Evidence Based Library and Information Practice

Research Article

An Examination of Academic Library Privacy Policy Compliance with Professional Guidelines

Greta Valentine
Data & Research Analyst
University of Kansas Libraries
Lawrence, Kansas, United States of America
Email: greta.valentine@ku.edu

Kate Barron
Research Data Curator
Stanford Libraries
Stanford, California, United States of America
Email: katebar@stanford.edu

Received: 15 Mar. 2022 **Accepted**: 7 June 2022

DOI: 10.18438/eblip30122

Abstract

Objective – The tension between upholding privacy as a professional value and the ubiquity of collecting patrons' data to provide online services is now common in libraries. Privacy policies that explain how the library collects and uses patron records are one way libraries can provide transparency around this issue. This study examines 78 policies collected from the public websites of U.S. Association of Research Libraries' (ARL) members and examines these policies for compliance with American Library Association (ALA) guidelines on privacy policy content. This overview can provide library policy makers with a sense of trends in the privacy policies of research-intensive academic libraries, and a sense of the gaps where current policies (and guidelines) may not adequately address current privacy concerns.

Methods – Content analysis was applied to analyze all privacy policies. A deductive codebook based on ALA privacy policy guidelines was first used to code all policies. The authors used consensus coding to arrive at agreement about where codes were present. An inductive codebook was then developed to address themes present in the text that remained uncoded after initial deductive coding.

Results – Deductive coding indicated low policy compliance with ALA guidelines. None of the 78 policies contained all 20 codes derived from the guidelines, and only 6% contained more than half. No individual policy contained more than 75% of the content recommended by ALA. Inductive coding revealed themes that expanded on the ALA guidelines or addressed emerging privacy concerns such as library-initiated data collection and sharing patron data with institutional partners. No single inductive code appeared in more than 63% of policies.

Conclusion – Academic library privacy policies appear to be evolving to address emerging concerns such as library-initiated data collection, invisible data collection via vendor platforms, and data sharing with institutional partners. However, this study indicates that most libraries do not provide patrons with a policy that comprehensively addresses how patrons' data are obtained, used, and shared by the library.

Introduction

The tension between upholding privacy as a professional value and the ubiquity of collecting patrons' data to provide online services is now a common one in libraries. Patrons who use digital library services are constantly providing the library with their personal data whether they know it or not. As stewards of this data, librarians are obligated to be transparent about the uses of patron data to provide services or make continuous improvements to these services. Privacy policies that detail the collection and use of patron data are one way to provide such transparency. While guidelines for creating comprehensive privacy policies exist, literature indicates that these guidelines are often not applied to actual policies. This study examines privacy policies from U.S. academic libraries and describes the content of these policies in relation to privacy guidelines. It also describes content contained in these policies which is not addressed by existing guidelines. Ultimately, this overview provides an environmental scan of recent policies across academic libraries in the US. The trends and gaps in this exploratory scan can inform the creation of robust policies that adequately address current privacy concerns in academic libraries.

Literature Review

Libraries and Privacy

Librarians have long been advocates of privacy, largely in ways that emphasize the protection of users' information and reading behavior. The American Library Association's (ALA) 1939 Code of Ethics outlines support for the privacy of users to pursue topics of interest without surveillance or punishment. Since that time, ALA's Office for Intellectual Freedom and the Intellectual Freedom Committee, along with the associated Privacy Committee, have produced a variety of documents expanding on circumstances in which user privacy should be protected. These include the *Policy on Confidentiality of Library Records* (1986), *Privacy: An Interpretation of the Library Bill of Rights* (2002, amended 2014 and 2019),

Resolution on the Retention of Library Usage Records (2006), and the *Policy Concerning Confidentiality of Personally Identifiable Information About Library Users* (2004) (Vaughan, 2020).

In addition to its ethical documents, the ALA first produced the *Privacy Tool Kit* in 2005, providing librarians with a set of practical resources focusing on privacy (ALA, 2014). The Tool Kit includes a list of privacy guidelines and checklists for navigating the exchange of data when using networked devices, assistive technology, public access computers, and other contexts that commonly require the collection or storage of patron data. Much of the content of the Tool Kit is included in the Issues and Advocacy section of ALA's website and is periodically updated (ALA, 2017a). The most recent version contains resources that support librarians in advocating for "the right to read, consider, and develop ideas and beliefs free from observation or unwanted surveillance by the government or others" (ALA, 2021a, para. 1), including a *Privacy Field Guide* that addresses privacy policies specifically (ALA, 2021c).

When this study was initiated, the *Privacy Field Guide* was not yet available. The primary documents referenced in this study are the *Privacy and Confidentiality Policy Checklist* (ALA, 2017b) and the page titled *Developing or Revising a Privacy Policy* (ALA, 2017c), which was available in the *Privacy Audits* section of the website. These resources were chosen since ALA's guidelines, ethical documents, and the *Privacy Tool Kit* are widely cited in library literature that addresses the creation or evaluation of library privacy policies (see Magi, 2010; Nichols Hess et al. 2014; and Vaughan, 2020 for a sampling).

Threats to Privacy in Libraries

Libraries today face a variety of external and internal threats to privacy. The primary tensions discussed in library literature tend to focus on digital technologies that proliferated over the past decade and have come to shape the way patrons interact with library resources. While libraries have always required some level of information about patrons to provide core services such as circulation and interlibrary loan (Coombs, 2005), early analog library accounts were established with information provided explicitly by the patron and used during discrete transactions, often face to face. Today, simply accessing a library's website can provide the site host with data such as a user's IP address, geolocation, cookies, and other potentially identifying information without notifying the user or requesting consent. Common cloud-based online platforms such as discovery layers, databases, and online public access catalogs (OPACs) are "largely based on the tracking, collection, and aggregation of user data" (Kritikos & Zimmer, 2017, p. 24). Libraries require this type of data collection when patrons discover and access proxied academic articles on digital platforms (O'Brien et al., 2018; Pekala, 2017), as well as any time they download materials, search the web, swipe an ID card, or log into a virtual environment (Jones & Salo, 2018).

