Licensing Privacy Project

Disrupting the Digital Status Quo: Why and How to Staff for Privacy in Academic Libraries

Eliza C. Bettinger
Meryl Bursic
Adam Chandler

June 2023

Project Statement

The Licensing Privacy Project seeks to use the power of library licensing agreements to effect change in third-party platform practices in order to bring them into alignment with library values of privacy, confidentiality, and respect for user control over their own data. It reflects an identified Pathway for Action from the IMLS-supported National Web Privacy Forum. 2 The goal is to develop model license language on user privacy that would support libraries in advocating for user privacy when contracting for services and content. By ensuring that user privacy is contractually protected in licensing agreements, service contracts, etc., libraries would be able to hold platforms accountable for their data practices. The project website is: https://publish.illinois.edu/licensingprivacy/.

Funding and License Statements

This white paper was written by Eliza C. Bettinger, Meryl Bursic, and Adam Chandler as a component of the Licensing Privacy Project led by Lisa Janicke Hinchliffe. This project was made possible in part by a grant from The Andrew W. Mellon Foundation.

Acknowledgements

We would like to acknowledge the generosity of Becky Yoose, LDH Consulting, William Marden, New York Public Library, and Amanda Ferrante, Jessica Winans, and Matt Nelson from EBSCO. They made time to share their insights with us over video conference and email. We also thank the many librarians, journalists, scholars, and organizations we mentioned and cited in this document. Their research and collective action shines light in places that would be invisible otherwise. Finally, we thank our colleagues at Cornell University Library, many of whom are invested in the work to make academic libraries safer, freer, and more equitable spaces, and without whose support this paper would not be possible.

Introduction

In this time of austerity in higher education, prioritizing privacy may sometimes seem like a vague aspiration that library leaders cannot afford (Cooper, 2021).

But we are here to argue the opposite: that staffing and strategizing for meaningful researcher privacy are in fact *essential* to the flourishing of academic libraries. Leaders who ignore these values will do harm to individual scholars and students. They will have surrendered their libraries' centrality to the intellectual life of the university, ceding ground instead to corporate data cartels whose interest is less to support rich discovery experiences than to enclose the research process within siloed and surveilled profit-making systems. They will have squandered the trust that students and scholars have placed in libraries' reputation for an ethical approach to information systems. And, they will have capitulated to a vision for academia, sometimes under pressure from their own institutions, that constrains freedom and reduces the human capacity for learning and imagination to a set of machine-made predictions based on prior behavior, which is a direct contradiction to the popular notion that libraries represent portals to unlimited exploration.

This paper is a roadmap to a shift of both mindset and practice in staffing academic libraries for privacy. Though the ALA Library Bill of Rights has encoded privacy as a central value since 1939, the reality of protecting privacy in the 21st century information environment is vastly more complicated than it was at that earlier time. In Section 1, we will discuss why privacy in libraries is such a thorny problem at this moment, which might help explain why little has been done so far to address the issues. In Section 2, we will explain why, despite these challenges, achieving meaningful reader privacy is an imperative goal, with a look at the harms currently impacting individual users and the scholarly enterprise collectively. Finally, in Section 3, we will recommend where to go from here.

21st-Century Challenges to Privacy in Libraries

The most well-known examples of library privacy practices historically had relatively straight-forward solutions. For example, during the McCarthy Era and other times of repression in the U.S., some libraries made policies to destroy print-based patron borrowing records so that they could not be subpoenaed by authorities. With this simple action alone, individual libraries could independently achieve major improvements to free inquiry protections for their patrons.

Today, protecting privacy in libraries is a much more "wicked problem" than in prior times – difficult to define, much less solve, and interdependent with multiple other aspects of the networked information age.

Patrons Are Not Protected by Laws and Regulations

First, contrary to what patrons and even library workers sometimes assume, U.S laws do not comprehensively protect readers from wide-ranging corporate and institutional surveillance, long-term data retention, and the use of this data by governmental and private customers around the world (Zuboff, 2020). Therefore, a passive approach to privacy, dependent on mere institutional compliance with law, is arguably equivalent to no approach at all.

Second, academic libraries do millions of dollars of business every year with information vendors that have morphed from a collection of publishing houses to a minimally regulated analytics industry, incentivized to monetize as much researcher and student data as possible as their central business model. This evolution presents a major vulnerability. Previous privacy risks could be mitigated by libraries' internal oversight of resources and infrastructure that they owned. Now, libraries pay subscription fees for access to both digital resources and infrastructures owned by the vendor, with

privacy terms governed externally by ongoing contracts. These terms, which include clauses pertaining to data use and sharing, can also change without reasonable notice given to the institution, and the onus is then on the institution to keep pace with and reconcile those changes.

Finally, just as corporate information vendors perform analytics on library patrons, academic institutions are increasingly intertwined with corporate data cartels that collect and retain personal data about their students for the purposes of e-marketing and online enrollment recruitment, EdTech platforms for course delivery and learning analytics, and "student success" and student retention initiatives (Hartman-Caverly, 2019).

