A Comprehensive Primer to Library Learning Analytics Practices, Initiatives, and Privacy Issues

Kyle M.L. Jones, Kristin A. Briney, Abigail Goben, Dorothea Salo, Andrew Asher, and Michael R. Perry

Universities are pursuing learning analytics practices to improve returns from their investments, develop behavioral and academic interventions to improve student success, and address political and financial pressures. Academic libraries are additionally undertaking learning analytics to demonstrate value to stakeholders, assess learning gains from instruction, and analyze student-library usage, et cetera. The adoption of these techniques leads to many professional ethics issues and practical concerns related to privacy. In this narrative literature review, we provide a foundational background in the field of learning analytics, library adoption of these practices, and identify ethical and practical privacy issues.

Introduction
In the wake of the ubiquitous computing infrastructure that higher education institutions have created, a “rising tide” of student data is overflowing as a result of interactions with digital systems. Institutions are motivated to mine this “gold,” to extract this “oil” as a social, political, and financial resource, and as a means to effect positive educational outcomes. This sociopolitical and sociotechnical movement to collect, describe, and act on student data has created momentum in the field of learning analytics (LA).

LA are defined as “the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs.” As a practice, LA focuses on student success and student
Comprehensive Primer to Library Learning Analytics Practices, Initiatives, and Privacy Issues

Background on Learning Analytics

The embrace of learning analytics (LA) echoes the 1970s rise of decision support practices and, more recently, business intelligence systems. However, LA are more accurately described as branching off of the academic analytics movement of the mid-2000s. Campbell, DeBlois, and Oblinger write that “thanks to enterprise-wide systems that generate massive amounts of data, data warehouses that aggregate disparate types of data, and processing power that sifts, sorts, and surfaces patterns,” analytics have become a viable set of tools.

LA researchers focus on digital learning environments because of the bountiful data collected by educational and other technologies in support of running complex academic institutions. One example is Unizin and its Common Data Model (UCDM). Unizin is an institutional consortium of more than 20 public universities and university systems serving a collective student body totaling more than 900,000. The UCDM maps a complex array of student-related data, which Unizin will aggregate in a centralized data warehouse for consortia members. The model contains demographic, economic, and academic information, among other things. It captures students’ academic work (such as assignments and discussion posts) and all their interactions with a learning management system (LMS)—logins, individual clicks, time-on-page—all the “digital footprints” expressed in an LMS’s logs and metadata.

Early LA efforts, such as those at Purdue University, used LMS data in combination with academic performance data to predict whether a student was at risk for academic failure in a specific course. Purdue’s predictive work continues to serve as a model for many LA
developments, including feature changes to common LMSs like Blackboard and Canvas. Importantly, critics later discovered significant issues with Purdue’s predictive model. Jones and McCoy explain that the researchers responsible for the model came to inaccurate conclusions by staking their statistically significant—but wrong—claim on a spurious correlation, seeing patterns in the process that did not exist.18

Other LA systems increasingly “nudge” students toward certain behaviors, such as communicating with instructors or seeking campus resources, based on these predictions.19 Some analytics systems adapt course materials and assessments based on students’ learning styles, past performance, and system behaviors.20 LA designers have also developed predictive analytics, nudging strategies, and adaptive learning algorithms for advising systems, which provide personalized course and program recommendations and suggest advisor interventions when the system predicts students will do poorly in their academics or they are not engaged with the university community.21

Beyond education technologies, institutions can and often do capture data related to interactions with virtual assistants, smartphones, wearable devices, tablets, personal computers, sensors, and ID card readers (example: card swipes). For example, Northeastern University, St. Louis University, Arizona State University, and the University of Texas at Dallas started initiatives using Amazon Echo Dots, whose Alexa personal assistant can retrieve campus information and, in some cases, get personalized information from SISs.22 Both Amazon and these institutions gained access to student data.23 These data were so valuable to UT-Dallas’s IT department that it proposed refusing to allow students to turn the Dots off.24 Purdue and Oral Roberts University pursued “quantified self” technologies to inform students of their learning behaviors and track their nonacademic activities; the former university uses mobile, desktop, and smartwatch applications, while the latter required its students to wear Fitbits.25 At the University of Arizona, a researcher analyzed students swiping their identification cards at more than 700 campus locations to study routines and social networks.26

