I/S: A JOURNAL OF LAW AND POLICY FOR THE INFORMATION SOCIETY

Reader Privacy in Digital Library Collaborations: Signs of Commitment, Opportunities for Improvement

ANNE KLINEFELTER *

CONTENTS

* Associate Professor of Law and Director of the Law Library, University of North Carolina. The author thanks research assistants Richard Chen and Laura Wright and law librarian Nick Sexton for their support. A number of individuals generously provided helpful consultation and insights including Deborah Caldwell-Stone, Kim Dulin, Dave Hansen, Chad Haefele, Michael Dorman, Tim Shearer, Luke Swindler, and Adam Ziegler. Margot Kaminski provided early editorial feedback that was significant and most appreciated. Thanks go also to my co-panelists, BJ Ard, and Seeta Peña Gangadharan for their ideas and inspiration, and to Peter Shane for the invitation to participate in the symposium on the Future of Libraries.

Abstract: Libraries collaborate to digitize collections large and small in order to provide information with fewer geographical, temporal, or socio-economic barriers. These collaborations promise economy of scale and breadth of impact, both for access to content and for preservation of decaying print source material. Some suggest this increased access to information through the digital environment comes at the expense of reader privacy, a value that United States librarians have advanced for nearly eighty years. Multiplying risks to digital reader privacy are said to weaken librarians' commitment to privacy of library use and to overwhelm libraries' ability to ensure confidential access to information. This article reviews some recent national and international organization statements on library privacy and finds continuing commitment to library privacy but varied approaches to balancing privacy with other goals and challenges in the digital environment. The article also evaluates privacy protections arising from libraries' digital collaboration work with Google Books and the related HathiTrust project, and finds a number of vulnerabilities to confidential library use of these resources. These reviews confirm that reader privacy is increasingly at risk even as librarians confirm their commitment to protecting reader privacy through organizational statements. The article concludes that libraries can use their collaborative traditions to develop better

approaches to protecting privacy as they develop digital collections. Even if libraries have limited success negotiating for or creating digital spaces for perfect digital reader privacy, much can be gained by making privacy an important feature of digital library design. Incremental but meaningful improvements can come from user authentication systems with privacy features, wider adoption of encryption, and innovations in website analytics tools. Reader privacy pressures and compromises are not new to libraries, and incremental solutions in the digital environment are worthy efforts that honor the tradition of libraries' commitment to reader privacy.

I. INTRODUCTION

A. *The Digital Collaboration Future and Its Reader Privacy Weaknesses*

Librarians are experienced collaborators, particularly with each other. Librarians across the country copy catalog records, expand local services through interlibrary loan, develop best practices for services through professional associations, and work together on advocacy for information law and policy. Librarians also collaborate to create digital collections with a goal of broad fee-free access to information, sometimes with other libraries and sometimes with commercial partners. The future of libraries will probably build on existing collaborations in order to pool resources for common goals, and because access to digital collections reduces the need for library users to make a trip to a particular library location.[1]

---

[1] *See* John Palfrey, *Hacking Libraries,* PUBLISHERS WKLY., June 15, 2015, at 34 (noting libraries' traditions of working together and envisioning a future of networked and interconnected libraries with individual libraries able to collaborate with others and serve local needs through shared platforms) (excerpted from his book, JOHN PALFREY, BIBLIO TECH: WHY LIBRARIES MATTER MORE THAN EVER IN THE AGE OF GOOGLE (2015)). *See generally* PETER HERNON & JOSEPH R. MATTHEWS, REFLECTING ON THE FUTURE OF ACADEMIC AND PUBLIC LIBRARIES 48 (2013); Fay Chadwell & Shan C. Sutton, *The Future of Open Access and Library Publishing,* 115 NEW LIBR. WORLD 225 (2014); Michelle M. Wu, *Building a Collaborative Digital Collection: A Necessary Evolution in Libraries*, 103 L. LIBR. J. 527 (2011).

The power of collaboration promises a much richer pool of materials than any one library can obtain and curate.[2] Research libraries have already contracted with Google to provide content for the Google Books project, while also developing the library preservation consortium known as HathiTrust for library-controlled access to copies of these digital records and other digital collections.[3] Visionary librarians have helped create the Digital Public Library of America (DPLA) to increase access to information in libraries of all types from across the country.[4] Librarians have contributed to the movement for open access to scholarship[5] and have shaped related efforts like Harvard Law Library's Caselaw Access Project, which

---

[2] Copyright and licensing restrain digital copies and distribution for a significant amount of recently created and published work, but librarians and others are focused on distribution of materials not so restricted and are advancing arguments for expanding digitization to cover works with debated status under the law. *See generally* David R. Hansen, *Copyright Reform Principles for Libraries, Archives, and Other Memory Institutions*, 29 BERKELEY TECH. L.J. 1559 (2014) (advocating copyright law reform to make traditional library exemptions applicable in technology-neutral ways); Julie L. Kimbrough & Laura N. Gasaway, *Publication of Government-Funded Research, Open Access, and the Public Interest,* 18 VAND. J. ENT. & TECH. L. 267 (2016) (explaining the movement towards fee-free access to publications, especially those based on government-supported research, and suggesting federal and state law changes may be required).

[3] *See Our Digital Library*, HATHITRUST, https://www.hathitrust.org/digital_library [https://perma.cc/5SES-KXDS] (describing HathiTrust as a "digital preservation repository" that provides preservation and access services for public and copyrighted material from sources including Google, etc.).

[4] *See History*, DIGITAL PUB. LIBR. OF AM., https://dp.la/info/about/history/https://dp.la/info/about/history/ [https://perma.cc/D5GQ-42K7] (explaining that the concept of DPLA was arranged by various leaders in an effort to develop a comprehensive, open network allowing access to a resource of information from all types of libraries across the nation). *See generally* Robert Darnton, *Digitize, Democratize: Libraries and the Future of Books*, 36 COLUM. J.L. & ARTS 1 (2012-13) (promoting the Digital Public Library of America as a more egalitarian alternative to the Google Books commercial product); John Palfrey, *A Digital Public Library of America?: Collective Management's Implications for Privacy, Private Use, and Fair Use*, 34 COLUM. J.L. & ARTS 837 (2011).

[5] PETER SUBER, OPEN ACCESS § 4 (2012) (defining and promoting new economic approaches to making academic publications widely accessible); Richard A. Danner, Kelly Leong & Wayne V. Miller, *The Durham Statement Two Years Later: Open Access in the Law School Journal Environment*, 103 L. LIBR. J. 39, 41-45 (2011) (evaluating progress towards not only fee-free legal scholarship, but also elimination of investment in print versions). Librarians also encourage each other to select openly available digital publications for their own scholarship. *See ACRL Policy Statement on Open Access to Scholarship by Academic Librarians*, ASS'N OF COLL. AND RES. LIBR. (June 2016), http://www.ala.org/acrl/standards/openaccess [https://perma.cc/D8LW-TBHL].

includes a partnership with the commercial legal research platform Ravel Law.[6] A number of other digital partnerships, small and large, are part of the evolving missions of libraries.[7]

This increased access to digital information is praised as "free,"[8] and these collaborative digitization efforts are promoted as egalitarian and democratizing.[9] But the addition of new privacy risks in the digital

---

[6] This project was originally titled *Free the Law*. *See* Adam Ziegler, *Caselaw Access Project*, ET SEQ: THE HARVARD LAW SCH. LIBR. BLOG (Aug. 8, 2016), http://etseq.law.harvard.edu/2016/08/caselaw-access-project/ [https://perma.cc/Z8W7-BEVT]; *Harvard and Ravel Collaborate*, RAVEL LAW, https://www.ravellaw.com/?modal=videos.hls-and-ravel [https://perma.cc/3K2D-8HRC].

[7] *Samples of Projects*, LYRASIS, http://www.lyrasis.org/LYRASIS%20Digital/Pages/Digitization%20Collaborative/Samples.aspx [https://perma.cc/2V3X-MLRH] (a nonprofit supporting digital collaborations among libraries, museums, and other cultural heritage organizations; it provides an index to sample successful projects, many involving smaller special collections). Larger projects that place the onus on the library to preserve and index born-digital content may stretch the resources of even the largest library, as the Library of Congress is finding with Twitter archive donations. *See* Andrew McGill, *Can Twitter Fit Inside the Library of Congress?*, THE ATLANTIC, Aug. 4, 2016; *see also* AXEL BRUNS & KATRIN WELLER, WEBSCI '16: PROCEEDINGS OF THE 8TH ACM CONFERENCE ON WEB SCIENCE, TWITTER AS A FIRST DRAFT OF THE PRESENT---AND THE CHALLENGES OF PRESERVING IT FOR THE FUTURE, 183-185 (2016), http://dl.acm.org/citation.cfm?id=2908174 [https://perma.cc/EU2G-TXCH] .

[8] *See, e.g.*, Robert J. Aalberts, Alexander Nill & Percy S. Poon, *Online Behavioral Targeting: What Does the Law Say?*, 37 J. OF CURRENT ISSUES & RES. IN ADVERT. 95, 97 (2016) ("consumers pay for the 'free' content by providing personal information---the basic building block for [online behavioral targeting]---which in turn leads to high ad revenues that allow publishers to keep the content free of charge"); David Hall, *Google, WestlawNext, LexisNexis and Open Access: How the Demand for Free Legal Research Will Change the Profession*, 26 SYRACUSE SCI. & TECH. L. REP. 53, 64-65 (2012) (describing the Google Scholar search engine as filling a need for free legal research but omitting to consider whether the business model is based on monetizing the individual's research trails and compromising confidentiality of the research); Chris J. Hoofnagle & Jan Whittington, *Free: Accounting for the Costs of the Internet's Most Popular Price*, 61 UCLA L. REV. 606, 608 (2014) (suggesting that the appeal of efficiencies through digitization, the concept of "free" will soon be seen as a norm).

[9] Darnton, *supra* note 4 (encouraging new economic models for digitized books and other intellectual content to be made widely available); Julie L. Kimbrough & Laura N. Gasaway, *Publication of Government-Funded Research, Open Access, and the Public Interest*, 18 VAND. J. ENT. & TECH. L. 267, 269 (2016) (evaluating the trend towards making publicly funded research available without monetary barriers and noting "[a]ccess to government information is often described as the hallmark of a democracy--only an informed citizenry can participate wisely in the democratic process."); *'Free the Law' Will Provide Open Access to All*, HARVARD GAZETTE (Oct. 29, 2015), http://news.harvard.edu/gazette/story/2015/10/free-the-law-will-provide-open-access-to-all/ [https://perma.cc/8JEL-CRCW] (promoting the increased access to court opinions through a collaborative digitization project between Harvard Law Library and the commercial legal research service Ravel Law).

environment could be seen as introduction of a privacy fee[10] that actually burdens democratic values.[11] This newer form of access may require each user to explicitly identify herself to an entity other than the library, or to leave enough digital bread crumbs to allow her online reading to be traceable by commercial or governmental tracking technologies. Individuals are identified or tracked in order to facilitate the digital library collaboration product and the overall system of access to information through the Internet. If the digital library product requires funding to cover intellectual property or technology hosting costs, access might be limited to a category of authorized users who must identify themselves through an "authentication" process in order to gain access.[12] If the product provides customizable features, a user might have to identify herself to some extent in order to avail themselves of those special settings.[13] In addition, the overall structure of online access to information is permeated with largely unregulated privacy vulnerabilities due to interest in tracking individuals for

---

[10] MICHAEL ZIMMER, ICONFERENCE,'12: PROCEEDINGS OF THE 2012 CONFERENCE, THE ETHICAL (RE)DESIGN OF THE GOOGLE BOOKS PROJECT 365-66 (2012) (describing privacy risks for users of Google Books). *But cf.* Palfrey, *supra* note 4 (advocating progress towards the Digital Public Library of America despite its potential to increase reader privacy risks that might emerge related to the need to license content and authenticate authorized users, or because of growth of online tracking online activity more generally).

[11] *See Privacy and Confidentiality*, AM. LIBR. ASS'N, http://www.ala.org/advocacy/privacyconfidentiality/privacy/privacyconfidentiality [https://perma.cc/TA43-8KSZ] ("The possibility of surveillance, whether direct or through access to records of speech, research and exploration, undermines a democratic society."); *see also* Neil M. Richards, *Intellectual Privacy*, 87 TEX. L. REV. 387, 399-407 (2008) (advancing intellectual privacy and freedom of thought as necessary for First Amendment search-for-truth and self-governance values); Alan Rubel & Mei Zhang, *Four Facts of Privacy and Intellectual Freedom in Licensing Contracts for Electronic Journals*, 76 C. & RES. LIBR. 427, 432-33 (2015) (identifying republican freedom as vulnerable to privacy loss); Paul M. Schwartz, *Privacy and Democracy in Cyberspace*, 52 VAND. L. REV. 1609, 1609 (1999) ("[W]idespread, silent collection of personal information in cyberspace . . . degrades the health of a deliberative democracy").

