Prioritizing Privacy: Training to Improve Practice in Library Analytics Projects

ABSTRACT

The **University of Illinois at Urbana-Champaign is the lead applicant**, in partnership with Indiana University-Indianapolis, for this Laura Bush 21st Century Librarian Community Catalysts Grant Proposal for *Prioritizing Privacy: Training to Improve Practice in Library Analytics Projects.*

Prioritizing Privacy is **a three-year continuing education program that will train academic library practitioners to comprehensively address privacy and other related ethical implications of learning analytics projects** (e.g., autonomy, agency, and trust). The training program will guide participants to explore learning analytics, privacy theory, privacy-by-design principles, and research ethics and then present participants with case studies. Participants will develop a plan for a learning analytics project prioritizing privacy protections.

The **project plan** will be carried out in three phases: Curriculum Design; Training and Evaluation; and, Dissemination. A team of content experts will contribute to the curriculum development and an advisory board will provide guidance and feedback. The **primary deliverables** of Prioritizing Privacy are: (1) face-to-face training for an estimated 200 participants; (2) online training for an estimated 200 participants; (3) an open educational resource packet consisting of the training curriculum, guidelines for facilitating the training, and recommendations for incorporating the materials into other training programs and library science courses; and (4) at least two peer-reviewed conference presentations and one peer-reviewed research publication.

As a result of *Prioritizing Privacy*, **academic library practitioners will be better prepared to consider fully the privacy implications of library analytics projects and to improve the design of such projects in order to strengthen personal data protection practices**. Participants will have expanded knowledge of the interplay and tensions between learning analytics and library values as well as improved ability to navigate these tensions. They will specific skills related to designing learning analytics projects with attention to privacy and be prepared to use various methods and tools that can be deployed to protect privacy and provide for better data management.

Prioritizing Privacy will **directly train up to 400 academic library professionals**. Each of these individuals will bring the knowledge and skills that they gain through the training to their workplace setting as they apply them in learning analytics projects. As library learning analytics work tends to involve teams as well as engagement with campus partners, an **estimated additional 1600-2000 people will be impacted indirectly** by *Prioritizing Privacy*. Ultimately, though, the impact of Prioritizing Privacy will be on how library learning analytics projects are designed and the resultant protections for students.

Given the “big data” nature of learning analytics projects, the impact of this is tremendous. For example, if each participant conducts a learning analytics project with only 2,500 students, which is relatively small size for a learning analytics project, that means that *Prioritizing Privacy* training will **impact the privacy protections offered to 1 million students**. The impact on students is only amplified once use of the OER packet and other deliverables from Prioritizing Privacy are considered.

Prioritizing Privacy: Training to Improve Practice in Library Analytics Projects

The University of Illinois at Urbana-Champaign, with Indiana University-Indianapolis, requests $249,198 for *Prioritizing Privacy: Training to Improve Practice in Library Analytics Projects*.

Project Description

Prioritizing Privacy is a three-year continuing education program that will train academic library practitioners to comprehensively address privacy and other related ethical implications of learning analytics projects (e.g., autonomy, agency, and trust). The training program will guide participants through materials related to learning analytics, privacy theory, privacy-by-design principles, and research ethics and then present participants with case studies. The case studies will explore the contours of ethical reasoning and decision-making, particularly with respect to challenges associated with the higher education accountability movement. Participants will finish training by strategically developing a plan for a learning analytics project prioritizing privacy protections.

The primary deliverables of *Prioritizing Privacy* are: (1) face-to-face training for an estimated 200 participants; (2) online training for an estimated 200 participants; (3) an open educational resource packet consisting of the training curriculum, guidelines for facilitating the training, and recommendations for incorporating the materials into other training programs and library science courses; and (4) at least two peer-reviewed conference presentations and one peer-reviewed research publication.

Statement of Broad Need

As a Community Catalysts project, the goal of *Prioritizing Privacy* is to strengthen the capacity of academic librarians to integrate privacy theory and principles with learning analytics practice.

Higher education institutions are facing significant accountability pressures to prove that their efforts produce valuable results and their resource expenditures are justifiable.[1] One result is that institutional administrators, and the stakeholders to whom they report, want quantifiable data to drive analytic solutions. These data include information on allocations, expenditures, etc. and are used to calculate investments and efficiencies.

In addition to traditional business intelligence strategies, colleges and universities have adopted learning analytics methods to investigate issues of student learning and success.[2] Learning analytics are the "measurement, collection, analysis, and reporting of [student and other data] for the purposes of understanding and optimizing learning and the environments in which it occurs."[3]

[1] Thaddieus Conner & Thomas Rabovsky, Accountability, Affordability, Access, *Policy Studies Journal*, 39(s1), 2011, https://doi.org/10.1111/j.1541-0072.2010.00389_7.x

[2] Phil Long & George Siemens, Penetrating the Fog: Analytics in Learning and Education, *EDUCASUE Review*, September/October 2011, https://er.educause.edu/~/media/files/article-downloads/erm1151.pdf

[3] George Siemens, Learning Analytics: Envisioning a Research Discipline and a Domain of Practice*, in Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 2012, http://dx.doi.org/10.1145/2330601.2330605

Learning analytics have helped institutions optimize advising, predict student retention, and increase student engagement.

