Internet users and online privacy: a study assessing whether Internet users' privacy is adequately protected.

Vasanti Patel[1]
Radmila Juric[1,2]

[1] South Bank University Business School
[2] Radmila Juric now works within the School of Informatics, University of Westminster

Internet Users and Online Privacy
A Study Assessing Whether Internet Users' Privacy is Adequately Protected

Vasanti Patel* and Radmila Juric**
South Bank University Business School, 103 Borough Road, London SE1 OAA, UK
Tel +44 (0)20 7815-7888, Fax +44 (0)20 7815-7793
*Email: vasanti_p@hotmail.com **E-mail juricr@sbu.ac.uk

Abstract

This paper examines the current state of Internet privacy. We assess the needs of UK Internet users in terms of online privacy protection, and determine the extent to which current privacy practices were satisfying those needs. Our work examines: (a) Internet users' attitudes towards online privacy; (b) 50 Web sites' privacy policies and practices and (c) existing privacy protection for users such as legislation and technological tools. The survey reveals a high level of concerns amongst Internet users related to their privacy in terms of (i) personally identifying information that they provide to Web sites, (ii) the information that Web sites collect through the use of cookies and IP addresses and (iii) the information derived by tracking users' on-line activities.

Keywords: Privacy on the Internet, Privacy policies, Internet interactions

1. Introduction

The World Wide Web has become an immensely popular and powerful medium in recent years. The role and importance of Internet technology is significant in all aspects of our lives. There are no geographic, political, social or racial boundaries, which enables ordinary individuals, professionals, and businesses to use this technology. To attract more users, many Web-sites offer personalised services where users are supposed to identify themselves and register their intention while using the Web-site. This means that this same technology that allows users to widely access and share information, may also violate users' privacy. Privacy was a sensitive issue far before the advent of computers. However, concerns have been magnified by the existence of large computer databases that compile individual data about different users and store them in many different forms [1]. Furthermore, if we take on how easy it now is to collect relevant information about users from the WWW and store them in such databases [2], we understand why Internet users are concerned about their privacy not only in terms of personal data, but information that Web sites may derive by tracking their on-line activities [3]. Many studies have been conducted (see section 5) which routinely report that privacy on the Internet is of great concern for everyone: the Internet is making it easier for businesses to collect more and more information from their customers than ever before.

The most common approaches to protecting privacy on the Internet is through **legislation/laws** such as the Data Protection Act [4] and technological **tools** that hide the identity of a user when data is being transmitted [5,6,7,8]. However, are these approaches adequate for protecting users' online privacy? In order to answer the question, we assess the needs of UK Internet users in terms of online privacy protection, and determine the extent to which current privacy practices were satisfying those needs. This was done through examining:

(a) Internet users' attitudes towards online privacy;
(b) the privacy policies and practices of Web-sites;
(c) ways of protecting users' privacy on the Internet.

The paper is organised as follows. The first part of our work, presented in section 2, focuses on Internet users' attitudes towards online privacy and was conducted through questionnaires distributed randomly to various participants/Internet users in North London areas, which resulted in 50 questionnaires returned and completed in late Summer 2000. Questions were designed to provide an insight into what aspects of the Internet users are most concerned about. The sample is certainly not large enough to statistically represent all UK Internet users, but it covers a population that we find in our close neighbourhoods, people that we meet every day on our streets when commuting to work, people that share a similar business, social and cultural environment. In section 3 we reported on a survey of 50 of the top 100 Web sites conducted in the summer

2000. The Web sites were randomly selected from a list provided by www.100hot.com. We applied the same set of questions to all chosen Web sites in order to examine their privacy policies and practices. In section 4 we focus on the offered online privacy protection through legislation and technological tools. We exclude issues related to Internet security as it is outside the scope of this paper. In section 5 we reflect on recent studies of online privacy, and our conclusions are in section 6.

2. Internet Users' Online Privacy

2.1. Internet Privacy and Business Behaviour

There are many different understandings and usage of the term 'Internet Privacy'. However, many definitions are centered on *'the right to be free from intrusion and interruption as well as the right to have control over one's personal information'* [9] which suggests that a user's personal privacy would be violated when that person was not specifically notified that their actions were not private. Thus every time a Web site collects information without a user's knowledge or consent, they are violating a user's privacy. The Internet makes it easier for businesses to 'generate, access, manipulate, and store information' than ever before [10]. However, information that is of a very high value to businesses can be now collected at a very low cost thus businesses are likely to take advantage of collecting data through the use of the Internet. It can be very alarming for a user to find that their online activities are continuously being monitored without their knowledge or consent. However some businesses do not seem to show any concern over a user's privacy. The article from [11] reported that the head of Sun Microsystems CEO, Scott McNealy, stated "you have zero privacy. Get over it". Many people would regard Sun Microsystems as a trusted organization and thus statements like this are really discouraging.

