



Librarians' Mental Models and Use of Privacy-Protection Technologies

Author _ **Monica G. Maceli** (mmaceli@pratt.edu), Pratt Institute, School of Information, New York City

Libraries of all kinds are uniquely positioned to educate and advocate for the use of privacy-protection technologies (PPT) by our patrons and in our communities; this naturally extends our decades-long mission to protect our patrons' privacy in both digital and physical library environments. In pursuit of building an understanding of the challenges to expanding such work, this research study assesses librarians' existing technical knowledge of the internet's function, and their current use and understanding of a series of popular PPT. This research study begins to close the gap in research by building our knowledge of practicing librarians' underlying assumptions about the function of the internet and how these internal models are employed in their understanding and use of PPT. Research study participants were asked to draw and explain their perception of the functioning of the internet, to use several PPT, and to explain the protections afforded by each technology. The findings indicate that participants maintained inaccurate mental models of the internet and PPT, leading to difficulties in understanding function, and that they struggled with organizational and technical barriers to integrating such technologies into their daily lives.

Protecting patron privacy has been a long-standing core value of librarians and libraries, one which has received growing attention as new digital tools are introduced and novel privacy threats revealed. Widely publicized privacy-related current events and large-scale data breaches have created a climate in which the average internet user is highly concerned with privacy and loss of control over the collection and use of their information (Madden 2014; Madden and Rainie 2015). Within the field of library and information science (LIS), through a variety of workshops, guides, and handbooks, the library community has disseminated information on privacy-protection technologies (PPT), both within their peer group and to their patrons.



Prior research highlights the difficulty of this task, finding that most internet users take little action to protect their privacy, either by making behavioral changes or through use of privacy-enhancing technologies, despite their privacy concerns (Malandrino et al. 2013). Though librarians are highly interested in and motivated to participate in such educational activities, there exists little research exploring librarians' needs and deeper knowledge in the area of PPT specifically. These technologies extend beyond the boundaries of the physical library and digital library systems, intended for broad use throughout one's online life.

The rapidly changing technological landscape has posed myriad challenges to librarianship, such as providing appropriate technology education, keeping pace with job requirements, and in posing new privacy-related threats and concerns. Though library-focused resources provide many specific privacy recommendations, they may assume preexisting knowledge or simply lack the time/space to cover the underlying technological concepts, many of which may have drastically changed since the practitioners received their library science degree.

Librarians' required knowledge in the realm of PPT go beyond that of the average information technology user. Librarians must be able to wield PPT effectively in their own personal and professional lives, and they must also be able to accurately explain and advocate for these technologies to their patrons, communities, and colleagues. Taken together, these responsibilities require a deeper understanding of such technologies than the casual user may naturally acquire.

As explored in prior research detailing the role of technology in librarianship, a significant percentage of practitioners may be engaging with technology in the workplace largely in the realm of office productivity and cataloging technologies (e.g., Maceli and Burke 2016). These offer little, if any, opportunity to deeply engage with fundamental technological concepts underlying current or potential privacy threats (such as networking, firewalls, server configuration, encryption, and many others). Knowledge of such concepts endows the individual with the necessary technical context to evaluate, explain, and assess new threats and PPT.

This research study begins to close the gap in research by building our knowledge of practicing librarians' underlying assumptions about the function of the internet and how these internal models are employed in their understanding and use of PPT. A common research technique employed in assessing pre-existing internalized knowledge of a concept is eliciting mental models (Norman 1983)

of a particular topic through sketching or by asking users to think aloud. Mental models yield another dimension of understanding of an individual's technical knowledge, which can complement their self-reported skills, which have been studied in ongoing large-scale surveys of library staff and librarians (such as Burke 2016). Many privacy researchers have explored the role of technical knowledge in the privacy choices and actions that users take. This research study applies a similar approach to the librarian population by describing participants' mental models of the internet, as well as several common PPT.

This research study explores librarians' mental models, or internalized understanding, of the internet and how such knowledge is applied in their use and perception of PPT. This research study pursues the following research questions:

1. How accurate are the mental models of the internet held by librarians?
2. How do these internet mental models inform the use and understanding of PPT?
3. What implications do these findings have for PPT education within library science?

Literature Review

Within the field of library and information science (LIS), handbooks, articles, toolkits, and guides focused on supporting the needs of current practitioners have covered topical privacy concerns going back several decades. These cover a wide range of topics, including: library record privacy (e.g., Bielefield and Cheeseman 1994), patron privacy (e.g., Murray 2003), assessments of novel threats or tools (Fortier and Burkell 2015), behavioral tracking and the Tor Browser (Macrina 2015), and general guides to privacy in the digital era (e.g., Woodward 2007). The Intellectual Freedom Committee of the American Library Association (ALA) maintains a significant set of website resources in its Privacy Tool Kit (American Library Association 2014) aimed at understanding existing law, crafting a privacy policy, and explaining related technology concepts (such as encryption, HTTPS, and Tor). Another influential organization, the Library Freedom Project, seeks to "create a privacy-centric paradigm shift in libraries and the communities they serve" through educating librarians about privacy and surveillance threats and suggests numerous PPT, through both its website and in-person workshops (Library Freedom Project 2018). PPT recommended to librarians and library staff in existing literature typically provide functionalities such as: encrypting data (either in storage or in transit across the



internet) to securely protect one's information, avoiding behavioral tracking of one's online activities through web browser plug-ins, creating and storing strong passwords, and employing virtual private networks (VPNs) to safely navigate insecure networks (e.g., Library Freedom Project 2018; Maceli 2018).

Taken as a whole, these publications and organizations tend to provide specific guidance and recommendations in putting privacy-related policy, procedures, and technologies immediately in practice, with an emphasis on protecting the patrons' data as it intersects with the library's physical space and digital services. On a deeper, conceptual level, understanding the function of the internet is critical to understanding electronic privacy choices and potential threats. However, little work has directly assessed practicing librarians' underlying assumptions about the function of the internet, and how these internal "mental models" are employed in their understanding and use of the suggested PPT. Originating within the field of human-computer interaction, the concept of a mental model refers to a user's internalized understanding of a system; this representation is employed when interacting with a system and informs the user's assumptions and actions (Norman 1983). Library and information science research has extensively studied mental models to understand use and perception of a variety of LIS-related systems going back several decades, including: information retrieval systems, online catalogs, and search engines (for example, Borgman 1986; Makri et al. 2007; Zhang 2008a; Holman 2011; among many others). A variety of methods have been used to elicit and describe users' mental models, the most common being verbal explanation (i.e., the think-aloud protocol) and sketching of concepts, often observed simultaneously.

In an early study of users' mental models of the internet, Thatcher and Greyling (1998) employed sketching to evaluate participants' concepts of the internet, organizing their drawings into six categories and finding that those categories demonstrating greater detail and completeness were associated with higher frequency of internet use in participants. These drawings fell into the categories of "simple modularity" or "modularity and networking" in which modular networks, including many users, transmission media, transmission methods, and both local and wide area networks were drawn. Though at the time, the world wide web was still in its nascent phase, the authors noted that graphical user interfaces failed to facilitate "the development of a broader understanding of the internet's structure" (304). Other research studying mental models of the internet reinforced the role that technical expertise

plays, finding that experienced, expert internet users maintained mental models that were more flexible and elaborate (Levin et al. 1999) and allowed them to overcome errors encountered when web browsing (Sheeran et al. 2000). Papastergiou (2005), in a study on Greek high school students' mental models of the internet, found that their mental models were overly simplistic, included many misconceptions, and generally failed to provide a sufficient explanation of the internet and its function. Papastergiou noted the participants' "difficulty in conceiving the existence and necessity of an—invisible to them—underlying physical infrastructure [of the internet]" (356).