Feeling the same pressures as their institutions, as well as the desire to contribute to student learning and success, academic libraries have begun to participate in learning analytics practices.[4] Learning analytics work extends long-standing library assessment and evaluation practices and they may help further demonstrate library impact on student learning, faculty productivity, and more.[5] Nonetheless, regardless of the benefits that could accrue, learning analytics unquestionably presents challenges to student privacy, thus straining the professional ethics commitments that librarians make to uphold user confidentiality, respect privacy in information seeking and use, and support intellectual freedom.[6] When facing these "privacy conundrums," [7] librarians may refrain from engaging with campus learning analytics projects, meaning that librarian values around privacy and confidentiality are missing from those campus conversations.

It is generally assumed that privacy is important and valuable – so much so that scholars and commentators may fail to argue why, exactly, this is so. Privacy is a complex, multifaceted concept at the center of a network of other values (e.g., autonomy, trust, secrecy, and relationship-building)[8] that is contextually situated and protected according to expected ways that information flows.[9] In the context of librarianship, privacy has been historically situated as an instrumental value in service to intellectual freedom.[10] Limited access to one's personal and intellectual behaviors enables individuals to pursue ideas, consider alternative value sets, and develop speech.[11] Specifically, "librarians argue that information is a good necessary for all individuals to

[4] Association of College and Research Libraries, *Academic Library Impact: Improving Practice and Essential Areas to Research*, 2017, http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/academiclib.pdf

[5] Megan Oakleaf, The Problems and Promise of Learning Analytics for Increasing and Demonstrating Library Value And Impact, *Information and Learning Science*, 119(1/2), 2018, https://doi.org/10.1108/ILS-08-2017-0080

[6] Kyle Jones & Dorothea Salo, Learning Analytics and the Academic Library: Professional Ethics Commitments at a Crossroads, *College & Research Libraries* 79(3), 2018, https://doi.org/10.5860/crl.79.3.304; Kyle Jones & Ellen LeClere, Contextual Expectations and Emerging Informational Harms: A Primer on Academic Library Participation in Learning Analytics Initiatives, in *Applying Library Values to Emerging Technology: Decision-Making in the Age of Open Access, Maker Spaces, and the Ever-Changing Library*, 2018, https://ssrn.com/abstract=2980784

[7] Megan Oakleaf, *Library Integration in Institutional Learning Analytics*, 2018, https://library.educause.edu/-/media/files/library/2018/11/liila.pdf

[8] Daniel J. Solove, *Understanding Privacy,* Harvard University Press, 2010

[9] Helen Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, Stanford University Press, 2009

[10] Alan Rubel, Libraries, Electronic Resources, and Privacy: The Case for Positive Intellectual Freedom, *Library Quarterly: Information, Community, Policy* 84(2), 2014, https://doi.org/10.1086/675331

[11] Neil M. Richards, Intellectual Privacy, *Texas Law Review* 87(2), 2008, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1108268

make rational decisions and participate in a democratic society. The profession comports itself accordingly and it has developed informational norms to respect those values."[12]

Librarians are aware of and trained in research methods and, to a lesser degree, research ethics; however, learning analytics presents new challenges to existing norms, values, and information policies at the local, state, and federal levels.[13] Learning analytics are employ predominantly quantitative methods and librarians are less likely to be trained in advanced statistical analysis than other campus experts.[14] There is a pressing need to train librarians to handle the particular data ethics issues that arise in learning analytics work—especially the privacy issues—before they begin pursuing learning analytics projects.

IMLS-funded projects such as *Assessment in Action*[15] have provided training in impact evaluation methods and generated a wealth of case studies of library impact on student learning and success.[16] Other IMLS-funded projects, such as *Library Values & Privacy in Our National Digital Strategies*[17] and *A National Forum on Web Privacy and Web Analytics*,[18] have convened the library community for important discussions of privacy issues related to a range of topics including library analytics, learning analytics, web design, library policy and practice, etc.

Lacking, however, has been any training focused on privacy and learning analytics. *Prioritizing Privacy* has its genesis in conversations at the *Library Values & Privacy in Our National Digital Strategies*, which identified the need for continuing education on privacy for academic librarians. The final report of that convening states that:

> Participants repeatedly expressed concern that library staff, professionals, and administrators all fell short in terms of receiving proper training and education around issues of patron privacy. Literacy gaps persist on issues of privacy law, new technological threats, possible technical solutions, and standard privacy best practices all threaten to limit the ability to sufficiently protect patron privacy.[19]

[12] Jones & LeClere, 2018

[13] Lisa Janicke Hinchliffe, Privacy in User Research: Can You?, *The Scholarly Kitchen,* https://scholarlykitchen.sspnet.org/2018/09/05/privacy-in-user-research-can-you/

[14] Juris Dilevko, Inferential Statistics and Librarianship, *Library & Information Science Research* 29(2), 2007, doi:10.1016/j.lisr.2007.04.003; Soyeon Park, The Study of Research Methods in LIS Education, *Library & Information Science Research* 26(4), 2004, doi:10.1016/j.lisr.2004.04.009

[15] Association of College and Research Libraries, *Assessment in Action: Academic Libraries and Student* Success, http://www.ala.org/acrl/AiA

