2015

Hidden online surveillance: What librarians should know to protect their privacy and that of their patrons

Alexandre Fortier

Jacquelyn Burkell
The University of Western Ontario, jburkell@uwo.ca

Citation of this paper:

Fortier, A. and Burkell, J. (2015) Hidden online surveillance: What librarians should know to protect their privacy and that of their patrons. Information Technology and Libraries, 32(3), 59-72.

Hidden Online Surveillance: What Librarians Should Know to Protect Their Own Privacy and That of Their Patrons

Alexandre Fortier
and
Jacquelyn Burkell

ABSTRACT

Librarians have a professional responsibility to protect the right to access information free from surveillance. This right is at risk from a new and increasing threat: the collection and use of non-personally identifying information such as IP addresses through online behavioral tracking. This paper provides an overview of behavioral tracking, identifying the risks and benefits, describes the mechanisms used to track this information, and offers strategies that can be used to identify and limit behavioral tracking. We argue that this knowledge is critical for librarians in two interconnected ways. First, librarians should be evaluating recommended websites with respect to behavioral tracking practices to help protect patron privacy; second, they should be providing digital literacy education about behavioral tracking to empower patrons to protect their own privacy online.

INTRODUCTION

Privacy is important to librarians. The American Library Association Code of Ethics (2008) states that "we protect each library user's right to privacy and confidentiality with respect to information sought or received and resources consulted, borrowed, acquired or transmitted," while the Canadian Library Association Code of Ethics (1976) states that members have responsibility to "protect the privacy and dignity of library users and staff." This translates to a core professional commitment: according to the American Library Association (2014, under "Why Libraries?"), "librarians feel a professional responsibility to protect the right to search for information free from surveillance."

Increasingly, information searches are conducted online, and as a result librarians should be paying specific attention to online surveillance in their efforts to satisfy their privacy-related professional responsibility. This is particularly important given the current environment of significant and increasing threat to privacy in the online context. Although many concerns about online privacy relate to the collection, use, and sharing of personally identifiable information, there is increasing awareness of the risks associated with the collection and use of what has been termed 'non-personally identifiable information' (e.g.: Internet Protocol addresses, pages visited, geographic location information, search strings, etc.; Office of the Privacy Commissioner of Canada

Alexandre Fortier (afortie@uwo.ca) is a PhD candidate and Lecturer, Faculty of Information and Media Studies, The University of Western Ontario, London, Ontario. **Jacquelyn Burkell** (jburkell@uwo.ca) is Associate Professor, Faculty of Information and Media Studies, The University of Western Ontario, London, Ontario.

2011, 12). This practice has been termed 'behavioral tracking', and recent revelations of government security agency collection of user metadata (Ball 2013; Weston, Greenwald and Gallager 2014) have heightened awareness of this issue. The problem, however, is not new, nor is the practice restricted to the actions of governmental agencies. Indeed, as early as 1996 commercial and non-commercial entities were practicing online behavioral tracking for purposes of website and interaction personalization and to present targeted advertising ("Affinicast unveils personalization tool" 1996; "AdOne Classified Network and ClickOver announce strategic alliance" 1997). Since these initial forays into behavioral tracking and personalization of online content the practice has proliferated, and many sites now use a variety of behavioral tracking tools to enhance user experience and deliver targeted advertisements (see, e.g., the "What they know" series from the Wall Street Journal 2010; Gomez, Pinnick and Soltani 2009; Soltani et al. 2009).

There can be no question that behavioral tracking is a form of surveillance (Castelluccia and Narayanan 2012), and the ubiquity of this practice means that users are regularly subject to this type of surveillance when they access online resources. In order to satisfy a professional commitment to support information access free from surveillance, librarians must therefore address two related issues: first, they must ensure that the resources they recommend are privacy-respecting in that those resources engage in little if any online surveillance; second, they must raise the digital literacy of their patrons with respect to online privacy, increasing understanding of online tracking mechanisms and the strategies that patrons can use to protect their privacy in light of these activities.

Addressing the first issue requires that librarians attend to surveillance practices when recommending online information resources. Privacy and surveillance issues, however, are notably absent from common guidelines for evaluating web resources (see, e.g., Kapoun 1998; University of California, Berkley 2012; John Hopkins University 2013), and thus librarians do not have the guidance they need to ensure that the resources they recommend are privacy-respecting. It is critical that librarians and other information professionals address this gap by developing an understanding of the surveillance mechanisms used by websites and the strategies that can be deployed to identify and even nullify these mechanisms. This same understanding is necessary to address the second goal of raising the privacy-related digital literacy of patrons. Librarians must understand tracking mechanisms and potential responses in order to integrate privacy literacy into library digital literacy initiatives that are central to the mission of libraries (American Library Association 2013).