2.2. How is Personal Information Collected Online and How Can It Be Used?

Personal information including hobbies, interests, preferences and even ways in which an user can be contacted such as an e-mail address or home address can be collected when a person is online. There are 3 main functions of the Internet that can enable the collection of information online from users:

Internet Protocol Address: through TCP/IP reveals the location and the software/hardware of the computer being used. The Web site from [12] provides a demonstration on how and what information are captured about users when they connect to the Internet.

Cookies: as 'a unique identifier that a Web server places on the users computer'[13] stores information on the hard drive about the sites that have been visited by the user. In [14] we find that "Web servers passively record transactional information in order to maintain the system". However cookies are also used for tracking users' online activities: which Web pages are visited and how long the user stayed at a particular Web page.

Registering Online: can enable a user to get access to the Web-sites special services, which are not possible without being registered. However registration means providing personal information such as full name, home address, telephone number, e-mail address, etc. There may be a registration fee, which means that credit card details may have to be provided, hence a detailed profile of users and their online activities is easily created.

When we use the Internet, most of the information collected is used for purposes which are necessary for the functioning and maintenance of the businesses and their interactions with customers. However, businesses, governments and even criminals can use this information for very different purposes. For example, by monitoring users online activities, it is possible to find out personal information such as hobbies, interests, and preferences, send users unsolicited mail and/or categorize or stereotype users according how repeatedly they visit these Web sites. Thus, 'a single piece of information about a user can support a tremendous range of activities' [14].

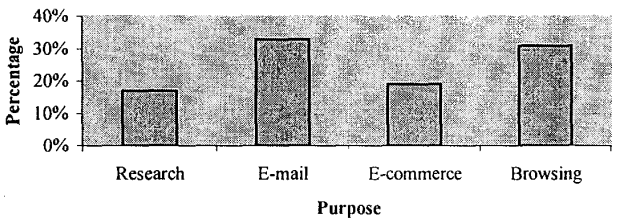
2.3. Internet User's Attitude towards Online Privacy

The Internet users' attitude towards online privacy is examined through the findings from questionnaires that were distributed randomly in the residential area of North London.. We received 51 replies to our questionnaire (28% questionnaires were returned) and all respondents were confronted with questions on

(a) the level of comfort when providing data to Web-sites,
(b) factors that are important for users when providing data,
(c) factors concerning a company's privacy practices on the Web,
(d) the use of unique identifiers and
(e) the invasion of personal privacy through the Web.

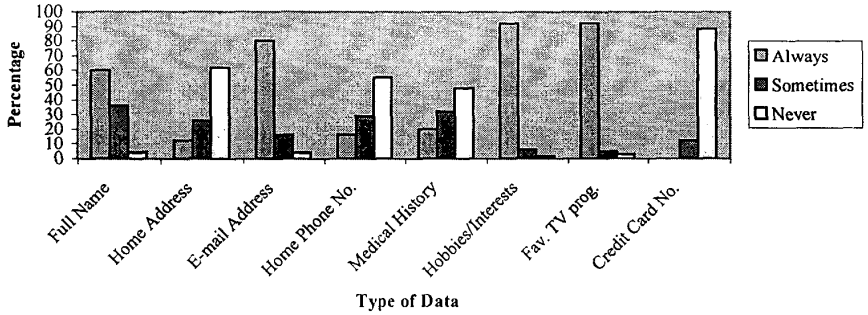
Our respondents were heavy Internet users with 68% of them using the Internet every day. However, for the majority of them, sending, receiving e-mail and browsing the WWW was the main usage of the Internet.