[16] Association of College and Research Libraries, *Reports from Assessment in Action,* https://apply.ala.org/aia/public; *Assessment in Action Bibliography,* https://www.acrl.ala.org/value/?page_id=980

[17] Michael Zimmer & Bonnie Tijerina, *Library Values & Privacy in our National Digital Strategies: Field guides, Convenings, and Conversations*, 2018, https://cipr.uwm.edu/2018/08/02/project-report-library-values-privacy/

[18] *National Web Privacy Forum: Achieving Privacy in the Age of Analytics,* 2018, https://www.lib.montana.edu/privacy-forum

[19] Zimmer & Tijerina, 2018

Project leaders for *A National Forum on Web Privacy and Web Analytics* have released draft versions of their "Action Handbook" and "Pathways for Action" for public comment. Many of these potential pathways document the need for privacy training, including the Toolkit for Values-Based Assessment, Privacy Research Institute, and Privacy Training for Leadership Institutes. Though these needs are articulated in the specific context of web analytics, the discussions during the convening identified that the need for such training is broader than web analytics and encompass learning analytics as well.

The report from the IMLS-funded *Library Integration in Institutional Learning Analytics* project makes clear the complexity of the kinds of considerations and decisions that librarians face when engaging with learning analytics:

> Finally, determining what data to include in learning analytics efforts requires deep reflection and thorough discussion of any data points that may raise security, privacy, or ethical concerns in order to decide whether each should be included or excluded from analysis. Librarians must take ownership of the decision-making process with regard to what library data is used to support student learning and success. How much data is needed? How detailed must it be? What data elements should never be used? On one end of the continuum is "all of the data;" at the other end is "none of the data." Librarians must decide: Is there a point on the continuum at which student learning and success support can be maximized while maintaining professional values and ethics? What does the point at which librarians can grasp the benefits and mitigate the risks look like? And what might it take to move libraries to that point?

The report goes on to observe that "by learning more about the issues involved in learning analytics data use, librarians can engage in productive discussions about potential risks."[20]

Indeed, there is no lack of information about the pressing need for training on privacy in learning analytics. *Prioritizing Privacy* responds to this widely-documented but not yet addressed need for continuing professional education on privacy in library learning analytics.

Project Design

Prioritizing Privacy is a three-year continuing education program that will teach academic library practitioners about privacy and other related ethical issues associated with learning analytics, provide them structured experiences to reflect on ethical issues intentionally and purposefully, and support the development of privacy protections for their learning analytics projects.

Project Personnel

The Project Team for *Prioritizing Privacy* consists of PI Lisa Janicke Hinchliffe and Co-PI Kyle Jones as well as a group of content area experts. In addition to the Project Team, *Prioritizing Privacy* will be benefit from the input of an Advisory Board.

Project Team

PI Lisa Janicke Hinchliffe and Co-PI Kyle Jones are leaders in the field of privacy and library analytics. They will both be involved collaboratively in aspects of Prioritizing Privacy, including co-

[20] Oakleaf, *Library Integration*, 2018

delivering the training workshops and online courses; however, they will take leadership in specific areas. PI Hinchliffe will be the lead for the overall administration of the grant, planning and scheduling logistics, curriculum design, and deposit of materials in an open repository. Co-PI Jones will be the lead for learning assessment, program evaluation, and the scholarly deliverables.

Lisa Janicke Hinchliffe (Principal Investigator) — As the 2010-2011 President of the Association of College and Research Libraries, Lisa led the launch the ACRL Value of Academic Libraries Initiative, which also kicked off a national conversation about academic libraries, learning analytics, and user privacy. Since that time she has presented and published widely on this topic, articulating recommended practice for academic libraries. Lisa served on the core working group for the *NISO Consensus Principles on Users' Digital Privacy in Library, Publisher, and Software-Provider Systems*. She co-led the design and implementation of ACRL's Assessment in Action program, funded in part by IMLS, and is currently the Co-PI on the IMLS-funded *CARLI Counts: Analytics and Advocacy* project. She is the author of the essay *Privacy in User Research: Can You?* Her website is at http://lisahinchliffe.com.

Kyle Jones (Co-Principal Investigator) — A professor of library and information science, Kyle is a leading researcher in privacy and learning analytics, with particular expertise in ethical systems and decision-making processes. Kyle is certified in the Quality Matters program evaluation framework. He is the PI on the IMLS-funded project *Data Doubles* that is investigating the view of college students on privacy and learning analytics. Kyle is the co-author of the recent *ARL SPEC Kit 360: Learning Analytics* as well as the editor of a forthcoming special issue of *Library Trends* titled "Learning Analytics and the Academic Library: Critical Questions about Real and Possible Futures." His website is at http://thecorkboard.org/research.

The content experts were recruited for their varied areas of theoretical and practical expertise, as well as to bring additional diversity to the project with respect to institutional type, educational background, etc.