In addition to campus-based data, some LA proponents have advocated for integrating social media data.27 A significant majority of 18- to 29-year-olds use social media; the enormous amount of these data may explain the desire to mine them for analytics.28 Institutions have already monitored student social media behaviors on a small scale, though not always for educational purposes. Georgia Tech University profiled a student activist,29 and the University of Virginia hired a social media monitoring company to surveil students’ social media for threats to campus safety.30 Nevertheless, as Chen, Vorvoreany, and Madhavan demonstrate, it is becoming increasingly easy for institutions to collect their students’ Twitter content and metadata wholesale via APIs using Boolean search terms and geolocation filters.31

Beyond internal usage of student data, institutions have begun to aggregate their student data collaboratively. Reyes notes that new technology enables “institutions to access, collaborate and contribute their data in an effort to build single platforms with the capacity for sharing multiple data sets.”32 Maturing interoperability standards, like Caliper, are enhancing aggregation initiatives.33 One example of mass aggregation is Carnegie Mellon University’s DataShop, which serves as “a central repository for data on the interactions between students and educational software.”34 primarily for research purposes.

While proponents often argue that LA are focused solely on students and their learning,35 the accepted definition of LA and ongoing rhetoric around them reveals a variety of motivations and goals. The ability to surface, analyze, and act on the “data explosion”36 has led some
to think that LA may serve as panaceas that can resolve seemingly intractable problems like recruitment, retention, and provide data useful for responding to external demands.\textsuperscript{37} Internal and external stakeholders believe LA can help demonstrate institutional “performance and cost-effectiveness” in ways that highlight solvency and increase public trust in and support of higher education as a costly investment.\textsuperscript{38} For instance, Mount St. Mary’s University planned to use predictive analytics to proactively dismiss students unlikely to be retained, arguing that such actions would improve its national ranking.\textsuperscript{39} Others have pushed back against the pressure to quantify student life, claiming that LA are leading institutions to count for the sake of counting without due consideration for ways such analytics and the infrastructures that support analytic practices can have “detrimental consequences.”\textsuperscript{40} Moreover, present evidence for the effectiveness of LA is mixed at best, raising a concern that the benefits are not worth the financial costs and/or possible harms.\textsuperscript{41}

### Academic Library Participation in Learning Analytics

Academic libraries, like the institutions they are part of, have identified many reasons to engage in LA, including a need to maintain alignment with greater institutional efforts and interests.\textsuperscript{42} Library administrators are facing pressure from their peers at the campus level to demonstrate their contributions to university-level goals.\textsuperscript{43} A survey of chief academic officers (CAOs) found that they desired quantitative evidence as a key requirement to support funding requests, arguing that correlational data “to retention, success, and even evidence of learning information literacy skills” is of the utmost importance to justify resource allocations. Twenty years of research focused on quantifying library value and impact (however defined) have resulted from these pressures to explain ever-increasing library budgets and an institution’s demand for a return on its investments in libraries.\textsuperscript{44} LA further advances these goals.

Information literacy instruction has seen an increased move toward outcomes-based assessment in the past two decades, which has normalized the capture of detailed and individual quantitative data for many academic librarians. The Association of College and Research Libraries (ACRL) \textit{Information Literacy Competency Standards}, adopted in 2000, encouraged a granular approach to data gathering about students and the analysis practices that support LA.\textsuperscript{45} These standards specifically called for 22 performance indicators, each with multiple outcomes. One of the goals indicated by the implementation of the standards is to identify how information literacy would “enhance the institution’s effectiveness.”\textsuperscript{46} The ACRL Framework for Information Literacy in Higher Education, which was adopted in 2016 and supplanted the \textit{Standards}, does not explicitly focus on assessment in the same quantitative way but does emphasize measuring impact.\textsuperscript{47}

ACRL’s long-standing Value of Academic Libraries (VAL) initiative has also expanded library assessment and evaluation practices.\textsuperscript{48} The initial report, \textit{Value of Academic Libraries: A Comprehensive Research Review and Report}, asserted that academic libraries need to “document and articulate”\textsuperscript{49} how a library contributes to its institution’s missions and goals, primarily due to ever-increasing accountability measures. The VAL report recognizes in the foreword that measuring library efficacy grates on some who feel such efforts signal a capitulation to neoliberal interests; additionally, the report acknowledged that attempts to quantify library efforts seem like further homogenization of academia. About these critiques, ACRL leadership commented that “this critical perspective [...] seems impractical, given the realities we face today in our institutions.”\textsuperscript{50} Funding agencies and many academic libraries agreed. The VAL
initiative has received several national grants in support of its work; notable related initiatives have emerged, including the “Assessment in Action: Academic Libraries and Student Success” program. Furthermore, VAL leaders have developed numerous reports demonstrating the success of and continued need for the research agenda. To date, hundreds of scholarly artifacts have been published and disseminated as a result of the VAL initiative.