Within privacy-related research, the users' baseline technical knowledge has been assessed through similar research means, including eliciting mental models of technical concepts, such as the internet (Kang et al. 2015) or home computer security (Wash 2010). Kang et al. (2015) conducted a study of mental models of the internet and their relationship to privacy and security-related knowledge and actions, finding that those with more articulated internet mental models had a greater awareness of privacy threats. Malandrino et al. (2013) found that users with greater levels of technology knowledge had a better understanding of privacy-related threats; however all users generally expressed a concern for privacy but little effort to take any protective actions. Less technology-savvy users reported greater concerns about their privacy but were generally unwilling to modify settings, change their behaviors, or install PPT (Malandrino et al. 2013). Kang et al. (2015) found no clear relationship between users' technical background and knowledge, and their privacy-protection actions. This "privacy paradox" noted by Bashir et al. (2015) in which users profess to care a great deal about protecting their privacy, yet in practice take little action, is a widespread, paradoxical finding throughout much privacy and security research.

Furthermore, though privacy concern is noted to be widespread for both the public and librarians (e.g., Zimmer 2014), concern alone would appear to have little impact on users' underlying understanding of what data might be collected, why, and through what technical means (e.g., Bergmann 2009; Schaub et al. 2016). Bashir et al. (2015) described several key knowledge gaps, demonstrating a problem of information asymmetry between users and internet service providers, in particular around users' understanding of cloud computing, online security, and the ability of companies to monetize and resell their users' personal data.

These knowledge gaps are of great relevance to LIS educational efforts in the context of digital literacy and the



more recently proposed concept of *privacy literacy*. In 2009, Rotman presented a privacy literacy framework consisting of: *understanding* how personal information is used online, *recognizing* where information may be shared, *realizing* the consequences of sharing, *evaluating* the benefits or drawbacks to sharing online, and *deciding* when it is appropriate to share information. Wissinger (2017) defined privacy literacy as focused on the “understanding of the responsibilities and risks associated with sharing information online” and thus distinct from digital literacy which focuses on one’s ability to conduct information tasks in a digital environment. In relation to protective actions such as use of PPT, Trepte et al. (2015) suggest that a lack of privacy literacy prevents users from effectively taking action to assuage their privacy-related concerns. Framed in this way, privacy literacy becomes a deeply personal and challenging critical thinking activity (Wissinger 2017) and one that is closely linked to privacy-protection actions (e.g., Trepte et al. 2015).

Several of the suggested privacy literacy dimensions, in particular understanding how personal information may be used and shared (Rotman 2009), are necessarily entwined with one’s technical knowledge and conception of the internet’s function, as the underlying technological infrastructure enables privacy threats and protection possibilities. Recent initiatives have begun to directly address the need for privacy literacy and deeper technical understanding in librarians and library staff, as a precursor to educating our patrons on such topics, complementing long-standing efforts (notably those of the Library Freedom Project mentioned earlier). An edited guide on protecting patron privacy (Newman and Tijerina 2017) highlighted two ongoing projects in this realm—the Data Privacy Project at Brooklyn Public Library and privacy training at the City University of New York (CUNY). Both projects emphasize widescale privacy literacy education for librarians, library staff, and patrons, as do the efforts of the ALA’s “Choose Privacy” resources website and annual week of related events.

Zimmer and Tijerina (2018) produced a report detailing their community-driven research to put forward a “national roadmap for a digital privacy strategy for libraries,” funded by a grant received from the Institute of Museum and Library Services (IMLS). Their findings emphasized that digital privacy themes are increasingly prominent in library events and conferences, and that librarians and library staff commonly engage in privacy-related outreach and education of their communities, but many knowledge and skill gaps exist. Zimmer’s (2014) survey research indicated that over 75 percent of librarians feel that libraries should

educate the public around privacy issues, though only 56 percent had participated in a privacy-themed training or information session in the past year. Zimmer and Tijerina’s (2018) culminating event—the “Library Values & Privacy Summit”—recommended future steps, notably: an increase in technical training and data literacy, enhanced focus on data privacy in LIS education, and the need to overcome the disconnect between perceptions and reality of systems’ functions. These findings, and those presented earlier, have direct connection to the outcomes of this research study and will be discussed further in later sections.

Method

The research study design consisted of a descriptive lab session, including: (1) a short initial survey, (2) a web browsing and sketching activity, and (3) the use of several PPT. This research study sought to evaluate the librarian participants’ understanding of the function of the internet, as well as their knowledge of the role and purpose of PPT. Librarians in the New York City area, working in a variety of types of libraries, were the focus of the research study. Participants were recruited from email solicitations sent to library-focused user groups in the region and received a gift card for their participation. An initial pilot study with six participants was completed; the subsequent research study included twenty-two librarian participants.

During each individual’s session, participants were first asked to read and sign the consent form, then complete a brief survey about their current use of PPT and understanding of related technical concepts (using a survey adapted from Kang et al. 2015). Participants were then asked to sketch “how the internet works” and explain their initial drawing to the researcher. Next, the researcher directed the participants to browse the web, visiting the websites of their choice, while using a series of PPT: (1) the DuckDuckGo search engine, (2) the Ghostery web browser extension, and 3) the web browser’s incognito mode.

The first technology studied, DuckDuckGo (DDG), is a privacy-protection search engine that emphasizes protecting searchers’ privacy by not collecting user’s personal information and not storing and tracking users’ searches (DuckDuckGo 2018). The second technology studied is the Ghostery web browser extension, available for all major browsers, which blocks tracking scripts that may be used to collect data on user behavior for a variety of purposes, such as advertising or marketing (Ghostery 2018). The last technology studied is incognito mode, a privacy feature in most web browsers which disables the storage of a user’s browsing history, copies of webpages visited, and cookies, which provides protection against later users



of the same computer being able to view the prior users' information. All of the technologies selected for use in the research study are freely available, popular within librarian communities (and often recommended to their patrons), require no technical expertise to activate, and need little customization before use, while providing a range of different types of privacy protection.

Participants were encouraged to explore the technologies' interfaces and documentation to further their understanding of the technologies' purpose, while browsing the web and using the think-aloud protocol to explain their findings to the observing researcher. At the conclusion of the web browsing sessions, participants expanded their original sketch to indicate how and where they perceived the technology to provide privacy protection(s), if any. All sessions were audio-recorded and the researcher took notes and observed.