Figure 1: The main purpose for using the Internet



(a) The Level of Comfort when providing the various types of information varied. The majority of respondents always felt comfortable when providing information about their own interests and preferences like hobbies and favourite television programme (see Figure 2). The comfort level in providing an e-mail address and full name was also very high: 80% of respondents always felt comfortable providing their e-mail address and 60% always felt comfortable providing their full name. Revealing their own interests and preferences can not personally identify any user. Home and e-mail addresses can both be used as a means of contacting, but an e-mail address does not always directly identify any person hence they may remain anonymous on the Web. This also suggests that users find it is easier to deal with a violation of privacy if it is through their e-mail addresses: it is easier to change an e-mail address than it is to change a home address. A few respondents stated that "they always felt comfortable" when providing information about their health or medical history (20%), home telephone number (16%), or home address (12%). ***None of the respondents stated that they felt comfortable when providing their credit card number***. This means that we still build a mistrusted relationship between Internet businesses and Internet users.

Figure 2: Internet users' comfort level when providing various types of data



(b) Factors that are Important when Providing Data are given in Table 1 which shows that the majority of respondents rated ***"whether the site will share their personal information with other companies"*** as the most important factor. Another important factor was "***if users can inspect their personal information that has been stored by a Web site***". However, a disappointingly low number of respondents was interested in whether a Web-site has a ***"privacy policy"***, which contradicts two previous findings and suggests that *Internet users may not be aware of the existence of privacy policies on Web sites*. There is also increasing publicity surrounding children's privacy on the Internet (see ***"parental consent"*** in Table 1), which may

have raised awareness about the dangers of children providing information to Web sites (publicity can determine a user's attitude towards certain issues!).

Table 1: Attitudes of Internet users Towards the Policies of Web Sites

Factors	Very Important	Quite Important	Not Important
Privacy policy	20	36	44
Access to data	68	24	8
Sharing data	76	12	12
Parental consent	48	28	24

(c) Factors Concerning a Company's Privacy Practices were collected through an open-ended question, which revealed that the 3 most popular factors for users were

- not collecting information from users without their knowledge,
- not being identified unless it was necessary and
- to know the purpose of Web sites collecting information.

This back up the previous findings indicating that users prefer not to be personally identified and that they require some degree of control over the use of the data they provide.

(d) The Use of Unique Identifiers revealed that 48% of respondents were concerned about *"cookies"*, but 36% said that they did not know what a *cookie* was. This was an unexpectedly high number if we bear in mind how important it is for users that Web-sites do not collect information without users' knowledge or consent. This suggests that many users are unaware of methods which Web sites can collect information through the use of *cookies* and what information in general users are making available to Web sites.

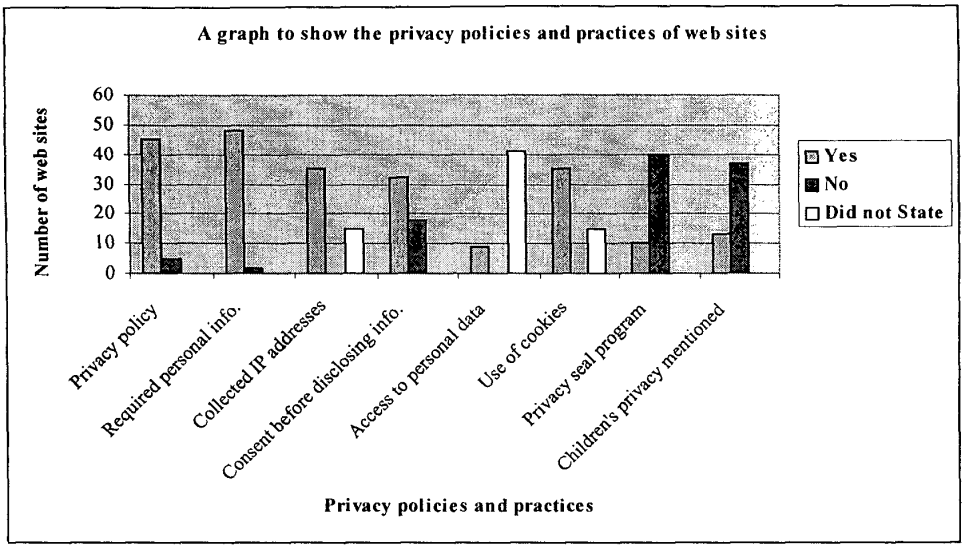
(e) Invasion of Personal Privacy revealed that 80% of respondents were very concerned about it because of uncertainty and mistrusting the Internet when operating with their personal information.