This paper provides an introduction to behavioral tracking mechanisms and responses. The goals of this paper are to provide an overview of the risks and benefits associated with online behavioral tracking, to discuss the various surveillance mechanisms that are used to track user behavior, and to provide strategies for identifying and limiting online behavioral tracking. We have elsewhere published analyses of behavioral tracking practices on websites recommended by information professionals (Burkell and Fortier 2015), and on practices with respect to the disclosure of tracking mechanisms (Burkell and Fortier 2015). This paper serves as an adjunct to

those empirical results, providing information professionals with background that will assist them in negotiating, on the part of themselves and their patrons, the complex territory of online privacy.

Consumer attitudes toward behavioral tracking

Survey data suggest that consumers are, in general, aware of behavioral tracking practices. The 2013 US Consumer Data Privacy Study (TRUSTe 2013), for example, reveals that 80 percent of users are aware of online behavioral tracking on their desktop devices, while slightly under 70 percent are aware of tracking on mobile devices (see also Office of the Privacy Commissioner of Canada 2013). Awareness, however, does not directly translate to understanding, and recent data indicate that even relatively sophisticated Internet users are not fully informed about behavioral tracking practices (McDonald and Cranor 2010; Smit et al. 2014). Moreover, attitudes about tracking are at best ambivalent (Ur et al. 2012), and many studies indicate a predominantly negative reaction to these practices (Turow et al. 2009; McDonald and Cranor 2010; TRUSTe 2013). Although it is not universally required by regulatory frameworks, many users feel that companies should request permission before collecting behavioral tracking data (Office of the Privacy Commissioner of Canada 2013). Finally, although some users take steps to limit or even eliminate behavioral tracking, many do not. For example, while one-third to three-quarters of survey respondents indicate that they manage or refuse browser cookies (TRUSTe 2013; comScore 2007; 2011; Rainie et al. 2013), at least one quarter reported no attempts to limit behavioral tracking. This may be attributed to the difficulty in using such mechanisms (Leon et al. 2011).

Behavioral tracking and its mechanisms

Tracking mechanisms transmit non-personally identifiable information to websites for different purposes. Originally, the information collected by these mechanisms was used to enhance user experience and to make these website interactions more efficient. Tracking mechanisms can record user actions on a web page and their interaction preferences. Using these data, websites can for example direct returning visitors to a specific location in the site, allowing those visitors to resume interaction with a website at the point where they were on the previous visit. Using the Internet Protocol (IP) address of a user, websites can display information relevant to the geographic area where a user is located. Tracking mechanisms also allow a website to remember registration details and the items users have put in their shopping basket (Harding, Reed and Gray 2007).

Tracking mechanisms are also of great use to webmasters, supporting the optimization of website design. Thus, for example, these mechanisms can inform webmasters of users' movements on their websites: what pages are visited, how often they are visited, and in what order. They can also indicate the common entry and exit points for a specific website. This information can be leveraged in site redesign to increase user satisfaction and traffic.

Website optimization and interaction personalization have potential benefit to users. At the same time, however, the detailed profile of user activities, potentially aggregated across multiple visits to different websites, presents potential privacy risks. The information gathered through tracking mechanisms can allow a website to identify browsing and information access habits, to infer user characteristics including location and some demographics, and to know what topics or products are of particular interest to a user.

Tracking mechanisms can be first-party or third-party, and the difference has implications for the detail that can be assembled in the user profile. First-party mechanisms are set by directly by the website a user is visiting, while third-party mechanisms are set by outside companies providing services, such as advertising, analysis of user patterns and social media integration, on the primary site. First-party tracking mechanisms collect information about a site visit and visitor and deliver that information to the site itself. Using first-party tracking, web sites can provide personalized interaction, integrating visit and visitor information both within a single visit and across multiple visits (Randall 1997). This information is available only to the web site itself, and thus neither includes information about visits to other sites nor is accessible by other websites, unless the information is sold or leaked by the first-party site (see Narayanan 2011).

