Physical-Equivalent Privacy

Dorothea Salo
Information School, University of Wisconsin-Madison, Madison WI, USA

4261 Helen C. White Hall, 600 N. Park St., Madison WI 53706. salo@wisc.edu
ORCID: 000-0002-6388-0311

Dorothea Salo is a Distinguished Faculty Associate in the University of Wisconsin at Madison's Information School. She has written and presented internationally on scholarly communication, copyright, privacy, linked data, and data curation. Her "Recover Analog and Digital Data" project rescues audio, video, and digital data from obsolete or decaying carriers. As co-investigator for the IMLS-funded Data Doubles project, she is helping investigate undergraduate students' perceptions of privacy relative to learning analytics practices. Salo holds an MA in Library and Information Studies and another in Spanish from UW-Madison.

Physical-Equivalent Privacy

This article introduces and applies the concept of "physical-equivalent privacy" to evaluate the appropriateness of data collection about library patrons' use of library-provided e-resources. It posits that as a matter of service equity, any data collection practice that causes e-resource users to enjoy less information privacy than users of an information-equivalent print resource is to be avoided. Analysis is grounded in real-world e-resource-related phenomena: secure (HTTPS) library websites and catalogs, the Adobe Digital Editions data-leak incident of 2014, and use of web trackers on e-resource websites. Implications of physical-equivalent privacy for the SeamlessAccess single-sign-on proposal will be discussed.

Keywords: privacy; confidentiality; equity

Problem statement

Information privacy, defined by the American Library Association (ALA) in its Privacy Interpretation as "the right to open inquiry without having the subject of one's interest examined or scrutinized by others, in person or online"[1] and the International Federation of Library Associations and Institutions (IFLA) as "personal privacy and the protection of personal data,"[2] is more important than ever given that surveillance of information use on the open Web has reached pandemic proportions. Just as on the open Web, it is difficult to know who is capturing how much data on patron use of library-purchased electronic resources, never mind how that data is used and who may access it for what reasons. Perhaps worse still, though, it is difficult to conceptualize privacy harms in the online milieu, partly because of the heavily-surveilled open Web habituating many to perpetual observation, partly because e-resource-related privacy harms are not routinely made evident to patrons or librarians such that they go largely unnoticed, much less prevented.

Available methods of systematic privacy evaluation tend toward the cumbersome and gappy. ALA privacy audits[3] cost serious time and effort, yet are limited to library-internal data practices; they do not evaluate patron privacy relative to third-party content, software, and service vendors. The corporate world, especially outside the

---

[1] American Library Association, *Privacy, An Interpretation of the Library Bill of Rights*, 2014, http://www.ala.org/advocacy/intfreedom/librarybill/interpretations/privacy.

[2] International Federation of Library Associations and Institutions, "IFLA Code of Ethics for Librarians and Other Information Workers (Full Version)," last updated December 27, 2016, https://www.ifla.org/publications/node/11092.

[3] Intellectual Freedom Committee, Privacy Subcommittee, "Developing or Revising a Library Privacy Policy," Text, Advocacy, Legislation & Issues, April 25, 2014, http://www.ala.org/advocacy/privacy/toolkit/policy.

United States, has embraced Privacy Impact Assessments.[4] These are designed to evaluate the privacy ramifications of a specific initiative, and are therefore far less cumbersome than privacy audits. They are compliance-based exercises closely tied to privacy law, however, which given the backward, patchwork state of United States privacy law does not presently capture the range and depth of library privacy concerns. The Digital Library Federation's Privacy and Ethics in Technology Working Group[5] has tried to fill in the privacy-assessment gap, but its efforts so far have shied away from clear normative recommendations and do not directly address e-resources. Librarians clearly need a nimbler, more intuitive privacy-evaluation heuristic, as well as a normative rationale to create urgency around renegotiating or even rejecting privacy-damaging e-resource commitments. Heuristics are themselves leaky and imperfect, but they can surface valuable, otherwise hard-to-articulate implicit insights and highlight situations in need of further investigation and action.[6]

Library norms around the privacy of patron use of physical information carriers such as print books, CDs, and DVDs have been codified and engineered into library processes and spaces over more than a century. Current practice is neither perfect nor uncontested, of course—common practices such as video surveillance and patron-accessible hold and reserve shelves have raised privacy debates, for example—but it does at least provide a familiar, stress-tested, time-refined baseline against which to compare the privacy of electronic information use. Moreover, societal privacy norms in physical spaces are typically far better articulated and observed than online privacy norms, often because violations of physical privacy are far easier for their targets to detect than are online privacy violations.

Indeed, not to compare information privacy between use of physical and electronic information carriers raises serious, troubling equity issues. Article I of the ALA Code of Ethics states (emphasis added) "We provide the highest level of service to all library users through appropriate and usefully organized resources; **equitable service policies; equitable access;** and accurate, unbiased, and courteous responses to all requests." A systematic decrease in the privacy of e-resource use relative to use of physical materials provides neither equitable service nor equitable information access to patrons with little or no choice but to use information in electronic form. These patrons include:

---

[4] Office of the Australian Information Commissioner, "Guide to Undertaking Privacy Impact Assessments," May 2020, https://www.oaic.gov.au/assets/privacy/guidance-and-advice/guide-to-undertaking-privacy-impact-assessments.pdf.

[5] "Privacy and Ethics in Technology," DLF Wiki, July 1, 2020, https://wiki.diglib.org/Privacy_and_Ethics_in_Technology.

[6] Gerd Gigerenzer, Peter M Todd, and ABC Research Group, *Simple Heuristics That Make Us Smart* (Oxford: Oxford University Press, 2001).

- those who must use e-resources because of a disability precluding use of an information-equivalent physical carrier;
- those who must use e-resources because of inability to travel to the physical library;
- those who rely on e-resources specifically to protect their own privacy from shoulder-surfers and other physical-world snoops, a group that includes many marginalized and minoritized populations, as well as those who prefer to enjoy socially-stigmatized entertainment genres without easily-perceptible evidence of those genres (such as book-cover art);[7]
- those who rely on e-resources because the information is unavailable in an information-equivalent physical carrier; and
- those who rely on e-resources because the information-equivalent physical carrier is unavailable, a highly salient group in the shadow of physical-library closures due to the 2020 coronavirus pandemic.

No plausible reading of librarianship's ethics codes or privacy guidelines stipulates that libraries may abridge, or allow to be abridged, the privacy of patrons because of the type of information carrier they use. The differential privacy harms to patrons with disabilities and patrons belonging to marginalized or minoritized populations are especially abhorrent; libraries have historically tuned their privacy practices to the needs of the most imperiled, and should do so again with respect to e-resources.