3. Privacy Policies and Practices of Web Sites

In this section we focus on 50 Web sites chosen from www.100hot.com which "filter out the seldom-visited sites and identify and rank the top 100 Web sites in such categories as technology, entertainment, finance, lifestyle, games, sports and news". We examine their privacy policies and practices through the following:

a) whether privacy policies are easy and quick to find
b) how information was collected from Internet users and used
c) whether it was possible for users to have access to their own data
d) how many sites made use of a unique identifier;
e) how many sites were members of a privacy seal program and
f) how many Web sites were concerned about children's privacy.

Figure 3 below shows the result of applying of a)-f) above to chosen Web sites. 90% of all Web-sites did ***have privacy policies*** that were relatively easy to find. However, they were difficult to understand and full of contradictory statements concerning the secondary use of data they collect. This gives Internet users ***an illusion and not assurance*** that their information and privacy are protected.



Web sites may ***collect personal information*** directly from the user on a voluntary basis through registration forms, surveys, competitions, etc and use this information mainly to fulfil other requirements and to respond to inquiries. 97% of examined Web sites do so, but often require users to provide personal information ***before they use all or some of the sites' services.*** For example the Sony web-site does not let users participate in some activities if they do not provide personal information and Yahoo does not provide a user with an email account if personal information is not provided. However, if a conventional store does not need to collect information from their customers when they visit and browse the store, then the same store- if it is online - should not feel the need to either.

70% of Web sites stated that they ***collect IP addresses.*** This is done before a user even has a chance to read a site's privacy policy. The reason for collecting IPs for 60% of them is "to perform routine system maintenance". There were some sites such as Sony, Etoys, IBM, CNN and Adforce that specifically stated that "personal information would not be collected without a user's consent or knowledge". However, it was difficult for us to determine whether a Web site collected personally identifiable information indirectly if it was not specifically stated by the Web-site.

We were also interested in the extent to which users are able to ***restrict Web sites from sharing, selling or renting their personal information to third parties.*** 62% of Web sites stated that they do not disclose any information without a user's consent. 22% stated that even with a user's consent, only aggregate information is shared. However most of the sites were not clear and often contradicted themselves about whether information was going to be disclosed. For example the Disney, Amazon and Realtor Web sites stated that they do not disclose information without a user's consent but then went on to state that *information will always be shared with their partner companies.* This implies that under some circumstances Web sites will be disclosing personal information without users' consent. However because the privacy policy has been made difficult to understand many users provide data without having an accurate idea of the secondary use of their personal data. Only a minority of the Web sites stated that users have the option of ***removing their names from the sites'*** marketing list (opt-out), but most of them stated that in order to opt-out, users must write to the company to notify them, hence making the whole process difficult. Web sites realize that users are more likely to opt-out if they can do so online rather than having to write to the company. Nearly all of the Web sites stated that they have no control over ***how third parties collect information,*** and will not be held responsible for the actions of third parties.

Only 18% of Web sites specifically stated that the ***user is able to access their personal data*** and can update their preferences. This suggests that if users are easily able to get access to their data in order to update their personal profiles, most of this information that is highly valuable to a business might be erased. Thus, if Web sites do not specifically state that users can have access to their personal data, users may forget the fact that they have the legal right to do so.

70% of Web sites surveyed stated in their policy that they ***enable cookies***. However only about half of these sites explained what a cookie was and how it was used. The majority of the sites stated that information would only be used to improve the sites' services. The Dash Web-site stated that if the user does not accept cookies, they could not use the services. This indicates that although companies state that any information they collect through the use of cookies will not identify the user, they may combine data with other personal information which can result in companies obtaining a detailed profile of users without their consent.

Only 20% of Web sites stated that it was a member of **TRUSTe** [8]. 26% of sites did acknowledge that children would be using their Web sites and that they should consult their parents or guardian before using the sites' services. However only 10% of them, such as Sony and Headbone stated that ***children can not use some services*** or will not accept any information without verifiable consent from parents. When analysing children's privacy it was important to take into account that not all the sites that were surveyed were targeted at children. Some sites such as Charles Schwab and Bloomberg which are stockbrokers only target those who are over 18 and thus will not mention children's privacy

4. Online Privacy Protection

Various technology tools, regulatory and self-regulatory frameworks, laws and industry guidelines can work together and help to protect users' privacy. The Data Protection Act of 1984 and technology tools are a major backbone. We relate the findings from chapters 2 and 3 to the legislation and technological tools to see if current users' privacy protection is adequate.