More recently, library LA literature is explicitly aligning with the methods and goals that higher education LA advocates espouse. Megan Oakleaf, in particular, has been an outspoken advocate of LA, highlighting how “many librarians have embraced the use of assessment and research to explore the links between student library interactions and student learning and success measures.” This argument moves beyond the documentation and assessment of library efforts to combining library data with identifiable student data from other sources to seek potential correlative trends. OCLC’s research with higher education administrators signaled that there is an ever-increasing interest in gathering identifiable student data. Furthermore, while those administrators did not explicitly identify data from the library as of interest, library LA advocates have mentioned it as a particular area of importance.

Library participation in institutional LA initiatives to date has been limited in part because library data are siloed from the rest of the campus. LA advocates argue that library data should be extracted from various library systems and integrated with centralized systems for cross-institution access and analysis, which some institutions have pursued. For example, the universities of Wollongong and Minnesota include sharing individual students’ library usage information with instructors/advisors to encourage them to nudge those students toward greater library use.

Library LA studies are increasingly possible due to a growing amount of accessible data, in part because data exhaust and metadata are inherent to library systems’ main function. Library LA studies demonstrate that analyzing data for secondary purposes (that is to say, not for their designed purposes) provides several analytical opportunities, as does the combination of library data sources with campus-captured student information. Library LA studies harvest data from existing library systems such as integrated library systems (ILS) and EZ-proxy logs, using spreadsheet or statistical software to run analyses (examples: Excel, SPSS, Tableau). Checkout counts are baseline usage statistics in LA studies, including interlibrary loans, equipment checkouts, and books. Additional facilities data are available, such as card-swipe entries into the library, library computer logins, and study room reservations. Proxy servers provide data about e-resource use.

Library LA additionally provide new possibilities to assess library interactions with students. Studies quantified transactions between students and library staff at the following service points: face-to-face reference services, asynchronous virtual reference services such as through e-mail, synchronous virtual reference services such as instant messaging, and instruction sessions. Others are mining text from chat reference transcripts. The scope of data collection varies from project to project; for example, course-based instruction allows for the aggregate analysis of students in an entire class section rather than individual-level analysis. Some libraries purchase Springshare’s LibAnalytics, OrangeBoy, Gale Analytics, WhoFi, and OCLC’s Wise to support analytic efforts. All these systems represent newer commercial library analytics tools, though vendors market these primarily to public libraries.

The majority of library LA studies also look to institutional sources for student data. Grade point averages (GPAs), obtained from an institutional research office, are one of the
most common metrics of student success. Kogut found that more than 40 percent of studies evaluated GPA when examining academic library contributions to undergraduate academic success,78 which is consistent with trends in broader LA practices.79 Occasionally, studies evaluated individual assignments or examined course grades instead of overall GPA.80 Other library LA studies include data gathered from student participation in first-year experience programs81 and uses of tutoring services.82 To this is added data gathered outside classroom environments, such as four-year graduation rates83 and year-to-year retention.84

Additional information often obtained from institutional sources includes demographic information such as age, gender, and ethnicity to create subgroups for comparison.85 A small number of library studies examine sensitive demographic data like socioeconomic status86 and disability status.87 Other studies captured a student’s standardized test scores,88 major,89 and various statuses like on/off campus residence status,90 full/part-time enrollment status,91 and first-generation status.92 The scope of demographic data also fluctuates. Some articles amass a large amount of demographic information,93 while others collect relatively few data or only look at demographics in aggregate.94

The research methods by which libraries conduct LA studies are varied. However, the majority of studies correlate the use of library services with measures of student success or make means-based comparisons. These studies are most frequently quantitative, though some use a mixed-methods approach.95 One form of analysis is propensity score matching, where target students are paired and compared to students with similar characteristics from a control group.96 Most studies run analyses at the individual student level, though a small portion of studies perform analytics on aggregated data such as from IPEDS or library association statistics.97

Library Learning Analytics, Ethics, and Privacy Issues

As academic libraries engage in learning analytics, questions arise surrounding the ethics of capturing, retaining, and using these data as well as concerns about user privacy. Student privacy issues ebb and flow according to new sociotechnical challenges, which are exemplified by privacy issues brought to the fore by LA’s rise. Higher education is struggling to achieve a privacy balance that can both protect students and, simultaneously, lead to new data-based insights. The same holds true for academic libraries. Library ethics may form a foundation from which to examine these new challenges.