For many sources and types of online surveillance, Article VI also comes into play: "We do not advance private interests at the expense of library users, colleagues, or our employing institutions." The thought of allowing, much less inviting, private interests to routinely record and analyze patron information behavior in the physical library—more, allowing them to share and sell what they learn, or use what they learn to manipulate those patrons' economic, health-related, educational, or political opinions—is deservedly repellent. Merely moving such surveillance and surveillance-fueled manipulation online cannot excuse it, yet many librarians have allowed, facilitated, encouraged, and even performed commercial surveillance on library patrons largely without challenge.

Arguments against library-based surveillance often start from Article III of the ALA Code of Ethics: "We protect each library user's right to privacy and confidentiality with respect to information sought or received and resources consulted, borrowed, acquired or transmitted" or the second article of ALA's Intellectual Freedom Principles for

---

[7] Vassiliki Veros, "Scholarship-In-Practice the Romance Reader and the Public Library," *The Australian Library Journal* 61, no. 4 (November 1, 2012): 298–306, https://doi.org/10.1080/00049670.2012.10739063.

Academic Libraries:[8] "The privacy of library users is and must be inviolable. Policies should be in place that maintain confidentiality of library borrowing records and of other information relating to personal use of library information and services." This article does not start with these because their deontological frame, positing privacy as a fundamental principle rather than a set of library practices embedded in a context of surveillance capitalism[9] and surveillance creep,[10] allows otherwise well-meaning librarians and scholars to abridge the privacy of patrons and colleagues in the name of competing fundamental principles—commonly assessment, advocacy for the library, or improved service—without considering, much less mitigating, associated harms. An example of language that permits such abridgement comes from the ALA Privacy Interpretation:

> Libraries should not monitor, track, or profile an individual's library use beyond operational needs. Data collected for analytical use should be limited to anonymous or aggregated data and not tied to individuals' personal data. Emerging biometric technologies, such as facial recognition, are inconsistent with the mission of facilitating access to library resources free from any unreasonable intrusion or surveillance.[11]

The latter two sentences are clear and strong, but the first contains an immense loophole in the undefined phrase "beyond operational needs." Defining "operational needs" to include surveillance is the same specious rationalization data brokers and web trackers typically use. Libraries should do better. The word "confidentiality," which appears often in ALA privacy guidance, can also be stretched past all sense simply by expanding the circle of people considered authorized to use patron data. Similarly, Megan Oakleaf's Educause piece on the Library Value Agenda[12] illustrated competing principles that disadvantage privacy more clearly than perhaps was intended when its table of contents posited privacy as not a requirement, not a desideratum, but as an actual *obstacle* to the surveillance-based practice of learning analytics in academic libraries. The more consequentialist approach of this article serves as a reminder that

---

[8] American Library Association, "Intellectual Freedom Principles for Academic Libraries: An Interpretation of the Library Bill of Rights," Text, Advocacy, Legislation & Issues, July 26, 2006, http://www.ala.org/advocacy/intfreedom/librarybill/interpretations/intellectual.

[9] Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for Human Future at the New Frontier of Power* (PublicAffairs, 2019).

[10] Andrew Hope, "Creep: The Growing Surveillance of Students' Online Activities," *Education and Society* 36, no. 1 (May 1, 2018): 55–72, https://doi.org/10.7459/es/36.1.05.

[11] American Library Association, *Privacy, An Interpretation of the Library Bill of Rights*.

[12] Megan Oakleaf, "Library Integration in Institutional Learning Analytics" (EDUCAUSE, November 2018), https://library.educause.edu/resources/2018/11/library-integration-in-institutional-learning-analytics.

librarians embraced privacy practices not merely as an unreachable deontological ideal fueled by vocational awe,[13] but to avoid *real, tangible harms* to our patrons and colleagues.[14]

This article therefore posits and applies a "physical-equivalent privacy" heuristic for librarians to assess patron privacy around e-resource-related vendors, services, and standards: the privacy of an e-resource may be considered *physical-equivalent* only when a patron using an information-equivalent physical resource would enjoy no more privacy than the same patron using the e-resource.

More colloquially, the concept of physical-equivalent privacy asks librarians to examine the personal and behavioral data collected, stored (and potentially leaked or hacked), aggregated and correlated with other personal and behavioral data, analyzed, shared, and/or sold about patron use of library-provided e-resources to determine what someone would have to do to gather and use the same amount of data about use of an information-equivalent physical carrier—a hypothetical one if need be, to acknowledge that not all information exists in physical carriers presently. If the necessary means of surveilling, profiling, or manipulating a physical-information user comparably are far too obtrusive or unethical for comfort, the e-resource 's privacy affordances must be considered inadequate at best, unethically inequitable or private-interest-favoring per the ALA Code of Ethics at worst.

A moment of pedantry

The phrase "physical-equivalent privacy" does not precisely capture the crucial distinction this paper draws between modes of information-use surveillance. This distinction is not between analog and digital information representations, but *online versus offline information access*: resources accessed via the internet carry vastly different, usually lesser, privacy affordances compared to offline resources accessed through physical-library processes such as shelf use and library checkout. A physical carrier of digital data such as a CD-ROM or DVD used wholly offline, for example, carries the privacy affordances of a print codex rather than an electronic journal article. This paper uses "physical-equivalent privacy" for its memorable alliteration and intuitive comprehensibility, both of which feel more important than exact precision.

---

[13] Fobazi Ettarh, "Vocational Awe and Librarianship: The Lies We Tell Ourselves," *In the Library with the Lead Pipe*, January 10, 2018, http://www.inthelibrarywiththeleadpipe.org/2018/vocational-awe/.

[14] Steve Witt, "The Evolution of Privacy within the American Library Association, 1906–2002," *Library Trends* 65, no. 4 (September 8, 2017): 639–58, https://doi.org/10.1353/lib.2017.0022.

Literature review

Clifford Lynch provides a substantial recent overview of "reading analytics" around ebooks, abstracting and indexing services, scholarly journals, e-textbooks, and citation-management platforms.[15] Eric Hellman has confirmed that advertising and behavioral-data-gathering trackers are common in library-provided research journals.[16] Vendor contracts with libraries rarely mention these, much less forbid them.[17] In addition, several devices on which library ebooks can be read surveil patron location, among other things.[18] Sarah Lamdan points out[19] that several library-content vendors also act as data brokers, whose business is aggregating and selling information about identified individuals' lives and habits to advertisers, governments, financial institutions, insurers, and other third parties. These vendors have offered no assurances that library-patron data is not swept into their data warehouses to be shared or sold, including to law-enforcement agencies, who traditionally constitute a threat against which libraries calibrate physical privacy practices.[20] Existing guidelines on e-resource and third-party service privacy such as NISO's Consensus Principles[21] do not clearly state what types and amount of patron tracking should be permitted or forbidden.