Survey results were analyzed through both quantitative and qualitative methods, to explore the closed and open-ended questions presented. The recordings of participant explanations were transcribed and analyzed using inductive qualitative analysis to code the transcripts and associated sketches, in pursuit of identifying themes and concepts of interest to the stated research questions. Sessions were conducted concurrent with data analysis until a saturation point was reached. A final coding scheme for the sketches and participants' explanations was developed, then the session data was coded and evaluated by two raters independently in an iterative fashion, until sufficient inter-rater agreement was measured (Cohen's kappa value of .77). A rubric was developed to rate each participant's sketches and associated transcript in their: (1) use of technical terminology, (2) technical accuracy, and (3) overall understanding of the technology. This three-dimension rubric was applied four times per participant—first to assess their understanding of each of the three PPT explored and lastly to assess their understanding of the internet's function more generally, based on assessing their diagram and associated think-aloud transcript. Participants' diagrams and verbal transcripts were rated on a 4-point ordinal scale ranging from "poor" to "excellent" against these measures, with the intention of yielding a dichotomous rating to group sketches as generally technically strong or weak. The researcher and a second information technology domain expert (both of whom regularly teach information technology courses within an ALA-accredited Master of Science in Library and Information Science program) assessed and rated each participant's response against the rubric, with an agreement level of .75, measured by Cohen's quadratic weighted kappa.

In the cases where there was rater disagreement, the researcher re-assessed the participant's response and ultimately assigned a final rating.

Results

A total of twenty-two librarian participants from New York City-area libraries completed the research study, which concluded in the spring of 2018. All participants worked in public, academic, or special libraries, with the exception of one museum librarian and one participant employed in an archive. Every participant had earned a Master's degree, such as the MSLIS or MLS, with approximately half having earned a related dual Master's degree. Librarian job titles were wide-ranging with many areas represented, including: serials, reference and instruction, cataloging, and young adult librarian. Participants are identified by number (i.e., [P1]) to protect their confidentiality.

Initial Survey Results

Most participants were mid to late-career with 64 percent having worked as a librarian for eleven or more years, with half falling into the 18-40 age range (Table 1 and 2).

Participants were asked how frequently they currently use PPT, with 55 percent reporting that they regularly or always used such technologies (Table 3).

Of the participants who reported using PPT, the self-reported technologies used included: incognito mode (13 participants); browser plugins, such as Ad-blocker, Privacy Badger, or Ghostery (9 participants); privacy setting changes to web browser and/or social media (3

Table 1. Participants' Age Demographics (N = 22)

Age Range	Participants	%
18-40	11	50
41-60	7	32
60+	4	18

Table 2. Participants' Career Stage (N = 22)

Years as Librarian	Participants	%
0-2 years	4	18
3-5 years	0	0
6-10 years	4	18
11-15 years	7	32
16 or more years	7	32



participants); encryption (3 participants); virtual private network (VPN) usage (2 participants), and DuckDuckGo usage (2 participants). One participant reported using Tor and the Signal messaging app.

Of the participants who did not use PPT or used them very infrequently, common reasons cited included: perceived difficulty of learning curve or lack of understanding (3 participants) and no underlying concerns of privacy-threats (2 participants). Participants with lengthier

careers in librarianship were no more likely to use PPT than those at an earlier stage in their careers. Participants noted several barriers to their use of PPT, including perceived dependency on information technology (IT) departments—“I’ll have to see if I can actually add that kind of stuff to my work computer without getting IT involved” [P1]—and lack of regular reminders to try out such technologies—“People talk about it in presentations and I always think I should use it, but then forget about it. So, I don’t use it” [P4].

Table 3. Participants’ Frequency of PPT Usage (N = 22)

Frequency of Use	Participants	%
Always (e.g., daily)	7	32
Regularly (e.g., a few times a week)	5	23
Occasionally (e.g., a few times a month)	4	18
Rarely (e.g., tried one once or twice)	4	18
Never	2	9

Lastly, within the initial survey, participants were asked to rate their familiarity with a series of technology terms and concepts, relating to privacy-protection (Figure 1, below) using a survey adapted from Kang et al. (2015). Participants professed to be most familiar with the concepts of: privacy settings, IP address, web browser plugins to block ads/trackers, incognito mode, and cookies. Less well understood, listed in order of decreasing familiarity, were: encryption, proxy server, privacy-protection search engine, virtual private network (VPN), Tor and SSL (secure sockets layer).

However, an exception to this linear relationship became apparent within a group of four participants who reported always using PPT yet rated their own knowledge of the technical concepts quite low. As will be detailed later on in this Results section, some of the most commonly used PPT, such as incognito browsing mode, were revealed to be quite poorly understood by many participants. This may have created a scenario where less technically-knowledgeable participants felt that they were protecting their privacy to a greater extent than they actually were in practice.

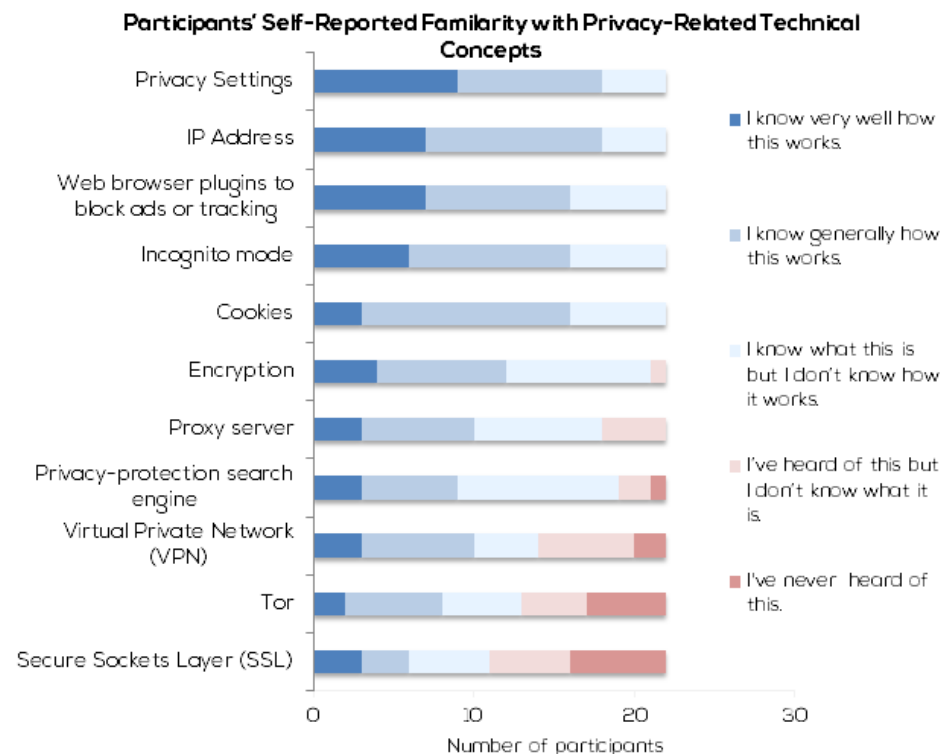


Figure 1. Participant responses to “How would you rate your familiarity with the following concepts or technologies?” on a five-point scale (N = 22). A general trend was noted in that the more frequently participants reported using PPT, the higher they rated their own knowledge of the technical concepts in the initial survey (Figure 2).

Sketching Exercise Results

Each participant then completed one baseline sketch of the function of the internet, responding



to the prompt of “sketch a diagram explaining how the internet works.” This diagram was subsequently modified to include the functions of the technologies studied: the DuckDuckGo search engine, the Ghostery browser plugin, and the web browser’s incognito mode. The initial diagram was created in black pen, with the subsequent additions (to explain DuckDuckGo, Ghostery, and incognito mode) sketched in red for clear differentiation. A sample participant diagram is included in Figure 3.