Third-party tracking mechanisms, by contrast, deliver information about a site visit and visitor to a third party. This transaction is often invisible to the user, and the information is transmitted typically without explicit user consent. Third-party tracking represents a greater menace to privacy, since third parties have a presence on multiple sites, and are able to collect information about users and their activities on all those sites and integrate that information across sites and across visits into a single detailed user profile (see Mayer and Mitchell 2012 for a discussion of privacy problems associated with third-party tracking). Research demonstrates that third-party tracking is a common and perhaps even ubiquitous practice (Gomez, Pinnick and Soltani 2009; (Burkell and Fortier 2013). It is not uncommon for websites to have trackers from more than one third party, and some websites, especially popular ones, have trackers from dozens of different organizations: Gomez, Pinnick and Soltani (2009), for example, found 100 unique web beacons on a single website. Furthermore, the same tracking companies are present on many different websites, allowing them to integrate into a single profile information about visits to each of these many sites. PrivacyChoice[1], which maintains a comprehensive database of tracking companies, estimates that *Google Display Network (Doubleclick)*, for instance, has a presence on 57 percent of websites. Thus, a user traveling the web is likely to be tracked by Doubleclick on more than half of the sites they visit, and Doubleclick has access to information about all visits to each of these many sites.

Worries about the potential privacy breaches that mechanisms for tracking a user's activities online can allow are not new. Even at their inception in the mid-1990s, HTTP cookies (also known as browser cookies) were generating controversy about the potential invasion of privacy

[1] http://www.privacychoice.org/.

(e.g. Randall 1997). Users, however, quickly realized that they could manage HTTP cookies using accessible browser settings that limit or even entirely disallow the practice of setting cookies. As a result, websites, advertisers and others who benefit from web audience segmentation and behavior analytics developed newer and more obscure tracking technologies including 'supercookies' and web beacons, and these technologies are now deployed along with HTTP cookies (Sipior, Ward and Mendoza 2011). Tracking technologies are constantly evolving in response to user behavior and advertiser demand, therefore keeping up to date is an ongoing challenge (see, e.g., Goodwin 2011).

HTTP cookies

HTTP cookies were originally meant to help web developers "invisibly" gather information about users in order to personalize and optimize user experience (Randall 1997). These cookies are simply a few lines of text shared in an HTTP transaction, and a typical cookie might include a user ID, the time of a visit, and the IP address of the computer. Cookies are associated with a specific browser, and the information is not shared between different browsers on the same machine: thus, the cookies stored by Firefox are not accessible to Internet Explorer, and vice versa. Cookies do not usually include identifying information such as name or address, and they are able to do so if and only if the user has explicitly provided this information to the website. When users want to access a web page, their browser sends a request to the server for the specific website and the server searches the hard drive for a cookie file from this site. If there is no cookie, a unique identifier code is assigned to the browser and a cookie file is saved on the hard drive. If there is a cookie, it is retrieved and the information is used to personalize and structure the website interaction (for a detailed description of the mechanics of cookies, see Kriscol 2001, 152–155).

Some HTTP cookies, called session or transient cookies, automatically expire when the browser is closed (Barth 2011). They are mainly used to keep track of what a consumer has added to a shopping cart or to allow users to navigate on a website without having to log in repeatedly. Other HTTP cookies, called permanent, persistent or stored cookies, are configured to keep track of users until the cookie reaches its expiration date, which can be set many years after creation (Barth 2011). Permanent HTTP cookies can be easily deleted using browser management tools (Sipior, Ward and Mendoza 2011). Studies have shown that approximately a third of users delete cookies once a month (e.g. comScore 2007; 2011). Such behavior, however, displeases advertisers, as it leads to an overestimation of the number of true unique visitors on a website and impede user tracking (Marshall 2005; see also comScore 2007; 2011).

Flash cookies and other 'supercookies'

To palliate this 'attack' on HTTP cookies, an online advertising company, United Virtualities, developed a backup system for cookies using the local shared object feature of Adobe's Flash Player plug-in: the persistent identification element (Sipior, Ward and Mendoza 2011). This type of storage, called Flash Player Local Shared Objects or, more commonly, Flash cookies, shares many similarities with HTTP cookies with regard to their tracking capabilities, storing similar

non-personally identifying information. Unlike HTTP cookies, however, Flash cookies do not have an expiration date, a characteristic that makes them permanent until they are manually deleted. They are also not handled by a browser, but are stored in a location accessible to different browsers and Flash widgets, which are thus all able to access the same cookie. They can hold much more data (up to 100 KB by default compared to 4 KB for HTTP cookies), and support more complex data types than HTTP cookies (see MacDonald and Cranor 2012 for a technical comparison of HTTP and Flash cookies). Moreover, it is estimated that Adobe's Flash Player is installed on over 99 percent of personal computers (Adobe 2011), making Flash cookies usable on virtually all computers.