Worse yet, libraries themselves are engaging in—and contributing patron information-use data to—surveillance-fueled Big Data practices. Some public libraries are signing on with and providing patron data, often fully identified, to customer-

---

[15] Clifford Lynch, "The Rise of Reading Analytics and the Emerging Calculus of Reader Privacy in the Digital World," *First Monday*, April 3, 2017, https://doi.org/10.5210/fm.v22i4.7414.

[16] Eric Hellman, "16 of the Top 20 Research Journals Let Ad Networks Spy on Their Readers," blog, Go To Hellman, accessed May 9, 2017, https://go-to-hellman.blogspot.ca/2015/03/16-of-top-20-research-journals-let-ad.html; Eric Hellman, "Reader Privacy for Research Journals Is Getting Worse," blog, Go To Hellman, March 2017, https://go-to-hellman.blogspot.com/2017/03/reader-privacy-for-research-journals-is.html.

[17] Trina Magi, "A Content Analysis of Library Vendor Privacy Policies: Do They Meet Our Standards?," *College & Research Libraries* 71, no. 3 (May 2010): 254–72, https://doi.org/10.5860/0710254.

[18] Stephen B Wicker, "EBook Readers, Location Surveillance and the Threat to Freedom of Association," in *Proceedings of the 14th International Conference on Availability, Reliability and Security*, 2019, https://doi.org/10.1145/3339252.3340501.

[19] Sarah Lamdan, "Librarianship at the Crossroads of ICE Surveillance," *In the Library with the Lead Pipe*, November 13, 2019, http://www.inthelibrarywiththeleadpipe.org/2019/ice-surveillance/.

[20] Witt, "The Evolution of Privacy within the American Library Association, 1906–2002."

[21] NISO, "NISO Consensus Principles on Users' Digital Privacy in Library, Publisher, and Software-Provider Systems," December 10, 2015, https://www.niso.org/publications/privacy-principles.

relationship management and analytics tools like OrangeBoy[22] and OCLC WISE.[23] This practice caused notable public embarrassment for the Santa Cruz Public Library (SCPL) when a citizen grand jury determined that SCPL "did not recognize the importance of" informing patrons and obtaining their consent to data collection and use related to SCPL's use of Gale Analytics on Demand.[24] Several public libraries and public-library consortia are building apps for voice-activated home assistants such as Alexa and Amazon Echo,[25] despite these devices' known affinity for collecting users' behavioral data[26] and frankly wretched track records for security.[27] Some academic libraries, consonant with recommendations in the Library Value Agenda,[28] are engaging (sometimes under significant pressure from the larger institution) in "learning analytics" initiatives that involve sharing many sources of identified data about student use of libraries and library-provided e-resources outside the library.[29] The move to open access via transformative agreements has raised the spectre of researcher surveillance and data sale by e-resource vendors hungry for new revenue streams.[30] Electronic textbooks have also come under scrutiny for their ability to capture reading and annotation behaviors, especially when embedded in learning-management systems.[31]

---

[22] "OrangeBoy, Inc.," OrangeBoy, Inc., accessed June 26, 2020, https://www.orangeboyinc.com.

[23] "OCLC Wise: Community Engagement System for Public Libraries," OCLC, June 12, 2020, https://www.oclc.org/en/wise.html.

[24] "Patron Privacy at Santa Cruz Public Libraries: Trust and Transparency in the Age of Data Analytics," June 24, 2019, http://www.co.santa-cruz.ca.us/Portals/0/County/GrandJury/GJ2019_final/SantaCruzPublicLibrariesReport.pdf.

[25] Miriam Sweeney and E. Davis, "Alexa, Are You Listening? An Exploration of Smart Voice Assistant Use and Privacy in Libraries," *Information Technology and Libraries* Forthcoming (2020), http://ir.ua.edu/handle/123456789/6783.

[26] Sidney Fussell, "Consumer Surveillance Enters Its Bargaining Phase," *The Atlantic*, June 4, 2019, sec. Technology, https://www.theatlantic.com/technology/archive/2019/06/alexa-google-incognito-mode-not-real-privacy/590734/.

[27] Dan Goodin, "Alexa and Google Home Abused to Eavesdrop and Phish Passwords," *Ars Technica*, October 20, 2019, https://arstechnica.com/information-technology/2019/10/alexa-and-google-home-abused-to-eavesdrop-and-phish-passwords/.

[28] Megan Oakleaf, "Value of Academic Libraries Report: A Comprehensive Research Review and Report," September 2010, http://www.ala.org/acrl/sites/ala.org.acrl/files/content/issues/value/val_report.pdf.

[29] Kyle M. L. Jones et al., "A Comprehensive Primer to Library Learning Analytics Practices, Initiatives, and Privacy Issues," *College & Research Libraries* 81, no. 3 (2020): 570–91, https://doi.org/10.5860/crl.81.3.570.

[30] Sam Popowich, "Proxying the Data Body: Artificial Intelligence, Federated Identity, and Machinic Subjection," *Journal of Contemporary Issues in Education* 15, no. 1 (June 28, 2020): 35–50, https://doi.org/10.20355/jcie29410.

[31] Kyle M. L. Jones, Alan Rubel, and Ellen LeClere, "A Matter of Trust: Higher Education Institutions as Information Fiduciaries in an Age of Educational Data Mining and Learning

Library efforts to stem these privacy compromises have been at best lukewarm and ineffectual so far; the problems have been fairly thoroughly theorized and documented, but few solutions have been proposed, much less adopted. Education of both librarians and patrons is one often-advanced suggestion,[32] but no systematic investigation of its prevalence or effectiveness has been done; the same is true of privacy advocacy.[33] Deborah Caldwell-Stone suggested addressing library ebook privacy through license negotiation, legislative advocacy, and standards work in 2012,[34] but there is little evidence libraries and consortia have adopted these recommendations. NISO's Shared Electronic Resource Understanding (SERU) Recommended Practice language addresses disclosure of "personal information about the user" only, not behavioral tracking or reidentification.[35] Licensing training for e-resource librarians rarely if ever mentions patron privacy as a negotiable element in licenses; for example, Lesley Ellen Harris's book on licensing only mentions privacy in the context of the library itself, not the e-resource vendor, tracking individual patrons' information use.[36]

Gauging physical-equivalent privacy

It is possible to estimate physical-equivalent privacy starting either from a physical/offline information carrier or its information-equivalent digital/online counterpart. Many librarians will no doubt find the thought process intuitive enough not to need explication. For clarity, however, the following procedure should suffice, starting from a digital/online carrier:

- Enumerate potential and actual information-privacy compromises when a library patron uses the digital/online information object. Remember to consider third-party network and web-based surveillance, as well as data leaks and breaches.

Analytics.," *Journal of the Association for Information Science and Technology*, 2019, 1–15, https://doi.org/10.1002/asi.24327.

[32] Sarah Hartman-Caverly, "Human Nature Is Not a Machine: On Liberty, Attention Engineering, and Learning Analytics," *Library Trends* 68, no. 1 (October 24, 2019): 24–53, https://doi.org/10.1353/lib.2019.0029.