Third-party platforms have introduced another layer of tension into libraries' attempts to provide quality service while safeguarding patron data. Some of these platforms implement their own data collection for analytics or provide options for customization and personalization of their services. When patrons provide their information to create a personalized browsing experience, libraries no longer have oversight of this data (Magi, 2010). In this complex data landscape, some literature posits that a library "can no longer be considered the sole gatekeeper of its patrons' private information, emphasizing the present reality that data privacy can be confusing, ambiguous, and opaque" (Vaughan, 2020). Opacity contributes to what Affonso and Sant'Ana (2018) define as information asymmetry, a power imbalance between information holders and users in which those who hold the data have more power. To access core library services such as checking out books, using interlibrary loan, or searching online databases, patrons must now participate in this invisible data collection – a threat to privacy insofar as privacy allows users a measure of control over releasing their data to others (Crawford & Schultz, 2014).

In addition to data collection for the provision of library services, it has become increasingly common for libraries themselves to mine or reuse patron data for purposes that go beyond providing a service. Academic libraries especially face pressure to collect and analyze student data to demonstrate the value of library services (Prindle & Loos, 2017). However, with their accompanying privacy concerns, projects that fall into the category of learning analytics have proven a contentious example of a library-initiated assessment (Rubel & Jones, 2016). Some literature advocates for such studies, contending that the benefits outweigh the risks (Beile et al. 2017; Jones, 2010; Oakleaf, 2010; Oakleaf 2018) or that these studies can be implemented in an ethical manner (Drachsler & Greller, 2016). There is widespread adoption of learning analytics in academic libraries. Of the 53 Association of Research Libraries (ARL) libraries that responded to a 2018 survey focused on learning analytics adoption, 83% indicated they were participating in learning analytics projects. However, there has also been a backlash against these types of library-initiated studies, pointing out ways they threaten patrons' privacy (Farkas, 2018; Jones & Salo, 2018; Perry et al., 2018; Prindle & Loos, 2017). This single but prominent example of library-initiated projects that collect and retain patron data outside the scope of providing a service exemplifies a relatively recent shift in attitude toward privacy within the library profession (Asher, 2017; Prindle & Loos, 2017).

The Role of Privacy Policies

Part of the concern regarding privacy violations in libraries deals with a lack of transparency and consent that would allow patrons the agency to control information about themselves or make informed decisions about how to interact with the library. One way that libraries deal with opacity surrounding data collection is to provide a privacy policy that details the purposes for which patrons' data are collected and used. This is only one purpose of a policy, there are many that focus on library values and guidance for staff. A comprehensive privacy policy allows the library to set privacy protection standards (Yoose, 2018), as well as foster appropriate conduct, consistency, and uniformity in how privacy is implemented while reducing confusion and empowering library workers (Magi, 2007). These policies also protect the organization and provide guidance should legal action arise (Magi, 2007).

However, much library literature argues that patrons are the primary audience for a library's privacy policy. Privacy policies can signal a commitment of integrity to a library's patrons (Vaughan, 2020), can represent "an enforceable guarantee to users" (Briney et al., 2018, p.15), and can act as "a first step in removing barriers to information and....closing the information divide" (Voeller, 2007, p. 18). Nichols Hess et al. (2014) document how 11 areas of library service encounter personally identifiable information, and how a privacy policy should address each of these scenarios to maximize appropriate handling of patrons' data. The ALA (2017c) states that a well-defined privacy policy should tell patrons how their data are being used and in what circumstances it might be disclosed. A policy that outlines how data are obtained, stored, and used is a baseline step towards allowing patrons to participate in informed consent, and has been shown to increase patrons' trust in the library (Nichols Hess et al., 2014; Sutlieff & Chelin, 2010).

The ALA (2017a, para. 4) stresses the importance of posting these policies where patrons can find them, indicating that patrons "have the right to be informed what policies and procedures govern the amount and retention of personally identifiable information, why that information is necessary for the library, and what the user can do to maintain his or her privacy." Further, literature indicates that patrons use these policies to guide their browsing and transaction decisions (Vaughan, 2020) and to increase awareness of how the library collects data about them (Affonso & Sant'Ana, 2018). The absence of a public policy deprives patrons of the ability to communicate discomfort with the use of their data or to withdraw consent (Asher et al., 2018). While simply posting a privacy policy is not an adequate method

of obtaining consent from a research perspective, it is the method many libraries employ, as there is often no means for patrons to indicate whether they want to supply or withhold data from the library beyond what is outlined in the privacy policy (Asher, 2017).

Issues With Library Privacy Policies

Despite the importance of privacy policies, several common threads in the literature indicate issues with current library policies. One of the most prevalent is the issue of locating policies to analyze. This issue is cited in studies across library types in both recent and more dated library privacy policy analyses (Affonso & Sant'Ana, 2018; Magi, 2007; Vaughan, 2020; Voeller, 2007). This does not necessarily mean that the libraries studied did not have a privacy policy in place, but the fact that researchers were unable to locate policies in multiple instances is a concern from the viewpoint of transparency to patrons. In 2018, 10% of 50 ARL libraries responding to a survey about learning analytics indicated they did not have a privacy policy in place (Perry et al., 2018).

A related issue is that some libraries may refer to the privacy policy of their parent institutions rather than a library-specific policy. Sturges et al. (2003) found that few of the 336 higher education and special libraries surveyed had a privacy policy separate from their parent organization. The same ARL survey mentioned above indicated that in 2018, 45 responding ARL member libraries (90%) had a privacy policy in place, but only 31 (62%) had a policy separate from that of their parent institution (Perry et al., 2018). The ALA guidelines (2017c) indicate that library privacy policies should comply with the policies of their parent institutions. However, in addition to this compliance there are cases where libraries may want to implement privacy practices more pertinent to the library itself, which may have goals and values distinct from those of the parent institution (Nichols Hess, et al. 2014).