Flash cookies represent a more resilient technology for tracking than HTTP cookies. Erasing traditional cookies within a browser does not affect Flash cookies, which needs to be erased in a separate panel (Sipior, Ward and Mendoza 2011). Flash cookies also have the ability to 'respawn' (or recreate) deleted HTTP cookies. A website using Flash cookies can therefore track users across sessions even if the user has taken reasonable steps to avoid this type of online profiling (Soltani et al. 2009), and although it is declining in incidence, this practice is still occurring, sometimes on very popular websites (Ayenson et al. 2011; MacDonald and Cranor 2012).

It should also be noted that other Internet technologies (e.g. Silverlight, JavaScript, and HTML5), which have so far attracted less attention from researchers, use local storage for similar purposes. One developer even created the 'evercookie', a very persistent cookie incorporating twelve types of storage mechanisms available in a browser that makes data persist and allows for respawning (Kamkar 2010), a method investigated by the National Security Agency to de-anonymize users of the Tor network, ('Tor Stinks' presentation 2013), a network which aims at concealing the location and usage of users.

Web beacons

Users' online behavior can also be monitored by web beacons (also called web bugs, clear GIFs or pixel tags), which tiny are image tags embedded within a document, appearing on a webpage or attached to an email, that are intended to be unnoticed (Martin, Wu and Alsaid 2003). The image tag creates a holding space for a referenced image residing on the Web, and beacons transmit information to a remote computer when the document (web page or email) is viewed. Web beacons can gather information on their own, and they can also retrieve information from a previously set cookie (Angwin 2010; see Martin, Wu and Alsaid 2003 for description of the different technological abilities of web beacons). Such capacity means, according to the Privacy Foundation (Smith 2000; quoted in Martin, Wu and Alsaid 2003), that beacons could potentially transfer to a third party demographic data and personally identifiable information (name, address, phone number, email address, etc.) that a user has typed on a page. Unlike cookies, beacons are not tied to a specific server and can track users over multiple web sites (Schoen 2009). Beacons, moreover, cannot be managed through browser settings. While blocking third-party cookies limit

their range of action, it does not preclude beacons from gathering information on their own, and users have to install extensions to their browser to efficiently limit the effects of web beacons.

Strategies for identifying behavioral tracking

In order to identify privacy-respecting online resources, librarians must learn to assess the behavioral tracking activities occurring on websites. The first step is to identify and review website privacy policies. Privacy guidelines regulating the collection, retention and use of personal information in the online environment usually require that users should be given notice of website practices (e.g., Fair Information Practice Principles[2] proposed in 1973 by the US Secretary's Advisory Committee on Automated Personal Data Systems, the *Convention for the Protection of Individuals with Regard to Automatic Processing* of Personal Data developed by the Council of Europe (1981), and the Organisation for Economic Co-operation and Development *Guidelines on the Protection of Privacy and Transborder flows of Personal Data*[3]). This notice is typically provided in privacy policies that identify what information is collected, how it is used, and with whom it is shared. Regulatory frameworks, however, did not originally contemplate the collection of non-personally identifiable information. While such disclosure would seem to be consistent with the Fair Information Practice Principles, the current mode of mode of control is in many cases self-regulatory[4][5], and full compliance with notice requirements is far from universal (Komanduri et al. 2011-2012). Thus, while disclosure of behavioral tracking practices in websites should be seen as diagnostic of the presence of these mechanisms, lack of disclosure cannot be interpreted to mean that the site does not engage in behavioral tracking (Komanduri et al. 2011-2012; Burkell and Fortier 2013b).

Furthermore, privacy policy disclosures, where they do exist, may be difficult to understand (Burkell and Fortier 2013b). Website privacy policies are often complex (Micheti, Burkell and Steeves 2010). They tend to be written with the goal of protecting a website owner against lawsuits rather than informing users (Earp et al. 2005; Pollach 2005). Pollach (2005), for example, details a variety of linguistic strategies that serve to undermine user understanding of website practices, including mitigation and enhancement, obfuscation of reality, relationship building, and persuasive appeals. Therefore, even if many websites acknowledge the collection of non-personally identifiable information, both from first- and third-party, the effectiveness of this disclosure is limited, making privacy policies a relatively ineffective tool to identify behavioral tracking practices.