From Publisher to Data Cartel: Shifting Roles and Responsibilities

Both libraries and academic institutions are faced with common challenges stemming from their reliance on third party vendors. First, market consolidation within the information services sector continues to narrow the choices of vendors available, where platform lock-in and switching costs are increasingly demonstrated through wielding of monopoly-like power (Giblin and Doctorow, 2022). With this power, non-negotiable contracts are becoming the norm. And, it is not always clear who owns the data collected or generated within these platforms.[1] These transformed data can then be, without explicit permission and implicitly permitted via contractual language, used internally or by any affiliates of the company, and either repackaged or used for other products and service offerings to provide additional competitive edges (Lamdan, 2023).

Another result of this changed landscape is that the boundary between the library and its vendors is becoming less and less clear, and regulatory and pre-digital protections that may have once been effective are no longer. Privacy, in the sense that it provides a barrier to unnecessary intrusion, must be reinforced as an approach that protects the integrity of the library. It can no longer be viewed as a passive option but an active requirement to preserve the library's ability to maintain its autonomy, lead its own decision-making processes, support the academic mission of colleges and universities, and maintain and manage its own direct relationships with patrons.

As Internet Archive founder Brewster Kahle wrote in an e-mail to a journalist recently: "If the library only negotiates access licenses for their students to view publishers' database products, is it a library anymore? Or is it a customer service department for corporate database products?" (Bustillos, 2023). Librarians should fight hard to maintain our independence.

Different Defaults for Data Collection: Libraries and Campus IT

At academic institutions, campus Information Technology (IT) administrators set standards optimized for service uptime and cost control. Left out may be library values like privacy.

For example, a library and campus IT might differ in their default implementations of Single Sign On (SSO), the technology that allows academic users to sign in to university-licensed products from off campus. Central IT administrations frequently set their default SSO attribute release (the amount of user metadata shared with the vendor) to pass on the user's personal details (such as name, email, department, and university status) to vendors because electronic services require it, and it is easier and

[1] Often, data is collected and used internally by companies in two common forms: aggregate (i.e., information that is collected, combined, and summarized at a high level) or "de-identified" (i.e., a term without popular consensus on its definition, but where it's broadly construed as the removal or manipulation of directly identifiable data and known indirect identifiers, such as name, date of birth, gender, contact information, national or state identification numbers, and physical addresses, but, as some privacy advocates observe, where other remaining data elements may still provide enough information to infer the identities of individuals).

faster to release everything for all requests than to customize the attribute release for each service. But, libraries should prefer not to pass all personal details to the vendors that provide library e-resources.

To address this mismatch in priorities, electronic resources staff from the Cornell Library have worked with identity management colleagues in IT to create a workflow removing personal details from the list of SSO attributes released when users connect to library-licensed databases (Chandler and Okerbloom, 2021). Consequently, staff in the library need to be vigilant to keep track of new resources and technology solutions because any InCommon Federation service provider automatically has permission to request attributes from the campus identity provider server.[2]

Another example is Google Analytics, a global personalization, marketing, and behavioral data tracking system. About 15 years ago, the "free" Google Analytics service quickly became the default web analytics software on higher education websites. Librarians who see the problematic nature of using Google Analytics in their library must convince local IT decision makers to use a different product. Like the SSO example, this requires skilled staff with knowledge of comparable alternatives where the library retains the data collected and can determine its use (Chandler and Wallace, 2016). Library staff must be prepared to defend their use case and insist on viable, alternative solutions.

Despite these competing needs and values, the library is uniquely positioned to partner with IT on connecting people to campus IT services through the promotion of digital and information literacy, tailored to the specific needs of various communities. IT can also benefit from learning how their technology choices and service offerings are having an impact on the very populations they are trying to serve. Point being IT is usually more concerned with continuous and highly-available delivery of a "solution" than its potential impact, whereas the library, in contrast, can function as an important sounding board for technologies used on campus.

This is why a meaningful partnership between the library and IT is needed. Librarians tend to have closer proximity to and regular interactions with the student and academic populations on campus, whereas IT is often more closely linked with administrative and operational functions of the institution. However, library staff need help. Even with their closer proximity to students and academics, librarians can fail to see or fully understand the harms that can result from these newer forms of digital privacy risk to which IT may have greater exposure or technical understanding. These harms, both current and accumulating, are real, and pose grave risks to both individual researchers and to the wider practices and values of the academy.

[2]The InCommon Federation, with over 1,000 members, creates standards and recommendations for networked services in higher education. Their default SSO attribute release policy includes personally identifiable information: "InCommon strongly recommends that each InCommon Identity Provider configure a default attribute release policy that permits releasing at least one non-reassigned and permanently unique user identifier. These can be chosen from eduPersonUniqueID, eduPersonPrincipalName, or eduPersonTargetedID" (accessed on May 21, 2023: https://perma.cc/RT33-YK4W).

Harms to Students, Scholars, and Free Inquiry

We find that at times, even librarians who support privacy for readers struggle to articulate clearly why privacy matters for researchers and the harms that can result from researcher surveillance. Here, we attempt to summarize those harms and provide a foundation for anyone who needs to make the case for why greater attention to privacy is warranted.