Studies also highlight that many current privacy policies are not comprehensive enough to adequately inform patrons about library privacy practices. According to Asher et al. (2018), "[f]ew to no institutions have established comprehensive policies and procedures around collection, retention, use, and reuse of student and employee data, in research or for any other reason" (p. 5). Some policies are simply not detailed enough. Voeller (2007) indicated that most of the 30 policies she examined contained less than 10 sentences, suggesting that they were unlikely to cover information in adequate depth.

Another concern is that many policies have not been updated to address current and emerging privacy concerns (Nichols Hess et al., 2014). A 2018 (Perry et al.) study surveyed ARL libraries about whether they had updated their privacy policies to reflect participation in learning analytics activities. Of the 53 libraries that responded, 43 indicated that they participated in learning analytics activities. However, only 7 (13%) had updated their privacy policies to account for this activity. Complexities in protecting patron privacy exist where such practices occur in libraries, yet policies across the profession do not provide adequate guidance in navigating these situations. This is one example of a myriad of cases where data collection has become prevalent in ways not conceived of by older privacy policies (Nichols Hess et al., 2014).

Aims

The library profession continues to tout privacy as a professional value despite the current challenges in implementing it. While a comprehensive, easily accessible privacy policy is certainly not the only way to uphold this value, it is a good first step.

This study aims to provide library policymakers with an overview of policy content across a subset of academic libraries in the US, and to compare these library privacy policies against professional guidelines. This study also examines policy content *not addressed* in professional guidelines, as the guidelines themselves may need to be updated or clarified. By analyzing policy content not addressed in the guidelines, the study results can describe content that librarians have deemed important enough to include in a policy beyond what professional guidelines have suggested.

This overview attempts to answer the following questions:

- Are the privacy policies of ARL libraries in the US readily available to patrons?
- Do these policies contain each of the elements specified in the ALA's *Privacy and Confidentiality Policy Checklist*?
- Do these policies contain additional elements not specified in the ALA checklist? What is the nature of these additional elements?

Methods

Target Population and Sampling

While libraries of all kinds face the challenge of dealing with patron data in an ethical manner, academic libraries face unique pressures to demonstrate and quantify their value, resulting in more retention of patron data (Tenopir, 2010). For this reason, the authors chose ARL institutions in the US as the study population and gathered privacy policies from the websites of these 99 libraries in the summer of 2019. The number of ARL libraries was manageable to examine within the scope of this study meaning that there was no sampling; the authors attempted to obtain policies from all 99 libraries.

Data Collection and Analysis

This study employed direct observation to locate privacy policies that were publicly available on the websites of ARL libraries. This method of collecting policies was important because it mimics the steps patrons must take to locate information about how the library uses their data. The authors located 78 privacy policies from the 99 libraries in the study population, for a success rate of 79%.

The authors employed document analysis, a subset of content analysis, to examine the text of these policies. Document analysis allows for the examination of material that has not been created or modified for the purpose of the study by either the researcher or the subject (Atkinson & Coffey, 1997; Bowen, 2009). This approach aims to limit the bias that a method like interviews or surveys may introduce, in which library staff may be inclined to respond to questions about the potentially sensitive topic of patron data stewardship in ways that reflect more favorably on their organization.

Document analysis is also ideal in that its systematic, iterative nature allows for detailed investigation of complex topics (Erlingsson & Brysiewicz, 2017). Privacy can be abstract, complex, and context-dependent, so a method that allows for a nuanced approach is ideal (Bengtsson, 2016). Document analysis can be implemented using both a deductive and an inductive approach. This study employed both, first using a deductive codebook based on professional guidelines to code all policies, followed by a second round of inductive coding to illuminate themes in text that remained uncoded after the initial round. Because the study's research aims focused on whether content was present or absent in each policy, the full policy was used as the unit of analysis rather than a line, paragraph, or other segment of text.

Deductive Codebook Development

The authors created an initial version of the deductive codebook based on the Policy Checklist document included in ALA's (2017b) online *Privacy Tool Kit*. Several documents in the Tool Kit provided guidance for shaping a privacy policy including the much more in-depth "Sections to Include in a Privacy Policy" (ALA, 2014). However, the "Sections to Include" document focused on privacy best practices and behavior in addition to policy content. In comparison, the Policy Checklist provided a concrete basis for an initial list of codes that captured the spirit of the "Sections to Include" document while focusing on policies specifically. The authors provided the draft list of codes derived from the checklist to four colleagues who were asked to apply them to the original text of the Checklist. Based on their responses, the authors refined the code definitions. When colleagues who had no familiarity with the *Privacy Tool Kit* documents could apply the codes with a high level of accuracy, the draft codebook was considered complete.

The authors then piloted the draft codebook on five policies outside of the study. They met to review this initial coding, resolve discrepancies, and adjust definitions in the codebook accordingly.

Deductive Codebook Application

Once the initial codebook was created, the authors each coded individual copies of half of the study policies. They did so by highlighting PDF documents of each policy and labeling highlighted text with the appropriate code. A spreadsheet was used to indicate whether each policy contained a given code, and the authors' results were compared to determine agreement at the policy level. The authors then met to resolve any discrepancies and arrive at a final consensus on which policies contained which codes. Initial agreement at the policy level was 88% after coding half of the policies. For individual codes where agreement was below 80%, the authors adjusted code definitions for clarity. They then proceeded to code the second half of the study policies with the updated codebook. This time, initial agreement before resolving discrepancies was 92%. They once again adjusted definitions and resolved discrepancies. Using this consensus-based method, the authors ensured that the codebook accurately described the policies to which it had been applied. The final version of the codebook is available in Appendix A.