[12] Julie Cohen, *What Privacy is For*, 126 HARV. L. REV. 1904, 1916 (2013) (noting that information flows are commonly used to authenticate individuals for access to databases); Al-Suqri, Mohammed Nasser & Esther Akomolafe-Fatuyi, *Security and Privacy in Digital Libraries: Challenges, Opportunities and Prospects*, 3 INT'L J. DIGITAL LIBR. SYSTEMS 54 (2012).

[13] Jean E. McLaughlin, *Personalization in Library Databases: Not Persuasive Enough?*, 29 LIBR. HI TECH 605, 612 (2011) (considering privacy implications of various requirements for customized features in library databases).

marketing and other commercial and even governmental purposes.[14] Consumer technologies and other strategies for avoiding these systemic privacy risks continue to play leapfrog with developments in tracking, resulting in a dynamic that is disadvantageous to the digital reader's privacy.[15]

The privacy-endangering online environment that digital library collaborations inhabit in order to reach a wide readership contrasts with the print-focused library tradition that offers confidential access to reading materials. Although much of published information, like a printed book, has been a commodity in modern culture, libraries, through the pooling of private or public funding, served as a communal intermediary to provide access to these products.[16] For many years, in this role as intermediary, libraries have had the opportunity and capacity to create a layer of confidentiality for research, reading and related ways of accessing print publications and other forms of creative expression. Library users could walk in, browse the stacks, and read books or listen to sound recordings without focused monitoring of their activities. Even library book circulation systems were designed to limit retention and sharing of records of individual library users' reading.[17]

During the Twentieth Century, prior to the rise of the World Wide Web and digital formats for publications and communication,

---

[14] At the time of this writing, the Federal Communications Commission was considering new rules to regulate Internet Service Providers' ability to monetize access to information about customers' Internet activity. John D. McKinnon, *Business News: FCC Tempers Broadband Proposal---Regulator Scales Back Tougher Privacy Rules After Backlash from Internet Providers*, WALL ST. J., Oct. 7, 2016, at B5.

[15] *See* Aalberts, *supra* note 8, at 95 (describing how websites, advertising networks, and Internet service providers create profiles on individuals because "[e]very online move a consumer makes, any search, any browsing, and any purchase can potentially be tracked down and analyzed . . ."); Chris Jay Hoofnagle et al., *Behavioral Advertising: The Offer You Cannot Refuse*, 6 HARV. L. & POL'Y REV. 273, 273 (2012) (reporting results of empirical studies showing dramatic expansion in the number of tracking technologies used by heavily visited websites, including the use of "new, previously unobserved tracking mechanisms that users cannot avoid even with the strongest privacy settings.").

[16] *See* Mary Murrell, *Digital + Library: Mass Book Digitization as Collective Inquiry*, 55 N.Y.L. SCH. L. REV. 221, 225-26 (2010-11) (relating concerns about the ongoing viability of publicly funded libraries as market forces disrupt recent traditions of information production and consumption).

[17] Marshall Breeding, *Issues and Technologies Related to Privacy and Security*, 52 LIBR. TECH. REP. 5, 7 (2016) (reviewing mechanisms and policies for privacy of circulation records).

lawmakers and librarians developed a framework to improve and protect the confidentiality of access to information through libraries. Privacy of library use remains presumptively protected by a combination of laws and library ethical commitments. Some degree of protection comes from state law in all fifty states, and by federal statute for some special libraries like the Library of Congress.[18] While a few states have more recently enacted legislation to protect the privacy of users of e-books and other online content, these laws have had limited impact because technologies and industries continue to change in ways that carry them outside of the scope of the law,[19] or the library user's consent is arguably too easily invoked.[20] Even before the

---

[18] *Privacy Laws Regarding Library Records*, AM. LIBR. ASS'N, http://www.ala.org/advocacy/privacyconfidentiality/privacy/stateprivacy [https://perma.cc/SA72-K4QS] (providing an index and text of state library privacy laws).

[19] California's Reader Privacy Act limits disclosure of book reading records of individuals whether that reading is in print or other formats, but it applies only to a book service that, "as its primary purpose, provides the rental, purchase, borrowing, browsing, or viewing of books." "Book service" does not include a store that sells a variety of consumer products when the book service sales do not exceed two percent of the store's total annual gross sales of consumer products sold in the United States. CAL. CIV. CODE § 1798.90(2) (2012); *see* B.J. Ard, *The Limits of Industry-Specific Law*, 51 IDAHO L. REV. 607, 609-611 (2015) (noting several weaknesses of the California law in meeting its sponsors' goals of protecting reading in the digital environment, including the potential for large book seller Amazon to diversify enough that book sales would not constitute two percent of annual gross and so escape application of the law). Arizona amended its library privacy law to add the phrase "including e-books" to its prohibition on libraries' sharing personally identifying reading records with a few exceptions. Arguably, one exception, "if necessary for the reasonable operation of the library," may be flexible enough to cover arrangements such as a library's lending of e-books that require individuals to register each use with the vendor if the library is unable to secure a more privacy-protecting set of terms from the vendor. A.R.S. § 41-151.22 (2016) (Westlaw current through the Second Regular Session of the Fifty-Second Legislature); *see also* DEL. CODE tit. 6 § 1206C (2016) (Westlaw current through 80 Laws 2016, ch. 345; effective as of January, 2016, limiting the disclosure of personal information about users of a commercial digital book service and requiring annual public reports of disclosures); MO. REV. STAT. §§ 182.815, 182.817 (adding to the library privacy statute language to cover an "e-book" or "digital resource or material" and requiring a court order or consent of the library user before disclosure of identifying information by any third party contracted by a library that receives, transmits, maintains, or stores a library record).

[20] Aalberts, *supra* note 8 (noting that courts generally find that consumers consent to online tracking if the service agreement provides notice, even if terms are "incomprehensible to the average consumer"); BJ Ard, *Confidentiality and the Problem of Third Parties: Protecting Reader Privacy in the Age of Intermediaries*, 16 YALE J. L. & TECH. 1, 26 (2012-13) (pointing out that state library privacy statutes generally yield with the library user's consent and may not apply to non-library actors at all); Julie Cohen, *DRM and Privacy*, 18 BERKELEY TECH. L. J. 575, 601-02 (2003) (describing consent components of intellectual privacy protection schemes as flawed because they involve

development of library privacy laws, librarians, through various library associations, articulated ethical commitments to the confidentiality of library use.[21] Librarians have also developed guidelines for protecting library use confidentiality such as discarding detailed records that link individuals with the titles of borrowed books soon after those books are returned, so that those records cannot be discoverable or used once their library purpose has been served.[22]

A range of rationales support protection of library use and of reader privacy more broadly.[23] Arguments have been advanced for recognizing reader privacy, or specifically library use privacy, under the First Amendment.[24] The American Library Association (ALA) promotes confidentiality of library use as a component of intellectual freedom and as a necessary support for an informed citizenry.[25] Private exploration of ideas is defended as a precondition to autonomy

---

tradeoff of incommensurable dignitary values that are not appropriate for market ordering).

[21] *See infra* Section II.

[22] *Library Privacy Guidelines for Data Exchange Between Networked Devices and Services*, AM' LIBR. ASS'N (June 24, 2016), http://www.ala.org/advocacy/library-privacy-guidelines-data-exchange-between-networked-devices-and-services [https://perma.cc/HLQ3-LK9M] (advising libraries to have methods for securely destroying personally identifying data that is no longer needed, including archived and backup copies).

[23] *See generally* Trina Magi, *Fourteen Reasons Privacy Matters: A Multidisciplinary Review of Scholarly Literature*, 81 LIBR. Q. 187 (2011) (reviewing and summarizing the literature of social sciences, law, and philosophy to bolster librarians' resolve to protect reader privacy in light of growing challenges).

[24] *See* Margot Kaminski & Shane Witnov, *The Conforming Effect: First Amendment Implications of Surveillance, Beyond Chilling Speech*, 49 U. RICH. L. REV. 465, 465-67, 475-78 (2015) (reviewing scholarly theories and mixed recognition of First Amendment protections for reader privacy by courts and legislatures). *See generally* Jonathan Marc Blitz, *Constitutional Safeguards for Silent Experiments in Living: Libraries, the Right to Read, and a First Amendment Theory for an Unaccompanied Right to Receive Information*, 74 UMKC L. REV. 799 (2006) (promoting public libraries as institutions meriting special First Amendment protections at least partly because of their commitment to confidential provision of access to information).

[25] AM. LIBR. ASS'N, INTELLECTUAL FREEDOM MANUAL 178 (Trina Magi & Martin Garner eds., 9th ed. 2015) (asserting that privacy is necessary for intellectual freedom); *Privacy and Confidentiality*, *supra* note 11 ("Lack of privacy and confidentiality chills users' choices, thereby suppressing access to ideas. The possibility of surveillance, whether direct or through access to records of speech, research and exploration, undermines a democratic society.")

and innovation, and as a support for a more tolerant and civil society.[26] Beyond abstract justifications, specific requests for records of individual's use of libraries by government and private actors have provided additional inspiration for the development of law and policy that protects confidentiality of library use.[27]

Despite these legal protections for library use privacy and development of librarian ethical commitments, as more information has been published in electronic format, libraries have had mixed experiences with adapting their privacy intermediary role to the digital environment. Digital economy trends have moved towards data mining the habits of online readers and researchers for an expanding array of purposes using methods that are difficult to trace, and government interest in surveillance of online activity has been revealed.[28] Digital intellectual property rights management,[29] sophisticated methods of targeting advertisements or otherwise

---

[26] JEFFREY ROSEN, THE UNWANTED GAZE: THE DESTRUCTION OF PRIVACY IN AMERICA 140-42 (2000) (addressing the benefits of when individuals with strongly held differing opinions can use privacy to create space for disagreement). *See generally* Julie Cohen, *What Privacy Is For*, 126 HARV. L. REV. 1904 (2013) (describing privacy as a necessary precondition for the dynamic process of self-definition which promotes social and political innovation and progress); Neil Richards, *The Perils of Social Reading*, 101 GEO. L.J. 689, 704-08 (2013) (examining the need for privacy in reading and thinking to encourage exploration of ideas outside of the mainstream that could support innovative intellectual activity).

[27] *See* Sarah Shik Lamdan, *Why Library Cards Offer More Privacy Rights Than Proof of Citizenship: Librarian Ethics and Freedom of Information Act Requestor Policies*, 30 GOV. INFO. Q. 131, 133 (2013) (reviewing historical origins of library patron privacy ethics). *See generally* HERBERT N. FOERSTEL, SURVEILLANCE IN THE STACKS (1991); Bruce S. Johnson, *A More Cooperative Clerk: The Confidentiality of Library Records*, 81 L. LIBR. J. 769 (1989); Bruce M. Kennedy, *Confidentiality of Library Records: A Survey of Problems, Policies, and Laws,* 81 L. LIBR. J. 733 (1989).

[28] *See* BERNARD E. HARCOURT, EXPOSED: DESIRE AND DISOBEDIENCE IN THE DIGITAL AGE (2015) (detailing and critiquing ways in which activity in the digital environment is surveilled); BRUCE SCHNEIER, DATA AND GOLIATH: THE HIDDEN BATTLES TO COLLECT YOUR DATA AND CONTROL YOUR WORLD § 1 (2015) (explaining how pervasive surveillance is conducted and the impacts of these practices, and recommending legal, corporate and individual solutions); Alan Rubel, *Libraries, Electronic Resources, and Privacy*: *The Case for Positive Intellectual Freedom*, 84 LIBR. Q. 183, 183-87 (2014) (identifying ways that information vendors seek identifying information about individual library users accessing content paid for by the library); David Gauvey Herbert, *This Company Has Built a Profile on Every American Adult*, BLOOMBERG BUSINESSWEEK, Aug. 5, 2016 (describing the growing data broker industry including how online activity is tracked and added to profiles).