Surveillance and Repression of Individual Readers and Scholars

One group of users in the U.S. for whom privacy is a particularly present concern are those who are nationals of countries whose home governments both intensely surveil their citizens and maintain repressive practices related to freedoms of speech and expression. Our highly networked global digital information system means that the ability to monitor one's citizens is not restrained by national borders.

As Eliza has been asked in her role as a reference and instruction librarian, "How can I make sure that my government at home doesn't know what I'm reading while I'm here at Cornell? Many of my friends want to know this too."

As a library with some commitment to privacy, we can give these students some partial answers: Check out physical books. (Circulation records are not saved or shared.) Use library computers when searching or reading on the open Internet, and do not sign into any identifying services while doing so. (Our library computers require no logins to access, and traffic logs are deleted with every hard restart, which happens at least daily.) We are glad to be able to offer these options. Central campus IT staff periodically pressure our library to begin to require logins on our public computer terminals. So far, the commitment of library staff and administration, and the existence of a written commitment[3] that we can point to have saved us from having to make that change.

But, if the concerned student wants to privately read current literature from most scholarly journals, we can offer no meaningful safety. Many of the databases where this literature is accessed requires readers to login individually, either with personal and direct accounts, or via institutional login. To a student accessing these resources, they may believe that they are still exploring within the realm of their institution and not a third party, especially if they are using their institution's single sign-on authentication. The information vendor is then capable of collecting and saving detailed information about what that particular user searches, reads, and downloads, perhaps under the guise of "personalization" features. As Sarah Lamdan (2023) and Cory Hanson (2019) have explained, we simply have no knowledge of or control over where those details then go, and one destination could be into the hands of a government that purchases them. Certainly, such details are valuable to governments that are known to go to great lengths to surveil the academic activities of students and researchers both at home and abroad, with serious consequences (Fan, 2022).

Indeed, even putatively democratic governments in the U.S. and elsewhere are currently moving with aggressive speed to exercise legislative control over what scholars may research, say, and teach, what kinds of knowledge they are permitted to pursue in state-supported universities (American Council of Learned Societies, 2023; Human Rights Watch, 2020; Robin, 2023). We do not currently know how far these legislative assaults on academic freedom will go, nor what sorts of surveillance concerns U.S.

[3] Cornell University Library's website has a "Commitment to Privacy" page, which specifically calls out the following position on public computing access: "To help maintain user privacy and confidentiality, we provide computer systems that have anonymous logins and that are programmed to return the kiosk to its original state when restarted. Our computers are also set up to restart after a period of inactivity to help ensure that no identifying information is left behind by the user" (accessed on May 22, 2023: https://perma.cc/U6EG-CCC7).

scholars will face in the future. However, it is highly relevant to note that currently no regulations or policies regulate how long the data cartels may retain data on the online behavior of scholars and students, what other datasets they may combine them with, or to whom these data may be sold or shared. Already, as Lamdan documents, we know that RELX sells data from its other surveillance analytics enterprises to law enforcement agencies.

Notably, the harm caused by this type of digital surveillance is not and will not be borne by all researchers equally. Researchers who study inequalities related to race, gender, sexual orientation, caste, religion, class, and other factors are more likely to bear harm, as are those whose work relates to evidence and interpretation of history, science, or social relations that are at odds with the preferred narratives of the state or of powerful capital interests. Nor does the harm of surveillance data for sale come solely from governments' abilities to buy this data. Large companies with a stake in scholars' scientific research are also known to go to great lengths to interfere with individual scholars' work as much as they can (Quinn, 2023; Oza, 2020). There is no reason to think they would stop short at purchasing legally obtained data about these scholars' reading, browsing, and saving histories.

Risk to Reputation and Trust

In terms of reputation, the last association that libraries should want in the public mind is that they are mere extensions of the Big Tech companies, which are seen increasingly as responsible for mass surveillance, threats to democracy, misinformation, emotional manipulation, and restrictions on free inquiry. But, if libraries continue to rely on scholarship-limiting data cartels to provide key resources and services, that is how their reputation will be built.

For the moment, students tend to be unaware of the extent of personally identifiable information (PII) collection on campus, their institutions' own uses of learning analytics, and the predictive implications of downstream linking of discrete sets of data collected in different times and places (Jones et al., 2020; Asher et al., 2022). Generally, students also do not realize that practices they plainly object to – such as universities sharing their information with commercial entities – are, "to a greater or lesser degree, routine in higher education" (Asher et al., 2022).

As these practices become more publicly known, students sharply object (Young, 2022). On some campuses, the rise of collaboration with companies that share data with government have led to protests and pressure campaigns that result in colleges backtracking on policies (Gurley, 2019; Gurley, 2020; Ongweso, 2020). Thompson Reuters and RELX have been singled out for criticism due to their contracts with U.S. Immigration and Customs Enforcement (ICE), and the "oceanic computerized view of a person's existence" they offer to federal law enforcement and others who may want it (Biddle, 2021).

Meanwhile, ironically, the same companies that collect, retain, analyze, and recombine data about students and researchers in an opaque, minimally regulated environment also impose tight restrictions on scholarly inquiry that makes use of the data that they license to libraries. Libraries are losing the trust of scholars who are dismayed to learn that despite the millions of dollars libraries spend on subscriptions to, for example, corpora of text for computational analysis, they may not be able to produce replicable results, share datasets with other researchers, or combine sets of data from multiple sources, all of which limit the questions they may ask, and the reliability of their results (McCracken and Raub, 2023).