A potential limitation of this study that bears mentioning here is the lack of a definition for which types of privacy policies to analyze in this study. Any policy labeled as a privacy policy or statement was included to accurately reflect where a given library did address privacy concerns in some form. However, policies pertaining only to the website (15% of the policies in this study), will naturally include fewer checklist items. Because the ALA guidelines were written to address policies that cover all library operations, codes that fall outside of these narrower policies' scopes will appear less frequently across all policies, perhaps resulting in a more negative assessment of the body of policies as a whole. If these are the only publicly available policies on a library's website, this lack of accessible privacy information remains a valid concern. However, future studies should define whether these service-specific or website-specific policies should be included, or whether only policies that address all library services should be evaluated.

Inductive Codebook Development

When deductive coding was complete, the authors examined the uncoded text that remained in the policies. With the intent of identifying themes that fell outside ALA's checklist, they followed Elo and Kyngas' (2008) and Guest et al.'s (2012) methodology to create an inductive codebook. While Guest et al.

(2012) focus on applied thematic analysis, the codebook creation process for document analysis is largely the same.

The authors began by reading through the policies several times and noting any concepts that might be good candidates for codes. They then compared and combined individual lists of concepts (126, initially) and grouped similar ideas together. The initial codebook identified 20 unique codes that could be expanded or split as coding progressed, depending on their prevalence in the data.

Inductive Codebook Application

Because creating the inductive codebook was a more complex process than creating the deductive codebook (which was based on pre-existing guidelines), the authors met more frequently during coding to determine whether code definitions should be updated. Policies were coded in four sessions, with the authors meeting each time to resolve discrepancies and update the codebook. Initial agreement after each round of coding ranged from 89% to 93%. Each time, the authors focused on clarifying definitions for codes where agreement was less than 80%. While no codes were added or removed throughout inductive coding, the definitions of several codes were updated. The final inductive codebook is available in Appendix B.

Results

Are Privacy Policies in United States ARL Libraries Readily Available to Patrons?

Of the 99 libraries in the study population, only 9 (9%) did not have an immediately discoverable privacy policy of any kind on their website. Another 11 libraries linked to their parent institutions' privacy policies rather than library-specific policies. One library posted a privacy policy that was restricted from public view. The remaining 78 libraries provided publicly discoverable library-specific policies which the authors analyzed in this study.

Most policies were available within three clicks of the libraries' main webpages. Thirty-five policies (45%) were in an "About" or similar section linked from the libraries' main webpages. The policy itself was most frequently located either on the "About" page or in a "Policies" sub-section of the "About" page. Another 30 policies (38%) were linked directly from the websites' footers, or a "Policies" link included in the footers. Most other policies were available via a "Policies" link located somewhere on the main webpages outside of the sites' footers. Eight policies (10%) were discovered only by doing a search for "privacy" or "privacy policy" in the websites' search function.

Do Policies Contain Each of the Elements Specified in the ALA's Privacy and Confidentiality Policy Checklist?

The following table shows the results of deductive coding. The full codebook is available in Appendix A.

Table 1
Deductive Coding Results[a]

Code	Policy Count (n = 78)	Percent
Laws	71	91
Limit-2	59	76
List	56	72
Principles	37	47
Purpose	30	38
Contact	27	35
Retention	23	29
Security	21	27
Need_to_Know	18	23
Vendors	17	22
Purge	16	21
Limit-1	16	21
Access	15	19
Unnecessary_Records	7	9
PII_in_Public	5	6
Mission	4	5
Review	2	3
Local_Server	1	1
Breach	0	0
Notify	0	0

[a]Full code definitions available in Appendix A.

Protecting Patron Records

The most prevalent codes dealt with limiting the degree to which patrons' data will be disclosed and distributed. Most policies referenced state or federal laws that dictate cases in which the library must disclose records (Laws, 91%). Three quarters of policies also included a statement that the library would limit the degree to which patron records would be disclosed (Limit-2, 76%). Though not a distinct code, it is notable that 17 policies (22%) specifically referenced the United States Patriot Act, which complicates libraries' protection of circulation records, and consequently the conditions for patrons' academic freedom (Asher et al., 2018; Magi, 2007).

Less than a quarter of policies contained statements that the library would limit collecting or monitoring patron data in the first place (Limit-1, 21%). A similar number of policies address data security (Security, 27%) and internal data governance (Retention, 29%; Purge, 21%). It was less common for policies to state that the library would avoid creating unnecessary records in the first place (Unnecessary_Records, 9%) or placing personally identifiable information (PII) on public view (PII_in_Public, 6%).

Only one policy explicitly stated that patron records will not be stored on a third-party server (Local_Server, 1%), a testament to the prevalence of third-party infrastructure used by academic libraries. Twenty-three percent of policies indicated that only library staff with a need to access data for the purpose of providing a service would be able to access patron data (Need_to_Know).

Transparency About Data Collected

Many policies described individual data points collected from patrons for use by the library (List, 72%). This code was applied in any case where specific data points were explicitly described. However, most policies did not contain a data dictionary or exhaustive list of the data their library obtain from patrons, and over a quarter of policies did not list any specific data points collected from patrons. It is likewise concerning that less than a quarter of policies indicated that the library ensures contracts with third-party vendors will reflect library policies and legal obligations concerning privacy (Vendors, 22%).

The language in the ALA (2017b, para. 7) Checklist suggests that libraries should "[n]otify users whenever the library collects their personally identifiable information...", implying a notification in real time. No policy examined for this study included a statement that patrons would be notified at the time of collection. This is an obvious challenge when library platforms or websites often collect data from or about patrons without their knowledge or consent.