4.1. Data Protection Act 1984

Although the Data Protection Act is the first major piece of legislation concerned exclusively with data protection, it is not the only law concerned with the handling of information, using the term in its widest sense [4]. We list the seven data protection principles and observe their presence within examined Web-sites:

'The information to be contained in personal data shall be obtained, and personal data shall be processed, fairly and lawfully'

It is unfair to collect data from users for one stated purpose and then to use that data for a completely different reason unless it has been clearly indicated. Collecting customer information for the purpose of fulfilling an order requirement and subsequently using it for marketing purposes which is against the interests of the customer would be regarded as unfair. Do businesses clearly state the purpose, for which data is collected? If data is going to be used for a secondary purpose is this clearly indicated on a Web-site?

'Personal data shall be held only for one or more specified and lawful purposes'

This means that any personal information that is provided to a Web site will not be used for an illegal purpose. How do we interpret the fact that Internet users do not feel **at all** comfortable when providing sensitive data? Is this solely because of 'security issues'?

'Personal data held for any purpose or purposes shall not be disclosed in any manner incompatible with that purpose or purposes'

If a business does intend to disclose data to a third party, this must be stated on the Web site hence users decide whether or not to provide personal information. Are Web sites clear about disclosures?

'Personal data held for any purpose or purposes shall be adequate, relevant and not excessive in relation to that purpose or purposes'

How often businesses collect data only because "it may be useful in the future"?

'Personal data shall be accurate and, where necessary, kept up-to-date'

Changes in collected personal data must be made in order to keep information updated and therefore accurate. Why then are businesses still making it difficult for users to get access to their personal data?

'A user shall be entitled –

(a) at reasonable intervals and without undue delay or expense –

(i) to be informed by any data user whether he holds personal data of which that user is the subject; and

(ii) to access any such data held by a data user; and

(b) where appropriate, to have such data corrected or erased'

Are businesses making it obvious to users that they have the ability to have access to their own personal data and update their preferences?

'*Appropriate security measures shall be taken against unauthorised access to, or alteration, disclosure or destruction of, personal data and against accidental loss or damage''.*

Businesses should take reasonable care over the security of personal data. In [15] the National Computing Center (NCC) reported that 'half of all businesses are risking prosecution because they do not have an information security policy'.

4.2. Technological Tools

A number of tools have been developed to help Internet users surf the Web anonymously. They mainly use anonymizing agents, which ensure that requests to a Web-site can not be linked to an IP address from which a user can be identified. *Crowds* anonymity agent [5] is based on the idea that people can be anonymous if they blend into a crowd, i.e. *Crowd* users submit their request through a group of Web- surfers running the *Crowd* software. In *Onion Routing* [6] users submit encrypted HTTP requests using a layered data structure that specifies symmetric cryptographic algorithms and keys to be used as data is transported to the intended recipient. In *Lucent Personalised Web Assistant* [7] a pseudonym agent helps users to build a persistent but anonymous relationship with Web-sites. Th *Platform for Privacy Preferences Project* (P3P) [16] provides a rich vocabulary for services to express user's privacy preferences and help users to make informed decisions about when and how to release their private data. However, users must be assured that when they release their data, services will use it only as they have promised. *TRUSTe* [8] is a self regulatory privacy initiative dedicated to building consumers trust and confidence on the Internet through a program in which Web-sites can be licensed to display a privacy seal or trustmark on their sites.

5. Other Studies on Internet Privacy

Finally, we briefly discuss some recent studies on Internet privacy from the US, to see how their findings overlap with the results of our survey.

A study undertaken by Pew Internet & American Life Project [17] found that although consumers are very concerned about their privacy on the Internet, many users are still not aware of exactly how dangerous it can be to share personal data. Moreover, they are even not unaware that their personal data is being shared. Susannah Fox, their research director stated that, "84 per cent of Internet users are concerned about businesses or strangers getting personal information about them or their family, and yet 56 per cent don't know that Web sites and advertisers can track a user's activities by placing cookies". Junkbusters .com [18] held a workshop recently where they found that companies believe that it is not their problem if consumers are not better educated about the types of information that is held about them. They state that it is obvious that these companies are trying to avoid informing users exactly what information is stored about them because 'if consumers could see half the information kept by companies, many would be horrified and demand its destruction'.