Implications of Datafication, Commodification, and the Future of Academia

In a critically important article, Gendron, Andrew, and Cooper (2022), the editors of a peer-reviewed journal on Elsevier's platform, have documented and contextualized the transformation they have witnessed in the academic journal peer review processes, drilling into how this change manifests itself in the manuscript management and reviewing workflow: "Here we are concerned with the gradual

removal of human involvement as journal editors and reviewers – as AI and automated expert systems take over diverse tasks and judgments historically carried out and performed by individuals" (p. 2).

Their article shines light on one set of the many, often invisible and little studied, ways that data cartels and corporate interests are deeply infiltrating the processes by which new knowledge is created, evaluated, and disseminated, the ways that learning and teaching are conducted and valued, and whose voices are heard, valued, and amplified.

When the journal these writers edit, *Critical Perspectives in Accounting* (*CPA*), was founded in 1990, the amount of clerical work required was significant. Mail was used instead of email, for example. Over time the technological systems created to mirror physical labor speed up the workflow and reduce the amount of drudgery. But each "improvement" on the front end of the editorial system came along with a background monitoring subsystem that feeds the array of clicks and mouse movements into the constantly growing "surveillant assemblage" (Haggerty and Ericson, 2000). All the pieces are in place to create machine learning models that can make predictions, labeled helpfully as "recommendations," about the editor behavior, such as sending a manuscript to a certain reviewer. With each click of the mouse or touchpad, "efficiency" is presumably increased while the variance, the spectrum of human decision making and affect, is reduced.

Most crucially, this process results in editorial decisions that "concentrate power, simplify and speed up academic production, and marginalize non-conformist thinking" – all trends that threaten the autonomy and freedom of scholars. Similarly, learning analytics projects initiated by universities and sometimes libraries themselves threaten the autonomy and liberty of students (Hartman-Caverly, 2019). When we consider adopting new technologies that datafy and commodify learning and research, it is important to recognize that changes of type and not just of quantity that we are bringing to the academic enterprise.

Staffing the Academic Library of the Future

WHY "Staff Up" for Privacy

A New Vision for Libraries: Our Key Role in Sustainable Scholarship

In the current moment, infrastructures for finding, sharing, and disseminating scholarship – the critical foundations of academic research – are undergoing profound transformations, with little check on implications for the future. In this paper so far, we have emphasized the role of corporate and institutional surveillance, privacy loss, and the automation of human intellectual activities. Elsewhere, librarians have identified other profound threats: unsustainable monetary costs demanded of academic institutions, paywalls and closed access that exacerbate global inequities and impede scholarly progress, walled gardens around data that cannot be analyzed or remixed outside of specific corporate platforms, subscription access models that remove long-term preservation from public hands, and institutional pressure to collect and retain behavioral data about students. All of these information harms stem from the same set of technological, social, and economic transformations.

Rather than watching the unfolding impacts from the sidelines, we urge academic libraries instead to assert our central role in recognizing and responding to these transformations, and to continue our core mission of ensuring free and equitable access to reliable information, now and in the future. To do so, we argue that libraries must embrace a new vision for staffing: one that places the safety of scholars and students at the center, where librarians are trusted consultants on navigating digital risks, and each library is a node in a network of institutions working together for equitable free inquiry.

Could Digital Transformations Center Values?

In early 2023, administrators of Vermont State University announced that its several member colleges' libraries would be eliminated as physical entities. Books, shelves, and service desks would be removed. A majority of library staff positions would be slashed. Backlash to the announcement was swift and outraged. Vermonters protested against losing a physical space dedicated to study and intellectual discovery and emphasized the value of physical browsing for the serendipitous discovery of new ideas. Administrators countered that check-out of physical materials had declined sharply since the pandemic began, that digital materials were more important than ever for access by remote and rural students, and, *above all,* that they must cut significant expenses in order for the schools to survive, implying that digital libraries would be less expensive to maintain than print.[4]

Unmentioned in the protests was the fact that by relying entirely on digital library materials, the great majority of which are owned by corporate tech companies, the university would subject its readers to total surveillance of their searches, downloads, and their online attention while they conducted research in this all digital "library." In the economic system academic institutions find themselves in today, if a library is "all digital," it is also nearly "all surveillance" and "all datafied."

Contrary to the cries of protestors in Vermont, it is not that a library is digital, in and of itself, that transforms its nature or betrays its values (D'Agostino, 2023). The betrayal comes when a library violates the privacy of its readers, facilitates the surveillance of their work, and offers up the whole process of knowledge creation to data cartels for extraction of profit from scholarly activities, and perversion of the system (Lamdan, 2023).

Indeed, there is a common misconception that the shift from physical to digital resources means that there is *less* need for independent university libraries and librarians on campuses. The Vermont State administrators appear to have believed this. But library leaders who accept this mindset are sure to become ever more marginal to the research process, as corporate database vendors happily privatize, datafy, and vertically integrate more and more of academic research and publishing.