[29] *See* sources cited *supra* note 28. *See generally* Cohen, *supra* note 20.

monetizing an individual reader's activity online,[30] and customization of content for the researcher[31] all depend on some level of identification of the individual reader, perhaps by a non-library content provider or collaborator. Even if access to the digital content is itself funded through a model that does not require individuals to identify themselves, some data security solutions introduce similar requirements that prevent confidentiality of access.[32] Additionally, most access to content through the Internet includes layers of privacy risk that are largely opaque, as individuals' activity online is tracked across the web for commercial purposes and as a governmental tool to investigate and perhaps prevent criminal activity.[33] As information is collected without legal restrictions on the use of that data, reading habits are likely to become more integrated into the growing profiling industry.[34] Some question whether library users actually care about confidentiality of reading anymore, given demonstrated interest in sharing this information through social media and in their use of

---

[30] *See* Lorrie Cranor et. al., *Panel I: Disclosure and Notice Practices in Private Data Collection*, 32 CARDOZO ARTS & ENT. L.J. 784, 791-92 (reporting on consumer confusion about how to opt-out of behavioral advertising); Christopher A. Summers, Robert W. Smith & Rebecca Walker Reczec, *An Audience of One: Behaviorally Targeted Ads as Implied Social Labels*, 43 J. CONSUMER RES. 156, 157 (2016) ("By placing data onto consumers' hard drives (i.e., cookies), firms are able to collect information about consumers' viewing and clicking patterns, web searches, purchase histories, and social media use, from both their personal computers and mobile devices . . . Advertising networks then create a user profile from this data and deliver ads for products that their software predicts will be appealing to the individual consumer.").

[31] *See* Karen Coombs, *Privacy vs. Personalization*, 132 LIBR. J. 28 (2007) (advocating libraries follow the lead of other user experiences on the web to create customized services library users can use on the basis of an opt-in approach).

[32] Results of a 2016 survey of college and university libraries show that identities of library users are generally protected through existing systems that authenticate authorized users of content paid for by libraries. *See* Clifford Lynch, *Report on the CNI Authentication & Authorization Survey*, COAL. FOR NETWORKED INFO. (Aug. 2016), https://www.cni.org/go/report-authentication-survey-2016 [https://perma.cc/LFW9-YUQ6]. *But see* David Crotty, *Coming Soon: Battles Over Academic Privacy---But Is This Fight Already Over?*, THE SCHOLARLY KITCHEN, Aug. 5, 2015 ("Methods like two-factor authentication [coming soon to scholarly publications] involve a much more granular identification of the user, rather than just knowing that someone at University X is looking at a paper. While journal marketers and advertisers are both very excited about the new possibilities this will open up, they are in opposition to policies of academic libraries").

[33] *See* sources cited *supra* note 28.

[34] Neil Richards, *Intellectual Privacy*, 87 TEX. L. REV. 387, 432-45 (2008) (reviewing examples of harmful uses of data about individuals' reading habits).

other online services that track and potentially share details of their activity.[35] Collectively, these privacy pressures lead many to view access to digital content as conditioned upon the transformation of the reader into another data subject for the modern data market,[36] and transformation of the library into a disempowered player within the information system.[37]

---

[35] AM. LIBR. ASS'N, AS LIBRARIES GO DIGITAL, PRIVACY ISSUES EMERGE, NEWSL. ON INTELL. FREEDOM 3 (Jan. 2013) (quoting David Weinberger, co-director of the Harvard Library Innovation Lab as opining, "[t]he privacy that libraries traditionally have been preserving is not always valued by their patrons, especially in an age of social networking."); Joseph Esposito, *Libraries May Have Gotten the Privacy Thing All Wrong*, THE SCHOLARLY KITCHEN (Jun. 23, 2016), https://scholarlykitchen.sspnet.org/2016/06/23/libraries-may-have-gotten-the-privacy-thing-all-wrong/ [https://perma.cc/752C-W72E] (asserting that library users already trade away reader privacy in other contexts, so librarians should give up on preventing the collection of individual user data and instead focus on how the data should be collected and used).

[36] *See* Thomas L. Reinsfelder, *E-books and Ethical Dilemmas for the Academic Reference Librarian*, 55 THE REFERENCE LIBR. 151, 160-61 (2014) (outlining ways that e-book services violate reader privacy and advising "[w]hen a significant level of privacy may not be possible, a choice must be made to either decline the services being offered or clearly explain to patrons how their data may be used or shared."); Andromeda Yelton, *The Ethics of Ebooks*, LIBR. J., Sept. 12, 2012, at 30-31 (warning that "[t]he future of ebooks in libraries is about trade-offs among deeply held values" and suggesting that as publishers negotiate or even refuse to sell ebooks to libraries, "privacy questions lurk"); Michael Zimmer, *Privacy on Planet Google: Using the Theory of Contextual Integrity to Clarify the Privacy Threats of Google's Quest for the Perfect Search Engine*, 3 J. BUS. & TECH. L. 109, 111-14 (2008) (describing the choice between using Google services and preserving privacy as a Faustian bargain); Deborah Caldwell-Stone, *A Digital Dilemma: Ebooks and Users' Rights: New Technology May Prove Inhospitable to Privacy*, 43 AM. LIBR. 18 (2012) (describing the status quo of library ebook options as conditioned upon comprising reader privacy). Some characterize the shift from print to digital access to library information as the time when libraries are transformed from public entities to commercial entities. *See* Carla Hesse, Dean of Social Sciences, Univ. of Cal., Berkeley, Remarks at Public Access and Google Books Settlement Conference (Aug. 28, 2009), http://www.ischool.berkeley.edu/newsandevents/events/20090828googlebooksconference [https://perma.cc/57MZ-QJPP] (session 4 at 14:30-24:49); *see also* Trina Magi, *Fourteen Reasons Privacy Matters: A Multidisciplinary Review of Scholarly Literature*, 81 LIBR. Q. 187, 188 (2011) ("In light of [digital technology trends] . . . some librarians may question whether the right to privacy is worth the trouble it takes to protect it.").

[37] Seeta Pena Gangadharan, *Who Is in Control of Your Library's Data?*, FUTURE TENSE (Nov. 10, 2015), http://www.slate.com/articles/technology/future_tense/2015/11/libraries_need_to_protect_patron_data_as_they_turn_high_tech.html [https://perma.cc/JC5X-66L4].

B. *Two Examinations of Digital Library Privacy*

Several studies have identified weaknesses in reader privacy, including library user privacy. Recent literature of librarianship is described as only minimally addressing privacy in articles recommending library implementation of interactive, customized, and other developing information technologies.[38] Large library licensing contracts have been examined for terms that would protect reader privacy, and the vast majority of these contracts were found to have failed to address the issue.[39] The privacy policies and practices of vendors who commonly supply content to libraries have been examined and found in many cases to be unclear or unsupportive of options for users to protect their privacy.[40] Some research shows that implementation of privacy protections and investment in privacy advocacy at the individual library level is likely not living up to the ethical commitments that librarians ostensibly continue to support.[41] Digitization projects, such as the Google Books initiative that draws on library research collections for much of its content, have been criticized for not extending the same protections for reader privacy that are afforded to traditional library users.[42] Additionally, the applicability of library privacy laws to modern information systems has been questioned.[43] On the other hand, an initial review of privacy and security protections for the integrated library systems that support acquisitions, cataloging, online catalogs and related services

---

[38] *See generally* Michael Zimmer, *Assessing the Treatment of Patron Privacy in Library 2.0 Literature*, 32 INFO. TECH. & LIBR. 29 (2013) (surveying professional literature of librarianship from 2005 to 2011).

[39] Alan Rubel & Mei Zhang, *Four Facets of Privacy and Intellectual Freedom in Licensing Contracts for Electronic Journals*, 76 C. & RES. LIBR. 425 (2015) (evaluating forty-two license agreements from libraries and finding them to inadequately protect patron privacy).

[40] *See* April Lambert, Michelle Parker & Masooda Bashir, *Library Patron Privacy in Jeopardy: An Analysis of the Privacy Policies of Digital Content Vendors*, 52 PROCEEDINGS OF THE ASS'N FOR INFO. SCI. AND TECH. 1, 7 (2015); Trina J. Magi, *A Content Analysis of Library Vendor Privacy Polices: Do They Meet Our Standards?*, 71 C. & RES. LIBR. 254 (2010).

[41] Michael Zimmer, *Librarians' Attitudes Regarding Information and Internet Privacy*, 84 LIBR. Q. 123, 147-48 (2014).

[42] ZIMMER, *supra* note 10.

[43] Ard, *supra* note 20, at 25-26.

demonstrated that many commonly used systems have privacy-protective functionalities.[44]

This article adds to these reviews of the state of library use privacy, particularly by focusing on evidence of librarians' commitments through recent library association reader privacy statements and guidelines, and by looking at the related Google Books and HathiTrust digitization projects to consider how reader privacy may or may not be protected in these contexts. Part II examines sample library association statements and guidance regarding reader privacy. Part III looks at the related Google Books and HathiTrust projects for reader privacy protections. Part IV concludes with optimism that library digital collaborations can foster both access and privacy through implementation of incremental protections already available and development of some new tools.[45]

II. LIBRARY PRIVACY STATEMENTS

A. *The American Library Association Statements on Confidentiality of Library Use*

The ALA, founded in 1876, is said to be "the oldest and largest library association in the world."[46] This grand Association has developed a robust, diverse, nuanced, and active approach to advancing the confidentiality of library use. The ALA has developed and updated a generous number of goals, statements, guidelines, and

---

[44] Marshall Breeding, *The Current State of Privacy and Security of Automation and Discovery Products*, 52 LIBR. TECH. REP. 13, 31 (2016).

[45] This article sometimes uses the terms "library use privacy," "library user privacy," and "reader privacy" interchangeably even though they each have distinct meanings. Reader privacy, for example, could apply well beyond the confines of library use and could address use of materials through a library's digital commercial partner. But, reader privacy might not properly describe confidentiality of a library user's listening to sound recordings or watching of films from a library collection. Similarly, confidentiality is the most apt term for protection of information a library user shares with a library rather than a sort of absolute secrecy sometimes associated with the term privacy. But, the term privacy is also used in this article because many state laws use this term, and because some library users might actually hope to make use of a library without creating or sharing any identifying trail.

[46] *About ALA*, AM. LIBR. ASS'N, http://www.ala.org/aboutala/ [https://perma.cc/ZRX2-BL4Y] (ALA describes its mission as "to provide leadership for the development, promotion and improvement of library and information services and the profession of librarianship in order to enhance learning and ensure access to information for all.").

programs for engagement on the issues relating to library privacy. These publications and efforts demonstrate an active commitment to reader privacy, as well as an awareness of the challenges posed by competing interests, particularly in the online environment. These ALA statements convey a deep engagement with developing pressures on reader privacy, an embrace of the advocacy role of the ALA, and a call to all individuals with control over the reality of library use confidentiality to integrate the ethic of privacy into practice.

At the highest level, the Strategic Plan for ALA includes only three areas of focus, and all three could in some way relate to a commitment to reader privacy. In June 2015, the ALA Council articulated three strategic directions of Advocacy, Information Policy, and Professional and Leadership Development.[47] At the same time, the Council highlighted nine core values, including ethics, professionalism and integrity; intellectual freedom (which the ALA asserts requires intellectual privacy);[48] and social responsibility and the public good. The challenges of promoting reader privacy fit neatly within this plan and within these core values, and other ALA statements and programs illustrate how library use privacy remains an important commitment of the ALA.

The *ALA Code of Ethics* is widely cited as evidence of librarians' commitment to protecting the confidentiality of library users. The 1939 ALA Code may be the first recognition of this ethical orientation with inclusion of the statement, "[i]t is the librarian's obligation to treat as confidential any private information obtained through contact with library patrons."[49] The most recent iteration of the ALA Code of Ethics, dated 2008, articulates a more explicit commitment to the privacy interests related to library use. This Code states, "[w]e protect each library user's right to privacy and confidentiality with respect to information sought or received and resources consulted, borrowed, acquired or transmitted."[50]

---

[47] *Id.*

[48] Magi & Garner, *supra* note 25 (asserting that privacy is necessary for intellectual freedom).

[49] AM. LIBR. ASS'N, CODE OF ETHICS FOR LIBRARIANS (1939) (included in Section II titled "Relation of the Librarian to His Constituency," as point number 11).

[50] AM. LIBR. ASS'N, CODE OF ETHICS OF THE AMERICAN LIBRARY ASSOCIATION, http://www.ala.org/advocacy/sites/ala.org.advocacy/files/content/proethics/codeofethics/Code%20of%20Ethics%20of%20the%20American%20Library%20Association.pdf

In 2014, the ALA Council reviewed and updated a more granular companion to the Code with *Privacy: An Interpretation of the Library Bill of Rights*. The statement includes an introduction with justifications for and history of librarians' protection of reader privacy, a section on the rights of library users, and an assertive description of responsibilities of library users and all persons involved in the provision of library services to respect others' privacy. The statement concludes with, "The [ALA] affirms that rights of privacy are necessary for intellectual freedom and are fundamental to the ethics and practice of librarianship."[51]

This 2014 ALA interpretation reveals both a fierce commitment to reader privacy and an awareness of the difficulties libraries face when attempting to manage the confidentiality of library use given new pressures to track individuals. The section on responsibilities begins with, "[t]he library profession has a long-standing commitment to an ethic of facilitating, not monitoring, access to information."[52] The ALA statement also includes the assertion, "[r]egardless of the technology used, everyone who collects or accesses personally identifiable information in any format has a legal and ethical obligation to protect confidentiality." [53]

This statement asserts that not only librarians, but also all those involved in providing access to information through a library have an ethical obligation to avoid compromising confidentiality of library use. This ethical assertion is, of course, aspirational, but it reveals an orientation to the culture of the library rather than to the profession of librarianship. Librarians themselves are not in a regulated profession such as law practice which requires lawyers to uphold rules of professional responsibility requiring confidential treatment of client matters or risk their license to practice or claims of malpractice.[54] So,

---

[https://perma.cc/J3Z2-Y73E] (adopted at the 1939 Midwinter Meeting by the ALA Council; amended June 30, 1981; June 28, 1995; and Jan. 22, 2008).