[2] The Privacy Act of 1974, 5 U.S.C. § 552a.
[3] C(80)58/FINAL, as amended on 11 July 2013 by C(2013)79.
[4] For instance, the new *Self-Regulatory Guidelines for Online Behavioral Advertising* identify the need to provide notice to users when behavioral data is collected that allows the tracking of users across websites and over time (United States Federal Trade Commission, 2009).
[5] Exceptions to this self-regulatory principle are increasing, including but not limited to the California Online Privacy Protection Act of 2003 (OPPA), and the EU Cookie Directive (2009/136/EC) of the European Parliament and of the Council.

As a result, librarians need to develop strategies and tools that allow them to assess directly the behavioral tracking practices of websites, in order that these practices can be considered in making websites recommendations. Different protocols can be followed in making this assessment, but they should be built around the following guiding principles (see Burkell and Fortier 2013a for a full discussion). The first important principle is that each website should be visited in an independent session to eliminate contamination. Each website under consideration should be visited in an independent session, beginning with the browser at an about:blank page, with clean data directories (no HTTP and Flash cookies, and an empty cache). The evaluator should ensure that browser settings are configured to allow cookies, tools to track web beacons (e.g., the Ghostery[6] browser extension) are installed in the browser, and Adobe Flash, via the Website Storage Settings panel is configured to accept data. The website should then be accessed directly by entering the domain name into the browser's navigation bar. Evaluators should mimic a typical user interaction with the website on many pages without clicking on advertisements or following links to outside sites. As they browse through the site, the evaluator should record the web beacons and trackers identified by the browser extension (e.g., Ghostery). At the end of the session, they should immediately review the contents of the browser cookie file and the Adobe Flash Panel via Website Storage settings, recording any cookies that are present. PrivacyChoice, as well as Ghostery, maintains a database of trackers that evaluators can use to identify associated privacy risk. While all third-party trackers raise some privacy issues, some of them put users at a greater risk than others, either because of their practices or their presence on a large number of websites. Evaluators should take that into account when making a decision.

Strategies for limiting behavioral tracking

Users may also take these steps to identify the presence of behavioral tracking, and digital literacy initiatives should provide this information along with tools and strategies that users can employ to limit tracking. It should be noted that elimination of all behavioral tracking may not be a desirable outcome from the perspective of users who benefit from the website personalization and optimization supported by these mechanisms. Targeted advertising can also be positive for many people, since it eliminates unwanted or 'useless' advertisements. Ultimately, a user must decide whether he or she wants to be tracked. Digital literacy initiatives should raise awareness of behavioral tracking and provide users with the tools they need to identify and control tracking should they choose to do so.

The easiest step is for users to learn how to manage HTTP cookies in every web browser that they use. Using browser settings, users can decide to refuse third-party cookies or even all cookies. The latter, however, will make the make the browsing experience much less efficient and may impede users from accessing some websites. Users should also learn how to delete cookies and they should be encouraged to think about periodically emptying the cookie file of each of their browsers. Controlling Flash cookies is more complex, yet crucial considering the capabilities of

[6] https://www.ghostery.com/.

Flash cookies. This is achieved through settings on the Adobe Website Storage Settings Panel. Browser extensions, such as Ghostery and AdBlock Plus[7], can be added to most browsers. Ghostery allows users to block trackers, either on a tracker-by-tracker basis, a site-by-site basis or a mixture of the two. Also customable, Adblock Plus allows users to block either all advertisements or only the ones they do not want to see. These extensions, however, may slow down Internet browsing.

Users can also change their Internet use habits. It is possible for user to use search engines that do dot store any non-personally identifiable information, such as Ixquick[8] and DuckDuckGo[9]. Ixquick returns the top ten results from multiple search engines. It only sets one cookie that remembers a user's search preferences and that is deleted after a user does not visit Ixquick for 90 days. DuckDuckGo, which returns the same search results for a given search term to all users, aims at getting information from the best sources rather than the most sources. While these search engines do not have all the functionality of the major search engines, both of them have received praise (e.g. McCracken 2011). The ultimate solution, one that allows a user to navigate online total anonymity, is to use the Tor[10] web browser, which impedes network surveillance or traffic analysis and which the U.S. National Security Agency has characterized as "the King of high secure, low latency Internet anonymity" (Schneier 2013). The anonymity afforded by Tor, however, comes at the price of reduced speed and limitations to available content.