Professional librarians address privacy and related ethical issues in part by reviewing guidelines in ethics codes, such as those by the American Library Association (ALA),98 the International Federation of Library Associations (IFLA),99 and the National Information Standards Organization (NISO).100 Each organization, to various degrees, emphasizes the importance and roles of privacy and confidentiality in intellectual processes associated with library use.

Where the ALA Code of Ethics is concerned, Principles 1 through 4 and 6 are particularly relevant to library LA projects, as they focus on equitable access, intellectual freedom, privacy, confidentiality, rights of information users and rights holders, and not advancing “private interests at the expense of library users, colleagues, or our employing institutions.”101 Similarly, the Library Bill of Rights’ seventh point explicitly calls out a user’s “right to privacy and confidentiality in their library use.”102 Similarly, IFLA’s Statement on Privacy in the Library Environment states that “library and information services should reject electronic surveillance and any type of illegitimate monitoring or collection of users’ personal data or information
behavior that would compromise their privacy and affect their rights to seek, receive, and impart information.”

Notably, NISO’s Consensus Principles on Users’ Digital Privacy in Library, Publisher, and Software-Provider Systems is intended as “a starting point” for pragmatic discussions on user privacy in library systems and electronic resources. In the past, ALA has historically put the onus on libraries to protect user privacy; but, more recently, it has turned to NISO’s principles, which takes a more balanced approach. The principles express the joint privacy obligations that libraries and third parties share. Its principles state that “[l]ibraries, publishers, and software providers have a shared obligation to foster a digital environment that respects library users’ privacy as they search, discover, and use those resources and services.” What is lacking from ALA, IFLA, and NISO’s documents is an explicit recognition that library data can be and are often used for research and organizational improvement purposes. Efforts to fill this gap have arisen from ALA itself, the Digital Library Federation, LITA, and the Library Freedom Project. Furthermore, the Institute of Museum and Library Services has distributed a significant amount of grant funding to support research into privacy and ethics issues associated with LA; see appendix B for a list of grants.

Outside of libraries, researchers and institutions have developed guiding principles for practice, policy, and technological design. The Asilomar conferences brought together researchers, technologists, and ethicists to develop “a framework to inform decisions about appropriate use of data and technology in learning research for higher education.” Other researchers, like Slade and Prinsloo, Sclater, and Pardo and Siemens, put forth their recommendations for principled practice.

Some universities have developed institutional guidelines. The Open University was among the first to develop principles explicitly addressing privacy and LA. More recently, respective faculty at the University of California and the University of Hawaii passed resolutions around student data. Both of these institutions’ stated principles address data access, control, and ownership; freedom of expression; ethical uses; and transparency, among other things. Policies from these and other universities are available in appendix C.

Despite these efforts, there is very little evidence that libraries have begun to specifically address LA in local policies. The only data suggesting a policy reaction to LA by libraries comes from Perry et al., whose survey of institutional members of the Association of Research Libraries found that “[m]ost respondents indicated that LA has not caused changes in their privacy policies. Only one respondent indicated that they created a ‘Responsible Use of Library Data’ statement in alignment with their learning analytics project.”

Since there is a notable gap in library policy and ethics vis-à-vis LA, several library-specific projects have sought to work through LA privacy ethics using a critical lens. Asher addresses the potential and demonstrated dangers of increasing data access, storage, and analysis of library users in higher-education environments. Building on Asher’s work, Jones and Salo analyzed how these issues clash with ALA’s Code of Ethics, arguing ways in which LA projects may inhibit intellectual freedom and privacy. In a recent piece, Jones, Rubel, and LeClere argue institutions are beholden to their students as “information fiduciaries,” that they have a moral obligation to use LA to benefit students and respect the trust students place in their university.