[51] *Privacy: An Interpretation of the Library Bill of Rights*, AM. LIBR. ASS'N, http://www.ala.org/advocacy/intfreedom/librarybill/interpretations/privacy [https://perma.cc/4YES-58G6].

[52] *Id.*

[53] *Id.*

[54] Am. Libr. Ass'n Committee on Prof. Ethics, *Enforcement of Code of Ethics of the American Library Association: Questions and Answer*, AM. LIBR. ASS'N (Jan. 2009), http://www.ala.org/advocacy/proethics/explanatory/enforcementfaq [https://perma.cc/

librarians' ethical commitments are largely a collective norm taught in Masters programs where librarians gain the credential generally required for employment. These norms are advanced by many librarians through library association structures, through local programs, and through some library and information science literature. Perhaps it should not be surprising that librarians, who make up the bulk of the membership in the ALA, have asserted this ethical obligation of library use privacy should apply to all persons involved in the provision of library services. Ethical obligations in the library context are not a matter of meeting licensing rules, but of advancing the collectively shaped culture of the library itself. This expansive view of library ethics for confidential library use mirrors the general approach of state library privacy statutes, which apply to the protection of the library user no matter who is in a position to control that outcome.[55]

Another way that the ALA promotes confidentiality of library use is through guidelines developed by the ALA Office of Intellectual Freedom (OIF). The *Privacy Tool Kit*, which was updated in 2014, contains a wealth of resources. One component is "Sections or Issues to Include in a Privacy Policy." The ALA privacy policy guidance suggests:

1. Notice & Openness;
2. Choice & Consent;
3. Access by Users;
4. Emerging Technologies;
5. Data Integrity & Security;
6. Enforcement & Redress;
7. Government Requests for Library Records;
8. Special Privacy Considerations (for different contexts and users).[56]

GFN3-94AN] ("Only those organizations with some kind of license or certification that can be withdrawn seem to have enforceable codes.").

[55] For links to state library privacy laws, *see* Privacy & Surveillance, AM. LIBR. ASS'N, http://www.ala.org/advocacy/privacyconfidentiality [https://perma.cc/ZGG5-C9XB].

[56] The Office of Intellectual Freedom of the American Library Association provides substantial annotations to this guide to library privacy policies. *See Privacy Tool Kit*, AM. LIBR. ASS'N, http://www.ala.org/advocacy/privacyconfidentiality/toolkitsprivacy/privacy [https://perma.cc/GDN2-YTB5] (revised by the IFC Privacy Subcommittee and approved by the Intellectual Freedom Committee January 2014).

The enumerated areas to address in library privacy policies overlap with Fair Information Practice Principles that have been influential in the privacy law of the United States and in other contexts and indicate an awareness among ALA leaders of how libraries' risks fit into common approaches to protecting privacy and how they may have special concerns.[57] Generally common to all such guidelines are the concepts of Notice and Openness, Choice and Consent, Access by Users, and Data Integrity and Security. Enforcement and Redress is not part of most articulations of fair information practices.[58] The remaining sections address particular challenges common to libraries with emerging technologies and government requests for library records, and special concerns for other contexts which address issues such as school libraries. Each of these proposed sections for library privacy policies is given more detailed treatment through the *Tool Kit*.

Consistent with the approach of the *Code of Ethics*, the *Privacy Tool Kit* asserts that all stakeholders with the authority to shape the privacy culture of the library have ethical obligations to protect

---

[57] SECRETARY'S ADVISORY COMM. ON AUTOMATED PERS. DATA SYS., DEP'T OF HEALTH EDUC. & WELFARE, RECORDS, COMPUTER, AND THE RIGHTS OF CITIZENS: REPORT OF THE SECRETARY'S ADVISORY COMMITTEE ON AUTOMATED PERSONAL DATA SYSTEMS (1973), https://aspe.hhs.gov/report/records-computers-and-rights-citizens [https://perma.cc/T7NP-LLNU]. This report outlines five Fair Information Practice principles that can be summarized as: (1) no secret data systems; (2) individuals must have access to data about themselves; (3) no secondary uses of data without consent; (4) individuals must be able to correct or amend data about themselves; and (5) collectors of data must ensure reliability and prevent misuse. The 1973 Fair Information Practice guidelines and other similar privacy principles and practices are linked in the American Library Association (ALA) *Privacy Tool Kit*. These Fair Information Practices were developed during the same period when most library use confidentiality laws were passed, and some components are evident in the way that these laws were framed. *See* ROBERT GELLMAN, FAIR INFORMATION PRACTICES: A BASIC HISTORY 1 (June 17, 2016), bobgellman.com/rg-docs/rg-FIPShistory.pdf [https://perma.cc/DB78-LZ65]. The Fair Information Practices are also incorporated into international agreements and have been described as a bridge between differing approaches to privacy and a common language for privacy. *See* Paula Breuning, *Fair Information Practice Principles: A Common Language for Privacy in a Diverse Data Environment*, POLICY@INTEL (Jan. 28, 2016), https://blogs.intel.com/policy/2016/01/28/blah-2/ [https://perma.cc/34E4-NL6V]. *But see* Julie Cohen, *DRM and Privacy*, 18 BERKELEY TECH. L.J. 575, 601-04 (2003) (critiquing fair information practices as a poor fit for intellectual consumption).

[58] *See* SECRETARY'S ADVISORY COMM. ON AUTOMATED PERS. DATA SYS., *supra* note 57; ORG. FOR ECON. COOPERATION & DEV., THE OECD PRIVACY FRAMEWORK 11 (1980) (revised 2013) [hereinafter OECD, *Guidelines*], http://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf [https://perma.cc/PRZ3-SR86]; Council Directive 95/46, art. 25, 1995 O.J. (L 281) (EC).

confidentiality of library use. The *Tool Kit* provides a section on "Implementation: A Checklist for Developing Privacy Procedures" that has subsections for governance bodies and policy makers, directors and supervisors, staff, and information technology services staff.[59] Directors and supervisors are advised to make sure contracts with library systems and other vendors are consistent with library privacy policy. They are also advised to ensure that subscription databases allow anonymous searching. The checklist for directors and supervisors also recommends a retention schedule for all personally identifying information regarding library users, presumably to encourage discarding of this information when it is no longer needed for core library uses. Information technology staff members are similarly advised to incorporate privacy into the selection of technologies and to provide notice to users when particular activities could put reader privacy at risk.

The *Privacy Tool Kit* links to a wealth of additional resources relating to library privacy, including suggested talking points,[60] information about ALA advocacy,[61] and an appendix with links to an array of legal, policy, and technology perspectives about library user privacy.

During 2015-2016, the OIF developed a series of specialized guidelines for particular areas of privacy risk. One addresses privacy issues related to e-book lending and digital content providers, which was approved by the OIF in June 2015.[62] Another focuses on privacy in the context of data exchange in networked devices and services and was approved in June 2016.[63] A third, also approved in June 2016,

---

[59] *Privacy Tool Kit*, *supra* note 56.

[60] *Library Privacy Talking Points: Key Messages and Tough Questions*, AM. LIBR. ASS'N, http://www.ala.org/advocacy/privacyconfidentiality/toolkitsprivacy/library-privacy-talking-points-key-messages-and-tough-questions [https://perma.cc/3UL4-HJNJ].

[61] *Advocacy at the Local, State, & National Levels*, AM. LIBR. ASS'N, http://www.ala.org/advocacy/privacyconfidentiality/toolkitsprivacy/advocacy-local-state-national-levels [https://perma.cc/XXP7-VA7V].

[62] *Library Privacy Guidelines for E-book Lending and Digital Content Vendors*, AM. LIBR. ASS'N, http://www.ala.org/advocacy/library-privacy-guidelines-e-book-lending-and-digital-content-vendors [https://perma.cc/JR6C-JMJV].

[63] *Library Privacy Guidelines for Data Exchange Between Networked Devices and Services*, AM. LIBR. ASS'N, http://www.ala.org/advocacy/library-privacy-guidelines-data-exchange-between-networked-devices-and-services [https://perma.cc/YH7D-PW4E].

addresses library websites and discovery systems.[64] These guidelines include suggestions for encryption and for regular audits of systems to make sure they remain privacy protective. They also promote the use of defaults that protect privacy while allowing library users to opt-in, with clear notice about privacy risks, to services that would preserve or share their personally identifying information.[65] The guidelines suggest users should have the opportunity to discontinue the collection and retention of their data and be able to have accumulated data destroyed.[66]

These most recent guidelines from the OIF are the most telling about librarians' ongoing commitment to reader privacy. Approaches follow general privacy management trends toward notice and consent, bolstered by some specific recommendations to give readers some choices to control collection and post-collection use.[67] The OIF guidelines retain strong reminders of library ethics of privacy and of state and other laws that may protect library use privacy. However, they also address serious problems of control and competing interests of third parties and libraries themselves. One example of this recognition of loss of control is the description in the guidelines addressing websites:

> "Library websites, OPACs, and discovery services may collect personal information about patrons for a variety of reasons including authentication, personalization, and user analytics. In addition, personal information is sometimes shared with third parties that provide content or other functionality for the website or service."[68]

As in the *ALA Code of Ethics*, as well as the *Privacy Tool Kit*, these guidelines impose the ethical burden of privacy protection on all who

---

[64] *Library Privacy Guidelines for Library Websites, OPACs, and Discovery Services*, AM. LIBR. ASS'N, http://www.ala.org/advocacy/library-privacy-guidelines-library-websites-opacs-and-discovery-services [https://perma.cc/U8VC-KKSM].

[65] *Id.*

[66] *Id.*

[67] *See* sources cited *supra* note 57.

[68] *See* sources cited *supra* note 64.

have an impact on library service. But, of course, that imposition is merely precatory.

Overall, the ALA conveys a strong ongoing commitment to the value of reader privacy and a fairly detailed practical approach to support the efforts of individual libraries. In addition, the ALA has active advocacy programs to promote strengthening of privacy laws and to promote awareness of privacy risks at the individual library level.

B. *International Federation of Library Associations and National Information Standards Organization Statements Regarding Library Use Confidentiality*

Librarians affiliate and are active through several different professional associations besides the ALA. This section looks at two new articulations of principles or practices for libraries and even their information system partners. In August 2015, The International Federation of Libraries (IFLA) Governing Board endorsed a *Statement on Privacy in the Library Environment*.[69] In December 2015, after a series of workshops with various stakeholders, the National Information Standards Organization (NISO) developed another set of principles to address the digital privacy of users of library, publisher, and software-provider systems.

The IFLA Statement describes the threats to library users' privacy from collection and sale of data about Internet users and their behavior.[70] The overall approach is at a broader level than some of the ALA guidelines and combines steadfast commitment and strong recommendations along with language conveying disappointment about the lack of control that libraries have over privacy of library use. The IFLA Statement asserts privacy as a human right, and references the *IFLA Code of Ethics*' articulation of respect for privacy.[71]

---

[69] INT'L FED'N OF LIBR. ASS'NS & INSTS., IFLA STATEMENT ON PRIVACY IN THE LIBRARY ENVIRONMENT, http://www.ifla.org/files/assets/hq/news/documents/ifla-statement-on-privacy-in-the-library-environment.pdf [https://perma.cc/4Y4R-AH8M].

[70] *Id.* at 1.

[71] *Id.* (stating that the IFLA Code of Ethics "identifies respect for personal privacy, protection of personal data, and confidentiality in the relationship between the user and library or information service as core principles."); *see IFLA Code of Ethics for Librarians and other Information Workers*, INT'L FED'N OF LIBR. ASS'NS & INSTS., http://www.ifla.org/news/ifla-code-of-ethics-for-librarians-and-other-information-workers-full-version [https://perma.cc/L4S7-Z8YD].