- **Kristin Briney** is the Data Services Librarian at the University of Wisconsin–Milwaukee and is a leading scholar on the intersection of privacy with data collection, analysis, management and reporting.
- **Christopher Gilliard** is a Professor of English at Macomb Community College whose scholarship focuses on privacy, institutional technology policy, digital redlining, and the re-inventions of discriminatory practices through data mining and algorithmic decision-making, especially as these apply to college students.
- **Sarah Crissinger Hare** is the Scholarly Communications Librarian at Indiana University, Bloomington, and has expertise in open pedagogy, designing for student agency and choice, and open educational resources.
- **Neil Richards** is Koch Distinguished Professor of Law at Washington University and expert on information privacy, including big data ethics, intellectual privacy, and trust.
- **Alan Rubel** is the Associate Professor in the Information School and Director of the Center for Law, Justice, and Society at the University of Wisconsin-Madison, and is an expert in privacy, intellectual freedom, and data analytics.
- **Michael Zimmer** is currently Associate Professor in the School of Information Sciences at the University of Wisconsin-Milwaukee with a research agenda focused on data and information ethics. Starting in the fall he will be faculty in the Department of Computer Science at Marquette University.

Advisory Board

Members of the Advisory Board were recruited for the broader but very-much related perspectives that they bring.

- **Andrew Asher** is the Assessment Librarian at Indiana University, Bloomington, and is a leading voice on issues related to the tensions between user privacy and library service development, particularly with respect to campus learning analytics projects.
- **Sol Bermann** is the Chief Privacy Officer and Interim Chief Information Security Officer at the University of Michigan and is a leading specialist on privacy and student data.
- **Deborah Caldwell-Stone** is the Deputy Director of the Office of Intellectual Freedom of the American Library Association and a national policy advocate and educator on privacy practices in libraries.
- **Debra Gilchrist** is Vice President for Learning and Student Success at Pierce College, an award-winning community college making extensive use of data to guide development, implementation, and assessment of policies, procedures, and strategic initiatives in order to advance student success.
- **Bonnie Tijerina** is a Researcher at the Data & Society Research Institute who has been a national leader in projects on privacy and libraries as well as privacy and research data.
- **Amelia Vance** is the Director of Education Privacy at the Future of Privacy Forum with extensive experience with FERPA and other governmental regulations related to student privacy.

The strength of the Project Team and Advisory Board will help ensure the successful completion of all grant activities and quality of the deliverables.

Project Workplan – Curriculum Design, Training and Evaluation, and Dissemination

Prioritizing Privacy will be carried out in three phases that generally correspond to the three years of the grant.

Year One: Curriculum Design

Prioritizing Privacy has been conceptualized through a needs assessment undertaken through a review of the literature and reports from previous IMLS grants related to privacy. Most of these reports include suggestions of different topics that could be addressed in such training. For example: "Areas for education and dialogue include: 1) anonymity, confidentiality and privacy; 2) personally identifiable information; 3) data privacy and security; 4) opt-in and opt-out choices; 5) institutional data sharing and storage; and 6) risk mitigation practices."[21] Throughout the grant, the literature will be monitored for additional relevant publications.

A needs assessment survey will also inform the curriculum of *Prioritizing Privacy*. Co-PI Jones will oversee the survey, which will gather information about the kinds of training academic library practitioners need and prefer with respect to both content and delivery. The literature review plus survey will serve as the foundation for the curriculum design, which will be guided by the

[21] Oakleaf, *Library Integration*, 2018

instructional design principles and processes in *Understanding by Design*,[22] a methodology that PI Hinchliffe has used for more than a decade in her work. A strength of this instructional design methodology is its central focus on identifying training outcomes and developing a robust assessment plan for determining if the outcomes have been achieved.

During the design phase, individuals who are representative of the intended training participants will work through prototypes of curricular materials as well to test their effectiveness using a participatory design process. They will be recruited through personal invitations from the Project Team. Librarians at PI Hinchliffe's institution who are aware of this grant application have already expressed interest in testing the prototypes. This kind of formative and iterative assessment process strengthens training materials by ensuring they are aligned with learner needs and background knowledge. The Advisory Board will also review the training curriculum as it is developed from their various lenses of experience and expertise.

It is anticipated that the curriculum will draw heavily upon the principles of respect for persons, beneficence, and justice, which are the underlying principles of human subjects research review as codified in the Common Rule[23] in the United States, as well as principles of privacy, quality service, etc. from the American Library Association's *Code of Ethics*[24] and the foundational principles of *Privacy by Design*.[25]

Year Two: Training and Evaluation

The *Prioritizing Privacy* training curriculum will be designed for both in-person and online delivery and will be delivered in both modes.

A minimum of four in-person workshops will be offered at relevant library conferences or as stand-alone events. Libraries or associations wishing to host these workshops will be solicited through an expression of interest call. The call will be shared broadly; however, targeted emails to recruit proposals will also be sent to the leaders of the American Indian Library Association, the Asian Pacific American Library Association, the Black Caucus of the American Library Association, the Chinese American Library Association, and REFORMA: The National Association to Promote Library Services to Latinos and the Spanish-Speaking. The call will also be sent to the listservs for the Chapters Council of the Association of College and Research Libraries and alumni of the *Assessment in Action* program.