Policy Purpose and Guiding Values

Slightly more than one third of policies stated the policy's purpose or scope (Purpose, 38%). Purpose statements tended to scope the policy by audience (campus, library branch, patron type) or by resource type (all library records, electronic records, website only). It was apparent from policy content that approximately 15% of the study policies applied only to the libraries' websites. Most purpose statements indicated that the policy addressed data collected for use in providing library services. However, some addressed patrons directly, informing them about choices they could make regarding sharing their data with the library.

The Principles code identified mentions of library or archives-specific documents that describe professional values. The document most frequently mentioned was the *ALA Code of Ethics*. While 47% of policies referenced at least one such document, very few went on to describe how professional values relate to the mission of that specific library (Mission, 5%).

Do Policies Contain Additional Elements Not Specified in the ALA Checklist?

The following table shows the results of inductive coding. For the full inductive codebook, see Appendix B.

Table 2
Inductive Coding Results[b]

Code	Policy Count (n = 78)	Percent
Library business	49	63
Values	48	62
Institutional policy	45	58
Assessment	40	51
Liability	35	45
Cookies	31	40
Third-party analytics	19	24
Definitions	15	19
Institutional data	14	18
Advising	11	14
Enforcement	11	14
Customization/Personalization	10	13
Extra-institutional policies	8	10
Workstation	7	9
Security cameras	6	8
Notify patrons of changes	5	6
Children's privacy	4	5
Do not notify patrons of changes	3	4
Video/Image capture	2	3
Social media	1	1

[b]Full code definitions available in Appendix B.

What is the Nature of These Additional Elements?

Collecting Patron Data to Provide Services

Only four of the codes identified in the inductive portion of this project appeared in more than half of the privacy policies. The most common code was Library Business (63%), which described how collecting patron data is necessary to provide certain library services. Some statements were general: "We use this information to maintain your library account and to provide services to you" (Duke University Libraries, 2013, para. 5), while other statements described the data points necessary to supply a specific service. The Library Business code often coincided with the deductive List code and provided additional context about how various data were used to provide a service.

Guiding Values

A high percentage of policies asserted the libraries' belief in or support for concepts such as confidentiality, intellectual freedom, academic freedom, or privacy itself (Values, 62%). These statements often coincided with the policy's scope and the ethical and legal documents guiding its creation. Most policies, however, did not provide definitions for these terms (Definitions, 19%).

Compliance With Institutional Policy

Just over half of the policies in this study referenced either another library policy, or a policy from the library's parent institution (Institutional Policy, 58%). Compliance with the parent institution's privacy policy was the most common occurrence. Other mentions included institutional policies that addressed acceptable use of electronic resources, information technology issues, or handling student records. Only 8 policies referenced a policy outside their own institution (Extra-institutional policy, 10%). Several of these addressed the handling of medical or academic records, while others referenced a privacy policy from another institution that the library had used as a model.

Limitations on Protecting Patron Data

Just under half of policies referenced contexts to which the policy did not extend (Liability, 45%). This most often coincided with information on third-party platforms that the library provided access to but did not maintain. Seventeen libraries stated that they negotiated with vendors to ensure privacy protection on third-party platforms. Somewhat surprisingly, 21 policies (27%) stated either that the library did *not* negotiate with vendors for these protections or stated that these external platforms were beyond the library's control. Twelve of the policies that included the Liability code (15% of all study policies) did not mention vendors specifically but stated that the library policy did not apply to linked external sites or platforms. They urged patrons to read these platforms' privacy policies for themselves.

Assessment and Invisible Data Collection

Assessment, Cookies, and Third-party analytics are discussed in 51%, 40%, and 24% of policies, respectively. All three codes were mentioned in the context of what was broadly termed "continuous improvement" and explained how data collected in the process of using the library may be used to improve library services. Contexts included troubleshooting technical issues, making purchasing decisions, and ensuring compliance with policies on the acceptable use of library resources. Statements addressing Third-party analytics most often referenced Google Analytics and indicated that this data were used to understand trends in library usage or to make improvements to the website. Ten policies stated that patron-provided data could be used to create a personalized web browsing experience (Customization/Personalization, 13%).

In a small number of cases, the Assessment code referenced data collected explicitly through surveys or focus groups. Four policies stated that they obtained permission from their Institutional Review Board (IRB) or similar oversight office to use patron data in assessment projects. Though references to assessment frequently indicated that data would be stored securely or de-identified, only two policies indicated that patron consent would be obtained during the data collection process. Additionally, despite 76% of policies indicating that the library would limit sharing patrons' data beyond the library, 14 policies stated that the library could provide patron data to or receive patron data from the library's parent institution (Institutional Data, 18%).

Educating Patrons About Privacy

Eleven policies recommended some form of action patrons could take to protect their privacy such as logging off public workstation computers, guarding personal information on shared software platforms, and creating sound passwords (Advising, 14%). In some cases, the policy acknowledged choices patrons could make to limit sharing their data: "If you do not want your email address released in response to a

public records request, do not send electronic mail to the University" (University of Florida George A. Smathers Libraries, 2016, para. 11); "If you are concerned about someone else seeing a list of what you are reading or searching for, the safest step is to not choose this option" (University of Arizona Libraries, 2016, para. 8).

Library Accountability and Policy Enforcement

Of the 78 policies analyzed, only 11 included a statement about how the policy would be enforced (Enforcement, 14%). While five policies indicated that patrons would be made aware of changes to the policy (Notify patrons of changes, 6%), another three explicitly stated patrons would *not* be notified of changes (Do not notify patrons of changes, 4%).