Somewhat in contrast with the ALA approach, IFLA identifies the "library and information services" as the entities responsible for protecting library use privacy. IFLA makes a distinction between library and information services and commercial content and service providers without taking the ALA stance that all of the entities involved in providing library services have a common ethical obligation for library use privacy.[72]

The IFLA Statement conveys a sense of resignation about the powerlessness of libraries to provide enough privacy to meet the aspirations of its members. The Statement acknowledges the library has control only over its own systems and practices.[73] The Statement describes options for the library in providing services that might have privacy-hostile features, including negotiation, refusal to acquire, and limitations on implementation of these services.[74] The IFLA Statement then adds, "[h]owever, library and information services' opportunities to influence, regulate or gain reliable knowledge of the data collection practices of commercial vendors or government institutions may be limited."[75]

Following this dour description of the abilities of libraries to protect reader privacy, the IFLA Statement next introduces eight broad recommendations for both libraries and information services:

---

[72] The IFLA inward focus on libraries and similar information services is also in contrast with the approach of another international library association, the International Coalition of Library Consortia (ICOLC). The ICOLC Guidelines speak directly to online vendors: "[T]he ICOLC issues these guidelines with respect to the privacy interests of our member libraries' users in the interest of informing the companies with which we do business about what is acceptable in the products and services that we license." The guidelines articulate the need for privacy policies and for access to products even when library users do not wish to allow personally identifying information to be shared with third parties. *Privacy Guidelines for Electronic Resource Vendors*, INT'L COAL. LIBR. CONSORTIA (July 1, 2002), http://www.icolc.net/statement/privacy-guidelines-electronic-resources-vendors [https://perma.cc/V2CF-QA2B].

[73] INT'L FED'N OF LIBR. ASS'NS & INSTS., *supra* note 69, at 2 ("Library and information services have the opportunity to make independent decisions about local system and data management. Library and information services can decide what kind of personal data they will collect on users and consider principles of data security, management, storage, sharing and retention.").

[74] *Id. See* Ard, *supra* note 20 (describing how libraries have been unable to secure privacy features in e-books and other digital content services because of their lack of bargaining power).

[75] INT'L FED'N OF LIBR. ASS'NS & INSTS., *supra* note 69, at 2.

1. Respect and advance privacy both at the level of practices and as a principle;
2. Support advocacy for and reflection on privacy and digital rights;
3. Reject electronic surveillance and limit data collection on library users' activities;
4. Ensure government access is necessary and proportionate to legitimate aims;
5. Educate library users regarding data and privacy risks with particular resources;
6. Support users' informed choices;
7. Develop media and training programs regarding privacy risks and protections;
8. Include data and privacy protection principles and practices in library and information professionals' education.[76]

The issues of library privacy have also caught the attention of the National Information Standards Organization, or NISO. In December 2015, after a series of workshops with various stakeholders, NISO developed a set of principles to address the digital privacy of users of library, publisher, and software-provider systems. The Preamble to the principles addresses the problem of libraries' lack of control over other information systems that monitor library users' activities. The Preamble states that all participants in this system "have a shared obligation to foster a digital environment that respects library users' privacy."[77] This call for the shared obligation is similar to the ALA emphasis on ethical commitments being tied to the library rather than to the profession of librarianship. The guidelines include some methodologies not unlike the ALA privacy policy and include some of the strategies addressed through the ALA OIF specialized guidelines.

These NISO Privacy Principles can be summarized as:

I. Shared Privacy Responsibilities (between libraries and those who operate through and for the library);

---

[76] *Id.*

[77] NAT'L INFO. STANDARDS ORG., NISO CONSENSUS PRINCIPLES ON USER'S DIGITAL PRIVACY IN LIBRARY, PUBLISHER, AND SOFTWARE-PROVIDER SYSTEMS (2015), http://www.niso.org/apps/group_public/download.php/16064/NISO%20Privacy%20Principles.pdf [https://perma.cc/BLT2-DCW6].

II. Transparency and Facilitating Privacy Awareness;
III. Security of Data;
IV. Data Collection and Use Balanced Against User Privacy;
V. Anonymization and Limited Retention of Data;
VI. Options and Informed Consent for Secondary Uses;
VII. Limited and Anonymized Sharing Data with Others Necessary to Service;
VIII. Notification of Privacy Policies and Practices;
IX. Accommodations for Anonymous Uses;
X. Access to One's Own User Data;
XI. Continuous Improvement of Privacy Protections;
XII. Accountability including Reviews, Reports, and Audits.[78]

The fourth NISO Privacy Principle frames data collection and use as practices to be balanced against user privacy, the core concern expressed about libraries' move to the digital environment. As libraries collaborate to improve access and preserve materials through digitization, the complaint is that user privacy is weakened. Given that these are "Consensus" principles drafted with libraries, publishers, and software-producers through workshops with all of these stakeholders in the current information system, a balancing approach would address tensions. However, the NISO Principles may also reflect receptivity in the library community to collecting, storing, and analyzing library user data in ways that typically would not have been encouraged under traditional approaches to library privacy.[79] Other NISO principles, though, promote privacy-enhancing protections that lean on the market-centered notice and choice model, with transparency, user control, and accountability.[80] The NISO principles also encourage data security, and offer several approaches to anonymized access.

---

[78] *Id.*

[79] *See* Ken Varnum, *Editorial Board Thoughts: Library Analytics and Patron Privacy*, 34 INFO. TECH. & LIBR., 2-3 (2015) (advocating for greater library reliance on user data and compromise on privacy protections, noting "As a profession, we have begun to realize that the straightforward (and arguably simplistic) approaches we have relied on for so long may no longer be appropriate or helpful.").

[80] Not unlike the Fair Information Practices. *See* sources cited *supra* note 57.

The ALA, IFLA, and NISO statements and guidelines are all very recent articulations of commitment to protecting privacy of library users. They have been developed concurrently with the explosion of data mining of individuals' access to information. So, while the challenges of addressing hidden privacy risks continue to expand, librarians and other library-focused stakeholders continue to update policies and guidelines instead of letting the reader privacy ethic die a quiet death.[81] All approaches embrace the value of reader privacy, but each includes some level of engagement with, frustration with, or embracing of the concept that in the digital environment, library privacy must be balanced somewhat with other interests.

III. GOOGLE BOOKS AND HATHITRUST AND READER PRIVACY PROTECTIONS

A. *Google Books*

1. *Google Books and Libraries*

In 2016, at the time of this writing, Google Books offered site visitors the opportunity to search a vast database of full text, scanned copies of books from cooperating library collections and from publisher and author "partners." Those books with existing copyright protections were only displayed through excerpts of content called "snippets," and books with expired copyright or otherwise not protected by copyright could be viewed in full through Google Books search.[82] Google links to sites where books found through the search engine could be purchased. Although Google suggests the idea for Google Books dates back to the beginning of Google in 1996, active

---

[81] Sarah Shik Lamden, *Why Library Cards Offer More Privacy Rights than Proof of Citizenship: Librarian Ethics and Freedom of Information Act Requestor Policies*, 30 GOV'T INFO. Q. 131, 134 (2014) (describing library association reader privacy guidelines and library reader privacy practices as constantly updated to address modern technologies in comparison with the lack of researcher privacy under the federal Freedom of Information Act).

[82] *What You'll See When You Search on Google Books*, GOOGLE BOOKS, https://www.google.com/googlebooks/library/screenshots.html [https://perma.cc/UFA4-CE4E].

collaborative planning to draw on the collections of major research libraries began in 2002.[83]

The game-changing digitization capabilities of Google were, and remain, compelling. Many major libraries in the United States and elsewhere were enthusiastic partners in the Google Books venture at its initiation in 2002 with a conversation between Larry Page at Google and librarians at his alma mater, the University of Michigan. As the early efforts gained steam, privacy was not a deal-breaker for librarians or library institutions, perhaps because the benefits were significant, the privacy risks for readers were unclear, and copyright issues were distracting. This section reviews some of the Google Books interactions between libraries and Google, and then looks to current Google Books privacy policies to assess whether reader privacy is diminished as access to content is expanded.

2. *From 1,000 Years to Six to Create Access – A Bargain at the Price of Reader Privacy?*

The first reason privacy may not have been high on librarians' list for Google Books was that the promised increase in access was a startling, transformative, and, at the time, unique opportunity. Libraries were attempting to digitize crumbling books to preserve their content, but their progress was frustratingly slow.[84] When Google entered the discussion, libraries' goals for preservation and for access to these materials were given a major boost. Google developed a plan to digitize selected research libraries' collections and give participating libraries digital copies with the returned print books.[85] In addition, Google was developing its own product that would be a

---

[83] *Google Books History*, GOOGLE BOOKS, https://www.google.com/googlebooks/about/history.html [https://perma.cc/6XDM-6SE2]; for the record, Google explains its name comes from a play on the word "googol" which is described as the mathematical term for a "1" followed by 100 zeros. *Company Overview*, GOOGLE CO., https://www.google.com/intl/en/about/company/ [https://perma.cc/LH95-TRPS].

[84] Kevin Bergquist, *Google Project About the Public Good*, THE UNIV. REC. ONLINE (Feb. 8, 2006), http://www.ur.umich.edu/0506/Feb06_06/22.shtml [https://perma.cc/2SWG-B6Q6]; Ron Chepesiuk, *Digitizing Rare Materials: Special Collections Go Global*, 32 AM. LIBR. 54 (2001) (writing before the Google Books project and describing the high costs of research libraries' efforts to digitize rare materials for preservation and access).

[85] *Google Books History*, *supra* note 83.

simple-to-use and massive database for access.[86] Just how that access would work was not clear in the beginning, but the collaboration held enormous promise for speeding up the digitization process that libraries were already undertaking.[87]

Library partners might have consciously decided or inadvertently acted as if the tradeoff of reader privacy for such advances in access and preservation was a good choice. Larry Page offered to reduce the University of Michigan Library's estimated time for digitizing its seven million volume collection from 1,000 years to six.[88] In the following year, the company developed new scanning technologies and search technologies for the project.[89] In 2006, Mary Sue Coleman, the University of Michigan President at the time, defended the partnership saying that the project advanced the university's highest ideals of promoting and sharing knowledge. She explained that prior to the collaboration with Google, the University Library was able to digitize between 5,000 and 8,000 volumes annually, not enough to protect decaying resources for future generations.[90] "I believe this venture with Google is one of the best answers we have to sharing knowledge on a global plane," she asserted.[91] "The soul of scholarship is research. From the current to the ancient, we must make all information discoverable to faculty, students, and the public."[92]

---

[86] *Id.*

[87] Barbara Quint, *The Day the World Changed: Google Takes Command*, 22 INFO. TODAY 7 (2005) (highlighting unknowns in 2005 including how long the project would take and how access might change the information market); Jonathan Band, *The Google Library Project Both Sides of the Story*, 10 INFO. OUTLOOK 35 (2006) (describing the state of the Google Books project in 2006, including some unsettled aspects of funding and access).

[88] *Google Books History*, *supra* note 83 ("When [Google co-founder Larry Page] learns that the current estimate for scanning the university library's seven million volumes is 1,000 years, he tells university president Mary Sue Coleman he believes Google can help make it happen in six."); *see also* Jessica Dye, *Scanning the Stacks: The Digital Rights Issues Behind Book Digitization Projects*, 29 ECONTENT 32, 34 (2006) (quoting Professor James Hilton, associate provost and interim university librarian for the University of Michigan, saying Google's help reduced the time required to digitize the library's seven million volumes from over 1,000 years to six).

[89] *Google Books History, supra* note 83.

[90] Bergquist, *supra* note 84.

[91] *Id.* (quoting Mary Sue Coleman).

[92] *Id.*

Legal scholar Pamela Samuelson, writing in 2010, characterized the Google expenditures as minimally thirty dollars per book to scan, and with a goal of twenty-million books to be scanned, the overall cost would be at least $600 million.[93] Google's ability to surmount the cost barrier to large-scale digitization was an amazing contribution to libraries' dreams of preserving crumbling collections and extending access.[94] Indeed, Google brought money, talent, and commitment to the project that transformed the way that these major research libraries could address core goals of preserving and providing access to information.

Contracts with participating libraries contain some terms that could relate to privacy of individuals using Google's interface for access to the scanned content. Although not all contracts with university and research libraries were made public, some are posted on an academic website, and some of the posted contracts have provisions relating to privacy.[95] The 2004 contract between the University of Michigan and Google has been characterized as having no provision for reader privacy.[96] However, it does include a provision requiring Google to post a notice of a privacy policy that governs the collection and use of information from individual users.[97] This requirement to post a privacy policy may satisfy the Notice, Openness,

---

[93] Citing Brewster Kahle's estimates on cost and Ken Auletta's estimate on the number of books to be scanned. KEN AULETTA, GOOGLED: THE END OF THE WORLD AS WE KNOW IT 258 (2009) (indicating that Google's goal for GBS is to scan twenty million books). Paul Courant estimated Google would scan 50 million books, so by his estimate the cost for Google could be $1.5 billion. *See* Pamela Samuelson, *Google Book Search and the Future of Books in Cyberspace*, 94 MINN. L. REV. 1308, 1311-12 (2010). An earlier estimate based on a presumed target of scanning thirty million books from the first five participating libraries was set at $750 million. Band, *supra* note 87, at 45.