A Values-Centered Digital Library Needs People

As the use of all types of digital resources grows, the need only becomes more acute for academic staff skilled in understanding them and explaining how they impact research and scholarly communication. The need for professional librarians to advise researchers not just where to find information, but how to do so safely and securely is ever greater.

Creating a digital library that embodies core academic values, ensures privacy and safety, enables quality discovery, and promotes the greater good would require that libraries marshal *more* staff and resources, not fewer. To realize such a vision for values-centered information systems, we need systems librarians to design, build, and maintain our own technology for private and open access. We need research and instruction librarians who can assist researchers to navigate these digital systems to their fullest advantage and collaborate closely with systems librarians to represent researchers' concerns and perspectives.

[4] Since subscriptions to digital materials are generally more expensive than purchases of physical materials, one supposes that the planned cost savings would come primarily by cutting staff who select, organize, curate, and maintain collections, manage borrowing and circulation, liaise with academic departments, teach information literacy and research skills, and answer reference questions. Instead, presumably, these services would either be outsourced to the corporate tech companies that provide digital library material for lease, or the college would go without.

HOW to "Staff Up" for Privacy

Define and Defend the Library's Values

1. **Write and publish a values statement. Articulate the library's commitment to privacy.** Does the library's mission and/or values statement mention privacy, intellectual freedom, democracy, and/or free inquiry?[5] Can library leaders and staff articulate the importance of digital privacy to researchers in the contemporary information landscape? As an example, 13 large research libraries co-authored a "Statement on Patron Privacy and Database Access" (2019).

   The methods by which researchers find and access information are different now than they were when the Library Bill of Rights was adopted in 1939 (American Library Association, 2019); however, the values codified there are just as critical as when they were first written. When we articulate our values, re-examined and re-interpreted for a new time, we create space in which to act, and build a guide for strategic planning. Statements of values and commitments do not change our conditions by themselves, but they can be a tremendous help in facilitating and supporting action.

2. **Stake out a set of technology ethics that will differentiate the academic library from Big Tech.** The fact that students and researchers trust libraries (Asher et al. 2022) is a key asset worth preserving. Articulate the ways that the library treats information ethically that is different from the commercial companies that people find less trustworthy.

3. **Build a narrative that frames enhancements to privacy as innovative.** Making changes that help research thrive is innovative. Finding ways to protect privacy in the current information landscape will require new experimentation and creativity. When one frames the library's commitment to privacy as a commitment to innovation within the digital space, one creates a narrative that is exciting and positive, and that can help build support internally and externally.

Develop Library Staff Privacy Skills and Competencies

4. **For all new hires across the library, make critical information literacy, information ethics, and/or privacy knowledge a job requirement.** The library may not be able to create new positions right away, but when hiring for *any* position, a library can look for new staff with the ability and interest to critically examine information infrastructures, and to incorporate privacy considerations into their work.

   Particularly for entry-level jobs, it is less important that candidates have specific privacy-related experiences already than that they have a curiosity and a critical orientation toward "looking under the hood," learning how digital information systems work and their impact on research and researchers (Beck et al., 2021). These are the staff who will be most likely to devise innovative strategies for teaching, vendor negotiations, and system designs in the future.

   Relatedly, candidates with these interests are often looking for jobs where their interests will be welcomed and supported. By including these requirements in a job ad, the library will attract candidates who may not apply to the library otherwise (Bettinger, 2022).

[5] Valentine and Barron (2022) found that most ARL libraries do not.

5. **Train existing staff.** Depending on the environment, create learning opportunities in the library or empower staff to seize opportunities that interest them.

   A wide range of opportunities exist. Professional organizations like the Library Freedom Project and private consulting services staffed by former librarians offer trainings and professional development on privacy topics for librarians. The ALA's Privacy Field Guides are designed to provide easily digestible information and practical action steps related to various privacy topics (Tijerina and Berman, 2018) and any one of them could form the centerpiece of a discussion in a staff meeting. A university's Computer Science, Information Science, Communication, Sociology, or Science and Technology Studies department might offer privacy-related courses relevant to different library roles. A growing number of library and digital scholarship conferences offer panels and workshops related to privacy and surveillance.

6. **Certify at least one library staff member as an IAPP Certified Information Privacy Technologist (CIPT) and empower them to act as a clearinghouse for privacy-related questions and analysis.** The International Association for Privacy Professionals (IAPP) offers several different certifications that are considered the gold standard in industry and government for managing privacy risk and evaluating technical systems for privacy. A librarian with the CIPT certification will be well-equipped to take on a part-time role as a privacy consultant for internal privacy concerns related to vendors and technical systems or for patron research concerns. A large organization may benefit from more than one CIPT certified staff member, with one housed in a reference/instruction unit, and another housed in a technical services, IT, or electronic resources unit.

   CIPT certification requires continuing education to stay current. Investing in and maintaining CIPT certification for staff brings with it the benefit of access to cutting-edge information, and the potential for the privacy specialist to offer new trainings internally. It is essential to empower a CIPT with a meaningful role in decision-making processes and time to contribute to privacy-related services in the library.

