Chapter 1

Introduction

Libraries regard the protection of the confidentiality and privacy of those who make use of their content and services as a core value. Yet libraries today face many obstacles in achieving optimal privacy and security in the implementation of their websites and other aspects of their technical infrastructure. Financial resources and technical expertise to implement the latest technologies for computer and network security may not always be available. While some libraries have deployed state-of-the-art systems, others struggle to gain access to even the most basic technologies.

The challenges are not all related to availability of resources. Libraries must also deal with the tensions, if not direct contradictions, between protecting privacy and their interest in providing services that meet the expectations of their patrons. In the context of the overarching concern to protect the privacy of their patrons, libraries also desire to implement tools and technologies that provide more personalized services and that might provide opportunities for better engagement with their patrons. The boundaries are not necessarily clear between the values of protecting privacy and the tools and technologies available for personalized services.

The commercial web today works on a business model of sales and advertising based on aggressive collection and sharing of personal data. The basic fabric of commercial technology, including the infrastructure of the web and in-person retail environments, has been honed to capture as much personal data as possible. This data powers a global advertising ecosystem designed to strengthen commercial sales through ever more finely targeted placement of ads. Libraries, in contrast, embrace a model of providing services based on privacy and confidentiality. For libraries to implement websites and other technologies that reflect their values of privacy in the context of a global infrastructure optimized for commerce and advertising invariably involves difficult choices and some compromise. While libraries may not be able to entirely isolate their web-based services from commercial technologies, they can implement measures that limit exposure and that meet their expectations for protection of privacy.

This issue of *Library Technology Reports* explores these issues and the technologies needed to deploy a library website with adequate protections for the privacy of those who visit. Without the implementation of standard encryption components, the online information-seeking activities of website visitors are vulnerable to exposure. Even when a site is properly encrypted, privacy can be circumvented through tracking agents placed on the site for analytics or advertising. In some cases, tracking mechanisms may be included inadvertently, such as when they are brought in through components used for desired features. Following discussion of the technical issues with implications for user privacy, this report includes the results of a broad study of the state of practice for these privacy-related technologies among public and academic libraries in the United States. This study reveals great progress among these libraries in the strengthening of privacy on their websites, though substantial gaps remain.

Libraries Value Privacy

This report is based on the fundamental concept that the values of the profession mandate that libraries implement technology systems able to respect confidentiality and protect privacy. The American Library Association provides a clear statement of the responsibilities of libraries related to this important topic in a document that was adopted by the ALA Council initially in 2002 and subsequently updated in 2014 and 2019. The following excerpts reinforce the aspects of privacy central to this issue of *Library Technology Reports*:

> The library profession has a long-standing ethic of facilitating, not monitoring, access to information. Libraries implement this commitment through

> the adoption of and adherence to library privacy policies that are consistent with applicable federal, state, local, and where appropriate, international law. It is essential that libraries maintain an updated, publicly available privacy policy that states what data is being collected, with whom it is shared, and how long it is kept. Everyone who provides governance, administration, or service in libraries, including volunteers, has a responsibility to maintain an environment respectful and protective of the privacy of all users. It is the library's responsibility to provide ongoing privacy education and training to library workers, governing bodies, and users in order to fulfill this responsibility.
>
> . . .
>
> The American Library Association affirms that rights of privacy are necessary for intellectual freedom and are fundamental to the ethical practice of librarianship. The rapid pace of information collection and changes in technology means that users' personally identifiable information and library-use data are at increased risk of exposure. The use of new technologies in libraries that rely on the collection, use, sharing, monitoring and/or tracking of user data may come into direct conflict with the *Library Bill of Rights* and librarians' ethical responsibilities. Libraries should consider privacy in the design and delivery of all programs and services, paying careful attention to their own policies and procedures and that of any vendors with whom they work. Privacy is the foundation upon which our libraries were built and the reason libraries are such a trusted part of every community.[1]