[33] Bohyun Kim, "Cybersecurity and Digital Surveillance versus Usability and Privacy: Why Libraries Need to Advocate for Online Privacy," *College & Research Libraries News* 77, no. 9 (October 1, 2016): 442–51, https://doi.org/10.5860/crln.77.9.9553.

[34] Deborah Caldwell-Stone, "A Digital Dilemma: Ebooks and Users' Rights," *American Libraries*, 2012, 20–23, https://www.jstor.org/stable/26197638.

[35] NISO SERU Standing Committee, "NISO RP-7-2012, SERU: A Shared Electronic Resource Understanding," May 2012, https://groups.niso.org/publications/rp/RP-7-2012_SERU.pdf.

[36] Lesley Ellen Harris, *Licensing Digital Content: A Practical Guide for Librarians, Third Edition* (American Library Association, 2018).

        - Use an ad- or tracker-blocking browser plugin to enumerate the domain names invoked when the e-resource is loaded. Depending on the plugin, it may give hints as to which domain names participate in tracking. If not, the privacy-focused search engine DuckDuckGo maintains a list of common web trackers, along with metadata about the exact tracking techniques they use.[37] (This list should be preferred to the lists most browser ad-blocking plugins use, as it is limited only to trackers and does not include advertisers and other organizations that do not employ tracking techniques.)
- Assess details of patron data collection, analysis, use, and sharing or sale based on the library's license with the vendor and the vendor's terms of service and privacy policies. Remember that personally-identifiable information is not the only information at issue here: revelation of the subject of patron inquiry, behavioral tracking, potential patron reidentification, and use of the data for attempted manipulation of the patron should also be assessed.
- Search the technology trade press for evidence of the vendor's known privacy or security problems, to estimate the vendor's general trustworthiness.
- Examine the vendor's ownership to shed light on threats to patron privacy from corporate data aggregation and sale as well as bankruptcies, mergers, or corporate acquisitions, as with the sale of Canvas learning-management system provider Instructure to Thoma Bravo.[38]
- For each potential and actual compromise noted, assess its legality and its conformance to libraries' existing professional ethics and norms, as well as local organizational policy. Daniel J. Solove's taxonomy of privacy compromises and harms should serve as excellent context for such an analysis.[39]

Next, perform the analogous analysis on an information-equivalent physical/offline resource:

- Estimate what an observer and/or the library would have to do to facilitate collection and use of the type and amount of information collected and used about the patron during e-resource use. Include actions that are plausible save

---

[37] *Duckduckgo/Tracker-Radar* (2020; repr., DuckDuckGo, 2020), https://github.com/duckduckgo/tracker-radar.

[38] Jeffrey R Young, "As Instructure Changes Ownership, Academics Worry Whether Student Data Will Be Protected.," *EdSurge*, January 17, 2020, sec. Education Technology, https://www.edsurge.com/news/2020-01-17-as-instructure-changes-ownership-academics-worry-whether-student-data-will-be-protected.

[39] Daniel J. Solove, "A Taxonomy of Privacy," *University of Pennsylvania Law Review* 154, no. 3 (2005): 477, https://www.pennlawreview.com/print/?id=12.

that they fall well outside the norms of librarianship, society at large, or both. For example, many privacy compromises in physical spaces involve physically following and monitoring a target to a socially-disapproved or even illegal degree. A compromise of e-resource usage privacy analogous to such behavior should be considered equally outside library or societal norms. Indeed, this sort of comparison is substantially the reason for positing physical-equivalent privacy in the first place!

- Estimate the difficulty of compromising the patron's information privacy unobserved by the patron, surrounding patrons, and library staff. An unobserved or even unobservable e-resource privacy compromise that would be easily observed in a physical environment raises serious concerns about exploiting patron ignorance; "the patron doesn't know!" does not justify surveillance under library ethics codes (or, for that matter, basic human decency).
- Examining the compromise's potential and actual harms is also warranted, and may be supplemented by searching for real-world examples or case studies. In the interest of equity, explicitly include marginalized and minoritized populations in this assessment.

Lacunae and ambiguity around privacy in license language and providers' privacy policies are quite likely. Privacy policies are written for legal compliance and protection of the provider itself, not for ethics elucidation or clarity to librarians and patrons. Librarians may of course request clarification from the provider, but should it not be forthcoming, under the present state of privacy law and organizational privacy and security practice it is advisable to assume the worst: any feasible and potentially-remunerative privacy compromise known to be common in online environments that is not explicitly forbidden by contract or written policy is likely taking place. Remuneration need not be monetary, of course; as already noted, much surveillance occurs under the auspices of safety or organizational self-improvement.

Example physical-equivalent privacy analyses

OPACs and HTTPS

HyperText Transfer Protocol (HTTP), the application-layer protocol underlying data transmission for the World Wide Web, was not originally designed to prevent the interception of transmitted website information by network observers. Realizing the need for secure and private Web communications, browser maker Netscape created the HTTPS ("secure HTTP") extension to the HTTP standard in 1994, and it has been further refined and updated since.

Many library websites and OPACs, despite efforts such as the HTTPS Pledge,[40] are still served over insecure HTTP connections. Marshall Breeding's automated tests in 2018 and 2019 revealed that 7.9% of US academic libraries and 18.3% of US public libraries were still serving their websites over unencrypted HTTP.[41] This practice leaves patrons' search histories, search results, and full content of browsed pages open to covert surveillance via packet-sniffing applications such as the free cross-platform Wireshark.[42] (Website content served insecurely is also vulnerable to alteration as it travels between website and patron device, but that danger is not privacy-relevant, so this article passes it by.) Patrons' exact vulnerability to surveillance in the scenario of a Wireshark user inside a physical library depends on how the library's network is set up. The worst case is all patron devices (library-provided and patron-owned) attaching to the same wifi network; a computer running Wireshark in promiscuous mode would capture all patron interactions with the library website and OPAC—including search terms, results pages, and individual items browsed—as well as all patron interactions with other insecurely-served websites. The best case is all library-provided computers attaching to a wired network and disallowing installation of novel software. In this situation, patrons using the web and OPAC on library-provided computers could not be easily surveilled with Wireshark, though patrons using their own devices, or library-provided devices (such as tablets) that cannot use a wired connection, could still be surveilled while attached to library-provided wifi.

The closest information-equivalent physical resource to the OPAC is the card catalog; the closest information-equivalent physical resource to the library website is the library's signage combined with library-related questions at the reference and circulation desks. To replicate the extent of Wireshark surveillance on a patron using a typical card catalog, a video camera or shoulder-surfer would have to record every drawer the patron opened, every card touched or glanced at, every library shelf subsequently browsed (analogous to some OPACs' call-number browse), and every book removed from the shelf (analogous to many OPACs' "save this" or "have this delivered" functions). To replicate Wireshark surveillance of library website use, patron perusal of signage would have to be video-recorded or physically surveilled, along with every question asked at the reference and circulation desks. A librarian performing such surveillance would be well outside library ethics and norms; a patron doing so would

---

[40] "HTTPS Pledge—Library Freedom," accessed July 7, 2020, https://libraryfreedom.org/index.php/https-pledge/.