Given these known ethical problems and quandaries related to LA, Asher et al. argue that librarians must update their sense of threats to user privacy by taking into account algorithm-
mic bias, reidentification, widespread data trading and hoarding, and surveillance creep. Briney homes in on research ethics and LA projects, noting that, even though institutional review boards (IRB) are responsible for identifying ethical issues, many miss violations of library privacy ethics. This situation is due to differences in disciplinary ethics and the fact that studies may be marked exempt from review.

For instance, for those projects that do not include obvious demographic data or are internal assessment and evaluation practices, they fail to trigger reviews. It is unclear from the literature if any academic libraries have established clear internal review processes for addressing data privacy concerns.

**Practical Library Ethical Privacy Issues**

Practical ethical privacy issues appear in the library LA literature because of the intersection of data-handling issues and the ethical implications of data practices. These may include informed consent, data collection, data retention, and ethical review, among others. At present, no professional library association has provided comprehensive guidance or best practices for ethical data handling, though the Technologies of Surveillance Working Group of the Digital Library Federation is drafting guidelines and has created a glossary explaining research ethics intersecting with the use of library user data.

A major ethical issue of data collection is the use of informed consent. In some instances, students may have the option to consent to or explicitly opt into LA practices and research projects. Alternately, LA initiatives may automatically include student data, putting the onus on students to opt out of data collection and analysis procedures to maintain their privacy. Consent is sometimes uneven across a project; it does not follow that, simply because students opt in to a focus group, the data representing those students should be included in other LA projects. For instance, Stone and Ramsden performed quantitative analytics without a reported consent process but then conducted a qualitative analysis via a focus group for which they obtained explicit consent. Very little information is available on opt-out procedures in academic libraries. One study was nearly unique in its description of how the studied library provided an, admittedly buried, opt-out mechanism. A review of data practices in library LA found that around 10 percent of published studies in this area explicitly mention consent, opt-in, or opt-out. This finding is similar to those from Perry et al., which found that only about 40 percent of respondent libraries inform students of analytics, and of those only 60 percent (approximately 20% overall) gave students the option to opt in or out of analytic projects. From this report, it seems that libraries are making most decisions about LA in the complete absence of student consent.

Data selection and retention remains a thorny practical problem for library LA practitioners. Arguments exist suggesting that libraries—and their institutions—should collect as much data as they can at the most granular level. Doing so opens up analytical avenues that would be prematurely shut down by collecting fewer data. Naturally, these decisions intersect with wider student privacy concerns. For instance, tracking that students checked out books, versus the number of books checked out, versus the type of books checked out leads to different student privacy implications. Similarly, electronic database usage can be measured as used/not used, the number of sessions, or even as timed increments within a specific database. There is no consistency in data granularity across the library LA literature and, as yet, no best practices or standards.
Where student data are not properly secured or are retained indefinitely instead of deleted, the potential for leakage, misuse, or a breach of privacy increases. Unfortunately, the library LA literature indicates that libraries keep data for long durations with little known about how data are secured. Briney found that only about 10 percent of published library LA studies mention data security practices. These findings are echoed in Perry et al., where just over half of the survey respondents claim that they have security protections (physical and/or digital) on LA data, and only a third secure data during transit. Additionally, fewer than half of respondents had a retention schedule for LA data, and a full 20 percent planned to hold onto these data “indefinitely.” Yoose provides a singular positive example of working through the tradeoffs between collecting adequate data for analysis and maintaining data securely using practices such as data minimization, de-identification, and limiting access to sensitive data. Overall, little is known about data security in library LA, with sufficient evidence of questionable data practices indicating that libraries must attend more carefully to these practical concerns. Best practices for dealing with sensitive data, prescribed limited data collection, robust security, and finite data retention and deletion would help in this area.

There is notable variance in how librarians participating in LA projects approach internal or external ethical reviews of their work, whether from the IRB or another party. Academic libraries frequently reported in their research that they only sought IRB review if planning formal publication and—as previously discussed—that IRBs may report “de-identified” data as exempt or as nonhuman subject data, requiring no formal review and no informed consent procedures. Additionally, for the United States, quality improvement projects may not be considered research requiring human subject protection, and the Family Educational Rights and Privacy Act (FERPA) allows institutional staff to access data when they have a “legitimate educational interest.” FERPA also includes a broad exemption for institutionally affiliated third-party actors, which may include library vendors such as Springshare or the pervasive use of Google Analytics on library websites. Because of these exceptions, LA studies may not ever undergo IRB review or receive any scrutiny around the standards of “beneficence” (respecting research subjects and protecting them from harm) and “justice” (ensuring that benefits and harms of research are equitably distributed).