CONCLUSION

It is widely understood that online privacy is at risk, threatened by the actions of governmental agencies and commercial entities. There is widespread awareness of and attention to the risks associated with the collection and use of personally identifiable information, but less attention is paid to an equally significant issue: the collection and use of information that is highly personal but nonetheless 'non-identifying'. This practice, termed 'behavioral tracking', is the focus of this paper. Other research demonstrates that behavioral tracking is widespread (Gomez, Pinnick and Soltani 2009; Burkell and Fortier 2013a), but users demonstrate only a limited knowledge of the practice and they do little to control tracking (comScore 2007; 2011; Rainie et al. 2013; TRUSTe 2013). We argue that librarians have a dual professional responsibility with respect to this issue: first, librarians should be aware of the surveillance practices of the websites they recommend to patrons and take these practices into account in making website recommendations; second, digital literacy initiatives spearheaded by librarians include a focus on online privacy, and provide patrons with the information they need to manage their own online privacy.

This paper presents an overview of online behavioral tracking mechanisms, and provides strategies for identifying and limiting online behavioral tracking. The information presented provides a basic understanding of tracking mechanisms along with practical strategies that

[7] https://adblockplus.org/.
[8] https://www.ixquick.com/.
[9] https://duckduckgo.com/.
[10] www.torproject.org/torbrowser/.

librarians can use to evaluate websites with respect to these practices and strategies that can be used to limit online tracking. We recommend that website evaluation standards be extended to include assessment of online privacy and especially behavioral tracking. We also recommend that librarians actively promote digital literacy by engaging in public education programs that take privacy and other digital literacy issues into account (American Library Association 2013). Finally, we note that protecting online privacy is an ongoing challenge, and librarians must ensure that they continually update their understanding of online surveillance mechanisms and the approaches that can be used to monitor and limit these activities.

ACKNOWLEDGEMENT

Support for this project was provided by the Office of the Privacy Commissioner of Canada through its Contributions Program. The views expressed in this document are those of the researchers and do not necessarily reflect the views of the Officer of the Privacy Commissioner of Canada.

REFERENCES

Adobe. 2011. "Adobe Flash Platform runtimes: PC penetration". http://www.adobe.com/mena_en/products/flashplatformruntimes/statistics.html.

"AdOne Classified Network and ClickOver announce strategic alliance". 1997. *Business Wire*, March 24.

"Affinicast unveils personalization tool". 1996. *AdAge*, December 4. http://adage.com/article/news/affinicast-unveils-personalization-tool/2714/.

American Library Association. 2008. *Code of Ethics*. http://www.ala.org/advocacy/proethics/codeofethics/codeethics.

———. 2013. *Digital literacy, libraries, and public policies: Report of the Office for Information Technology Policy's Digital Literacy Task Force*. http://www.districtdispatch.org/wp-content/uploads/2013/01/2012_OITP_digilitreport_1_22_13.pdf.

———. 2014. *Choose Privacy Week*. Accessed April 8. http://chooseprivacyweek.org.

Angwin, Julia. 2010. "The web's new gold mine: Your secrets". *The Wall Street Journal* July 31. http://online.wsj.com/news/articles/SB10001424052748703940904575395073512989404.

Ayenson, Mika, Dietrich James Wambach, Ashkan Soltani, Nathan Good and Chris Jay Hoofnagle. 2011. "Flash cookies and privacy II: Now with HTML5 and ETag respawning". *Social Science Research Network*. http://ssrn.com/abstract=1898390.

Ball, James. 2013. "NSA stores metadata of millions of web users for up to a year, secret files show". *The Guardian*, September 30. http://www.theguardian.com/world/2013/sep/30/nsa-americans-metadata-year-documents.

Barth, Adam. 2011. "HTTP State Management Mechanism". *Internet Engineering Task Force*, RFC 6265. http://tools.ietf.org/html/rfc6265.

Burkell, Jacquelyn and Alexandre Fortier. 2013. Privacy policy disclosures of behavioural tracking on consumer health websites. *Proceedings of the 76th Annual Meeting of the Association for Information Science and Technology*, edited by Andrew Grove. doi: 10.1002/meet.14505001087.

Burkell, Jacquelyn and Alexandre Fortier. 2015. Could we do better? Behavioural tracking on recommended consumer health websites. *Health Information and Libraries Journal* 32 (3): 182–194.