The coding scheme developed is shown below in Figure 4, organized into categories in a tree diagram. The most frequent technology terms used by participants in their sketches and think-aloud transcripts were: *information, computer, browser, server, data, IP address, and search history*.

The sketches and associated transcripts were then rated on the participant’s (1) amount of technical terminology used, (2) technical accuracy, and (3) overall understanding of the concept, across the four systems (1) the internet, (2) the DuckDuckGo search engine, (3) the Ghostery web browser plugin, and (4) incognito mode. Both the transcript and associated sketch were assessed as one for each participant, to ensure that participants’ scores were not influenced by their drawing abilities, but rather focused on the underlying ideas and concepts being expressed.

Function of the Internet

The initial sketch explaining the participant’s perception of the function of the internet demonstrated the highest ratings of all the systems across all three dimensions—knowledge and use of technical terms, as well as overall understanding. However, though this was the highest rated concept, the majority of participants were still scored poorly on their ability to accurately describe the function of the internet, as detailed in Figure 5. The participants who could accurately describe the internet’s function illustrated the inter-connected nature of the internet, consisting of many computer and network devices, and described the request and resulting response needed to transmit a webpage from a user’s client computer to a web server across the internet. Several detailed the TCP/IP set of network protocols allowing for addressing and routing

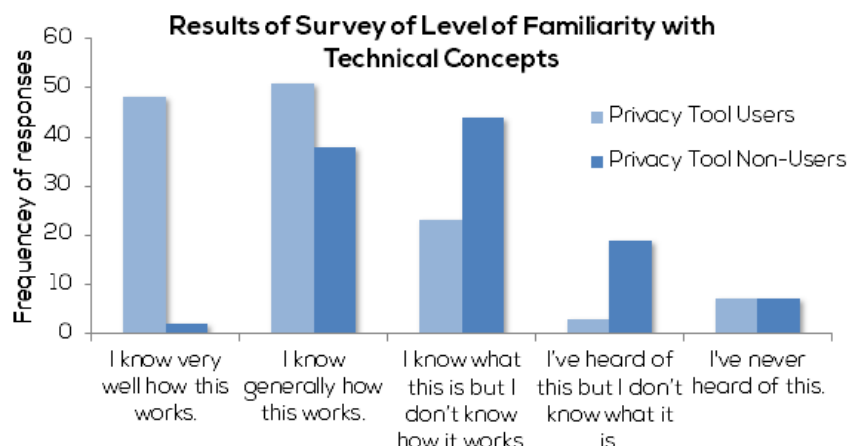


Figure 2. Participants’ professed familiarity with internet and privacy-related technical terms and concepts (listed in Figure 1 on page 23), across frequent and infrequent users of PPT (N = 22, 11 terms rated by each participant, for a total of 242 ratings).



Figure 3. Sample participant diagram detailing the function of the internet (black pen), plus DuckDuckGo, Ghostery, and incognito mode (all labeled and in red pen).

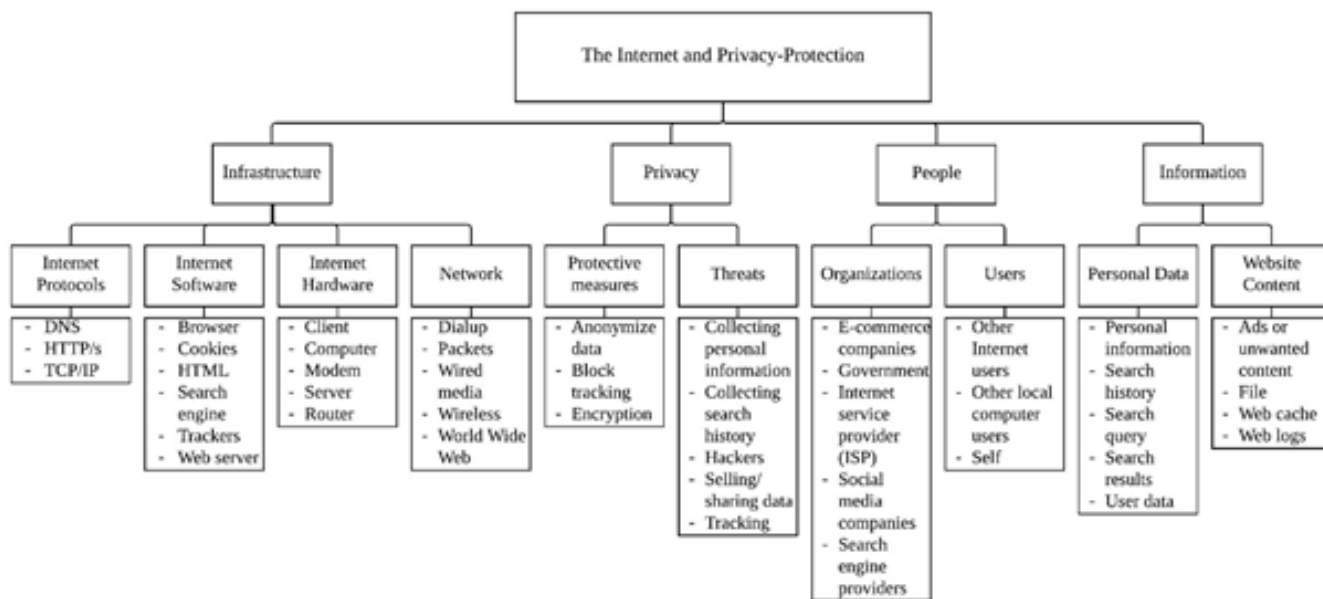


Figure 4. Librarian participants' collective mental model of the internet and privacy-related concepts, expressed as a tree diagram based on coded data.

Rating of participants' understanding of the internet

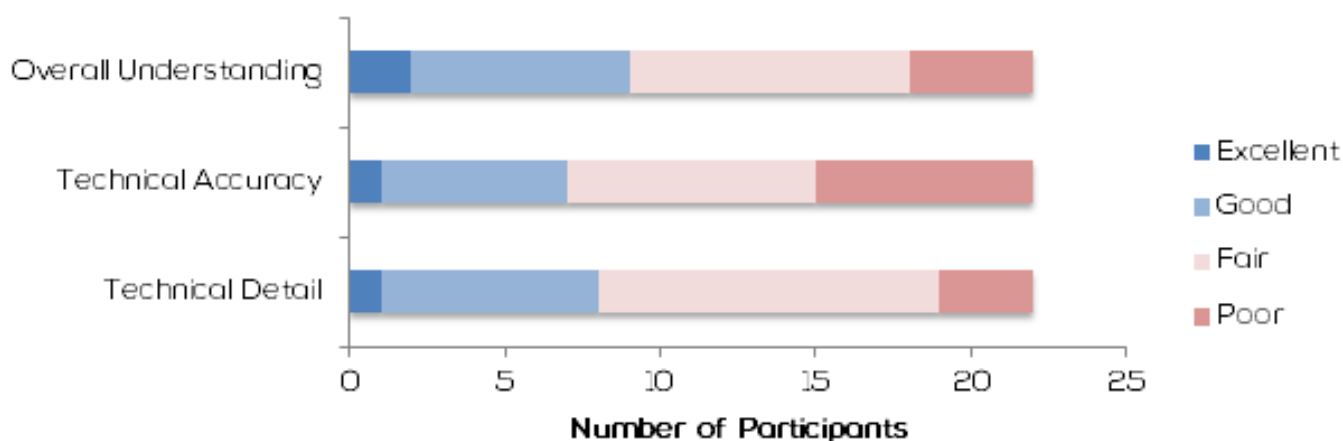


Figure 5. Ratings of participants' sketches demonstrating their understanding of the function of the internet (N = 22).

online, including the need to structure data into packets for transmission.