[41] Marshall Breeding, *Protecting Privacy on Library Websites : Critical Technologies and Implementation Trends* (Chicago, IL: ALA TechSource, October 2019.), https://www.alastore.ala.org/content/protecting-privacy-library-websites-critical-technologies-and-implementation-trends.

[42] "Wireshark · Go Deep.," accessed June 26, 2020, https://www.wireshark.org/.

likely be stopped, possibly escorted from the premises or even banned. The behavior of both would be painfully obvious and most likely intolerable to the surveilled patron.

A patron using an insecurely-served website or OPAC outside the library would also be vulnerable to surveillance by a Wireshark user attached to the same local network—an open wifi network in a coffeeshop or airport, for example. It is harder to establish a physical analogue for this scenario, but physically surveilling a patron's browsing in a library's bookmobile comes close. This, too, would not be acceptable under societal and library privacy norms.

Library websites and OPACs served over insecure HTTP, then, do not provide physical-equivalent privacy to a decidedly problematic degree.

Adobe leak, 2014

Adobe Digital Editions is software for reading ebooks, including library-provided ebooks. In 2014, security researchers determined that Adobe Digital Editions was transmitting fine-grained data about ebook use by identified individual users over the internet to Adobe without encryption.[43] For ebooks in the EPUB (rather than PDF) format, transmitted data included:

- user and device identifiers
- each ebook accessed
- length of time spent reading the ebook
- percentage of ebook read
- exact pages viewed

To accumulate analogous data about patron use of an information-equivalent print book, the patron would have to be surveilled at all times while reading, both inside and outside the library, either by video cameras or shoulder-surfing human beings. Neither scenario is acceptable under library ethics codes—guidance on video surveillance in the library routinely recommends that cameras be made unable to capture exact information use—suggesting that Adobe's surveillance of ebook use should also be unacceptable to librarians. To the extent that e-resource purchase models in libraries require this level of identifiable surveillance of patron information use—charges per page read per patron, for example, or percentage of book read per patron—they, too, are ethically unacceptable. Moreover, Adobe's failure to encrypt the data during transmission is analogous to posting the data on a public bulletin board for perusal or

---

[43] Sean Gallagher, "Adobe's e-Book Reader Sends Your Reading Logs Back to Adobe—in Plain Text," *Ars Technica*, October 7, 2014, https://arstechnica.com/information-technology/2014/10/adobes-e-book-reader-sends-your-reading-logs-back-to-adobe-in-plain-text/.

even copying by anyone and everyone, a scenario library privacy ethics certainly would not allow.

E-resource usage recording that eschews patron and device identifiers as well as amount read—"some patron read this ebook"—is acceptably close to physical-equivalent privacy. Libraries commonly record uses of physical information carriers without heed to who used them; Integrated Library Systems keep checkout counts on items, for example, and many reshelving processes involve scanning barcodes on items that patrons have taken from the shelves but not checked out, so that in-library use can be roughly tallied. Recording exact amount of use is a less clear case, requiring some imagination to analyze as it is not a common practice around physical library materials. Imagine, however, a library recording that six pages of a returned print book were damaged, without regard to which patron damaged them; this seems ethically unproblematic, as no patron comes to harm or is associated with their specific information use. The ethical acceptability of this scenario suggests that exact-use recording from an e-resource without identifying the patron is physical-equivalent privacy.

Care must be taken, however, to eliminate the scenario of patron reidentification through recording, storage, and behavioral analysis of individual deidentified patrons' e-resource use over time, which is not physical-equivalent privacy. ("Pseudonymous" use recording, which records an individual's use over time without attaching a direct identifier to that individual, is deeply vulnerable to this type of reidentification.) Except under subpoena or in rare cases of severe, repeated, and/or deliberate damage to physical materials, librarians would not try to associate a patron's checkout history with that patron, much less reidentify a patron through retaining their checkout history, to avoid harm to the patron's autonomy and intellectual freedom.

Another relevant cautionary tale is that of Aaron Swartz,[44] whose attempt to download large swathes of the JSTOR database, access to which had been purchased by Massachusetts Institute of Technology (MIT) Libraries for the MIT campus, from the MIT network led to his federal prosecution under the Computer Fraud and Abuse Act. Shutting down excessive e-resource downloading by individual patrons is a common e-resource chore; the MIT report on Swartz, for example, notes that the MIT Libraries handled 65 excessive-downloading incidents in academic year 2010-11, and that such incidents are generally resolved with a warning email to the downloader. The closest physical analogue to stopping excessive downloads is imposing limits on the number of physical materials that may be simultaneously checked out by the same patron. That such limits are not ethically problematic suggests that library-internal monitoring for

---

[44] Hal Abelson, "Report to the President: MIT and the Prosecution of Aaron Swartz" (Massachusetts Institute of Technology, July 26, 2013), http://swartz-report.mit.edu/.

excessive downloads by specific individuals constitutes physical-equivalent privacy. Disclosing excessive checkout attempts to law enforcement or to the publishers or vendors of the physical materials involved would be well beyond the ethical pale, however, except under highly rare and bizarre circumstances. So would disclosure to law enforcement, publishers, or vendors of attempted checkouts (excessive or not) by someone without a valid library card, to extend the physical analogy far enough to encompass the Swartz incident.

Adobe resolved the matter by patching Adobe Digital Editions to encrypt the use data in transit. As far as anyone knows, however, the data is still being collected, transmitted, stored for an indeterminate time, analysed, and shared, all of this identifiably. Adobe's longtime market position as a data broker[45] storing and using device identifiers[46] easily matched to individuals bolsters the suspicion that library-patron information fattens its saleable data portfolio. (Even if it does not, Adobe's past and present lack of transparency on this point should unsettle librarians.) That an Adobe subsidiary recently gathered and analyzed data of protesters against police violence via geolocation of their mobile devices[47] should give libraries even more pause about Adobe surveillance of library patrons' reading behavior. Patrons who use Adobe Digital Edition to read library ebooks clearly do not enjoy physical-equivalent privacy.

Web trackers in e-resources

This is a difficult situation to analyze holistically, since practices differ across platforms and providers, and the maze of data brokers and data sales and sharing is near-impenetrable generally.[48] This paper therefore analyzes the New England Journal of Medicine's website because of Eric Hellman's preliminary analysis of its tracker and advertiser complement and its privacy policy.[49]

---

[45] Adobe, "Adobe Audience Finder | Mobilewalla," accessed June 26, 2020, https://www.adobe-audience-finder.com/data_partner/mobilewalla/.