Reidentification remains a significant issue as data sets increase in size and are combined with other data. This commonly cited issue outside of the context of higher education is just as applicable to colleges and universities and their libraries. The literature often describes student data used for LA purposes as anonymous (and thus no longer the private information of a single individual); yet, so-called anonymization procedures are not sufficient. Anonymization procedures are often described only superficially, with the argument made that the removal of direct identifiers (such as names and identification numbers) is sufficient while ignoring the fact that other identifiers that include educational information can easily allow for reidentification. Several publications also failed to fuzz, combine, or redact populations too small to protect individual subjects (“small n’s”), heightening those subjects’ vulnerability to reidentification. At least one study reported an easily reidentifiable outlier—an 83-year-old person—making that individual’s use of the library no longer private.

**Recommendations**

Ethical transgressions do not arise solely due to a lack of concern for the harms LA may create. The aforementioned issues are likely due to accessibility to larger swaths of data and librarians...
needing to reskill and reconsider data practices in light of new methodologies. As van Dijk and Kitchin argue, the datafication of human life has created a new scientific research paradigm with accompanying changes in the study of value, knowledge, and being. The current state of the literature reveals skill gaps in the profession that require addressing to alleviate the ethical issues around LA. Once the profession addresses these gaps, some of the privacy concerns may decrease.

Institutional dynamics complicate the ethics of LA. Institutional administrators may push to override the library profession’s ethical principles in an effort to pursue data-mining practices and demonstrate oft-cited outcomes such as increases in retention and graduation. Librarians may feel an obligation to collect as much data as possible, especially as a political tool to curry more favor when budget requests are due; Jones has documented such pressures in interviews with librarians.

How to navigate these tricky political conditions will depend on local conditions, but there exist a few guiding principles for moving forward. First, ethical positions should be clear to librarians and made clear to those pressuring them to pursue LA. Librarians should reflect on their ethics and discuss their positions with their peers to enter into campus LA discussions with a clear sense of direction; see the “Action Handbook” by Young, Clark, Mannheimer, and Hincliff for more ideas in this area. Second, while some interpret ethical principles as black/white directives or bright lines that they should not cross, they should instead be pragmatically used as guideposts that point to optimal outcomes and need interpretation at the local level; doing so will lead to a more fruitful discussion with administrators and others pursuing an LA agenda. Within academic libraries, practitioners and researchers have provided examples of how to balance ethical concerns with practical data needs.

Librarians are not rigorously prepared to conduct ethical qualitative or quantitative research. Previous research has found that there is no relationship between taking a research methods course in a library and information science graduate program and feeling prepared to conduct research as a professional librarian. Librarians are likely to employ self- and continuing education to fill the research methods gap in their formal education. For LA specifically, more than 15 percent of survey respondents in Perry et al. said that staff receive no training at all for analytics projects. Even where training is occurring, it is not uniform or thorough and most often consisted of IRB and FERPA training.

LA require practitioners to increase their methodological capacities if they are to pursue analytic-based practices and conduct LA research in rigorous and valid ways. It does no good to the profession or the students it serves by taking up LA to appease administrators—and using precious resources and time—only to do it poorly. Moreover, as Robertshaw and Asher highlight, applying analytic practices incorrectly and drawing conclusions without appropriate statistical support is an unethical practice.

A complicating factor for librarians is that institutions expect professional librarians to conduct research yet do not provide sufficient support. Hoffman, Berg, and Koufogiannakis found that “Institutional Structures and Supports” were one of three broad categories that contribute to research productivity. Intra-library mentoring also contributes to research success, but not all libraries offer formal mentoring; libraries that grant tenure are more likely to provide this type of support. Lacking resources like time, institutional support, or professional mentorship may cause librarians to 1) not be prepared methodologically, 2) not fully consider the ethical implications of their study design and implementation, and/or 3) not have time to conduct their research in an ethically rigorous way.
Librarians, and particularly library administrators, should push for institutional support from their administrative peers and colleagues in other offices. Ethically pursuing LA requires time and effort to establish goals, reflect on practices, improve infrastructures, and develop strategies. Training specifically for librarians in this area is beginning to be available from IMLS-funded projects and professional organizations. Moreover, librarians should seek the support of those whose expertise is complementary to their own, such as seeking input from statistical experts, getting guidance from institutional researchers, and conferring with philosophy professors to help them navigate thorny ethical issues. These collaborative approaches have the potential to lead to more fully informed and ethically sound LA practices.