Discussion and Limitations

In response to this study's aims, the authors found that a majority of ARL institutions (79%) do provide some type of privacy policy to their patrons. However, these policies demonstrate low compliance with ALA guidelines. While all but two deductive codes were present in at least one policy, none of the 78 policies in this study contained all 20 checklist items, and only 6% contained more than half. No policy contained more than 75% of the content suggested by ALA. It should not be assumed that because a policy contains more checklist items, that institution better implements or enforces privacy. However, if one purpose of a policy is to inform patrons about how their data are collected and used, it is concerning that more policies do not follow the guidelines in a comprehensive way.

This lack of guideline coverage may be due in part to policy length. The median number of sentences in the policies analyzed was 16, with 36% containing 10 sentences or fewer. Short policies may contain valuable details pertinent to patrons, but a policy 10 sentences in length is unlikely to address data collection and handling in depth. As mentioned in the methodology section, this may be due in part to the fact that some of these policies do not cover all library operations. However, there is little difference between providing patrons with a privacy policy that only covers certain services or the website and providing them with a library-wide policy that does not adequately cover uses of their information in depth.

This study's final aim was to uncover whether policies contain themes beyond what ALA guidelines recommend. The significant amount of uncoded text that remained when deductive coding was complete indicated that there was more to uncover in these policies, and inductive coding revealed that there were indeed common threads across policies which were not explicitly addressed in the guidelines. While some themes simply provided additional detail on items included in the Policy Checklist, several codes touched on entirely new topics. For example, the Assessment and Institutional Data codes addressed library-initiated data collection that occurred in many cases without the patrons' knowledge; and Security Camera and Video/Image capture addressed explicit collection of patron data in physical library spaces.

Since this study's methodology was created in 2019, ALA has replaced the 2017 Policy Checklist (and several of its other privacy-related documents) with an updated *Privacy Policy Field Guide* (ALA, 2021c). The 2021 Field Guide does address some of the areas mentioned above, including cookies, data encryption, network security, and facial recognition software, as well as how to spot red flags in vendor policies. It urges policy makers to "[r]emind users that their information is confidential, but also tell them who has access to it at your library" (ALA, 2021c, p. 18).

While the *Privacy Policy Field Guide* includes promising updates, based on the policies analyzed in this study, there is still room to improve the guidelines by addressing assessment projects and institutional data sharing in more detail. This is particularly true of library-initiated analyses that use patron data for purposes other than the one data was originally collected for. This lack of explicit transparency contributes to what Affonso and Sant'Ana (2018) have referred to as information asymmetry, a power imbalance in which the patron has no ability to consent to or control uses of data about themselves, since they have no knowledge that data collection or data sharing is occurring.

A limitation that should be addressed in future research on this topic relates to content analysis as a methodology. As with many qualitative approaches, document analysis is typically used alongside additional methods to triangulate the validity of findings (Bowen, 2009). Ideally, another research method such as interviews could provide additional context for the creation and implementation of privacy policies that a static document does not allow. Future research could include interviews or focus groups that would address the decisions that go into policy creation, as well as discussing whether alternate means may be appropriate for informing patrons about the uses of their data.

Conclusion

Both academic library privacy policies and professional guidelines appear to be evolving to address emerging concerns such as surveillance technology, data collection via vendor platforms, and data sharing with institutional partners. While the breadth of detail included across all policies in this study was encouraging, very few individual policies addressed most of the content suggested by professional guidelines in depth. Even with an abundance of privacy-related guidelines available from sources like ALA, many policies still leave patrons without the detail required to understand how the library collects and uses their data. Additionally, the inclusion of assessment projects and sharing data with institutional partners in the policies analyzed indicates that current guidelines may benefit from expanding to address library-initiated projects in more detail.

Author Contributions

Greta Valentine: Conceptualization, Methodology, Formal analysis, Writing – original draft
Kate Barron: Formal analysis, Writing – review & editing

Acknowledgements

The authors wish to thank the Institute for Research Design in Librarianship (IRDL) for valuable input that shaped the research design of this study.

References

Affonso, E. P., & Sant'Ana, R. C. G. (2018). Privacy awareness issues in user data collection by digital libraries. *IFLA Journal*, *44*(3), 170–182. https://doi.org/10.1177/0340035218777275

American Library Association. (2014). *Privacy Tool Kit*. ALA Intellectual Freedom Committee. https://alair.ala.org/handle/11213/16714

American Library Association. (2017a). *Privacy.* https://web.archive.org/web/20171101171848/https://www.ala.org/advocacy/privacy

American Library Association. (2017b). *Privacy and Confidentiality Policy*. https://web.archive.org/web/20170322042558/http://www.ala.org/tools/challengesupport/privacypolicy

American Library Association. (2017c, April 28). *Developing or Revising a Library Privacy Policy*. https://web.archive.org/web/20210428022310/http://www.ala.org/advocacy/privacy/toolkit/policy#sectionstoinclude

American Library Association. (2021a). *Privacy*. https://www.ala.org/advocacy/privacy

American Library Association. (n.d.). *Privacy Field Guides for Libraries*. *Privacy Policies.* https://libraryprivacyguides.org/privacy-policies/

Asher, A. D. (2017). Risk, benefits, and user privacy: Evaluating the ethics of library data. In B. Newman & B. Tijerina (Eds.), *Protecting patron privacy: A LITA guide.* (pp. 43–56). Rowman & Littlefield. https://hdl.handle.net/2022/22035

Asher, A., Briney, K., Gardner, G., Hinchliffe, L., Levernier, J., Nowviskie, B., Salo, D., & Shorish, Y. (2018). *Ethics in Research Use of Library Patron Data: Glossary and Explainer*. https://osf.io/xfkz6/

Atkinson, P. A., & Coffey, A. (1997). Analysing documentary realities. In D. Silverman (Ed.), *Qualitative research: Theory, method and practice*. (pp. 45-62). Sage Publications.