[94] Samuelson, *supra* note 93.

[95] The Public Index project website contains copies of some of the agreements between libraries and Google for the Google Books project. This project, led by Professor James Grimmelmann, was undertaken by the Public-Interest Book Search Initiative and the Institute for Information Law and Policy at New York Law School. *See Library Documents*, THE PUBLIC INDEX, http://www.thepublicindex.org/filings/libraries [https://perma.cc/4FLA-XLKG].

[96] *See generally* SIVA VAIDYANATHAN, THE GOOGLIZATION OF EVERYTHING: (AND WHY WE SHOULD WORRY) (2012) (noting this absence and citing the contract).

[97] Coop. Agreement Between Google & U. of Mich., § 4.5.2 (2004) http://www.lib.umich.edu/sites/default/files/services/mdp/um-google-cooperative-agreement.pdf [https://perma.cc/A78Q-6H5P].

and Transparency principles that are a part of the ALA, IFLA, and NISO library privacy statements.[98] Another section of the contract somewhat oddly describes the ability of each contracting party to have access to confidential information of the other.[99] It is unclear what this statement was intended to cover. If Google gained access to the reading records of library partners in the Google Books project, that disclosure was not revealed.[100] Nor have libraries reported gaining access to proprietary algorithms or detailed business plans of Google.

The University of Michigan contract also contains a provision titled "Searching Free to the Public" that requires Google and its successors to make any Internet access to the scanned content available to site visitors "a display of search results that shall have no direct costs to end users."[101] The language of the discussions at the time points to a definition of "direct costs" as likely meaning financial costs.[102] As the next section discusses, the idea of tracking individuals' reader activity across the web and over time and then monetizing that data was not reported in the early years of the Google Books project.

3. *Unimagined Future Abilities to Track and Trade Reader Data – Why Contract Against Unknown Risks?*

A second reason that libraries collaborating in the development of Google Books did not raise reader privacy issues may be because the privacy issues were not obvious and likely were less serious at the time. The risks may have only been seen as a minimal intrusion of some sidebar advertisements that related to the Google Books search statement or the content of the selected book displayed by a

---

[98] These concepts of Notice/Openness/Transparency are also part of the Fair Information Practice Principles that have been influential in United States privacy law. *See* sources cited *supra* note 57.

[99] Coop. Agreement Between Google & U. of Mich., *supra* note 97, at § 6.

[100] One might expect privacy advocates would have raised this point in amicus briefs that were submitted in copyright litigation over Google Books. When the fairness of a proposed settlement was considered, privacy advocates did not suggest that libraries had actually opened up their patron records to Google. *See* sources cited *infra* note 115.

[101] Coop. Agreement Between Google & U. of Mich., *supra* note 97, § 4.3.

[102] Dye, *supra* note 88, at 33 (describing the Google Books plan as an aspiration to make "the biggest, most widely accessible library ever" through efforts to bring collections "online through its free search engine.").

researcher.[103] The idea that loss of reader privacy was deeply connected to funding for Google Books access does not appear to have been part of the early conversation. In 2006, the framework for the project included an expectation that Google would not display any advertisements next to snippets of books scanned from collaborating libraries but would rely on the appeal of the Books project to draw more users to Google in general, distinguishing itself from competitors in a broader search engine market and generating revenue indirectly.[104] At the same time, if publishers became part of a proposed Partner program allowing for full display of a book, the expectation was that publishers and Google might share revenue from contextual advertisements that simply matched the content of the displayed text with presumably related advertisements.[105] In 2010, the business model for Google Books was assumed to be based on online advertisements or subscription fees, although some fairly mysterious alternative business models were anticipated.[106] These early business model discussions focused on subscription fees and advertisements based on the content displayed rather than on a complex profile of the individual reader's more extensive viewing or other habits.[107] In fact, in 2009 Paul Courant, Director of the University of Michigan Libraries at the time, advanced the idea that Google's business interests in drawing many viewers who could see advertisements would be a positive option because it would help ease pressure for subscription fees that many feared could become exorbitant.[108]

Early in the project, commodification of identifiable individuals' online reading habits was not featured regularly in public discourse.

---

[103] Samuelson, *supra* note 93, at 1337-38 (2010). *See* Darnton, *supra* note 4, at 15 (describing "discreet advertisements").

[104] Band, *supra* note 87, at 45.

[105] *Id.* at 35.

[106] Samuelson, *supra* note 93, at 1330-44 (reviewing cost recovery possibilities for Google Books and noting that the service of ads beside scholarly reading might offend academics as transforming the scholarly enterprise into a shopping mall).

[107] *See generally id.*

[108] *Id.* at 1337; citing Paul N. Courant, *What's at Stake in the Google Book Search Settlement?*, ECONOMISTS' VOICE, Oct. 2009, at 5, http://www.bepress.com/ev/vol6/iss9/art7/ [https://perma.cc/GCH2-ZCXE] ("[I]t seems likely that Google is more interested in attracting people to its site than it is in profiting directly from sales of books, and hence would prefer prices to be low.").

Behavioral advertising is said to have developed during the time the Google Books project was launched, but this practice was largely invisible to consumers.[109] Google's reliance on this practice came to the forefront in 2007 with the announcement that Google and online advertising company Doubleclick intended to merge.[110] Several privacy organizations challenged the plan, but the Federal Trade Commission ultimately approved the merger, which was finalized in 2008.[111] The growth of the data broker industry to encompass data from online reading was also not much discussed in the early 2000's. The Federal Trade Commission report on Data Brokers in late 2014 may mark the time when the growing industry finally garnered widespread attention.[112]

So, libraries that partnered with Google Books early in the development of the project may not have understood that library user

---

[109] FED. TRADE COMM'N, F.T.C. STAFF REPORT: SELF-REGULATORY PRINCIPLES FOR ONLINE BEHAVIORAL ADVERTISING 1-2 (2009), https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-staff-report-self-regulatory-principles-online-behavioral-advertising/p085400behavadreport.pdf [https://perma.cc/69LX-8TSW] (both advocating a market approach to regulation of behavioral advertising and describing the practice as generally invisible to consumers); newspaper reports of e-book readers that track individuals' reading at a granular level did not reach a wide audience until 2012, when *The Wall Street Journal* ran an article titled, "Your E-Book is Reading You." Alexandra Alter, *Your E-Book is Reading You*, WALL ST. J., (July 19, 2012), http://www.wsj.com/articles/SB10001424052702304870304577490950051438304 [https://perma.cc/U27D-445E].

[110] FED. TRADE COMM'N, STATEMENT OF FEDERAL TRADE COMMISSION CONCERNING GOOGLE/DOUBLECLICK 1 (2007), https://www.ftc.gov/system/files/documents/public_statements/418081/071220googledc-commstmt.pdf [https://perma.cc/Q3QP-KMEB].

[111] *See* James Schedwin, Note, *Behavioral Targeting: Issues Involving the Microsoft-aQuantive and Google-DoubleClick Mergers, and the Current and Proposed Solutions to Those Issues,* 4 I/S: J.L. & POL'Y FOR INFO. SOC'Y, 709, 717-22 (2008) (chronicling the merger of Google and DoubleClick).

[112] *See generally* FED. TRADE COMM'N, DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY (2014), https://duckduckgo.com/l/?kh=-1&uddg=https%3A%2F%2Fwww.ftc.gov%2Fsystem%2Ffiles%2Fdocuments%2Freports%2Fdata-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014%2F140527databrokerreport.pdf [https://perma.cc/6PSR-GMCK]; *see also* Saranga Komanduri, Richard Shay, Gerg Norcie, Blase Ur & Lorrie Faith Cranor, *Adchoices: Compliance with Online Behavioral Advertising Notice and Choice Requirements,* 7 I/S: J.L. & POL'Y FOR INFO. SOC'Y 603, 603 (2012) (noting that online behavioral advertising had been tracking users across websites, often without their knowledge, and noting industry collaboration as early as 1999).

privacy, perhaps transformed into reader privacy through Google access, could be a major concession for users of that service. Again, paired with the vast expansion in access and possibilities for preservation, librarians might not have been willing to let the modest intrusion of contextual advertising stand in the way of the digitization partnership. So what if while reading about the history of Singapore a researcher were to see an ad for travel services for Singapore in the sidebar? That experience might be no different from advertisement placement in print publications, similarly preserving the disconnect between the advertisement placement and the individual reader.

4. *Privacy Raised in Litigation, But Other Issues Dominated*

Copyright challenges have been the most prominent legal challenge for the Google Books project. Books both in and out of copyright were scanned, made searchable, and content would have been made viewable in a few short excerpts related to search terms or in full text if the work were in the public domain or if Google had permission for full display. Rights holders were invited to submit a request if they did not want their books to be scanned.[113] But privacy concerns were also raised during litigation challenging Google Books.[114] Amicus briefs submitted by both privacy advocacy organizations and a group of library associations raised the alarm that reader privacy was endangered.[115] Privacy was raised as part of fairness considerations in a proposed settlement, and the federal district court opinion reviewing the settlement concluded "[t]he privacy concerns are real."[116] Nonetheless, the judge did "not believe they were a basis in themselves to reject the proposed settlement." In

---

[113] *See generally* Band, *supra* note 87.

[114] Copyright litigation resulting in a proposed and rejected settlement also addresses other issues such as antitrust and adequacy of class representation. *See* Authors Guild v. Google, Inc., 770 F. Supp. 2d 666 (S.D.N.Y. 2011).

[115] *See* Privacy Authors and Publishers' Objection to Proposed Settlement at 8, *Authors Guild*, 770 F. Supp. 2d 666 (No. 05 CV 8136-DC); Brief for the Center for Democracy & Technology as Amicus Curiae in Support of Approval of the Settlement and Protection of Reader Privacy at 7-11, *Authors Guild*, 770 F. Supp. 2d 666 (No. 05 CV 8136-DC); Memorandum of Points and Authorities in Support of EPIC's [Electronic Privacy Information Center] Motion to Intervene at 14-15, *Authors Guild*, 770 F. Supp. 2d 666 (No. 05 CV 8136-DC).

[116] *Authors Guild*, 770 F. Supp. 2d at 683.

defense of this position, the opinion refers to promises made by Google in their brief to the court to prevent sharing of personal information of copyright holders or of readers of Google Books. [117] The judge conceded, however, that these promises were "voluntary undertakings only" and added that he thought "certain, additional privacy protections might be incorporated [into the settlement], while still accommodating Google's marketing efforts."[118]

A 2011 court filing from Google, responding to privacy objections to the proposed settlement, provides insights into Google's interest in reader data, as well as their perspective on the existence of legal obligations to protect reader privacy. Google argued that demands for specific promises for future designs or features of the Google Books product were not a normal part of product development, and so were unreasonable.[119] Google also characterized the call to purge all logging data or other information related to individual users of Google Books after 30 days as a diminution in the service's capacity to support a user who might rely on the service to track an ongoing research project.[120] In addition, the purging would prevent Google from "recommending Books to users on the basis of an analysis of their long-term preferences."[121] These features of long-term retention and analysis of individually identifying reader habits appear to have been part of the product plan in 2011.

The same Google filing asserted other arguments that reader privacy law claims were either not valid given case law, or that the scope of protection generally was spotty and narrow. First, Google argued that any reader privacy protections that might be available under the First Amendment would not apply to Google as a private actor, and the court's approval of the copyright-focused settlement would not create state action necessary to trigger that Constitutional protection.[122] Google also asserted that reader privacy issues were

---

[117] *Id.* at 683-84.

[118] *Id.* at 684. Of course, since the settlement was not approved, the suggestion of adding more privacy protection to the terms was hypothetical.

[119] Brief of Defendant at 54, *Authors Guild*, 770 F. Supp. 2d 666 (No. 05 CV 8136-DC).

[120] *Id.* at 57.

[121] *Id.*

[122] *Id.* at 53.

beyond the scope of the copyright-focused pleadings, and so should not be part of the settlement review.[123] In addition, Google stated that strong reader privacy laws with limited jurisdiction or scope could not be imposed in a general fashion to Google Books,[124] and that Google should not be held to a higher standard for reader privacy than other Internet services.[125]

Google further suggests it made privacy-protecting concessions during settlement negotiations by applying the general Google Privacy Policy to Google Books and incorporating a reference to the potential applicability of special "books laws" with limited jurisdiction.[126] As the next subsection describes, in 2012 Google merged its privacy policies into one policy that describes how the company merges data collected from all of its services. However, Google Books retains a supplemental privacy policy that makes reference to these special "books laws."