Hosts and locations for the workshop will be selected to ensure geographic and participant diversity. The Advisory Board will provide input on the selection process. Enrollment in each workshop will be capped at 25 participants to ensure a productive training environment that includes attention to individual learner needs. Hosts will be responsible for recruiting individual participants in the workshop and arranging for registration, training space, and other logistical

[22] Grant Wiggins & Jay McTighe, *Understanding by Design*, Association for Supervision and Curriculum Development, 1998

[23] *49 CFR Part 11: Federal Policy for the Protection of Human Subjects*, 2017, https://www.govinfo.gov/content/pkg/FR-2017-01-19/pdf/2017-01058.pdf

[24] American Library Association, *Code of Ethics*, 2008, http://www.ala.org/tools/ethics

[25] Privacy by Design Centre of Excellence, *The Seven Foundational Principles*, https://www.ryerson.ca/pbdce/certification/seven-foundational-principles-of-privacy-by-design/

considerations. The grant budget reflects the costs of putting on workshops (instructor expenses and training materials); however, if the locations chosen prove less expensive than anticipated, the number of workshops will be expanded (provided IMLS approves).

The online training will be offered as a structured, multi-week online course using Canvas Free, a learning management system that is available for use without charge.[26] Given the focus of this grant, in selecting an online training system, special attention was paid to the privacy policies of the available options. A review of the Canvas Free Privacy Policy found it to be typical of online learning management systems with no extreme levels of data capture.[27]

The online course will be offered four times with enrollments of up to 25 participants per course. A call for participation will again be distributed broadly as well as with the targeted recruitment mentioned above. The baseline requirement to apply to participate in the online training will be having an interest in academic libraries and learning analytics and an interest in privacy and other ethical considerations in learning analytics. In selecting participants, the goal will be to include those with a range of experiences with learning analytics (none, developing, and extensive). There will be no "litmus test" for whether an applicant is already committed to any particular approach to privacy and other ethical considerations. Consideration will also be given to geographic, career stage, and participant diversity. The Advisory Board will provide input on the selection process.

Summative and formative learning assessments will be built into the delivery of both in-person and online learning environments in order to evaluate the effectiveness of the training and identify opportunities for improvement. In addition, program evaluation will include a pre/post survey related to knowledge, skills, and confidence with respect to privacy and library learning analytics; assessment of learning activities that demonstrate achievement relative to each of the primary training concepts; and rubric evaluation of each participant's final project (a plan for a learning analytics project prioritizing privacy protections). Approximately one month after the training ends, participants will also be invited to participate in an interview to provide additional insights about the impact of the training and participants experiences. If participants publish or blog about their experiences or if the training is covered by the press (e.g., *Library Journal*), these materials will also be used in evaluating the impact of the program.

The workshops and online training will be repeated in Year Three as well using the same criteria for selection of sites and participants. As such, this training will directly reach up to 400 participants.

Year Three: Dissemination

Distributing the training materials for the in-person workshop and for the online course is a key component of the sustainability of the investment in *Prioritizing Privacy*. By making the materials available through a digital open educational resource (OER) packet, other individuals and organizations can use and adapt the materials for their own needs. The OER packet will include the training curriculum itself for the workshops and online course, guidelines for facilitating the two types of training, an explainer of the underlying instructional design model, and recommendations for incorporating the materials into other training programs and library science courses. The

[26] Instructure, *Try Canvas,* 2019, https://www.canvaslms.com/try-canvas
[27] Instructure, *Instructure Privacy Policy*, 2017 https://www.instructure.com/policies/privacy

materials will carry a CC-BY-NC license to enable re-use. The team will place the OER packet in IDEALS, the institutional repository at the University of Illinois at Urbana-Champaign.

In order to raise awareness availability of these materials, they will be marketed to the library community. The Project Team will offer a webinar providing an overview of the packet and how it can be used. The webinar will be recorded and made available through IDEALS as well for later viewing. In addition, presentations about the project will be made at conferences in order to foster further use of the OER packet. Possible conferences that will be targeted for these presentations include ACRL, CNI (Coalition for Networked Information), ALISE (Association for Library and Information Science Education), and ASIS&T (Association for Information Science and Technology).

Finally, in addition to the report on the grant itself, the results of *Prioritizing Privacy* will be published in a peer-reviewed article based on findings developed from the data gathered in Year Two. To ensure maximum access to the findings, the goal will be to place the manuscript in a high quality open-access journal such as *College & Research Libraries.*

Diversity Plan

Participation by a diverse community of practitioners across all phases of the *Prioritizing Privacy* is understood to be critical to the success of the project, particularly as underrepresented and marginalized individuals are at greater risk of harm when privacy is not prioritized in library assessment and impact evaluation projects.

Both the Project Team and the Advisory Board are comprised of individuals with a diversity of perspectives, experiences, institution types, race/ethnicity, and gender. The training materials themselves will be developed using culturally responsive pedagogy checklists disseminated at INDABA: Conquering Racism,[28] a weekend workshop sponsored by the School of Information Sciences at the University of Illinois at Urbana-Champaign.

Most central, however, to the diversity plan for *Prioritizing Privacy* is the careful attention given to soliciting hosts for the in-person workshops and participants for the online training. In particular, the plan centers reaching out to the Asian Pacific American Library Association, the Black Caucus of the American Library Association, the Chinese American Library Association, and REFORMA: The National Association to Promote Library Services to Latinos and the Spanish-Speaking.