Reflecting these dichotomous results, some participants were able to easily and accurately explain the technical functions of the internet, for example:

So as a user I'm on the client—I make a request and it's a host name and there are all these DNS servers that understand

that that host name is equal to a certain IP. . . . So once the request knows where it's going the TCP/IP packet is broken down and if it's encrypted as HTTPS, it's encrypted in one way, otherwise the packet is just sent down into bytes that the internet is able to handle. That these devices are able to handle until it gets to the target website where the information is reconstructed and the packets sent back and there's a lot of communication back and forth. And packets from here



[client] . . . could go different routes until it gets to the ISP. [P5]

Whereas other participants struggled to explain on a deeper technical level and used vague or magical adjectives: “It’s [the internet] a wonderful, wonderful mystery and how it happens we don’t know” [P16]. Though frequent PPT users generally self-assessed their technical skills higher than non-users, as detailed earlier in this Results section, their internet mental model diagrams were rated with similar proportions of excellent/good to fair/poor rated diagrams (Figure 6).

Function of DuckDuckGo Search Engine

As a privacy-protection search engine, DuckDuckGo appeared to have achieved a fair amount of name recognition within the library community and many participants related that they had heard of the technology before, either within formal trainings or from colleagues. Relatively few participants were aware of DuckDuckGo’s functionality on a deeper level and the concept of a “privacy-protection search engine” was unclear to several participants and prompted further questions. As the diagram ratings in Figure 7 illustrate, approximately one third of participants could articulate an accurate overall understanding of DuckDuckGo’s privacy-related functions (namely, the fact that DuckDuckGo does not store users’ search history) but few were able to explain in deeply technical terms.

During the think-aloud portion while using DuckDuckGo, many participants expressed confusion around how such a service would be monetized via inclusion of ad-networks and how that would impact their privacy policies (despite each participant exploring DuckDuckGo’s documentation and mission statement) and their descriptions of the service’s benefits were inconsistent. One participant stated “I was actually surprised that there were advertisements appearing at the top because I didn’t know that there were paid ads in DuckDuckGo in a way that there are in Google” [P2], while another hypothesized that “It’s [DuckDuckGo] selling my information to advertisers in a really limited way, but not in a way that tracks me across the entire internet”

Internet Function Diagram Ratings

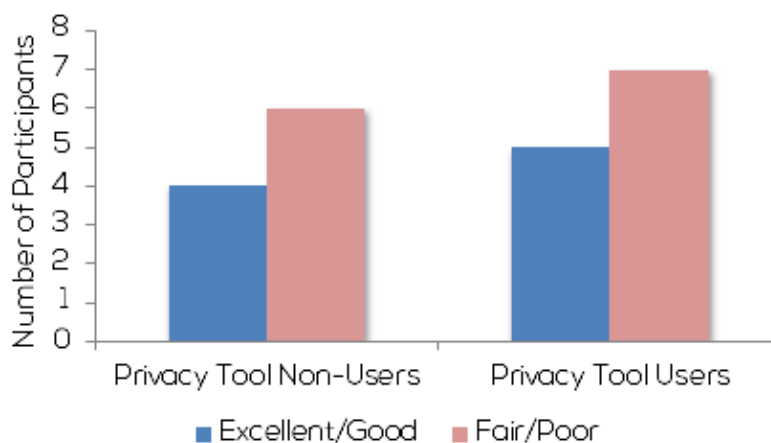


Figure 6. Ratings of participants’ overall understanding of the function of the internet, across frequent and infrequent users of PPT (N = 22).

Rating of participants’ understanding of DuckDuckGo

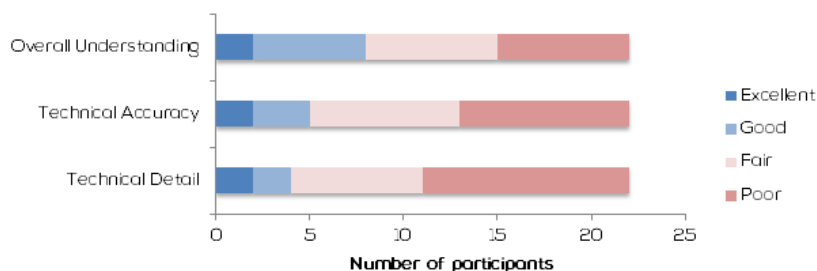


Figure 7. Ratings of participants’ sketches demonstrating their understanding of the function of the DuckDuckGo search engine (N = 22).

[P3]. A participant succinctly summed up the difficulty in demonstrating DuckDuckGo’s unique functionality, saying: “I guess it works by not doing things that everybody else does” [P1].

Function of Ghostery Web Browser Plugin

Many participants had heard of Ghostery, through training sessions or from colleagues, though few were regular users. After using and exploring Ghostery, few participants were able to convey the technical functionality it provided (Figure 8), though many could infer from Ghostery’s visual interface that a form of “blocking” of tracking was taking place.



Participants praised Ghostery's visual interface as a distinct advantage over the other technologies studied, describing it as "an awareness tool... to let you know when you're being tracked and what information is being tracked about it" [P14] as well as a technology that "educates you at the same time" [P20]. Even with these perceived advantages though, few participants could express an accurate and detailed understanding of the technological function of Ghostery, as evidenced by the bulk of them rating "poor" for technical accuracy and detail in Figure 8.

Rating of participants' understanding of Ghostery

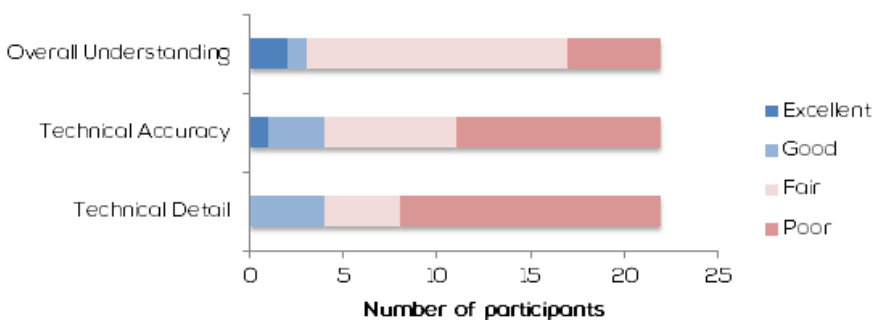


Figure 8. Ratings of participants' sketches demonstrating their understanding of the function of the Ghostery web browser plugin (N = 22).