Beile, P., Choudhury, K., & Wang, M. C. (2017). Hidden treasure on the road to Xanadu: What connecting library service usage data to unique student IDs can reveal. *Journal of Library Administration*, *57*(2), 151–173. https://doi.org/10.1080/01930826.2016.1235899

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, *9*(2), 27–40. http://dx.doi.org/10.3316/QRJ0902027

Briney, K., Yoose, B., Ockerbloom, J. M., Swauger, S., Harper, C., Levernier, J., & Shorish, Y. (2018). *A Practical Guide to Performing a Library User Data Risk Assessment in Library-Built Systems*. https://doi.org/10.17605/OSF.IO/V2C3M

Coombs, K. A. (2005). Protecting user privacy in the age of digital libraries. *Computers in Libraries*, *25*(6), 16–20.

Crawford, K., & Schultz, J. (2014). Big data and due process: Toward a framework to redress predictive privacy harms. *Boston College Law Review*, *55*(1), 93–128. https://lawdigitalcommons.bc.edu/cgi/viewcontent.cgi?article=3351&context=bclr

Drachsler, H., & Greller, W. (2016, April 25). Privacy and analytics – it's a DELICATE issue. A checklist for trusted learning analytics. *LAK '16: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, 89-98. https://doi.org/10.1145/2883851.2883893

Duke University Libraries. (2013, September 6). *Duke University Libraries Privacy Statement.* https://library.duke.edu/about/privacy

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, *62*(1), 107–115. https://doi.org/10.1111/j.1365-2648.2007.04569.x

Erlingsson, C., & Brysiewicz, P. (2017). A hands-on guide to doing content analysis. *African Journal of Emergency Medicine*, *7*(3), 93-99. https://doi.org/10.1016/j.afjem.2017.08.001

Farkas, M. (2018). We can, but should we? When trends challenge our professional values. *American Libraries Magazine*, *49*(3/4), 46-47. https://americanlibrariesmagazine.org/2018/03/01/learning-analytics-we-can-but-should-we/

Guest, G., MacQueen, K. M., & Namey, E. E. (2012). *Applied thematic analysis*. Sage Publications.

Jones, J. L. (2010). Using library swipe-card data to inform decision making. *University Library Faculty Presentations, 21*, 1-9. https://scholarworks.gsu.edu/univ_lib_facpres/21

Jones, K. M. L., & Salo, D. (2018). Learning analytics and the academic library: Professional ethics commitments at a crossroads. *College & Research Libraries, 79*(3), 304-323. https://doi.org/10.5860/crl.79.3.304

Kritikos, K. C., & Zimmer, M. (2017). Privacy policies and practices with cloud-based services in public libraries: An exploratory case of BiblioCommons. *Journal of Intellectual Freedom & Privacy*, *2*(1), 23-37. http://dx.doi.org/10.5860/jifp.v2i1.6252

Magi, T. J. (2007). The gap between theory and practice: A study of the prevalence and strength of patron confidentiality policies in public and academic libraries. *Library & Information Science Research*, *29*(4), 455-470. https://doi.org/10.1016/j.lisr.2007.07.001

Magi, T. J. (2010). A content analysis of library vendor privacy policies: Do they meet our standards? *College & Research Libraries*, *71*(3), 254-272. https://doi.org/10.5860/0710254

Nichols Hess, A., LaPorte-Fiori, R., & Engwall, K. (2015). Preserving patron privacy in the 21st century academic library. *The Journal of Academic Librarianship*, *41*(1), 105-114. https://doi.org/10.1016/j.acalib.2014.10.010

Oakleaf, M. (2010). *The value of academic libraries: A comprehensive research review and report.* Association of College & Research Libraries. http://www.ala.org/acrl/sites/ala.org.acrl/files/content/issues/value/val_report.pdf

O'Brien, P., Young, S. W. H., Arlitsch, K., & Benedict, K. (2018). Protecting privacy on the web: A study of HTTPS and Google Analytics implementation in academic library websites. *Online Information Review*, *42*(6), 734-751. http://dx.doi.org/10.1108/OIR-02-2018-0056

Pekala, S. (2017). Privacy and user experience in 21st century library discovery. *Information Technology and Libraries*, *36*(2), 48-58. https://doi.org/10.6017/ital.v36i2.9817

Perry, M. R., Briney, K. A., Goben, A., Asher, A., Jones, K. M. L., Robertshaw, M. B., & Salo, D. (2018). *SPEC Kit 360: Learning analytics*. https://publications.arl.org/Learning-Analytics-SPEC-Kit-360/

Prindle, S., & Loos, A. (2017). Information ethics and academic libraries: Data privacy in the era of big data. *Journal of Information Ethics*, *26*(2), 22-33. https://www.proquest.com/scholarly-journals/information-ethics-academic-libraries-data/docview/2027533656/se-2?accountid=10920

Rubel, A., & Jones, K. M. L. (2016). Student privacy in learning analytics: An information ethics perspective. *The Information Society*, *32*(2), 143-159. https://doi.org/10.1080/01972243.2016.1130502

Sturges, P., Davies, E., Dearnley, J., Iliffe, U., Oppenheim, C., & Hardy, R. (2003). User privacy in the digital library environment: An investigation of policies and preparedness. *Library Management*, *24*(1/2), 44-50. http://dx.doi.org/10.1108/01435120310454502

Sutlieff, L., & Chelin, J. (2010). 'An absolute prerequisite': The importance of user privacy and trust in maintaining academic freedom at the library. *Journal of Librarianship and Information Science*, *42*(3), 163-177. https://doi.org/10.1177/0961000610368916

Tenopir, C. (2010). Measuring the value of the academic library: Return on investment and other value measures. *The Serials Librarian*, *58*(1-4), 39-48. https://doi.org/10.1080/03615261003623005

University of Arizona Libraries. (2020, August 24). *Privacy statement*. University Libraries. https://web.archive.org/web/20201021081113/https://new.library.arizona.edu/policies/privacy