5. *Google Books Privacy Policy, Funding Model*

As of 2016, Google placed "Privacy" as the first link at the bottom of the Google Books page, assuming reading from left to right. From there, a curious reader could spelunk for a long time to explore how the Google Books privacy policy might differ from the general Google privacy policy, how advertising was conducted, what choices a Google services user has, etc. Even though Google's privacy policy was

---

[123] *Id.* at 54.

[124] *Id.* at 56. These laws would likely refer to state library confidentiality statutes, and state constitutional claims.

[125] Plaintiff's Supplemental Memorandum Responding to Specific Objections at 53, *Authors Guild*, 770 F. Supp. 2d 666 (No. 05 CV 8136-DC) ("Google should not be required to make detailed privacy commitments in the [terms of the settlement] for services that have not even been designed yet."). But see subsequent calls from the Federal Trade Commission for "Privacy by Design" in which privacy commitments are built in to the design of products and services and considered at every stage in development of a product or service. FED. TRADE COMM'N, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: RECOMMENDATIONS FOR BUSINESSES AND POLICYMAKERS 22 (2012), https://www.ftc.gov/reports/protecting-consumer-privacy-era-rapid-change-recommendations-businesses-policymakers [https://perma.cc/8EJA-F38Q]; *see* Tattered Cover. Inc. v. City of Thornton, 44 P.3d 1044 (Colo. 2002) (as modified on denial of rehearing (Apr. 29, 2002)) (finding that the Colorado Constitution protects reader privacy in bookstore records of purchase and may represent a higher level of protection than under the First Amendment).

[126] Brief for the Defendant at 55, *Authors Guild*, 770 F. Supp. 2d 666 (No. 05 CV 8136-DC).

substantially consolidated across the various Google services in 2012, the amount of information available or necessary to understand the full scope of risks to personal privacy was layered in 2016. Some of that information was highly technical.[127] Commenters have interpreted the policy information in various ways, ranging from compliments on its transparency[128] to complaints of deliberate obfuscation.[129]

At the time of this writing, a visitor to the Google Books site who clicked on the webpage privacy policy would be linked to a 2011 "Google Play-Privacy Policy for Books." Google Play was the name for a marketplace of Google services and applications, and most of the language of the linked policy was for Google Play. A section of the policy did apply specifically to "[p]ractices specific to books on Google Play product," but the description appeared to cover the purchase of books.[130] This Google Play-Google Books policy began with a link to archived policies and then the general Google Privacy Policy. So, the curious reader would then leave this page to read the current general policy, which at the time of this writing was dated August 29, 2016. The general privacy policy was lengthy, and contained links to more

---

[127] Some technologies used by Google are linked to terms used in the Privacy Policy. A page with "Key Terms" defines such technologies as pixel tag, server logs, HTTP referrer, and unique device identifier. *See Key Terms*, GOOGLE, https://www.google.com/intl/en/policies/privacy/key-terms/ [https://perma.cc/X7X4-PSDJ]. Another page explains types of cookies used by Google. *See Types of Cookies Used by Google*, GOOGLE, https://www.google.com/intl/en/policies/technologies/types/ [https://perma.cc/GK3K-U4ZU] (using language such as "Our main advertising cookie on non-Google sites is named 'id' or 'IDE' and is stored in browsers under the domain doubleclick.net.").

[128] *See* Derek S. Witte, *Privacy Deleted: Is it Too Late to Protect Our Privacy Online?*, 17 J. INTERNET L. 1, 17 (2014) (describing the consolidated Google Privacy Policies as "easier to read and understand," if frightening, in its disclosures about collection and use of data).

[129] *See* Lorie Cranor, et al., *Panel 1: Disclosure and Notice Practices in Private Data Collection*, 32 CARDOZO ARTS & ENT. L.J. 784, 800-01 (2014) (Helen Nissenbaum describing the inadequacy of simple privacy policies to describe complex tracking practices like those of Google and the difficulty of most consumers in understanding a policy that is transparent because it is too complex); Omer Tene, *What Google Knows: Privacy and Internet Search Engines*, 2008 UTAH L. REV. 1433, 1467-69 (discussing the disadvantages to customers from Google's "self-serving" privacy policy being multi-layered, containing language that could be interpreted in various ways, and being subject to change that could reduce users' privacy protections).

[130] *Practices specific to books on Google Play product*, GOOGLE PLAY: PRIVACY POLICY FOR BOOKS (Oct. 13, 2011), https://books.google.com/intl/en/googlebooks/privacy.html [https://perma.cc/8F2W-DT6T].

detailed information. The PDF version was just over nine pages long. The policy was organized around what information was used and collected, choices available to a site visitor, compliance, information Google shares, information security, and several other topics including links to specific product policies. Google Play was one of the links for other product policies, and that link took the reader to a page dated December 16, 2015 with "Practices specific to Google Play Books."[131] This 2015 page contained text that was mostly the same as the 2011 text linked from the Google Books search page. Both contained notice that Google stores the last five pages of viewed book text for those with a Google account.[132]

To summarize the general Google privacy policy, a lot of information was collected from individual users of Google services generally, and Google explained that it merged information collected across its services and potentially through activity on other sites and apps and through Google Analytics.[133] As of this 2016 review, the company collected information about devices, including hardware model, operating system version, unique device identifiers, and mobile network information, including phone number. Log information including details of use such as search terms, IP address, cookies that uniquely identify the web browser software, and pixel tags was also collected.[134] The general policy also stated "[w]e may share non-personally identifiable information publicly and with our partners – like publishers, advertisers or connected sites."[135] The link

---

[131] *Practices specific to Google Play Books*, GOOGLE PLAY (Dec. 16, 2015), https://play.google.com/books/intl/en/privacy.html [https://perma.cc/8BDD-F75Y].

[132] *See* sources cited *supra* notes 130-31.

[133] *How We Use Information We Collect*, GOOGLE PRIVACY & TERMS, https://www.google.com/policies/privacy/#application [https://perma.cc/R9ZT-TE8F]. For a definition of "your activity on other sites and apps", *see* GOOGLE PRIVACY & TERMS, https://www.google.com/policies/privacy/ [https://perma.cc/MM83-4EMK] (providing notice that if a website uses Google advertising services or Google Analytics, "[t]hese products share information about your activity with Google, and depending on your account settings and the products in use (for instance, when a partner uses Google Analytics in conjunction with our advertising services), this data may be associated with your personal information.").

[134] Google provides definitions of technologies it uses to collect data. *See* sources cited *supra* note 127.

[135] *Information We Share*, GOOGLE PRIVACY & TERMS, https://www.google.com/policies/privacy/#application [https://perma.cc/4LJD-UFZ9].

to "non-personally identifiable information" defines that category somewhat tautologically as "information that is recorded about users so that it no longer reflects or references an individually identifiable user."[136] While Google did not share "sensitive personal information" unless an individual opted-in to that sharing, Google did not consider search queries or reading material accessed using Google Books to fit in this category of highly protected information.[137]

Another Google webpage, not linked from the general Google Privacy Policy or the privacy policy page for Google Books, shed light on the kind of book reader information that might be collected and shared as of 2016. "Best Practices for Authors and Publishers" included a promotion of advertising and website analytics services. Google recommended "[w]ith AdWords you can run targeted ad campaigns that put each of your books in front of the readers most likely to buy them. Target by keyword, geography, subject, and/or website."[138] Another invitation was to use the Google Analytics tool that "shows you which sites, search engines, and keywords refer your traffic and how visitors interact with your site."[139] These tracking and targeting services did not describe connecting authors and publishers with identified individuals searching Google Books. But the data collected through advertising and analytics of user activity would become part of Google's vast store of data.

---

[136] The issues of robust anonymization are debated as re-identification is increasingly possible through the linking of data sets. *See* Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86 N.Y.U. L. REV. 1814 (2011) (recommending that information be evaluated along a continuum relating to the ease with which it can describe a particular individual given known risks of links with other data sets that could add identifying information to presumptively anonymized data.); *Google Books Privacy Policy*, GOOGLE BOOKS, https://www.google.com/googlebooks/privacy-sep09.html [https://perma.cc/4RGR-QQS6].

[137] *Information We Share*, GOOGLE PRIVACY & TERMS, https://www.google.com/intl/en/policies/privacy/ [https://perma.cc/4LJD-UFZ9] (linking to a *Key Terms* page which explains "This is a particular category of personal information relating to confidential medical facts, racial or ethnic origins, political or religious beliefs or sexuality.").

[138] *Best Practices for Publishers and Authors*, GOOGLE BOOKS PARTNER PROGRAM, https://www.google.com/googlebooks/partners/resources.html [https://perma.cc/NL8Y-C5U4].

[139] *Id.*

Returning to the 2011 Google Play-Privacy Policy for Books linked from the Google Books site, the 2016 spelunker could find language that suggested there was support for those who wished to segregate their reading activity from their Google Account. If a Google user had a Google Account, required by some Google services, the Google Books user could take advantage of other related features, such as saving search history and the ability to purchase books and other content through Google. This data would become part of the individual's Google Profile. But some level of reader privacy may be promised through the language that stated unless a user is logged in to Google, activity on Google Play "will not be associated with your Google Account."[140] The two archived Google Books Privacy Policies, dated 2009 and 2010, offer similar reassurances for "activity on Google Books."[141] On the other hand, the 2015 privacy policy relating to Google Books, linked from Google's general privacy policy, omits the text that refers to this ability to segregate Google Book activity from an existing account profile.[142]

Google's business practices and market dominance suggest that, even without signing in to a Google account, the information collected through use of Google Books might easily be linked with other data collected, stored, and analyzed to track and even identify individuals. Google Books users may have to exercise hyper-vigilance to allow their research and reading habits to escape this profiling.[143] Whether the detailed profile capabilities of Google are shared with others depends on the meaning of the promise to share personal information only with "affiliates or other trusted businesses or persons to process it for [Google]" based on Google's instructions and in compliance with

---

[140] *Key provisions form the Google Privacy Policy*, GOOGLE PLAY: PRIVACY POLICY FOR BOOKS, https://books.google.com/intl/en/googlebooks/privacy.html [https://perma.cc/H4XP-8JEQ].

[141] For a section on "Key provisions from the Google Privacy Policy," *see Google Books Privacy Policy, Archived Version: September 3, 2009*, GOOGLE PRIVACY & TERMS, https://www.google.com/googlebooks/privacy-sep09.html [https://perma.cc/4RGR-QQS6].

[142] *Practices specific to Google Play Books*, *supra* note 131.

[143] *See* Kathryn J. Strandburg, *Free Fall: The Online Market's Consumer Preference Disconnect*, 2013 U. CHI. LEGAL F. 95, 159-65 (describing a hypothetical experience of an online consumer attempting to make privacy-protecting choices in selecting and using online services and demonstrating the difficulties consumers face weighing pros and cons of moving to a no-data-collection site or service).

Google's Privacy Policy and any other appropriate confidentiality and security measures.[144] Google also promised not to share data with government entities absent sufficient process under the law, and they provide some reports on requests received and Google's responses.[145] These assurances could suggest that Google treats its profiles as protected proprietary data it will use to sell advertisements, but perhaps not to provide granular data-broker-style reports or even details of individuals' reading to publishers or others.

To address data security concerns, Google promised that it works hard to protect data from unauthorized access or use. The Google Privacy Policy included a section on "Information security" that outlined several steps Google takes to protect its customer data. Google offered encryption for most of its services. The company also promised to regularly review its practices and to levy penalties on contractors that violate their privacy obligations to the company. Google also described options for protecting users who choose to log in to Google accounts.[146]

Financial statements provide insights that go beyond the disclosures Google made in its general Privacy Policy and Google Books privacy policy. Before 2016, in filings to the United States Securities and Exchange Commission, Google has stated that its mission was "to organize the world's information and make it universally accessible and useful."[147] This mission is consistent with that of libraries collaborating on digitization projects. And, like libraries, Google identified data security as a risk given the threat of

---

[144] *Information we share*, GOOGLE PRIVACY & TERMS, https://www.google.com/intl/en/policies/privacy/ [https://perma.cc/4LJD-UFZ9]; *see* GOOGLE TRANSPARENCY REPORT, https://www.google.com/transparencyreport/userdatarequests/ [https://perma.cc/4LJD-UFZ9] (providing a variety of reports and charts on governmental requests for Google user information).

[145] *Id.*

[146] *Information security*, GOOGLE PRIVACY & TERMS, https://www.google.com/intl/en/policies/privacy/?fg=1#infosecurity [https://perma.cc/HT54-AQ4Y].