There exists an unquestionable power imbalance between students and institutional actors, including librarians. Institutions can analytically dig into student lives by accessing, analyzing, and acting on granular student data. Jones and McCoy argue that the power disparity enables institutions to render students as data subjects whose lives can be molded according to the interests of educational-technology designers and institutions. Rubel and Jones caution that institutional interests should not be conflated with student interests, noting that, while powerful administrators may want to increase retention and graduation rates, the question remains whether the analytical means employed to achieve those ends are justifiable.

Whose interests are served by LA is a paramount question. Where researchers asked students about their perceptions of learning analytics, their positive views partially depended on an ability to manage their privacy. Students may not feel that they can freely opt out of analytics without damaging their academic success, even when they are informed about data practices. Inside the library, this is likely to add to students’ existing library anxiety, and it may abruptly and negatively affect the way students view the library as a safe intellectual space. If a power balance is to occur when librarians pursue LA, they are obligated to work hand-in-hand with students; engaging a student advisory board may resolve this concern.

**Conclusion**

As LA initiatives continue to evolve, librarians will need to understand the changes in the field and the potential impact on local and national practice. Privacy considerations will continue to emerge as potential and actual threats to student privacy arise. Library ethics do not directly inform the Common Rule and ethical interpretations by IRBs, so such things cannot guide library LA projects as well as librarians might wish. While many librarians are undertaking and publishing about LA initiatives, current research has identified several gaps in practice, including the need for adoption of sensitive data-management best practices and the need for more training and interdisciplinary collaboration to improve methodological and ethical practices. Additionally, libraries must integrate student participation to identify their preferences and needs. Further research is also warranted to understand LA projects and practices at nonresearch libraries and academic libraries outside the United States.

Academic libraries’ goal should be to balance the potential rewards from LA practices with the risks these practices pose to users. To do this successfully, librarians first need to know which practices endanger user privacy so clearly and dangerously that they are never allowable under any circumstances. Next, librarians need guidelines for balancing concerns with benefits appropriately. Finally, librarians need to ensure that their data-collection and data-management practices accord responsibly with best practices. By identifying and understanding these risks, librarians may begin to approach LA initiatives from an informed and critical perspective to best serve their students.
APPENDIX A. Abbreviations
ACRL = Association of College and Research Libraries
ALI = Academic Library Impact
ALA = American Library Association
ARL = Association of Research Libraries
CAO = Chief academic officer
GPA = Grade point average
IFLA = International Federation of Library Associations
ILS = Integrated library system
IRB = Institutional review board
LA = Learning analytics
LMS = Learning management system
NISO = National Information Standards Organization
OCLC = Online Computer Library Center
SIS = Student information system
UCDM = Unizin Common Data Model
VAL = Value of Academic Libraries
### APPENDIX B. Institute of Museum and Library Services Grants on Learning Analytics

<table>
<thead>
<tr>
<th>Title</th>
<th>Amount</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student Privacy in the Datafied Classroom: Identifying Instructional Student Privacy Practices to Facilitate Librarian/Faculty Conversations</td>
<td>$306,682</td>
<td><a href="https://imls.gov/grants/awarded/lg-18-19-0032-19">https://imls.gov/grants/awarded/lg-18-19-0032-19</a></td>
</tr>
<tr>
<td>Getting to Know Their Data Doubles</td>
<td>$514,484</td>
<td><a href="https://www.imls.gov/grants/awarded/lg-96-18-0044-18">https://www.imls.gov/grants/awarded/lg-96-18-0044-18</a></td>
</tr>
<tr>
<td>Library Integration in Institutional Learning Analytics (LIILA)</td>
<td>$99,876</td>
<td><a href="https://www.imls.gov/grants/awarded/lg-98-17-0019-17">https://www.imls.gov/grants/awarded/lg-98-17-0019-17</a></td>
</tr>
<tr>
<td>Privacy in Libraries</td>
<td>$249,504</td>
<td><a href="https://www.imls.gov/grants/awarded/re-95-17-0076-17">https://www.imls.gov/grants/awarded/re-95-17-0076-17</a></td>
</tr>
</tbody>
</table>
APPENDIX C. Learning Analytics Codes of Ethics and Institutional Policies