Canadian Library Association. 1976. *Code of Ethics*. http://www.cla.ca/Content/NavigationMenu/Resources/PositionStatements/Code_of_Ethics.htm.

Castelluccia, Claude and Arvind Narayanan. 2012. *Privacy considerations of online behavioural tracking*. Heraklion, Greece: European Union Agency for Network and Information Security. http://www.enisa.europa.eu/activities/identity-and-trust/library/deliverables/privacy-considerations-of-online-behavioural-tracking.

comScore 2007. *The impact of cookie deletion on the accuracy of site-server and ad-server metrics: An empirical comScore study*. https://www.comscore.com/fre/Insights/Presentations_and_Whitepapers/2007/Cookie_Deletion_Whitepaper.

———. 2011. *The impact of cookie deletion on site-server and ad-server metrics in Latin America: An empirical comScore study*. http://www.comscore.com/Insights/Presentations_and_Whitepapers/2011/Impact_of_Cookie_Deletion_on_Site-Server_and_Ad-Server_Metrics_in_Latin_America.

Council of Europe. 1981. *Convention for the protection of individuals with regard to automatic processing of personal data*. http://conventions.coe.int/Treaty/en/Treaties/Html/108.htm.

Earp, Julia B., Annie I. Antón, Lynda. Aiman-Smith and William H. Stufflebeam. 2005. "Examining Internet privacy policies within the context of user values". *IEEE Transactions on Engineering and Management* 52 (2): 227–237.

Gomez, Joshua, Travis Pinnick and Ashkan Soltani. 2009. *KnowPrivacy*. http://ashkansoltani.files.wordpress.com/2013/01/knowprivacy_final_report.pdf.

Goodwin Josh. 2011. Super cookies, ever cookies, zombie cookies, oh my. *Ensighten*, blog entry. http://www.ensighten.com/blog/super-cookies-ever-cookies-zombie-cookies-oh-my.

Harding, William T., Anita J. Reed and Robert L. Gray. 2001. Cookies and web bugs: What they are and how they work together. *Information Systems Management* 18 (3): 17–24.

Johns Hopkins University Sheridan Libraries. 2013. *Evaluating information found on the Internet*. http://guides.library.jhu.edu/evaluatinginformation.

Kamkar, Samy. 2010. "evercookie". http://samy.pl/evercookie/.

Kapoun, Jim. 1998. "Teaching undergrads web evaluation: A guide for library instruction". *College & Research Libraries News*, July/August: 522–523.

Komanduri, Saranga, Richard Shay, Greg Norcie, Blase Ur and Lorrie Faith Cranor. 2011-2012. "AdChoices? Compliance with online behavioral advertising notice and choice requirements". *I/S: A Journal of Law and Policy for the Information Society* 7: 603–638.

Kristol, David M. 2001. HTTP Cookies: Standards, privacy, and politics. *ACM Transactions on Internet Technology* 1 (2): 151–198.

Leon, Pedro Giovanni, Blase Ur, Rebecca Balebako, Lorrie Faith Cranor, Richard Shay, and Yang Wang. 2012. "Why Johnny can't op out: A usability evaluation of tools to limit online behavioral advertising". *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. http://dl.acm.org/citation.cfm?id=2207759.

Marshall, Matt. 2005. "New cookies much harder to crumble". *The Standard-Times*, May 15. http://www.southcoasttoday.com/apps/pbcs.dll/article?AID=/20050515/NEWS/305159957.

Martin, David, Hailin Wu and Adil Alsaid. 2003. Hidden surveillance by web sites: Web bugs in contemporary use. *Communications of the ACM* 46 (1): 258–264.

Mayer, Jonathan R. and John C. Mitchell. 2012. Third-party web tracking: Policy and technology. *Proceedings of the 2012 IEEE Symposium on Security and Privacy*. https://cyberlaw.stanford.edu/files/publication/files/trackingsurvey12.pdf.

McCracken, Harry. 2011. "50 websites that make the web great. *Time*, August 16. http://content.time.com/time/specials/packages/0,28757,2087815,00.html.

McDonald, Aleecia M. and Lorrie Faith Cranor. 2010. "Beliefs and behaviors: Internet users' understanding of behavioral advertising". *Social Science Research Network*. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1989092.

———. 2012. "A survey of the use of Adobe Flash Local Shared Objects to respawn HTTP cookies". *I/S: A Journal of Law and Policy for the Information Society* 7 (3): 639–687.