Develop Privacy-Enhancing Services and Practices in the Library

We have found that once we began with small steps toward privacy services in the library, the needs of our patrons began to present themselves, and new opportunities for services arose. We suggest starting small and following where the ideas of library staff and feedback from patrons lead. Here are some ideas for getting started.

7. **Model the risks faced by library patrons, librarians, and libraries.** Privacy risks exist for students, researchers, and scholars in the course of doing their work, and institutions risk losing trust and reputation if they mismanage private information or fail to provide relevant guidance and safety measures. Encourage staff across units and functions in the library to learn about and map out the risks faced by the patrons they serve, and/or the units they are part of. This process can help the harms of privacy loss to feel more concrete and less vague. Start learning about risk modeling at the Electronic Frontier Foundation (Electronic Frontier Foundation, 2021) and the Global Cyber Alliance (Global Cyber Alliance, n.d.).

8. **Adopt Privacy by Design and Physical Equivalent Privacy frameworks for making decisions about vendor contracts and designing library systems.** Privacy by Design, a set of principles conceived in 2009 by a then official in the provincial government of Ontario, Canada, and since widely adopted by privacy-concerned regulators, requires that organizations treat privacy as a

default, instead of a principle to consider after other key decisions have been made (Cavoukian, 2011; Yoose, 2019). After a library articulates its values, and staff have modeled risks faced by patrons, a Privacy by Design framework can guide the design of systems and contracts that build in privacy and safety from the start.

The concept of "physical-equivalent privacy" is that any data collection practice that causes users of electronic resources to enjoy less privacy than users of the resource's physical equivalent should be avoided. This concept is useful for thinking through potential harms of vendor contracts and digital library systems (Salo, 2021).

9. **Embed privacy consultations and instruction into information literacy, reference, and research programs.** Privacy should be among the digital information equity issues that immediately bring the library to researchers' minds. When they are concerned about privacy, they should know that they can turn to trusted liaison librarians for help, just as when they have questions or concerns about open access, copyright, algorithmic awareness, understanding artificial intelligence, or data management.

   Becoming a trusted resource in this area requires reference librarians to understand privacy concerns related, not only to library resources, but also to the wider world of information resources on the open web and the multiple networks that researchers will connect to during the research lifecycle.

   Two useful instruction examples to get started are "Teaching the Right to the Future Tense," a privacy literacy workshop by two reference and instruction librarians (Hartman-Caverly and Chisholm, 2023) and the "Online Harassment Field Manual" from PEN America that presents resources for scholars, students, writers, or activists targeted for harassment (2023).

10. **Flag products with mandatory personalization for risk analysis at the beginning of licensing negotiations.** Good stewardship of behavioral data begins with a careful accounting of which library vendors collect it when patrons search and read licensed content. Some vendors require individual library patrons to create personal accounts to use their system, while others gather data without explicit consent, utilizing methods such as cookies or browser fingerprinting.

    At start of negotiations, library staff should ask the vendor to what extent personalization features in their product offering or conduct their own investigation to determine the personalization status of the product. In cases where patrons are not given a choice to opt-in or opt-out of personalization with clear notice provided, a thorough review should be conducted by the library to understand the full extent of the risk to patrons. The Vendor Contract and Policy Rubric developed by Becky Yoose as part of the Licensing Privacy project is designed to enable this assessment (LDH Consulting Services, 2022).

    If library leadership makes the decision to license a resource that includes mandatory personalization, the catalog record for this service should be annotated and a warning from the library to patrons about known risks should posted on the resource's service or login page as well as on the SSO connection page.

Advance Leadership, Policy, and Collective Action

11. **Work toward the goal of creating a high-level position with ultimate responsibility for navigating privacy-related decisions in the library.** In some organizations, like the New York Public Library, this position has the title of "Privacy Officer;" in organizations governed by the GDPR, the title is "Data Protection Officer." Sometimes, the person in this role is a lawyer. Regardless of the title, this person should have a seat at the table for strategic operations and negotiations, and function with some degree of independence for their privacy responsibilities, , preferably reporting directly to the library dean or university librarian.

12. **Join collective efforts.** The privacy challenges libraries face are collective ones, and ultimately, they can only be addressed collectively. Organizations like SPARC, Library Futures, ALA and ACRL, and regional resource-sharing consortia are all pursuing or have the potential to pursue collective work in areas like negotiating with vendors, creating and supporting library-controlled infrastructure, and lobbying for public policy reform. Libraries should investigate the collective work currently underway and invest library staff time and membership dues strategically to help aid these efforts.

    Here we echo Sarah Lamdan's call to envision public information and its search and retrieval infrastructure as a shared public good (2023) and Rebecca Giblin and Cory Doctorow's urgent appeal for collective action (2022). The companies from whom libraries license content have built tremendous power. For libraries to respond effectively, it will be useless to act only as atomized units; libraries must find and use the power that comes from acting together. Moreover, well-resourced institutions will need to learn about and act in consideration of the needs and constraints of less-resourced institutions in order to create an information environment that serves all.

Conclusion

In preparing this paper, we have aimed to take seriously the warning made by researcher Erin Glass: "If the academic believes in democracy, then the academic should recognize that the digital status quo represents a serious threat to its development and survival" (Glass, 2021).