Function of Incognito Mode

In the initial survey, incognito mode was the most common self-reported PPT used by research study participants. Incognito mode serves to protect against later users of the same local computer viewing one's stored history and data, but does not prevent internet service providers or websites from tracking or collecting data from the user. Despite the widespread use of this tool, participants' understanding of incognito mode was quite polarized with some participants able to accurately articulate its functions, while many others struggled to do so (Figure 9).

The participants who clearly demonstrated their understanding of this technology and the inherent limitations of incognito mode, were able to articulate this well, even though for some they may just have made the connection during the session:

I actually thought this was a more significant privacy service that it actually prevented information from being passed.

That it was anonymizing, but it seems like from this it's not actually anonymizing anything, it's just preventing stuff from being deposited locally. [P20]

Other participants had difficulty explaining the exact functionality, interpreting how the functionality was described within the browser, or had formed an inaccurate understanding. In many cases this had led participants

Rating of participants' understanding of incognito mode

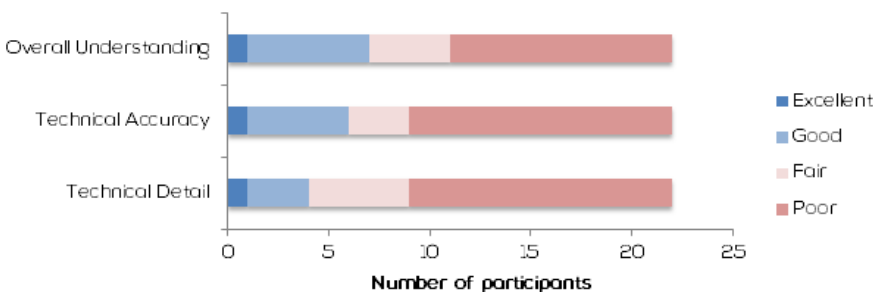


Figure 9. Ratings of participants' sketches demonstrating their understanding of the function of the web browser's incognito mode (N = 22).

to regularly use incognito mode with the assumption that it was providing a much greater level of anonymity than it in fact did. Participant comments illustrated some of the confusion surrounding incognito mode's benefits and were inaccurate as to the privacy benefits provided, for example: "I think I would still show up as an entity that visited the site so the analytics would still be there, but maybe they wouldn't know from where" [P6].

Overall Understanding of All Concepts Studied

Looking at the overall understanding rating across each of the concepts studied shows that the general function of the internet and the DuckDuckGo search engine were best understood overall (Figure 10), followed by incognito mode and, lastly, Ghostery. However, across all the technologies and concepts studied, the majority of participant responses were rated fair or poor in their ability to



accurately describe the technical functionality, with incognito mode having the most extreme variations in understanding. One librarian participant nicely illustrated these gaps in understanding:

I can't see what parts of my information are coming and going so it's hard for me to tell what I need or what these do in terms of [privacy] . . . it's hard for me to wrap my mind around the privacy part of that. [P9]

Other general themes of interest to these topics emerged during the participants' unstructured think-aloud discussion with the researcher. These included: difficulty in raising privacy awareness in their peers and the physical library as the main site of privacy concerns. One participant mentioned that "I've taught some privacy workshops in my library before, but without a whole lot of success in getting people to them; and I mean people come, but those are the people who are already interested" [P3]. Another participant stated that "a lot of the privacy talk I hear around libraries is about protecting the patron privacy, which is great, but I'm not as concerned about my own" [P14].

Discussion

This research study first sought to understand: *How accurate are the mental models of the internet held by librarians?* Mental models are typically somewhat incomplete and potentially inaccurate, while constantly changing in response to new information; this is expected and only becomes an issue when the existing mental model impedes the efficient use of a system (Norman 1983). Thus, one would *not* expect a large portion of the librarian participants to express a highly accurate and detailed mental model of any of the technical concepts studied, and that was in fact the case. Collectively, as a group, the technical terms and concepts the participants expressed (as detailed in the coding scheme presented in Figure 4) covered important aspects—including technical, human, and organizational dimensions—of the modern internet and related privacy concerns. In contrast to prior research exploring mental models of the internet, the coding scheme representing the participants' holistic view of the internet and privacy-related technologies and threats had similarities to prior work (e.g., Zhang 2008b), particularly around aspects of

Rating of participants' overall understanding of all four technologies studied

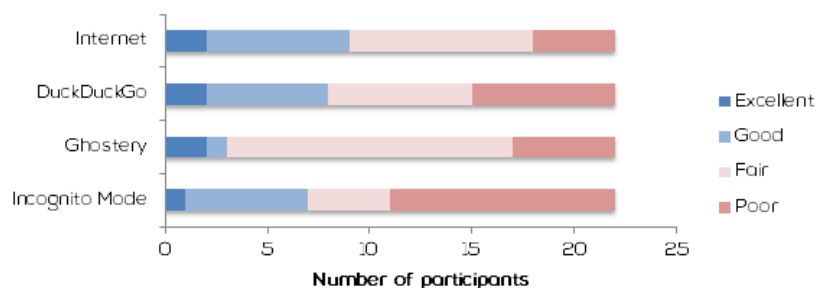


Figure 10. Ratings of participant sketches demonstrating their overall understanding of the function of the internet, DuckDuckGo search engine, Ghostery web browser plugin, and the web browser's incognito mode (N = 22).

internet infrastructure, which have remained largely consistent since the early days of the web.

On an individual participant basis, though, the expert raters found the number of technical terms used, the accuracy in use of such terms, and the overall ability to describe participants' technical understanding to be weak in many responses. Of the total of 22 participants' mental models studied, 41 percent were rated good or excellent in their overall understanding of the internet's technical function, while 59 percent were rated poor or fair by the expert raters. As in prior work (e.g., Kang et al. 2015), surveying participants on their prior experiences with technologies and techniques was used to assess baseline technical knowledge and relate this to their subsequent mental models. A similar survey was used in this research study (adapted from Kang et al. 2015), finding that simple recognition of many privacy-related technical concepts existed, but deeper understanding was missing across more technical concepts, such as: Tor, virtual private networks, or SSL as used in encrypted web browsing (Figure 1). The level of existing technical knowledge was tied to the ability to articulate a more detailed and accurate mental model of the internet, and this relationship was mirrored in the presented results with a larger portion of excellent/good rated diagrams associated with higher self-assessment of technical concept knowledge. One noticeable difference in self-professed technology knowledge was observed between participants who did and did not use PPT frequently, with frequent users citing greater technical familiarity across all concepts surveyed (Figure 2).



Though frequent PPT users generally self-assessed their technical skills higher than non-users, the resulting internet mental model diagrams were rated with similar proportions of excellent/good to fair/poor rated diagrams (Figure 6), which was unexpected. This suggests that frequent users of PPT have better recognition and understanding of related technical terms but may not have integrated this understanding into their conception of the function of the internet in a deeper fashion.

Though the survey focused on privacy-related technical terms, broader surveys of technology-related skills and usage of librarians and library staff have been conducted regularly within LIS research work, with recent findings indicating that most practitioners primarily utilize office productivity and cataloging technologies, with relatively small numbers engaging in deeply technical systems-related work (e.g., Maceli and Burke 2016). Library practitioners who are motivated to seek out additional information about PPT may have increased familiarity with more technical terms and concepts, but still lack the opportunity to build deeper understanding through hands-on work.