<table>
<thead>
<tr>
<th>Name</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asilomar II: Student Data and Records in the Digital Era</td>
<td><a href="https://asilomar.sites.stanford.edu">https://asilomar.sites.stanford.edu</a></td>
</tr>
<tr>
<td>Jisc: Code of Practice for Learning Analytics</td>
<td><a href="https://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics">https://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics</a></td>
</tr>
<tr>
<td>The Open University: Ethical Use of Student Data for Learning Analytics</td>
<td><a href="https://help.open.ac.uk/documents/policies/ethical-use-of-student-data">https://help.open.ac.uk/documents/policies/ethical-use-of-student-data</a></td>
</tr>
<tr>
<td>University of Hawaii: Resolution Supporting Learner Data Privacy Principles and Practices</td>
<td><a href="https://docs.google.com/document/d/1bqJTZ4tCK3SFDsS4rXVNK5xIS87g0ajfE3EYk3u0Zh8/edit">https://docs.google.com/document/d/1bqJTZ4tCK3SFDsS4rXVNK5xIS87g0ajfE3EYk3u0Zh8/edit</a></td>
</tr>
</tbody>
</table>

Notes


utdmercury.com/alexa-where-are-you-here/ [accessed 3 March 2020].


34. PSLC DataShop, “About / Frequently Asked Questions (FAQ),” DataShop@CMU, available online at https://pslcdatashop.web.cmu.edu/about/faq.html [accessed 25 February 2020].


39. Paula Young Lee, “Drown the Bunnies’: Mount St. Mary’s President Fires Faculty for Backlash against His ‘Put a Glock to Their Heads’ Freshman Retention Plan,” Salon (Feb. 10, 2016), available online at https://www.salon.com/2016/02/10/drown_the_bunnies_mount_st_marys_president_fires_faculty_for_backlash_against_his_put_a_glock_to_their_heads_freshman_retention_plan/ [accessed 3 March 2020].


51. Malenfant, “Assessment in Action.”


57. Connaway et al., Academic Library Impact.


62. Lisa Massengale, Pattie Piotrowski, and Devin Savage, “Identifying and Articulating Library Connections...


65. Massengale, Piotrowski, and Savage, “Identifying and Articulating Library Connections to Student Success.”


67. Massengale, Piotrowski, and Savage, “Identifying and Articulating Library Connections to Student Success.”


82. Soria et al., “Library Use and Undergraduate Student Outcomes.”


92. Soria, Fransen, and Nackerud, “The Impact of Academic Library Resources on Undergraduates’ Degree Completion.”


96. Soria, Fransen, and Nackerud, “The Impact of Academic Library Resources on Undergraduates’ Degree Completion.”

97. Kogut, “Academic Library Services and Undergraduate Academic Success.”


101. ALA. “Code of Ethics of the American Library Association.”

102. ALA. “Library Bill of Rights.”

103. IFLA. “IFLA Statement on Privacy in the Library Environment.”

104. NISO. “NISO Consensus Principles on Users’ Digital Privacy in Library, Publisher, and Software-Provider Systems.”


106. Connaway et al., Academic Library Impact.


118. Asher et al., “Ethics in Research Use of Library Patron Data.”


121. Asher, “Risk, Benefits, and User Privacy.”


124. LeMaistre, Shi, and Thanki, “Connecting Library Use to Student Success.”


126. Perry et al., “Learning Analytics, SPEC Kit 360.”


129. Becky Yoose, “Patrons, Privacy, and Technology in the Library” (2017), available online at https://docs.google.com/presentation/d/14gaycdLUCiD7I1Y2BAsNggsG3D6fmVQFDnk5S6d1XA/edit#slide=id.p [accessed 3 March 2020].

130. Fransen and Peterson, “Graduate in Four Years? Yes, the Library Can Help with That!”


132. Alison Pepper and Margie Jantti, “The Tipping Point: How Granular Statistics Can Make a Big Differ-


134. Perry et al., “Learning Analytics, SPEC Kit 360.”


142. Perry et al., “Learning Analytics, SPEC Kit 360.”


146. Jones, “Just Because You Can Doesn’t Mean You Should.”


152. Perry et al., “Learning Analytics, SPEC Kit 360.”

153. Robertshaw and Asher, “Unethical Numbers?”


Jones and McCoy, “Reconsidering Data in Learning Analytics.”