Micheti, Anca, Jacquelyn Burkell and Valerie Steeves. 2010. "Fixing broken doors: Strategies for drafting privacy policies young people can understand". *Bulletin of Science, Technology, and Society*. 30 (2): 130–143.

Narayanan, Arvind. 2011. "There is no such thing as anonymous online". Blog entry, July 28. https://cyberlaw.stanford.edu/blog/2011/07/there-no-such-thing-anonymous-online-tracking.

Office of the Privacy Commissioner of Canada. 2011. *Report on the 2010 Office of the Privacy Commissioner of Canada's Consultations on Online Tracking, Profiling and Targeting, and Cloud Computing*. https://www.priv.gc.ca/resource/consultations/report_201105_e.pdf.

———. 2013. Survey of Canadians on privacy-related issues. http://www.priv.gc.ca/information/por-rop/2013/por_2013_01_e.pdf.

Pollach, Irene. 2005. "A typology of communicative strategies in online privacy policies: Ethics, power, and informed consent". *Journal of Business Ethics* 62 (3): 221–235.

Rainie, Lee, Sara Kiesler, Ruogu Kang and Mary Madden. Anonymity, privacy, and security online. *Pew Research Internet Project*. http://www.pewinternet.org/2013/09/05/anonymity-privacy-and-security-online/.

Randall, Neil. 1997. "The new cookie monster". *PC Magazine* 16 (8): 211–214.

Schneier, Bruce. 2013. "Attacking Tor: How the NSA targets users' online anonymity". *The Guardian*, 4 October. http://www.theguardian.com/world/2013/oct/04/tor-attacks-nsa-users-online-anonymity.

Schoen, Seth. 2009. "New cookie technologies: Harder to see and remove, widely used to track you". Blog entry, September 14. https://www.eff.org/deeplinks/2009/09/new-cookie-technologies-harder-see-and-remove-wide.

Sipior , Janice C., Burke T. Ward and Ruben A. Mendoza. 2011. Online privacy concerns associated with cookies, Flash cookies, and web beacons. *Journal of Internet Commerce* 10 (1): 1–16.

Smit, Edith G., Guda Van Noort Hilde A. M. Voorveld. 2014. Understanding online behavioural advertising: User knowledge, privacy concerns, and online coping behaviour in Europe. *Computers in Human Behavior* 32 (1): 15–22.

Smith, R. M. 2000. "Why are they bugging you?" *Privacy Foundation.* http://www.privacyfoundation.org/resources/whyusewb.asp.

Soltani, Ashkan, Shannon Canty, Quentin Mayo, Lauren Thomas, Chris Jay Hoofnagle. 2009. "Flash cookies and privacy". *Social Science Research Network*. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1446862.

"'Tor Stinks' presentation". 2013. *The Guardian Online*, October 4. http://www.theguardian.com/world/interactive/2013/oct/04/tor-stinks-nsa-presentation-document.

TRUSTe. 2013. *US 2013 Consumer data privacy study – Advertising edition*. http://www.truste.com/us-advertising-privacy-index-2013/.

Turow, Joseph, Jennifer King, Chris Jay Hoofnagle, Amy Bleakley and Michael Hennessy. 2009. "Americans reject tailored advertising and three activities that enable it". *Social Science Research Network*. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1478214.

United States Federal Trade Commission. 2009. *FTC staff report: Self-regulatory principles for online behavioral advertising*. http://www.ftc.gov/os/2009/02/P085400behavadreport.pdf.

University of California, Berkley Library. 2012. "Finding information on the Internet: A tutorial" http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/Evaluate.html.

Ur, Blase, Pedro Giovanni Leon, Lorrie Faith Cranor, Richard Shay, and Yang Wang. 2012. "Smart, useful, scary, creepy: Perceptions of online behavioral advertising". *SOUPS '12 Proceedings of the Eighth Symposium on Usable Privacy and Security.* http://dl.acm.org/citation.cfm?id=2335362.

Weston, Greg, Glenn Greenwal and Ryan Gallagher. 2014. "CSEC used airport Wi-Fi to track Canadian travelers: Edward Snowden documents". *CBC News,* January 30. http://www.cbc.ca/news/politics/csec-used-airport-wi-fi-to-track-canadian-travellers-edward-snowden-documents-1.2517881.

"What they know". 2010. *The Wall Street Journal Online*. http://blogs.wsj.com/wtk/.