[46] Adobe, "About the ID Service," accessed June 26, 2020, https://docs.adobe.com/content/help/en/id-service/using/intro/about-id-service.html.

[47] Caroline Haskins, "Almost 17,000 Protesters Had No Idea A Tech Company Was Tracing Their Location," *BuzzFeed News*, accessed June 26, 2020, https://www.buzzfeednews.com/article/carolinehaskins1/protests-tech-company-spying.

[48] Matthew Crain, "The Limits of Transparency: Data Brokers and Commodification," *New Media & Society* 20, no. 1 (January 1, 2018): 88–104, https://doi.org/10.1177/1461444816657096.

[49] Eric Hellman, "Suggested Improvements for a Medical Journal Privacy Policy," *Go To Hellman* (blog), April 1, 2015, https://go-to-hellman.blogspot.com/2015/04/suggested-improvements-for-medical.html.

Loading the home page at https://www.nejm.org/ calls several domains and subdomains; domains appearing on DuckDuckGo's tracker list have been marked with an asterisk:

- nejm.org: csvc.nejm.org, files.nejm.org
- adobedtm.com*: assets.adobedtm.com
- b2c.com*: api.b2c.com
- cloudflare.com*: cdnjs.cloudflare.com
- demdex.net* (a subsidiary of Adobe): dpm.demdex.net
- doubleclick.net* (a subsidiary of Google): securepubads.g.doubleclick.net
- edgesuite.net*: use-stls.adobe.com.edgesuite.net
- google-analytics.com*
- googleadservices.com*
- moatads.com*: z.moatads.com
- nejmcareercenter.org: apps1.nejmcareercenter.org
- nejmgroup-production.org
- nejmgroup.org: qow.nejmgroup.org
- pingdom.net*: rum-static.pingdom.net
- resultspage.com*: nejm.resultspage.com
- sli-systems.net: svip-usa1.sli-systems.net
- typekit.net*: p.typekit.net, use.typekit.net

An exhaustive survey of the surveillance techniques, data use and sales, and privacy policies of the trackers noted above is well beyond the scope of this paper. At minimum, though, several trackers in the list sell data to or share data with data brokers. The following domains in the list, by their own admission in their own services' marketing materials, actively attempt to reidentify users and add information-use data to dossiers on those users held by the domains, sometimes as part of personalization or real-time bidding advertising processes: demdex.net, doubleclick.net, googleadservices.com, pingdom.com, and sli-systems.net. Patrons logged in via the library to access NEJM's materials encounter no fewer trackers and ads. It is important to realize that website-based trackers attempting patron reidentification do not require that the library identify the patron, though doing so would of course make reidentification immensely faster and more reliable for the tracker.

Analogous behavior with respect to physical resources must be imagined from whole cloth, as the level of information-use surveillance considered routine online would be utterly beyond the pale in a physical library. Imagine, then, third-party-installed video cameras that track each patron for the duration of the library visit, recording and storing every information object the patron examines as well as the patron's route through the library, nominally so that the patron can be presented with the advertising

that the advertiser believes most likely to result in a purchase. (In reality, such data gathered online is used for a broad spectrum of purposes by insurers, loan originators, educational institutions, law enforcement, political organizations seeking to sway voters, conspiracy theorists looking for new adherents, and more.) Behind the scenes and without patrons' knowledge or genuine consent, each patron is reidentified by name via matching their appearance and behavior with stored artifacts of prior surveillance. Footage of each patron's visit and conclusions inferred from it about the patron are then added to the existing dossier on the patron, which is sold to and shared with other advertisers as well as data brokers. This is favoring private interests over patron welfare with a vengeance! It demonstrates rather starkly that NEJM's website readers do not enjoy physical-equivalent privacy. NEJM is not an outlier, of course; many e-resource vendors carry similar or even greater loads of trackers and surveillance-based advertisers.[50]

Looking forward: SeamlessAccess

Most academic libraries provide authorized access to e-resources with a combination of IP-address recognition on campus and proxy-server use off-campus (with some on-campus proxy-server use in specialized situations). Neither provides full physical-equivalent privacy, though IP-address recognition sometimes come close. Because some IP addresses on college campuses are static (that is, persistently assigned to a specific device typically in use by a single patron), Wireshark users on the relevant local campus network can sniff that patron's interactions with e-resource websites (obtaining URLs at minimum, and full data from providers whose websites are not served securely). E-resource vendors can build information-use dossiers on the relevant library patron through standard web-server logs, as well as the web-surveillance practices discussed with reference to the NEJM website. Patrons whose devices receive temporary IP addresses (typically through the DHCP protocol fueling most wifi connections) do enjoy privacy that could be very close to physical-equivalent, leaving aside web-tracker use or device-identifier collection by the e-resource vendor. Access through proxy servers provides physical-equivalent privacy—analogous to checkout records on a physical copy of a journal issue—if and only if the proxy-server logs that associate individual library patrons with exactly which e-resources they access are routinely and promptly deleted, and identified information-use data is never disclosed outside the library. Prompt and routine log deletion is not current practice in at least some academic libraries,[51] however, and at least one library learning-analytics project disclosed identified

---

[50] Hellman, "Reader Privacy for Research Journals Is Getting Worse."

[51] Scarlet Galvan, personal communication, n.d.

e-resource use to researchers in the student-services office,[52] albeit gathered via "click-through script" rather than specifically proxy-server logs.

In 2019, the International Association of Scientific, Technical, and Medical Publishers (STM) and the National Information Standards Organization (NISO) published for public comment a draft of "Recommended Practices for Improved Access to Institutionally-Provided Information Resources: Results from the Resource Access in the 21st Century (RA21) Project."[53] The report's main technical recommendation involved the replacement of IP-address recognition and proxy servers with authentication and authorization technologies built around the Security Assertion Markup Language (SAML), in use in many (though far from all) academic libraries via the OpenAthens and InCommon federations, as well as Shibboleth-based single-sign-on technologies.

Responses to the draft collected by NISO during the public-comment period[54] were mostly critical. Privacy-related objections brought up repeatedly by various commenters included:

- the danger to patron privacy of ceding decisions about which identifiers and other patron metadata are revealed to the e-resource vendor to campus IT units, which do not have the same privacy commitments as libraries;
- the impracticality of "end-to-end [user] traceability… for detecting fraud," stated as a guiding principle of the RA21 draft, without compromises to patron privacy;
- references to a "minimal request" of patron metadata from an e-resource vendor without this term being defined, leaving the door open to such requests including direct patron identifiers;
- the permitted use of persistent pseudonymous identifiers (that is, identifiers specific to a patron that do not directly identify that patron), which combined

---

[52] Krista M. Soria, Jan Fransen, and Shane Nackerud, "The Impact of Academic Library Resources on Undergraduates' Degree Completion," *College & Research Libraries* 78, no. 6 (September 2017): 812–23, https://doi.org/10.5860/crl.78.6.812.