University of Florida George A. Smathers Libraries. (2020). *Privacy Policy*. https://web.archive.org/web/20200909053228/https://uflib.ufl.edu/about/user-policies/privacy-policy/

Vaughan, J. (2020). Library privacy policies. *Library Technology Reports*, *56*(6), 5–53. https://doi.org/10.5860/ltr.56n6

Voeller, S. (2007). Privacy policy assessment for the Livingston Lord Library at Minnesota State University Moorhead. *Library Philosophy and Practice, 151*. https://digitalcommons.unl.edu/libphilprac/151/

Yoose, B. (2018, February 15). *Data Analytics and Patron Privacy in Libraries: A Balancing Act*. https://osf.io/xb4mf/

Appendix A
Deductive Codebook

Label	Definition	Qualifications/Exclusions
Access	States that patrons have the right to access, see, or update personally identifiable information (PII) that the library collects about them	Deals with patrons accessing data about themselves. Does not apply to the library releasing information to third parties.
Breach	States that the library will notify patrons in the event of a data breach	Deals with the inadvertent release of patron PII or records.
Contact	Provides a means for the patron to contact the library regarding the policy	Must include actual contact information (phone number, email address, chat link) rather than just a name or title for patrons to contact. We did not count contact information in the footers of websites.
Laws	Policy includes references to any federal, state, or local laws that impact the policy	Can include generic reference to being in compliance with "state and federal laws", or similar
Limit-1	States that the library will limit the degree to which patrons' PII will be monitored/collected	Deals with how data are obtained FROM the patron
Limit-2	States that the library will limit the degree to which patrons' PII will be disclosed/distributed	Deals with how data are disclosed by the library. Can include instances of aggregating data (see NC State)
List	Lists the personally identifiable information (PII) the library will be collecting from patrons when they use library services	Can apply to individual instances of PII collection – does not have to reference a comprehensive list. Must refer to PII. Can list specific PII or library records that include PII. Applies to sections that list electronic information such as IP address, browsing history etc., even if the policy does not refer to this information as PII.
Local_Server	States that patron records will remain on a local server rather than being exported to the cloud or a third-party server	
Mission	Explains how protecting user privacy and confidentiality relates to the mission of the library	Relates only to the specific mission of that particular library, not ALA/professional ethics
Need_to_Know	States that only authorized library staff will access patron records	Can refer to library employees or university employees as long as the case for accessing the records is clear. Use only for blanket statements, not particular instances.
Notify	States that the library will notify patrons when data are being collected from them	Refers to or states that the library will notify patrons in real time when information is being collected from them. Should not be used for policy content that only addresses opting in or out of using cookies.

PII_in_Public	States that the library will avoid placing patron records in public view	
Principles	Refers to principles on which the library's commitment to protecting privacy is based	Relates to the principles of librarianship (i.e., references to the *ALA Code of Ethics*, intellectual freedom, etc.) Can include generic reference to library principles or professional documents.
Purge	States that the library will regularly purge identifiable patron records	Can be a general statement about purging records after a set period of time, or an example of removing specific records. Must indicate a REMOVAL or deletion of records, as opposed to Retention, which indicates cases where records are kept.
Purpose	States the purpose of the policy	States the purpose of the policy. May include scope: "This policy applies to X Campus, X Patrons, X Resources etc."
Retention	States that the library will not retain patron records that are not needed for efficient operation of the library	Must indicate keeping records and the purpose for keeping those records, as opposed to Purge, which indicates a removal or deletion of records
Review	States how often the policy will be reviewed	Must state actual cycle or dates when policy is updated. Does not apply to date policy was last updated, or statements that the policy is updated periodically.
Security	States that patron PII will be stored securely	Must refer to patron records
Unnecessary_Records	States that the library will avoid creating unnecessary records about the patron	
Vendors	States that the library will ensure contracts and licenses with vendors will reflect library policies and legal obligations concerning privacy	Must refer to licenses negotiated with library vendors

Appendix B
Inductive Codebook

Label	Definition
Advising	Policy advises users regarding privacy best practices
Assessment	Policy refers to library-initiated analysis for continuous improvement
Children's privacy	Policy describes special privacy protections for children
Cookies	Policy describes or defines cookies or their implications for privacy and personalization
Customization/Personalization	Policy describes options for submitting personal information in order to customize the use of a digital library service
Definitions	Policy includes definitions of terms relating to patron privacy, confidentiality and patron rights. Also includes statements informing the patron how these concepts relate to one another.
Do not notify patrons of changes	Policy states that privacy policy may change without notice to patrons
Enforcement	Policy describes how the libraries will enforce the policy, including how they will deal with violations or prevention measures such as privacy audits
Extra-institutional policies	Policy refers to non-university policies or other library policies. Codes of Ethics should be coded Principles.
Institutional data	Policy describes situations in which the library receives or shares patron data with other campus units
Institutional policies	Policy refers to library's other policies or parent institution's policies
Liability	Policy describes contexts to which the policy does not extend. Includes statements that library operations or systems may not be 100% secure.
Library business	Policy describes data collected or used to provide services or facilitate smooth operations
Notify patrons of changes	Policy states library will notify patrons of changes to privacy policy. Includes cases where the library notifies individual patrons, or where they post changes in a public place before the changes take effect.
Security cameras	Policy describes library use of security cameras
Social media	Policy describes use of patron information or images on their social media accounts
Third-party analytics	Policy mentions use of Google analytics or other third-party web analytics on the library website. Policy must identify the tool as a web analytics tool.
Values	Policy asserts belief in or support for professional library values such as privacy, confidentiality, intellectual freedom, or academic freedom. Policy states clear position or attitude towards upholding these values.
Video/Image capture	Policy describes restrictions on capturing image or video of patrons in the library
Workstation	Policy describes use of on-site computers or workstations