case of group one, a lot of domestic violence goes undetected, there are a lot of people suffering from it who have good reason not to discuss it or divulge it. Therefore by talking to the agencies that provide support the reality of the situation can be explored without putting people at risk or under duress.

4.1.1 Group One

So far two semi-structured interviews have been carried out with two managers of Refuges. Refuges provide safe accommodation to women and their children. Families are often located there in the middle of the night having fled from highly abusive situations from other parts of the country. For these women, it is imperative that their locations are kept secret as they are at their most vulnerable (Women's Aid, 2002).

4.1.2 Group Two

The aim of the experiment was to explore the impact of the current World Wide Web upon the personal privacy of individuals. Ten respondents were selected through personal contacts and referrals from two of the initial contacts. The experiment took part in three stages; the first step was a semi-structured interview to ascertain the respondent’s knowledge and concerns on personal privacy. The second step was to search, with their permission, the Internet for publicly accessible information. The aim was not to carry out a hugely in depth search but to discover how much information might be discovered with a little effort and cost. The third stage was to present the findings and to explore their feelings about those findings.

4.1.3 Group Three

A pilot survey was carried out exploring the amount of information that teenagers who use Microsoft Messenger were prepared to divulge in the public directory. This looked at what information was divulged and how safe they felt the information was.

4.2 Findings

This is a very brief selection of some of the most pertinent findings from each of the three groups.

4.2.1 Group One

Mobile phones were used as tools of abuse though text harassment, photo messages and calls. Conversely they were also life savers as the means to summon help or escape. Changing the phone number was not an easy option as it meant isolating themselves from the people who were most needed. “Social Engineering” where manipulating people within utility companies or government agencies circumvented the technical security to give out names and addresses was identified as another threat. Websites such as www.upmystreet.com and www.multimap.co.uk that show exactly where a postcode is situated which has caused difficulty in one case because the PO Box address clearly states the correct postcode for the property.

4.2.2 Group Two

People were unsure of what information could be gleaned about them through the Internet. Mostly they were relieved that there was not more available about them and participating in the study seemed to allay their fears. The majority of the people interviewed were not concerned about the information found.

The date of birth was found to be the key to finding out more public information, for example mothers maiden name and access to certificates. Some of the social networks and genealogical sites helped to find this information, but on the whole this was unobtainable. Concern was raised about finding the mother's maiden name easily, the possibility of identity theft and the use of jargon throughout the Internet. Two people expressed concern that the Internet enabled an intrusive society. One person felt it necessary to take action by ringing their bank and changing their identifying data from their mother's maiden name to something less likely to be discovered.

Public records caused concern where the electoral roll could be combined with the Land Registry to calculate that a woman lived on her own. This made the person concerned feel vulnerable to attack or burglary.

4.2.3 Group Three

Only 32 responses were received for the third group of people. The average age was 17 with the youngest being 15 and the oldest being 25. 17 of those people had put photographs on the web. Overall, most people were happy to make their email address public, but address and phone number depended on who it was that it would be given to. Overwhelmingly, 30 people were happy to have their gender known. 19 people did not regret putting personal information into their profile.

5. Semantic Web

> "*Semantic Web* – a web of data that can be processed directly or indirectly by machines" (p191 Berners-Lee with Fischetti, 2000)

The Semantic Web is a way of exposing data through common markup languages and ontologies to increase the usefulness and range of uses that the data can be put to. This gives machines the ability to process the data and to glean the context by using ontologies to reason with. Protocols and standards are layered with trust at the top of the stack.

The Oxford English Dictionary defines trust as a belief in the reliability, or truth of a matter. Trading and e-commerce needs trust to function, along with non-repudiation, authentication and verification.

Guerra (2003) proposes that privacy plays a role in trust when it is combined with identity and security in the e-commerce context. However, the current focus of trust for the Semantic Web is towards trustworthiness and credence given to data on the Internet (Golbeck et al, 2003).

Other work concentrates on the manipulation and representation of privacy policies (W3C, 2005), or making links between individuals to determine whether to place trust in them or not (Foaf, 2006).

These approaches presuppose that an individual has a valid choice, that they can make use of a different website or service, or that they had an explicit choice in being able to release their information. There is nothing here that stops a company changing their policy, or an individual not having any choice but to fill in the information, or even to acknowledge that somebody else has released their information without their consent or knowledge.

What is seen here in the emerging design of the Semantic Web is that privacy together with the use and control of personal information is not a specific or integral part of the semantic web infrastructure, but as a sideline to trust. This leads to the possibility that privacy issues may be overlooked, or not properly implemented.

In the past there was a rudimentary protection for privacy because information was held in many different places. There was a high cost involved in combining that information in terms of time, effort and money. Now the Semantic Web is proposing that all pieces of information whether they are pictures, sounds or text could be combined. Information held in an unrecognisable format is no longer a defence with the move towards interoperability. Software agents will infer, reason and combine the information they find and as yet, no indication of the implications for privacy is given.

This would appear to confirm the view taken that technology not only ignores personal privacy, but also puts individuals at risk without giving them the ability to do anything about it (Solove, 2003; Garfinkel, 2000)

6. Conclusion

The findings above have illustrated different areas where technology has caused vulnerability. The ease with which information is obtained has created a situation where people are anxious about the technology and feel powerless to do anything about it. The examples are the location based tracking on mobile phones; postcodes being shown on web sites and inferences being made about single females.

The teenage respondents were willing to share information easily and felt their context protected them, for example signing up to the social networking sites such as www.myspace.com, www.bebo.com, www.faceparty.co.uk allowed them to post personal details about themselves. However, not all of these sites had guarantees about who could access them and so perhaps the implications of who could obtain the personal information were not completely explored.