The internet mental model results were then considered in relation to the PPT studied, questioning: *How do these internet mental models inform the use and understanding of PPT?* A handful of participants were able to confidently express and demonstrate their technical knowledge in the area of PPT, but the majority of participants struggled with the activities and received low diagram ratings across all technologies. The small number of participants who had internet understanding diagram ratings of “excellent” continued to demonstrate their technical understanding, with high understanding ratings for each of the PPT studied, but ratings across the remaining participants were otherwise low.

However, participants of all levels of technical knowledge were able to articulate the general privacy threats they may be vulnerable to (as expressed collectively in the coding scheme in Figure 4) around the storing and collecting of personal data and search history, and the possibility for that data to be resold to other parties. Two participants (9 percent of the total research study participants) noted that they were not unduly concerned with privacy threats and thus did not use PPT. This small percentage is generally in line with the findings of larger-scale studies such as in Zimmer (2014), which surveyed librarians' concerns around privacy and personal information collection, finding that 3 percent were relatively unconcerned about data collection from companies and 7 percent were unconcerned about government data collection.

The gap between the relatively abstract knowledge of privacy threats and tangible technical knowledge,

nevertheless, meant that most participants could not express how PPT might intervene in the privacy-threatening processes of browsing a website or conducting a search. There was no clear relationship observed between the participants' technical knowledge and their length of time in the field or their area of focus within librarianship. Their technical knowledge appeared to be motivated much more by a particular individual's interest in this area and willingness to pursue the topic further, often in their leisure time.

The intention of this research study is not to emphasize the specific feedback generated by each tool; PPT will no doubt change in the future in response to new threats. Rather, the aim is to understand the deeper dimensions that improve or impair user understanding in the librarian population. In the case of technologies such as incognito mode, participants had a hard time reconciling what the technology said it did (when exploring the interface and documentation) with their pre-existing assumptions, even if they were regular users. A particular issue seemed to be the visual feedback offered by the tool, or the lack thereof, combined with the challenge of indicating that a technology or service is protecting users' privacy by “not doing things that everybody else does” [P1].

Lastly, the research study considered: *What implications do these findings have for PPT education within library science?* Though not the explicit focus of the research study, many participants volunteered details of an initiating event that sparked their interest in privacy and protection technologies, such as the Snowden revelations. Most participants reported attending at least one privacy-related technology training either within the workplace or at professional development events, such as offered at conferences. This is higher than Zimmer's (2014) findings, which reported that 56 percent of respondents had participated in a privacy educational event within the past year. It is likely that participants in this research study self-selected for having a pre-existing interest in privacy and thus were more likely to seek out such opportunities. Few of the participants, however, had attended multiple trainings.

A resounding theme in participant responses was the challenge encountered in carrying over what was learned into their daily lives, despite clear recognition of the importance of such technologies. Though interest and awareness of privacy were generally high in participants, many described significant barriers to their own use of PPT, as well as their effectiveness in conveying their importance to patrons. Many felt disempowered to use such technologies in the workplace without support from their IT departments, or to employ what they learned in daily practice. These findings suggest that many librarians



may demonstrate the privacy paradox, evident in prior research results (Bashir et al. 2015), of experiencing privacy concerns but taking relatively little action, similar to challenges in encouraging patron adoption in this area (e.g., Maceli 2018).

Those who had taken action, and integrated the PPT into their lives and workplaces, reported a lack of deeper understanding of the technical functionality and struggled with the many tradeoffs in convenience. Those that actively worked towards awareness and use of PPT, either in other staff members or their patrons, were often met with less enthusiasm than they would have liked.

Many participants discussed the current privacy and security initiatives within their libraries that *were* receiving significant attention. These largely addressed the long-standing concerns of protecting patron privacy in the physical library space (such as patrons' information on public access computers) and records-keeping policies of the library itself. There was general agreement that the patron privacy focus was largely bounded by the physical library and the library's technology services (e.g., electronic resources), but little about privacy concerns in other aspects of the patrons' lives or for the librarians and library staff themselves.

Clearly, despite the proliferation of one-time workshops and conference presentations to disseminate privacy information, the results suggest that this information tends not to have lasting behavioral impact on librarians, who may be constrained by organizational factors or their technical knowledge. Furthermore, many PPT work invisibly, without obvious educational benefit to users as to what protections are employed and how they fit into the larger internet infrastructure. This barrier to deeper learning of systems and infrastructure was noticed decades ago in prior research (such as Thatcher and Greyling 1998; Papastergiou 2005), where it was suggested that graphical user interfaces prevented users from naturally learning the deeper technical concepts at work during use.

The themes emerging from participant responses during this research study suggest struggles with conceptualizing the baseline technical activities that take place as their data traverses the internet, as well as difficulty in "seeing" what changes or differs when utilizing PPT. This aligns with the findings of Zimmer and Tijerina (2018) who emphasized the need to overcome the disconnect between perceptions of systems' function and actual system function. Not only does this potentially impact librarian use of such technologies, but also their ability to effectively explain and advocate for the use of such technologies by our patrons.

Zimmer and Tijerina's (2018) report also advocates for additional focus on data privacy in MLS graduate programs. It appears that most library practitioners are receiving the bulk of their privacy literacy education in the workplace and thus continuing education programs may be most effective in reaching active librarians with the most current technical information. To complement educational efforts both within and after the MLS program, these issues may potentially be tackled with further attention to the design of educational PPT interfaces, perhaps in intentionally reducing the invisibility of systems design, that has been a long-time pursuit of good design, in order to expose more of the underlying functionality at work and encourage deeper engagement.

Conclusions and Future Work

Though many library science-focused educational initiatives have increased privacy awareness and concerns, largely in the continuing education realm, the privacy-protection actions of librarians have lagged behind. A great deal of focus still remains on privacy within the physical library and our own library records, as was described by the participants in this research study. Less emphasis appeared to be placed on educating patrons (and librarians themselves) to protect their privacy as they browse the web more generally. Nearly every research study participant had exposure to privacy training of some kind in the workplace or other professional development opportunities, but this approach failed to create lasting behavioral change. This research study indicates that the information science field is in need of educational and teaching technologies with greater impact on one's privacy choices and behaviors. Specifically, the findings presented suggest future work in building educational technologies that can assist users in making the connection between underlying internet infrastructure and their own information as it traverses the network. And as new educational technologies are introduced, regularly-conducted and large-scale library survey research (such as Burke 2016) investigating the technologies commonly employed by librarians and library staff can be expanded to question the use of PPT, such that changes over time and effectiveness of educational initiatives in this area can be assessed.

Acknowledgements

The author would like to thank the Faculty Development Fund of Pratt Institute for their generous support of this research project during academic year 2017–2018.