For individuals seeking to participant, the baseline qualification (i.e., interest in learning analytics and academic libraries) is intentionally minimal to remove any barriers to application that would be raised by requiring a particular job type, employment status, level of experience, or stance on how to implement ethical principles. The intention is to enroll participants who are positioned to benefit from and apply the training. The online training option will also enable participation by those who are not able to travel due to family or work obligations or otherwise limited resources.

Broad Impact

As a result of *Prioritizing Privacy*, academic library practitioners will be better prepared to consider fully the privacy implications of library analytics projects and to improve the design of such

[28] University of Illinois School of Information Sciences, *INDABA: Conquering Racism - Culturally Responsive Pedagogy Resources*, 2018, https://publish.illinois.edu/ischoolincolor/culturally-responsive-pedagogy-resources/

projects in order to strengthen personal data protection practices. Participants will have expanded knowledge and depth of understanding of the interplay and tensions between learning analytics and library values as well as improved ability to navigate these tensions. They will have specific skills related to designing learning analytics projects with attention to privacy concerns and be prepared to use various methods and tools that can be deployed to protect privacy and provide for better data management.

Prioritizing Privacy is designed to directly train up to 400 academic library professionals. Each of these individuals will bring the knowledge and skills that they gain through the training to their workplace setting as they apply them in learning analytics projects. As library learning analytics work tends to involve teams as well as engagement with campus partners, the impact of the training is multiplied as it is applied in practice. A conservative estimate, based on information about the size of the campus teams who participated in *Assessment in Action*, is that each participant will likely work directly with at least 4-5 others their campus. In other words, a total of an estimated 1600-2000 people will be impacted indirectly by *Prioritizing Privacy*.

In addition, an unknowable number of library professionals will benefit from the OER packet of curriculum materials and related webinar, conference presentations, and scholarly publications; however, a high-level of engagement is predicted. For example, PI Hinchliffe's essay "Privacy in User Research" has been accessed more than 2000 times since it was published in September 2018.[29] Similarly, "Learning Analytics and the Academic Library: Professional Ethics Commitments at a Crossroads," co-authored by Co-PI Jones and Dorothea Salo is currently the second most accessed article in *College & Research Libraries* with almost 6000 views.[30] Based on these metrics, it is likely that accesses to the OER packet, conference presentation materials, and publications from *Prioritizing Privacy* will quickly reach into the thousands of views and downloads.

Ultimately, though, the impact of *Prioritizing Privacy* will be on how library learning analytics projects are designed and the resultant protections for students. Given the "big data" nature of learning analytics projects, the impact of this is tremendous. For example, if each participant conducts a learning analytics project with only 2,500 students, which is relatively small size for a learning analytics project, that means that Prioritizing Privacy training will impact the privacy protections offered to 1 million students. The impact on students is only amplified once use of the OER packet and other deliverables from Prioritizing Privacy are considered.

As participants in the *Prioritizing Privacy* in-person workshops and online courses begin to disseminate the results of their learning analytics projects though their own conference presentations and scholarly publications, they will be models for the kinds of data management practices that put privacy considerations at the center for learning analytics work. As a result, the dismal results reported in a study of current practices with respect to data privacy in library learning analytics projects[31] will be displaced by those that reflect careful attention to privacy in data collection, analysis, management, and reporting.

[29] Hinchliffe, 2018 – Statistics from WordPress analytics module on March 4, 2018.

[30] Jones & Salo, 2018 – Statistics from *College & Research Libraries* website on March 4, 2018.

[31] Kristin Briney, Data Management Practices in Academic Library Learning Analytics: A Critical Review. *Journal of Librarianship and Scholarly Communication*, 7(1), 2019, https://jlsc-pub.org/articles/abstract/10.7710/2162-3309.2268/

Prioritizing Privacy: Training to Improve Practice in Library Analytics Projects

SCHEDULE OF COMPLETION – YEAR ONE

Year One: Curriculum Design												
Activity	**9/19**	**10/19**	**11/19**	**12/19**	**1/20**	**2/20**	**3/20**	**4/20**	**5/20**	**6/20**	**7/20**	**8/20**
Identify background information for content experts	■	■										
Establish curriculum development workspace	■	■										
Conference calls with content experts to discuss content and plan in person meeting			■	■	■							
In person curriculum development meeting						■						
Produce draft curriculum for comment by advisory board and potential participants							■					
Review by advisory board and potential participants								■				
Finalize curriculum									■	■	■	■
Recruit and select hosts for first round of in-person workshops and online course participation							■	■	■	■		
Conference calls with advisory board		■			■			■			■	

Prioritizing Privacy: Training to Improve Practice in Library Analytics Projects

SCHEDULE OF COMPLETION – YEAR TWO

Year Two: Training and Evaluation												
Activity	**9/20**	**10/20**	**11/20**	**12/20**	**1/21**	**2/21**	**3/21**	**4/21**	**5/21**	**6/21**	**7/21**	**8/21**
Deliver in-person workshops (dates to be determined with host organizations)	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨		
Deliver online courses		▨	▨			▨	▨					
Conference calls with content experts to discuss any possible revisions to curriculum				▨				▨				
Assessment of training – in-person workshops and online courses	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨		
Recruit and select hosts for in-person workshops and participants for online course												
Recruit and select hosts for second round of in-person workshops and online course participation							▨	▨	▨	▨		
Revise curriculum											▨	▨
Conference calls with advisory board		▨			▨			▨			▨	