We believe that the digital status quo represents a serious threat to democracy as well as to the universities and libraries that strive to nurture free inquiry and intellectual autonomy. We urge the leaders of academic libraries to actively stake out a role for their institutions and their consortia in creating a new digital status quo for research and learning, one that centers privacy, safety, autonomy, free inquiry, and equity.

About the Authors

Eliza Bettinger

Eliza leads Cornell University Library's research and teaching services for digital scholarship in the humanities and related social sciences. She also teaches students and consults with researchers across the university about issues related to digital privacy, surveillance, and digital harassment, and is a member of the Library Freedom Project. Uniting all her work is a concern for helping people make meaning from, and exercise control over, data, information, and narratives.

Meryl Bursic

Meryl is a Senior Security Engineer within the Cornell University IT Security Office. She supports various IT governance, risk, and compliance efforts across campus, in addition to providing privacy and security consultations. Meryl's areas of focus are research security, data ethics, privacy and security-related regulatory monitoring, risk-based security assessments and mitigation plans, and NIST standards implementation guidance.

Adam Chandler

Adam is Director of Automation, Assessment, and Post-Cataloging Services at Cornell University Library. He has over two decades of experience in academic libraries, leading projects locally and nationally, on topics ranging from licensing standards for electronic resources, metadata quality, usability, to patron privacy. Adam's current areas of focus are privacy and the relationship between ebook metadata quality and discovery.

References

American Library Association. 2019. “Library Bill of Rights.” https://www.ala.org/advocacy/intfreedom/librarybill. Accessed May 17, 2023.

American Council of Learned Societies. 2023. “The Effort to Undermine Academic Freedom in Florida House Bill 999.” 2023. March 3, 2023. https://perma.cc/86LG-SNZT.

Asher, Andrew D., Kristin A. Briney, Kyle M. L. Jones, Mariana Regalado, Michael R. Perry, Abigail Goben, Maura A. Smale, and Dorothea Salo. 2022. “Questions of Trust: A Survey of Student Expectations and Perspectives on Library Learning Analytics.” *The Library Quarterly* 92 (2): 151–71. https://doi.org/10.1086/718605.

Beck, Estee, M. Ellen Goin, Andrew Ho, Alexis Parks, and Stephen Rowe. 2021. “Critical Digital Literacy as Method for Teaching Tactics of Response to Online Surveillance and Privacy Erosion.” *Computers and Composition*, Rhetorics of Data: Collection, Consent, & Critical Digital Literacies, 61 (September): 102654. https://doi.org/10.1016/j.compcom.2021.102654.

Bettinger, Eliza C. 2022. “Eliminating Alibis.” *Inside Higher Ed*. May 9, 2022. https://perma.cc/JG49-D8TC.

Biddle, Sam. 2021. “LexisNexis to Provide Giant Database of Personal Information to ICE.” *The Intercept*. April 2, 2021. https://perma.cc/A6S7-SZC9.

Bustillos, Maria. 2023. “Just Because ChatBots Can’t Think Doesn’t Mean They Can’t Lie.” *The Nation*, March 17, 2023. https://perma.cc/G9UE-7QPA.

Chandler, Adam, and Melissa Wallace. 2016. “Using Piwik Instead of Google Analytics at the Cornell University Library.” *The Serials Librarian* 71 (3–4): 173–79. https://doi.org/10.1080/0361526X.2016.1245645 .

Chandler, Adam, and John Mark Ockerbloom. 2021. “Practical Protection of Library Patron Privacy in Single-Sign On.” Presented at the Digital Library Federation 2021 Forum, Virtual Conference, November 1. https://hdl.handle.net/1813/113073.

Cavoukian, Ann. 2011. “Privacy by Design: The 7 Foundational Principles.” Information and Privacy Commissioner of Ontario. https://perma.cc/RGT4-M4QB.

Cooper, Danielle. 2021. “Licensing Privacy: Views from Library Leadership.” Licensing Privacy Project. https://perma.cc/M2KE-4LCX.

D’Agostino, Susan. 2023. “Vermont State University’s ‘All-Digital’ Library Fiasco.” *Inside Higher Ed* (blog). February 16, 2023. https://perma.cc/VBK5-XCDH.

Electronic Frontier Foundation. 2021. “Your Security Plan: Surveillance Self Defense.” February 21, 2021. https://ssd.eff.org/module/your-security-plan.

Fan, Jiayang. 2022. “How Dissent Grows in China.” *The New Yorker*, December 7, 2022. https://perma.cc/67UU-TP6L.

Global Cyber Alliance. n.d. “GCA Cybersecurity Toolkit for Journalists.” Accessed May 22, 2023. https://gcatoolkit.org/journalists/.

Gendron, Yves, Jane Andrew, and Christine Cooper. 2022. “The Perils of Artificial Intelligence in Academic Publishing.” *Critical Perspectives on Accounting* 87 (September): 102411. https://doi.org/10.1016/j.cpa.2021.102411.