[53] National Information Standards Organization, "Document Details - NISO_RP-27-2019_RA21_Identity_Discovery_and_Persistence-Public_comment.Pdf," accessed July 23, 2020, https://groups.niso.org/apps/group_public/document.php?document_id=21376&wg_abbrev=niso-ra21.

[54] National Information Standards Organization. https://groups.niso.org/apps/group_public/document.php?document_id=21376&wg_abbrev=niso-ra21

with behavioral tracking and standard web surveillance make patron reidentification and re-association with their e-resource use a near-certainty;
- no strictures on e-resource vendor employees and their service providers against seeking, using, sharing, or selling identified or reidentifiable information-use data; and
- the acceptability per the draft of an e-resource vendor insisting via contract on non-anonymous, non-pseudonymous patron identifiers.

After the comment period closed, NISO released the RA21 document as a Recommended Practice substantially unaltered while marking all comments "Addressed (Unresolved)." NISO Executive Director Todd Carpenter subsequently wrote in defense of RA21 that "technology in and of itself is neither good nor bad" (a stance forcefully and repeatedly discredited by too many scholars in LIS, science and technology studies, philosophy, law, gender and race studies, and other disciplines to cite), bewilderingly citing India's continued use of the poisonous and environmentally devastating chemical DDT as both justified in context (despite negative externalities) and an analogue to using RA21 despite its challenges to patron privacy (among other difficulties). The concluding paragraph of Carpenter's piece begins "It would be absurd to avoid the use of a technology, simply because there may be misuse," a statement that both ignores the lengthy, well-established real-world history of surveillance misuse and inexplicably posits that misuse is not adequate reason to eschew a technology, especially when alternate technologies are available and in widespread use already.

Examining both the RA21 recommendation and Carpenter's defense of it strongly suggests that under an RA21-driven authentication regime, library patrons would not enjoy anything close to physical-equivalent privacy. In the worst case, an RA21 system would permit the level of surveillance already noted with respect to NEJM, even adding to it the direct identification of patrons to the e-resource vendor by the institution. Individual perpetrators of excessive downloading (or other "fraud" as the RA21 document phrases it, again without defining or limiting the term) would be far more easily identifiable to the e-resource vendor, raising the spectre of additional prosecutions similar to that of Aaron Swartz instead of current discreet and largely non-punitive institution-internal practices. Even in the most privacy-favoring implementation of RA21 possible, libraries will be substantially less able to assert control over transmission of user metadata to e-resource vendors because libraries do not run campus authentication systems. This will degrade patron privacy compared to the present (itself imperfect) system of IP-address and proxy-server use, and will not provide patrons physical-equivalent privacy.

NISO and STM subsequently joined with GÉANT, Internet2, and ORCID to launch a successor effort called SeamlessAccess.[55] Its technological recommendations are under review at time of writing, but the base technologies are SAML and single-sign-on via Shibboleth, as with RA21. More effort is underway to clarify which patron metadata may be transmitted between institution and e-resource vendor, in the form of "federated identity entity categories," listed by SeamlessAccess[56] as:

- Authentication Only - this use case covers authentication only; the Service Provider does not want any attributes (specific pieces of data about an authenticated user) from the Identity Provider, only a confirmation that the authentication was successful.
- Anonymous Authorization - this use case supports authorization decisions through the sharing of additional information such as entitlement data (e.g., faculty versus student), while keeping the user completely anonymous [sic] to Service Providers.
- Pseudonymous Authorization - this use case supports authentication, authorization, and allows for personalization per Service Provider through the sharing of a per-service user identifier without requesting any personal information such as name or email address.

A patron pulling a physical issue of a journal from the library shelf for perusal would not (except under highly unusual circumstances) be identified, much less recorded, as a user of that journal. A patron checking out such an issue (or a bound volume of a journal) would be identified as using it only until the physical item was returned and checked back in, though most library systems would subsequently keep a record of the checkout without reference to the patron. Taking that into consideration, the Authentication Only category provides physical-equivalent privacy compared to the in-library patron, and actually improves on physical-equivalent privacy compared to the patron who checks out a physical issue or volume. (This analysis only examines the privacy of the authentication process, of course; it does not assess use of web trackers or log-based surveillance by the vendor.)

Unfortunately, the Authentication Only category contains a significant loophole that librarians desiring physical-equivalent privacy for patrons will have to close when negotiating e-resource contracts. The documentation states (emphasis added), "By asserting this attribute, Identity Providers [institutions] are indicating that they will not

---

[55] "SeamlessAccess.Org - True Single Sign On," accessed July 24, 2020, https://seamlessaccess.org.

[56] SeamlessAccess, "Public Comments Now Open for New Federated Identity Entity Categories - SeamlessAccess," SeamlessAccess, accessed July 24, 2020, https://seamlessaccess.org/posts/2020-07-08-entitycategories/.

release any user attributes to Service Providers [vendors] who assert support for this category **unless bilateral arrangements are in place**." Considering vendor past practices and incentives as described in this paper's literature review and examples, it appears highly likely that vendors will add language to e-resource contracts demanding that institutions supply user attributes not necessary to provide services, without associated restrictions on how vendors will use those attributes, in hopes that librarians will accept the language without thinking. Unless SeamlessAccess closes this loophole, libraries will have to seek and strike out such contract provisions during negotiations.

For a physical analogue to e-resource-related services that require at least some knowledge of patron identity to function, consider a journal-alert service in which librarians notify patrons who have asked to know when new print issues of certain journals arrive. Under common library-privacy norms, librarians would consider the list of patrons and their desired journals confidential. They would not reveal outside the library which patrons have asked for, much less received, which issues of which journals except under highly rare and coercive circumstances such as a judicial warrant. The list would not be contributed to campus assessment, human-resources, or other analytics processes, though the fact of use would commonly be tallied and communicated without reference to any specific patron. Librarians would also neither seek to infer nor share patron characteristics based on their journal choices, except implicitly as part of the existing professional relationship between a librarian and an individual patron. That is, a reference librarian who maintains the list and therefore already knows a given patron's taste in journals might consider that taste while answering that patron's reference question. A reference librarian unaware of the patron's journal choices, however, would not read the list to learn them.