Prioritizing Privacy: Training to Improve Practice in Library Analytics Projects

SCHEDULE OF COMPLETION – YEAR THREE

Year Three: Dissemination												
Activity	**9/21**	**10/21**	**11/21**	**12/21**	**1/22**	**2/22**	**3/22**	**4/22**	**5/22**	**6/22**	**7/22**	**8/22**
Deliver in-person workshops (dates to be determined with host organizations)	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨		
Deliver online courses		▨	▨			▨	▨					
Interviews with participants in Year 2 training programs to evaluate impact	▨	▨	▨									
Propose conference presentations	▨	▨	▨									
Present conference presentations					▨	▨	▨	▨	▨			
Draft scholarly article				▨	▨							
Submit scholarly article and then follow processes for review, revise, etc. through to publication.						▨	▨	▨	▨	▨	▨	▨
Prepare OER packet and deposit in IDEALS									▨	▨	▨	▨
Conference calls with advisory board		▨			▨			▨			▨	

DIGITAL PRODUCT FORM

Introduction

The Institute of Museum and Library Services (IMLS) is committed to expanding public access to federally funded digital products (e.g., digital content, resources, assets, software, and datasets). The products you create with IMLS funding require careful stewardship to protect and enhance their value, and they should be freely and readily available for use and re-use by libraries, archives, museums, and the public. Because technology is dynamic and because we do not want to inhibit innovation, we do not want to prescribe set standards and practices that could become quickly outdated. Instead, we ask that you answer questions that address specific aspects of creating and managing digital products. Like all components of your IMLS application, your answers will be used by IMLS staff and by expert peer reviewers to evaluate your application, and they will be important in determining whether your project will be funded.

Instructions

All applications must include a Digital Product Form.

□ Please check here if you have reviewed Parts I, II, III, and IV below and you have determined that your proposal does NOT involve the creation of digital products (i.e., digital content, resources, assets, software, or datasets). You must still submit this Digital Product Form with your proposal even if you check this box, because this Digital Product Form is a Required Document.

If you ARE creating digital products, you must provide answers to the questions in Part I. In addition, you must also complete at least one of the subsequent sections. If you intend to create or collect digital content, resources, or assets, complete Part II. If you intend to develop software, complete Part III. If you intend to create a dataset, complete Part IV.

Part I: Intellectual Property Rights and Permissions

A.1 What will be the intellectual property status of the digital products (content, resources, assets, software, or datasets) you intend to create? Who will hold the copyright(s)? How will you explain property rights and permissions to potential users (for example, by assigning a non-restrictive license such as BSD, GNU, MIT, or Creative Commons to the product)? Explain and justify your licensing selections.

The University of Illinois will hold rights and ownership to the resulting intellectual property, but will not use them for commercial gain. All *Prioritizing Privacy* content including training curriculum, reports, etc. will be made available for reuse through a Creative Commons license. Applying a Creative Commons license will maximize potential for reuse of the materials.

A.2 What ownership rights will your organization assert over the new digital products and what conditions will you impose on access and use? Explain and justify any terms of access and conditions of use and detail how you will notify potential users about relevant terms or conditions.

The University of Illinois will hold rights and ownership to the resulting digital products, but will not use it for commercial gain. The *Prioritizing Privacy* content will be provided openly and freely in IDEALS (University of Illinois at Urbana-Champaign's institutional repository) for anyone to use. The project staff will notify potential users of content through public announcements, postings on the site, and through social media and other communication channels.

A. 3 If you will create any products that may involve privacy concerns, require obtaining permissions or rights, or raise any cultural sensitivities, describe the issues and how you plan to address them.

We do not anticipate creating any products in *Prioritizing Privacy* that would involve privacy concerns. Projects eventually undertaken by librarians at their home institutions as a result of the training may involve data that has privacy concerns; however, that data will belong to the institutions and not become part of the grant deliverables.

While it is the intention of project staff to create all materials related to the project, if we find it necessary to include any third-party content, we will focus on publicly available materials, and seek and obtain all necessary permissions for additional content.

Part II: Projects Creating or Collecting Digital Content, Resources, or Assets

A. Creating or Collecting New Digital Content, Resources, or Assets

A.1 Describe the digital content, resources, or assets you will create or collect, the quantities of each type, and the format(s) you will use.

We intend to create a project website to promote *Prioritizing Privacy* and to disseminate materials resulting from the grant. These materials might include, but are not limited to, an open educational resource packet consisting of the training curriculum, guidelines for facilitating the training, and recommendations for incorporating the materials into other training programs and library science courses as well as grant reports.

A.2 List the equipment, software, and supplies that you will use to create the content, resources, or assets, or the name of the service provider that will perform the work.

Project staff will use existing server space and software (Microsoft Office, web editing suite, etc.) for this project. MAXQDA software will be used to analyze the feedback from participants in the training programs. Canvas will be the platform for delivery of the online training course.

A.3 List all the digital file formats (e.g., XML, TIFF, MPEG) you plan to use, along with the relevant information about the appropriate quality standards (e.g., resolution, sampling rate, or pixel dimensions).

Project materials will likely be in PDF, Word, JPG or other standard document formats, as applicable to the type of material. All documents will adhere to quality standards relevant to the document's format.