Giblin, Rebecca, and Cory Doctorow. 2022. *Chokepoint Capitalism: How Big Tech and Big Content Captured Creative Labor Markets and How We'll Win Them Back*. Boston: Beacon Press.

Glass, Erin Rose. 2021. "Reprogramming the Invisible Discipline: An Emancipatory Approach to Digital Technology through Higher Education." In *People, Practice, Power: Digital Humanities Outside the Center*, edited by Anne McGrail, Angel David Nieves, and Senier Siobhan. Debates in the Digital Humanities. Minneapolis, London: University of Minnesota Press. https://perma.cc/V5NZ-XPS6.

Gurley, Lauren Kaori. 2019. "Students at 16 Universities Are Protesting Palantir's Presence on Campuses." *Vice* (blog). November 19, 2019. https://perma.cc/362A-QXV.

Gurley, Lauren Kaori. 2020. "Students and Maintenance Staff Protest Surveillance at Wesleyan University." *Vice* (blog). March 5, 2020. https://perma.cc/FBM8-FCU9.

Haggerty, K. D., and R. V. Ericson. 2000. "The Surveillant Assemblage." *British Journal of Sociology* 51 (4): 605–22.

Hanson, Cody. 2019. "User Tracking on Academic Publisher Platforms." Presented at the Coalition for Networked Information Spring 2019 Member Meeting, St. Louis, MO, April. https://perma.cc/BUL9-6B4A.

Hartman-Caverly, Sarah, and Alexandria Edyn Chisholm. 2023. "Teaching the Right to the Future Tense: A Privacy Literacy Workshop." https://scholarsphere.psu.edu/resources/3cd44531-a18e-42c2-8c46-fac22f4c3fba.

Hartman-Caverly, Sarah. 2019. "Human Nature Is Not a Machine: On Liberty, Attention Engineering, and Learning Analytics." *Library Trends* 68 (1): 24–53. https://doi.org/10.1353/lib.2019.0029.

Jones, Kyle M. L., Andrew Asher, Abigail Goben, Michael R. Perry, Dorothea Salo, Kristin A. Briney, and M. Brooke Robertshaw. 2020. "'We're Being Tracked at All Times': Student Perspectives of Their Privacy in Relation to Learning Analytics in Higher Education." *Journal of the Association for Information Science and Technology* 71 (9): 1044–59. https://doi.org/10.1002/asi.24358.

Lamdan, Sarah. 2023. *Data Cartels: The Companies That Control and Monopolize Our Information*. Stanford, California: Stanford University Press.

LDH Consulting Services. 2022. "Developing the Vendor Contract and Policy Rubric." Licensing Privacy Project. https://perma.cc/HRL6-8S9E.

McCracken, Peter, and Emma Raub. 2023. "Licensing Challenges Associated with Text and Data Mining: How Do We Get Our Patrons What They Need?" *Journal of Librarianship and Scholarly Communication* 11 (1). https://doi.org/10.31274/jlsc.15530.

Ongweso, Edward. 2020. "Shareholders Push Thompson Reuters to End Intimate Ties with ICE." *Motherboard.* May 26, 2020. https://perma.cc/38DK-QGHY

Oza, Anil. 2020. "Two Professors Faced Years of Harassment for Defying the Fossil Fuel Industry. Now, They Are Reframing the Discussion Around Fracking." *The Cornell Daily Sun*. November 16, 2020. https://perma.cc/NQ6L-EB4Y.

PEN America. 2023. "Online Harassment Field Manual." Online Harassment Field Manual. 2023. https://onlineharassmentfieldmanual.pen.org/.

Quinn, Ryan. 2023. “Professor Says He Was Barred from Campus After FOIA Inquiry.” *Inside Higher Ed*. February 23, 2023. https://perma.cc/W7U2-VDZN .

Robin, Ron. 2023. “Why Israel’s Universities Stood Up for Democracy.” *Inside Higher Ed*. March 30, 2023. https://perma.cc/5VKQ-JEW8.

Salo, Dorothea. 2021. “Physical-Equivalent Privacy.” *The Serials Librarian* 81 (1): 20–34. https://doi.org/10.1080/0361526X.2021.1875962.

“Statement on Patron Privacy and Database Access.” 2019. Stanford Libraries. https://perma.cc/2WJ3-39W9.

Tijerina, Bonnie, and Erin Berman. 2018. “Privacy Field Guides for Libraries.” https://libraryprivacyguides.org/.

Valentine, Greta, and Kate Barron. 2022. “An Examination of Academic Library Privacy Policy Compliance with Professional Guidelines.” *Evidence Based Library and Information Practice* 17 (3): 77–96. https://doi.org/10.18438/eblip30122.

Yoose, Becky. 2019. “Baking Privacy Into Your Library: The What, How, and Why of Privacy by Design.” *LDH Consulting Services* (blog). April 1, 2019. https://ldhconsultingservices.com/baking-privacy-into-your-library-the-what-how-and-why-of-privacy-by-design/.

Young, Stephen. 2022. “GUEST ROOM. Hail Cornell! Patron of Digital Sovereignty?” *The Cornell Daily Sun*. September 27, 2022. https://perma.cc/WFE8-5EB9.

Zuboff, Shoshana. 2020. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: Public Affairs.