The Anonymous Authorization category may provide physical-equivalent privacy for the patron desiring a service analogous to physical new-issue notifications, but present documentation suggests this might not be so. As specified in the current documentation for the entity category, the institution provides the e-resource vendor entitlement data for the patron in question that could allow access to the material, possibly including departments with which the patron is affiliated as well as on- and off-campus services the patron is entitled to use. This combination of entitlements may itself uniquely identify the patron, especially when compared with other information available from data brokers or vendor-internal datastores obtained by web surveillance. ("Completely anonymous" in the description of the category does not consider reidentifiability and should therefore be removed from the category description for inaccuracy. Additional research on the reidentifiability of individuals from their university-affiliate entitlement lists would be highly desirable.) Moreover, the burden of choosing which entitlements to reveal to the vendor is left on the institution, raising the very real spectre that campus IT may lazily reveal all available entitlements rather than

choosing carefully which entitlements to reveal to which vendors; reidentifiability for at least some patrons under this scenario is all but guaranteed. If, on the other hand, the protocol were designed such that the e-resource vendor provides the institution a list of entitlements, any of which would authorize the patron to use the e-resource, and the institution only responds with whether the patron holds any of those entitlements, that would indeed constitute (and even improve on) physical-equivalent privacy.

The Pseudonymous Authorization category typically will not by itself provide physical-equivalent privacy because of the high likelihood of reidentification (that is, the reassociation of the pseudonym with the patron), and poor defenses against data misuse by vendors. Any patron whose device has a static campus IP address is immediately reidentifiable, as is any patron using a device whose identifier is known to the vendor (as with the Adobe leak). Log analysis, web tracking, and behavioral tracking provide additional avenues to reidentification. The documentation for this category states that "a Service Provider [vendor] claims that it will not use attributes for purposes that fall outside of the service definition as presented at the time of registration to its users and referred to in metadata." This still allows vendors to sign patrons up to privacy-destroying services by disclosing further attribute use during registration for the service, in the typical and heavily-discredited notice-and-consent dark pattern wearily familiar to scholars of internet privacy and law.[57] Only provisions (in the SeamlessAccess standard itself or in contracts between libraries and vendors) that outright forbid attempted reidentification of a pseudonymous identifier and use of patron data outside the provided service can provide physical-equivalent privacy for a patron authorized under the Pseudonymous Authorization category.

Conclusions, cautions, and suggestions for future work

Librarians cannot trumpet our commitment to patron privacy without hypocrisy when we do little or nothing to protect it. Librarians genuinely interested in protecting the privacy of information use must be far more active than we have heretofore been in assessing, challenging, and ameliorating privacy and intellectual-freedom threats from e-resource use.

Part of our difficulty therein lies in the outdatedness of the list of privacy harms and consequences libraries have historically sought to prevent. Certainly law enforcement is still a looming threat, as are nosy neighbors, library employees with poor professional ethics and boundaries, and domestic stalkers. Library privacy guidance has not,

---

[57] Robert H. Sloan and Richard Warner, "Beyond Notice and Choice: Privacy, Norms, and Consent," *Journal of High Technology Law* 14, no. 1 (2014): 370–412, https://heinonline.org/HOL/LandingPage?handle=hein.journals/jhtl14&div=11&id=&page=.

however, fully come to grips with surveillance normalization, surveillance creep, the failings of notice-and-consent as a basis for surveillance, behavioral tracking, reidentification, data brokers, electronic redlining, predictive analytics, or surveillance-based behavior manipulation as yet. Indeed, Bawden and Robinson[58] in explicating Floridi[59] go a step further than merely revising privacy guidance and associated processes by challenging the notional basis on which that guidance has been predicated:

> Fundamental to Floridi's model is the belief that, because personal information plays a crucial constitutive role in who I am and who I can become, protection of privacy should be identified as protection of personal identity and a breach of informational privacy as an aggression against personal identity and self-development. Protection of privacy should be based directly on protection of human dignity, rather than on secondary considerations, such as a right to property, to freedom of expression, or to privacy per se.

Revision of existing library privacy guidance that takes cognizance of modern privacy-endangering phenomena and models of privacy will not be easy but is deeply needed. Legislative advocacy to extend patron physical-records protection to e-resource use is also warranted.

One flaw in the concept of physical-equivalent privacy is what Helen Nissenbaum calls "the tyranny of the normal" and discusses vis-à-vis video surveillance:

> The case of video surveillance may be one such lost cause, so commonplace now that objections are increasingly difficult to carry against the force of the reasonable expectation, against what I regard as the "tyranny of the normal." As long as contextual integrity is tied solely to actual practice, as long as it merely defines a heuristic for detecting effectively when novel practices deviate from entrenched norms, it can be judged an instrument of this tyranny.[60]

Similarly, physical-equivalent privacy is vulnerable to the temptation to violate the spirit of the concept by weakening privacy protections around use of physical spaces and materials—perhaps even via video surveillance!—to justify analogous weakening of

---

[58] David Bawden and Lyn Robinson, "'The Dearest of Our Possessions': Applying Floridi's Information Privacy Concept in Models of Information Behavior and Information Literacy," *Journal of the Association for Information Science and Technology* n/a, no. n/a (May 19, 2020), https://doi.org/10.1002/asi.24367.

[59] Luciano Floridi, "The Ontological Interpretation of Informational Privacy," *Ethics and Information Technology* 7, no. 4 (December 2005): 185–200, https://doi.org/10.1007/s10676-006-0001-7.

[60] H Nissenbaum, *Privacy in Context : Technology, Policy, and the Integrity of Social Life* (Stanford, CA: Stanford Law Books, 2010).

e-resource privacy protections. Nissenbaum's suggestion for resisting such temptations is "to compare entrenched normative practices against novel alternatives or competing practices on the basis of how effective each is in supporting, achieving, or promoting relevant contextual values." In a library context, relevant contextual values from library ethics codes include (but are not limited to) service equity and accessibility, intellectual freedom, and elevating patron welfare over private gain.

Physical-equivalent privacy may be a usable privacy-testing heuristic in more situations that contrast physical with virtual services and spaces, though it can only work when suitable privacy norms and controls already exist for either the physical or virtual side of the comparison. In higher education, comparing the privacy protections of physical classrooms to learning-analytics régimes could curb some of the latter's significant excesses, for example. Comparing management and disclosure of managed physical records to electronic records and to burgeoning workplace surveillance could also be fruitful. Because K-12 school buildings (outside the classroom itself) and public campus spaces are presently surveillance-ethics battlegrounds both physically and virtually, however, comparing the physical situation with the virtual is unlikely to produce useful ethical insight because of Nissenbaum's "tyranny of the normal."

Heuristics are only approximations, often leaky ones. Further work refining and applying the concept of physical-equivalent privacy would of course be welcome, as would additional ethical analysis of specific situations, and pragmatic suggestions for using physical-equivalent privacy and the insights it delivers to assess e-resource license negotiations and terms.

Acknowledgments and declaration of interest

The author wishes to extend her appreciation to NASIG for the invitation to deliver a keynote that sparked the idea at the heart of this paper.

Declaration of interest statement: The author is a co-investigator on the IMLS-funded Data Doubles research project, which studies undergraduate-student perspectives on privacy issues associated with academic library participation in learning analytics initiatives. No Data Doubles data, analyses, or publications past or future were used in this paper, except as cited.