[147] *See, e.g.*, GOOGLE INC., ANNUAL REPORT (FORM 10-K) 3 (2014). *But see* ALPHABET INC. AND GOOGLE INC., ANNUAL REPORT (FORM 10-K) 2 (2016) (omitting this mission and describing more diversification of business interests). This mission statement still appears on the Google website at *About Google,* GOOGLE, https://www.google.com/intl/en/about/ [https://perma.cc/LH95-TRPS].

data breach.[148] But, unlike libraries, Google did not suggest it would achieve data security through destruction of data after some initial use. The filings also highlight the differences in funding models for libraries and for Google. While most libraries are funded by parent institutions, public monies, and private donations, Google's business model is based on "generat[ing] revenues primarily by delivering relevant, cost-effective online advertising."[149] Google has stated that advertising revenue is critically important to its ability to operate, reporting, "[w]e generate a significant portion of our revenues from advertising, and a reduction in spending by or loss of advertisers could seriously harm our business."[150] Google reported that in 2013, ninety-one percent of the company's revenues were generated from their advertisers,[151] and in 2015, even after some diversification, ninety percent of Google revenues came through advertising.[152] Google also reported that if users were able to employ new technologies to block ads, this development "would harm our business."[153] At the same time, the company indicated that privacy concerns about their data practices, merited or not, could damage the company's reputation and "deter current and potential users" from using Google products and services.[154]

Google's privacy policies and financial filings may not directly or simplistically answer the question of whether access to Google Books comes at the expense of reader privacy, but the scales tip towards access and away from privacy. Google is collecting and merging data on all sorts of reading of digital material through its services and elsewhere on the web. And, even if Google Books data can be segregated from Google Accounts, it is possible for this separation to be less meaningful if other collected data provides enough detail to

---

[148] ALPHABET INC. AND GOOGLE INC., *supra* note 147.

[149] GOOGLE INC., *supra* note 147.

[150] *Id.* at 12.

[151] *Id.*

[152] ALPHABET INC. AND GOOGLE INC., *supra* note 147.

[153] *Id.* at 18.

[154] *Id.* at 14 ("Concerns about our practices with regard to the collection, use, disclosure, or security of personal information or other privacy related matters, even if unfounded, could damage our reputation and adversely affect our operating results.").

reveal an identity of the Google Books reader. Advertising is based on reader data and combined with other data, even if Google Profiles are not the home for that linked information. While profiles may not be sold, the information is discoverable by the government through proper process, and even though Google promises to challenge insufficient requests, retention of this information means it remains available.

B. *HathiTrust*

1. *Structure and Purpose of HathiTrust*

The HathiTrust collaboration was created in 2008 when a California library consortium joined forces with a dozen university libraries in the consortium known as the Committee on Institutional Cooperation.[155] The ambition of HathiTrust is to preserve material from library collections through digitization, and to make as much of it accessible online as copyright law allows.[156] The partnership at the time of this writing in 2016 had 100 member libraries and was open to members from around the world.[157] HathiTrust claimed that it seeks to sustain the enterprise as a "public good," while also serving member institutions.[158] Its mission was expressed as "to contribute to research, scholarship, and the common good by collaboratively collecting, organizing, preserving, communicating, and sharing the record of human knowledge."[159] The name HathiTrust derives from the Hindu word for elephant, chosen because of the animal's connotations of "memory, wisdom, and strength."[160] HathiTrust described the selection of "Trust" as a reflection of the core value and

---

[155] *Launch of HathiTrust*, HATHITRUST DIGITAL LIBR. (Oct. 13, 2008), https://www.hathitrust.org/press_10-13-2008 [https://perma.cc/PH25-XGYK].

[156] *Id.*

[157] *Welcome to the Shared Digital Future*, HATHITRUST DIGITAL LIBR., https://www.hathitrust.org/about [https://perma.cc/K54J-QUN6].

[158] *Mission and Goals*, HATHITRUST DIGITAL LIBR., https://www.hathitrust.org/mission_goals [https://perma.cc/6GUN-GJLV].

[159] *Id.*

[160] *Help-General*, HATHITRUST DIGITAL LIBR., https://www.hathitrust.org/help_general [https://perma.cc/S2VQ-EFL7].

one of the greatest assets of research libraries that form the shared digital repository.[161]

HathiTrust describes the repository as coming "from a variety of sources, including Google, the Internet Archive, Microsoft, and in-house partner institution initiatives."[162] Much of the HathiTrust corpus comes from member libraries' deposit of collection content that was digitized by Google as part of the Google Books project.[163] The HathiTrust collaboration preserves, provides access, and offers research guidance for users. The project also supports nonprofit and educational uses of the corpus to conduct advanced computational research.[164]

2. *HathiTrust Privacy Policies, Funding Model*

As of 2016, HathiTrust linked to its privacy policy on the bottom right of its website.[165] The bottom banner consistently stayed with each page display. This policy provided some detail as to the types of data collected, uses intended, retention periods and anonymization, and some privacy risks for users of the website, including HathiTrust's reliance on Google Analytics. Google Analytics has been a widely implemented service for websites that tracks visitors to the site and various identifiers that are not directly identifying of the individual.[166]

HathiTrust does have at least two reasons to require some identification of site visitors. If a site visitor is affiliated with a member institution that has digitized a decaying book, copyright law allows that member institution to provide access to a digitized version, but only to its affiliated users.[167] Another reason to require some

---

[161] *Id.*

[162] *Our Digital Library,* HATHITRUST DIGITAL LIBR., https://www.hathitrust.org/digital_library [https://perma.cc/36X9-LCUW].

[163] *Id.*

[164] *Our Research Center*, HATHITRUST DIGITAL LIBR., https://www.hathitrust.org/htrc [https://perma.cc/8YRW-Z374].

[165] *Privacy Policy,* HATHITRUST DIGITAL LIBR., https://www.hathitrust.org/privacy [https://perma.cc/HQT9-MBQJ].

[166] *Id.*

[167] 17 U.S.C. § 108 (2016).

identification is so that a person can customize access to keep track of her research in the HathiTrust collection. At the time of this writing, the saving of research trails was not tied to site users by default. Researchers would have to opt-in to this feature. Either of these options allows some anonymization of the site visitor. In the first instance, a HathiTrust website reader would be asked to choose her home institution, which would then provide its own login relying on a system that shares login details only with the home institution. The system used to support this confidentiality layer is called Shibboleth.[168] In the second instance, the returning HathiTrust user could create a "Friend" account managed through the University of Michigan (home to HathiTrust) with a non-University of Michigan email account. Links to webpages regarding the Friend account process did not address privacy.

The HathiTrust Privacy Policy as of 2016 indicated that all visitors to the website were subject to the monitoring features of Google Analytics, a service that transmits information such as IP address, unique browser identifiers, referring URLs, and website use information back to Google for analysis before reporting results to HathiTrust. HathiTrust offered a link to Google's general Privacy Policy. The HathiTrust site indicated that it did not share user information with any other third party. However, reliance on Google Analytics service means HathiTrust users would have to take further steps to achieve some privacy protection that is superior to using Google Books itself. HathiTrust site visitors were able to opt out of the Google Analytics service, but that opt out required the user to accept and retain a browser add-on from Google to remind the Analytics process to avoid its normal collection of information from the user each time she visits the HathiTrust website. At the time of this writing, the add-on required the user to enable third-party cookies on the browser, a setting which disables this browser feature designed to provide some broad privacy protection online. The add-on for reader privacy was not the default, and though it enhanced privacy in regard to Google Analytics, it would leave the reader vulnerable to third-party cookies that may invade privacy.

Whether Google Analytics addresses library reader privacy concerns sufficiently is not clear. The company offers some accommodations, such as an option to limit reporting of full IP

---

[168] *Shibboleth Login*, HATHITRUST DIGITAL LIBR., https://www.hathitrust.org/shibboleth [https://perma.cc/T7PN-X7GP].

addresses of site visitors to the website manager, in this case, HathiTrust, but the language describing this feature did not make clear whether Google itself collects and retains the full IP address before sending the obscured details to the website manager.[169] HathiTrust's privacy policy does not disclose whether the site has implemented these additional if perhaps marginal privacy protections.

Google Analytics is widely adopted across the Internet, and the aggregate information Google collects itself is seen as a rich trove of data to be merged with other information Google acquires.[170] Some librarians have found limited implementation of Google Analytics combined with notice and the cookie-based opt-out to be sufficiently privacy protecting.[171] But, others have raised concerns that web analytics such as Google Analytics could be used to contribute data to digital profiles well beyond the control of libraries.[172]

HathiTrust does not address data security in its privacy policy, but the site is encrypted so that information traveling across the web is highly protected from view.[173]

As of 2016, funding for HathiTrust came from fees paid by partners, academic, and research institutions from around the world.[174] The fees were based on two calculations. First, all partners would pay an equal amount to support public domain volumes in

---

[169] *Safeguarding Your Data,* GOOGLE, https://support.google.com/analytics/answer/6004245?hl=en&ref_topic=2919631 [https://perma.cc/PR8L-GPHC] (explaining that a "method known as IP masking gives website owners using Google Analytics the option to tell Google Analytics to use only a portion of an IP address, rather than the entire address, for geolocation.").

[170] Marshall Breeding, *Data Opens New Opportunities for Libraries*, INFOTODAY, Apr. 2016, at 15 (advocating libraries use analytics to inform website design decisions, but noting that "[t]he data accumulated from all the sites that participate in Google Analytics represents a massive resource that Google may be able to tap directly or indirectly for its other services.").

[171] Wayne Loftus, *Demonstrating Success: Web Analytics and Continuous Improvement*, 6 J. OF WEB LIBRARIANSHIP 45, 54 (2012).

[172] Raizel Liebler & Keidra Chaney, *Google Analytics: Analyzing the Latest Wave of Legal Concerns for Google in the U.S. and the E.U.*, 7 BUFF. INTELL. PROP. L.J. 135, 143 (2010).

[173] *See* Marshall Breeding, *Protecting Patron Privacy: Libraries are failing to use HTTPS*, AM. LIBR. MAG., June 2016, at 78 ("The use of a secure communication protocol (HTTPS) provides the best approach available today for protecting patron privacy" on the web).

[174] *Cost*, HATHITRUST DIGITAL LIBR., https://www.hathitrust.org/cost [https://perma.cc/QTJ3-WL3S].

HathiTrust.[175] In addition, partners would pay a portion of the cost of in-copyright volumes that overlap with volumes held in their print collections, allowing participating libraries to pay according to the benefits they receive through HathiTrust.[176] This large-scale digitization collaboration of libraries has a very different funding model from Google Books, which relies on behavioral advertising to sustain its services. HathiTrust's privacy protections are less strained by the need to target advertisements and more a function of website analytics, and potentially a function of authentication of users needing to make full use of HathiTrust services.

IV. CONCLUDING THOUGHTS ON VIABILITY OF LIBRARY PRIVACY IN THE COLLABORATIVE DIGITAL FUTURE

Libraries are improving access to information for current and future researchers through collaborative digitization initiatives, but library privacy is left vulnerable by technologies and practices that make digital reading more exposed to tracking attempts that are not possible in the traditional print library environment. Examination of national and international library privacy statements reveal an ongoing commitment to library use privacy but show a range of responses to the expanding risks in this digital age. Major collaborations to digitize books demonstrate that reader privacy may be incompatible with behavioral advertising and perhaps even with use of commercial tools to manage and evaluate library-funded projects if those tools feed into commercial data collection and analysis.

However, libraries may yet frame the digital future with both access and privacy. A library-led collaboration to enable digital access is better situated to protect privacy than a similar project funded through data mining of the reader. Projects like HathiTrust can compete effectively with Google Books and provide access to materials while offering enhanced privacy protection. As these projects become linked with other library-centered digitization efforts, the expanded network could create a safer place for confidential exploration of ideas.

---

[175] *Id.*

[176] *Id.*

Ultimately, funding models matter. HathiTrust draws on member libraries for funding, while other digital library collaborations like the Digital Public Library of America look to foundations and other noncommercial sustainability plans that can preserve the values of libraries. Both access and privacy can flourish in these library-centric public-interest collaborations.

But libraries will also need to collaborate to implement and develop systems that provide access and incorporate privacy. Scanning and indexing are important contributions, but digital library projects must also update privacy-protecting tools for authentication of users like the Shibboleth system. Encryption needs to be extended to all library digital environments. Also, noncommercial, privacy-protecting website analytics tools should be a focus of a creative community of library-friendly programmers.

These types of reader privacy protections are just parts of the puzzle, but they are nonetheless significant. Libraries have developed a reputation as privacy advocates, and that reputation may be part of the trust that HathiTrust claims is integral to the value of libraries. This reputation comes despite a history of imperfect achievements. Not long ago, libraries used a signature card that stayed tucked in a book pocket until a new reader chose to add his name to the exposed list in order to check out the book. And yet, at the same time, librarians and libraries continued to advance the value of protecting reader privacy. New challenges and setbacks similarly do not have to derail commitments to privacy of library use, especially when libraries can work together to develop solutions. Libraries have traditions of access, privacy, and collaboration, and all three should be joined in our digital future.