References

- American Library Association Office for Intellectual Freedom. 2014. "Privacy Tool Kit." Accessed December 10, 2018. <http://www.ala.org/advocacy/privacy/toolkit>.
- Bashir, Masooda, Carol Hayes, April D. Lambert, and Jay P. Kesan. 2015. "Online privacy and informed consent: The dilemma of information asymmetry." *Proceedings of the Association for Information Science and Technology* 50, no. 1: 1–10. <https://doi.org/10.1002/pra.2015.145052010043>.
- Bielefield, Arlene and Lawrence Cheeseman. 1994. *Maintaining the Privacy of Library Records: A Handbook and Guide*. New York: Neal-Schuman.
- Bergmann, Mike. 2009. "Testing privacy awareness." *The Future of Identity in the Information Society, IFIP Advances in Information and Communication Technology* 298: 237–53.
- Borgman, Christine L. 1986. "The user's mental model of an information-retrieval system: An experiment on a prototype online catalog." *International Journal of Man-Machine Studies* 24, no. 1: 47–64.
- Burke, John. 2016. "Survey says . . . : How library staff members are using technologies." In *Neal-Schuman Library Technology Companion: A Basic Guide for Library Staff, 5th edition*, edited by John J. Burke, 15–30. Chicago: Neal-Schuman.
- DuckDuckGo. 2018. "About DuckDuckGo." Accessed November 29, 2018. <https://duckduckgo.com/about>.
- Fortier, Alexandre and Jacquelyn Burkell. 2015. "Hidden online surveillance: What librarians should know to protect their own privacy and that of their patrons." *Information Technology & Libraries* 43, no. 3: 59–72.
- Ghostery. 2018. "About Ghostery." Accessed October 21, 2018. <https://www.ghostery.com/about-ghostery/>.
- Holman, Lucy. 2011. "Millennial students' mental models of search: Implications for academic librarians and database developers." *The Journal of Academic Librarianship* 37, no. 1: 19–27.
- Kang, Ruogu, Laura Dabbish, Nathaniel Fruchter, and Sara Kiesler. 2015. "'My data just goes everywhere': User mental models of the internet and implications for privacy and security." In *Symposium on Usable Privacy and Security (SOUPS)*, 39–52. Berkeley, CA: USENIX Association.
- Levin, James A., Matthew J. Stuve, and Michael J. Jacobson. 1999. "Teachers' conceptions of the internet and the world wide web: A representational toolkit as a model of expertise." *Journal of Educational Computing Research* 21, no. 1 (July 1999): 1–23. <https://doi.org/10.2190/KV7J-PVAG-2WWM-TTWJ>.
- Library Freedom Project. 2018. "Library Freedom Project—Making real the promise of intellectual freedom in libraries." Accessed on December 17, 2018. <https://libraryfreedomproject.org/>.
- Maceli, Monica. 2018. "Encouraging patron adoption of privacy-protection technologies: Challenges for public libraries." *IFLA Journal—Special Issue on Privacy* 44, no. 2: 195–202. <https://doi.org/10.1177/0340035218773786>.
- Maceli, Monica and John J. Burke. 2016. "Technology skills in the workplace: Information professionals' current use and future aspirations." *Information Technology and Libraries* 35, no. 4: 35–62.
- Macrina, Alison. 2015. "The Tor browser and intellectual freedom in the digital age." *Reference & User Services Quarterly* 54, no. 4: 17–20.
- Madden, Mary. 2014. "Public perceptions of privacy and security in the post-Snowden era." Last modified November 12, 2014. <http://www.pewinternet.org/2014/11/12/public-privacy-perceptions/>.
- Madden, Mary and Lee Rainie. 2015. "Americans' attitudes about privacy, security and surveillance." Last modified May 20, 2015. <http://www.pewinternet.org/2015/05/20/americans-attitudes-about-privacy-security-and-surveillance/>.
- Makri, Stephann, Ann Blandford, Jeremy Gow, Jon Rimmer, Claire Warwick, and George Buchanan. 2007. "A library or just another information resource? A case study of users' mental models of traditional and digital libraries." *Journal of the American Society for Information Science and Technology* 58, no. 3: 433–45.
- Malandrino, Delfina, Vittorio Scarano, and Raffaele Spinelli. 2013. "How increased awareness can impact attitudes and behaviors toward online privacy protection." In *Proceedings of the 2013 International Conference on Social Computing*, 57–62. Washington, DC: IEEE Computer Society.
- Murray, Peter E. 2003. *Library Patron Privacy: SPEC Kit*. Washington, DC: Association of Research Libraries, Office of Leadership and Management Services.
- Newman, Bobbi, and Bonnie Tijerina, eds. 2017. *Protecting Patron Privacy: A LITA Guide*. Lanham, MA: Rowman & Littlefield.
- Norman, Donald A. 1983. "Some observations on mental models." In *Mental Models*, edited by Dedre Gentner and Albert L. Stevens, 7–14. New York: Psychology Press.
- Papastergiou, Marina. 2005. "Students' mental models of the internet and their didactical exploitation in informatics education." *Education and Information Technology*, 10, no. 4: 341–60.
- Rotman, Dana. 2009. "Are you looking at me? Social media and privacy literacy." Poster presented at the iConference, Chapel Hill, North Carolina, February 8–11.
- Schaub, Florian, Aditya Marella, Pranshu Kalvani, Blase Ur, Chao Pan, Emily Forney, and Lorrie Faith Cranor. 2016. "Watching them watching me: Browser extensions' impact on user



- privacy awareness and concern.” In *NDSS Workshop on Usable Security*. San Diego: Internet Society.
- Sheeran, Louise, M. A. Sasse, Jon Rimmer, and Ian Wakeman. 2000. “Back to basics: Is a better understanding of the internet a precursor for effective use of the web?” In *Proceedings of the 1st Nordic Conference on Computer Human Interaction*. Stockholm.
- Thatcher, Andrew, and Mike Greyling. 1998. “Mental models of the internet.” *International Journal of Industrial Ergonomics* 22, no. 4–5: 299–305.
- Trepte, Sabine, Doris Teutsch, Philipp K. Masur, Carolin Eicher, Mona Fischer, Alisa Hennhöfer, and Fabienne Lind. 2015. “Do people know about privacy and data protection strategies? Towards the ‘Online Privacy Literacy Scale’ (OPLIS).” In *Reforming European Data Protection Law*, 333–65. Netherlands: Springer.
- Wash, Rick. 2010. “Folk models of home computer security.” In *Proceedings of the Sixth Symposium on Usable Privacy and Security*, 1–16. New York: ACM.
- Wissinger, Christina L. 2017. “Privacy literacy: From theory to practice.” *Communications in Information Literacy* 11, no. 2: 378–89.
- Woodward, Jeannette A. 2007. *What Every Librarian Should Know about Electronic Privacy*. Westport, CN: Libraries Unlimited.
- Zhang, Yan. 2008a. “The influence of mental models on undergraduate students’ searching behavior on the web.” *Information Processing & Management*, 44, no. 3: 1330–45.
- Zhang, Yan. 2008b. “Undergraduate students’ mental models of the web as an information retrieval system.” *Journal of the American Society for Information Science and Technology* 59, no. 13: 2087–98.
- Zimmer, Michael. 2014. “Librarians’ attitudes regarding information and internet privacy.” *The Library Quarterly*, 84, no. 2: 123–51.
- Zimmer, Michael and Bonnie Tijerina. 2018. “Library Values & Privacy in our National Digital Strategies: Field guides, Convenings, and Conversations.” Milwaukee, WI: Center for Information Policy Research. https://cpb-us-w2.wpmucdn.com/people.uwm.edu/dist/b/524/files/2018/08/LibraryValuesAndPrivacy_Report-28qqhttp.pdf.