Combining these vulnerabilities through shared data and reasoning tools causes great concern. The Semantic Web has the potential to create a bigger vulnerability in achieving this combination. However, there is merit to be had in addressing this issue face on, to hand control back to the individual, for them to be able to mitigate the risks for themselves.

The next phase of the research is to create a Semantic Web tool for use by the individual to protect their own privacy. The prototype will be designed, implemented and evaluated to explore some of the issues raised earlier in this paper.

The objectives for the Semantic Web application will be as follows:

- to discover how to protect personal information,
- to interact with the user to provide an understanding of the implications of the release of the personal data
- to discover for the user how the information about to be given may be combined with other information found.

7. References

Atkinson, T, (2004), Technology Safety Project, *Washington State Coalition Against Domestic Violence*, www.wscadv.org/projects/Tech/index.htm (Accessed 23 November 2005)

Berners-Lee, T, 2000, "Semantic Web on XML", www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html, (Accessed 23 November 2005)

Berners-Lee, T. with Fischetti, M. 2000 *Weaving the Web*. Texere Publishing. London.

Blunkett, D, (2004), "ID card pilot scheme under way", 10 Downing Street, 26 April 2004 www.number-10.gov.uk/output/page5701.asp, (Accessed 23 November 2005)

Byers, S., Cranor, L., Kormann D and McDaniel, P, (2004), "Searching for Privacy: Design and Implementation of a P3P-Enabled Search Engine", In *Proceedings of the 2004 Workshop on Privacy Enhancing Technologies (PET 2004)*, Toronto, Canada

Carins, R, (2005), "Credit Checks and Your Job", Fresh Finance, www.freshfinance.net/articles-creditchecksjob.htm , (Accessed 23 November 2005)

Colvin, M, (2002), *Developing Key Privacy Rights*, Hart Publishing, Oregon

Cranor, L.F. and Garfinkel, S, (2005), *Security and Usability*, O'Reilly, USA

Davies, S, (2000), "The Death of Privacy: A Personal View", BBC, Video Cassette

Dinev, T. and Hart, P, (2004), "Internet Privacy Concerns and their Antecedents - Measurement Validity and a Regression Model", *Behaviour and Information Technology*, Volume 23, Issue 6, November 2004 pages 413-422

Foaf, (2006), "The Friend of a Friend Project", http://www.foaf-project.org/ (Accessed 16 February 2006)

HiSPEC, (2002), "Privacy Enhancing Technologies State of the Art Review", www.hispec.org.uk/public_documents/7_1PETreview3.pdf (Accessed 23 November 2005)

Garfinkel, S, (2000), *Database Nation*, O'Reilly Associates, Sebastopol, CA

Golbeck, J., Parsia, B., Hendler, J, (2003), "Trust Networks on the Semantic Web", In *Proceedings of Cooperative Intelligent Agents 2003*, Helsinki, Finland, August 2003

Guerra, A.G., Zizzo, D.J., Dutton, W.H and Peltu, M, (2003), "Economics of Trust in the Information Economy: Issues of Identity, Privacy and Security", *Oxford Internet Institute, Research Report No 1*, April 2003

Katyal, S, (2004), "Privacy vs. Piracy", *Yale Journal of Law & Technology*, Vol 7. p222

Lecomte, J., Clarke, N.M., Furnell. S.M, (2005), "Artificial Imposter Profiling for Keystroke Analysis on a Mobile Handset", In *Proceedings of 5th International Network Conference*, University of the Aegean and University of Plymouth

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 , (Accessed 23 November 2005)

Margulis, S.T, (1977), “Conceptions of Privacy: Current Status and Next Steps”, *Journal of Social Issues*, 33. 5-10

Microsoft , (2005), “Hector the Protector”, http://www.microsoft.com/nz/athome/security/children/hector.mspx, (Accessed 23 November 2005)

Oppliger, R, (2005), “Privacy-enhancing technologies for the world wide web”, *Computer Communications*, 28 (16) pp1791-1797

Power, M, (2006), “How to catch a cheating partner”, *The Independent*, 17th February, 2006

Phillips, D.J, (2004), “Privacy Policy and PETs,”, *New Media and Society*, 6 (6) pp691-706

Raab, C.D and Bennett, C.J, (1998), “Distribution of Privacy Risks: Who Needs Protection”, *Information Society*, Vol 14, Issue 4, pp263-274

Schwartz, P. M., (1999), “Privacy and Democracy in Cyberspace”, *Social Science Research Network*, www.paulschwartz.net/pdf/VAND-SCHWARTZ.pdf, (Accessed 23 November 2005)

Security Focus, (2005), “Hidden DRM code's legitimacy questioned”, *The Register*, 3 November 2005 www.theregister.co.uk/2005/11/03/secfocus_drm/ , (Accessed 23 November 2005)

Solove, D.J, (2003), “Identity Theft, Privacy, and the Architecture of Vulnerability”, *Hastings Law Journal*, Vol 54 1227, 1232

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

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 , (Accessed 23 November 2005)

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

W3C, (2005), “Platform for Privacy Preferences (P3P) Project”, http://www.w3.org/P3P/, (Accessed 16 February 2006)

Warren, S and Brandeis, L, (1890), “The Right to Privacy”, Quoted in Developing Key Privacy Rights, Colvin, M., Harvard Law Review, 4 (1890), 193

Westin, A. F., (2003), “Social and Political Dimensions of Privacy”, *Journal of Social Issues*, Vol 59, Issue 2, pp431-453

Womens Aid Federation of England, (2002), “Domestic Violence Statistical Factsheet 2002”, www.womensaid.org.uk/dv/dvfactsh2002.htm , (Accessed 23 November 2005)