Libraries provide access to information both through their physical facilities and through their websites. Within their physical premises, libraries take great care to ensure that information about the resources and services accessed by a patron remains private and is not shared with other individuals or organizations. The integrated library systems used to manage the lending of materials are configured to maximize patron privacy. While it is necessary to maintain a link between bibliographic records and patron records when an item is borrowed, extensive measures are taken to ensure the privacy of the transaction. While the loan is active, the connection between the item and patron data is needed to support operational tasks such as sending notices when the item is past its loan period. But once the item has been returned, libraries routinely remove all traces of the transaction from the systems involved. It is common for libraries to retain only anonymized data for concluded circulation transactions so that no records are available that reveal what items any given patron has borrowed or consulted. Log files that may otherwise hold data related to these transactions are likewise scrubbed or anonymized. Even during the interval of an active loan transaction, precautions are implemented to ensure that only specifically authorized personnel are able to view patron data, including the items on loan.

The removal of data describing completed loan transactions is only one example of the measures libraries take to ensure that no traces remain regarding the specific resources that patrons may have accessed. When asked for information regarding the resources any given patron may have accessed, even by law enforcement agencies, libraries want to be able to truthfully respond that the information is not available. Even in the event of a security breach, there should be the least possible personally identifiable information or data regarding information access. This approach toward the privacy of access to resources enables patrons to use information provided by the library without fear of judgement or reprisals.

Privacy versus Personalized Services

These measures taken to protect privacy can be seen as a constraint on the ability of the library to engage in personalized services or to enable social features. Removing data related to completed loan transactions, for example, eliminates the ability of patrons to view items they have previously borrowed, a feature most persons would expect to be available. Any e-commerce site would track all previous purchases and use data collected on items bought or viewed to present recommendations. These environments would also use data from other customers to inform recommendations: "Others who purchased this item also purchased these."

To support these kinds of personalized services, recommendations, and social sharing features, many libraries enable the collection of the associated personalized data. This collection of personalized data would usually be enabled through specific patron consent. Patrons would have the ability to opt in to retention of data on items borrowed or other types of interactions in order to receive enhanced personalized services. Whether opt-in or opt-out options are selected by default would be determined by library policy, reflected in the organization's stated privacy policies.

Protection of Online Information-Seeking Activities

Libraries use their websites to provide information regarding their facilities and services and as portals

through which their patrons can explore information resources. Libraries provide extensive collections of electronic resources and other digital content for access to the general public and to their websites. A typical interaction includes a patron typing a topic of interest into a search box, viewing results, making selections, and viewing or downloading content. These transactions include data regarding the topics of interest and items accessed by a given individual at a specific time. Even when the patron accessing the information isn't signed into a library account, technical information from the network and browser may identify, or at least imply, a specific individual.

Patron use of a library website involves data at least as sensitive as data related to physical items borrowed. For many—probably most—libraries, the quantity of information accessed by patrons through the website exceeds loans of physical materials. Achieving the same level of privacy for information access by library patrons through the website as for transactions representing physical loans requires attention to some technical details relating to the library's website and any related systems or services. Privacy for web-based transactions requires that no one can listen in on the network in a way that reveals patron information-seeking activities and that the data related to the transaction not be shared with other individuals or organizations.

This report explores issues relating to the privacy and security of data that represents the online information-seeking activities of individuals through a library website. For the purposes of this report, the term *library patron* means any person who uses library resources, including the general public. The report focuses on the technical and functional characteristics of the main library website. Online catalogs, discovery services, and the extensive portfolio of information resources that may be accessed through a library website are all subject to the same concerns but are not directly addressed in this study.

Note

1. American Library Association, *Privacy: An Interpretation of the Library Bill of Rights*, (Chicago: American Library Association, 2002, amended 2014 and 2019), http://www.ala.org/advocacy/intfreedom/librarybill/interpretations/privacy.