Consumers International [19] conducted a study on 751 sites and found that 'over two-thirds of these sites collected some sort of personal information about their visitors'. The majority of these sites asked visitors for information that can easily identify the individual. Furthermore, most of these sites did not make it possible for users to choose whether or not they wanted to be on a mailing list or to have their information shared with third parties. The majority of these sites were not good at explaining to users how they were going to be using their personal data. Thus, since users are not being educated, they remain ignorant as to how their personal data will be used.

Many high-profile businesses use P3P, and even Microsoft is integrating P3P in the upcoming release of Internet Explorer. However privacy advocates say that the P3P system is ineffective and does not protect the privacy rights of individuals [20]. Furthermore, a report conducted by the Electronic Privacy Information Centre found that 'P3P fails to comply with baseline standards for privacy protection'. The report recommends that privacy standards should be built on Fair Information Practices. Simple rules on how personal information will be collected and used will also help to gain consumer trust [21].

President George W. Bush has decided not to use the email to communicate, at least while he is in the White House. He has discovered what many Internet users already fear, and that is a loss of privacy [22].

Finally a study conducted by Buchanan Email Ltd. [23] found that 'the majority of British Web sites do not have an email management strategy in place'. Currently the email culture is one where the consumer provides a lot of personal information to Web sites, but receives a minimal amount of feedback from them.

6. Conclusions

We are still witnessing a mistrusted relationship between Internet businesses and Internet users. We need to be able to trust the Web when providing sensitive data and without this trust users will not feel comfortable when conducting business transactions via the Internet. However, when sensitive information is exchanged on the Internet, it is not enough that we are "just concerned with security issues": encryption technology may protect data to a great extent. We have to address our on-line privacy and make sure that it is protected better in future than it is now. We need strong security that bonds together tools and legislation. Our survey revealed that neither are users satisfied at how their privacy is compromised during Web interactions, nor are Web-sites building trustworthy relationships with their users. It seems that we have a long way to go if we want to protect users' on-line privacy and enforce their anonymity in Web interactions.

References

[1] Cranor, L. (1998) "Internet Privacy: A Public Concern", *netWorker 2*, pp 13-18

[2] Chan, D. (1981) "Untraceable electronic mail, return addresses and digital pseudonyms," CACM 24(2), February 1982, pp.84-88

[3] Wang, H., Lee M., and Wang, C (1998) "Consumer privacy concerns about Internet Marketing", CACM 41(3), March 1998, pp 63-70

[4] Savage N., and Edwards C. (1985) "A Guide to the Data Protection Act – Implementing the Act", London Financial Trading Publications (2nd edtition).

[5] Reiter M.K., and Rubin A.D.,(1999) "Anonymous Web Transactions with Crowds", CACM 42(2), February 1998, pp 32-38.

[6] Goldschlag D., read M., and Syverson P. (1999) "Onion Routing for Anonymous and Private Internet Connections", CACM 42(2), February 1998, pp 39-41.

[7] Gabber E., Gibbons P.B., Kristol D.M., Matias Y. and Mayer A. (1999)"Consistent, yet Anonymous, Web Access with LPWA", CACM 42(2), February 1998, pp 42-47.

[8] Benassi P. (1999) TRUSTe: An Online Privacy Seal Program, CACM 42(2), February 1998, pp 56-59.

[9] http://www.media-awareness.ca/eng/issues/priv/privacy.htm

[10] Cate F.H. (1997) Privacy in the Information Age, Brookings Institution Press, Washington DC.

[11] Radcliff D. (1999) " A Cry for Privacy", Computerworld, 17-May-1999, pp 46-47

[12] Consumer.net (1998) http://privacy.net/analyze

[13] [http://www.junkbusters.com/nt/en/cookies.html]

[14] .The Centre for Democracy and Technology http://cdt.org/privacy/guide/basic/totpen.html

[15] Campos G. (2000) "Companies Risk fines over Data Security", Computer weekly, 17-February 2000, page 16

[16] Reagle J., and Cranor L.F. (1999) "The Platform for Privacy Preferences", CACM 42(2), February 1998, pp 48-55

[17] http://www.crmdaily.com/perl/story/?id=6876

[18] http://junkbusters.com/profiling.html

[19] http://www.nua.ie/surveys/?f=VS&art_id=905356395&rel=true

[20] http://www.privacy.org/article.php?sid=671

[21] http://www.epic.org/Reports/prettypoorprivacy.html

[22] http://www.newsfactor.com/perl/story/6892.html

[23] http://www.nua.ie/surveys/?f=VS&art_id=905354656&rel=true