B. Workflow and Asset Maintenance/Preservation

B.1 Describe your quality control plan. How will you monitor and evaluate your workflow and products?

Resources and materials will be compiled by the project staff, who will also be responsible for maintaining the quality and integrity of materials through regular review of the content and documents.

B.2 Describe your plan for preserving and maintaining digital assets during and after the award period of performance. Your plan may address storage systems, shared repositories, technical documentation, migration planning, and commitment of organizational funding for these purposes. Please note: You may charge the federal award before closeout for the costs of publication or sharing of research results if the costs are not incurred during the period of performance of the federal award (see 2 C.F.R. § 200.461).

Project staff will be responsible for maintaining the digital content for the duration of the grant and depositing the materials into IDEALS, the institutional repository of the University of Illinois at Urbana-Champaign, for long-term preservation and access.

C. Metadata

C.1 Describe how you will produce any and all technical, descriptive, administrative, or preservation metadata. Specify which standards you will use for the metadata structure (e.g., MARC, Dublin Core, Encoded Archival Description, PBCore, PREMIS) and metadata content (e.g., thesauri).

The project will rely on underlying metadata structures of IDEALS, the institutional repository of the University of Illinois at Urbana-Champaign.

C.2 Explain your strategy for preserving and maintaining metadata created or collected during and after the award period of performance.

Any necessary metadata created of collected during and after the award period of performance will tracked in IDEALS, the institutional repository of the University of Illinois at Urbana-Champaign, for long-term preservation and access.

C.3 Explain what metadata sharing and/or other strategies you will use to facilitate widespread discovery and use of the digital content, resources, or assets created during your project (e.g., an API [Application Programming Interface], contributions to a digital platform, or other ways you might enable batch queries and retrieval of metadata).

The project will rely on underlying metadata structures of IDEALS, the institutional repository of the University of Illinois at Urbana-Champaign, and its existing mechanisms that facilitate discovery and use of content in the repository.

D. Access and Use

D.1 Describe how you will make the digital content, resources, or assets available to the public. Include details such as the delivery strategy (e.g., openly available online, available to specified audiences) and underlying hardware/software platforms and infrastructure (e.g., specific digital repository software or leased services, accessibility via standard web browsers, requirements for special software tools in order to use the content).

Prioritizing Privacy content will be made available freely and openly to everyone through DEALS, the institutional repository of the University of Illinois at Urbana-Champaign. Access will be through a standard web browser and files will be in PDF, Word, JPG or other standard document formats.

D.2 Provide the name(s) and URL(s) (Uniform Resource Locator) for any examples of previous digital content, resources, or assets your organization has created.

The University of Illinois Library has an extensive web presence and a long record of developing digital content.

As two primary examples:

IDEALS Institutional Repository - https://www.ideals.illinois.edu/

Digital Collections Portal - https://digital.library.illinois.edu/

Part III. Projects Developing Software

A. General Information

A.1 Describe the software you intend to create, including a summary of the major functions it will perform and the intended primary audience(s) it will serve.

A.2 List other existing software that wholly or partially performs the same functions, and explain how the software you intend to create is different, and justify why those differences are significant and necessary.

B. Technical Information

B.1 List the programming languages, platforms, software, or other applications you will use to create your software and explain why you chose them.

B.2 Describe how the software you intend to create will extend or interoperate with relevant existing software.

B.3 Describe any underlying additional software or system dependencies necessary to run the software you intend to create.

B.4 Describe the processes you will use for development, documentation, and for maintaining and updating documentation for users of the software.

B.5 Provide the name(s) and URL(s) for examples of any previous software your organization has created.

C. Access and Use

C.1 We expect applicants seeking federal funds for software to develop and release these products under open-source licenses to maximize access and promote reuse. What ownership rights will your organization assert over the software you intend to create, and what conditions will you impose on its access and use? Identify and explain the license under which you will release source code for the software you develop (e.g., BSD, GNU, or MIT software licenses). Explain and justify any prohibitive terms or conditions of use or access and detail how you will notify potential users about relevant terms and conditions.

C.2 Describe how you will make the software and source code available to the public and/or its intended users.

C.3 Identify where you will deposit the source code for the software you intend to develop:

Name of publicly accessible source code repository:

URL:

Part IV: Projects Creating Datasets

A.1 Identify the type of data you plan to collect or generate, and the purpose or intended use to which you expect it to be put. Describe the method(s) you will use and the approximate dates or intervals at which you will collect or generate it.

A.2 Does the proposed data collection or research activity require approval by any internal review panel or institutional review board (IRB)? If so, has the proposed research activity been approved? If not, what is your plan for securing approval?

A.3 Will you collect any personally identifiable information (PII), confidential information (e.g., trade secrets), or proprietary information? If so, detail the specific steps you will take to protect such information while you prepare the data files for public release (e.g., data anonymization, data suppression PII, or synthetic data).

A.4 If you will collect additional documentation, such as consent agreements, along with the data, describe plans for preserving the documentation and ensuring that its relationship to the collected data is maintained.

A.5 What methods will you use to collect or generate the data? Provide details about any technical requirements or dependencies that would be necessary for understanding, retrieving, displaying, or processing the dataset(s).

A.6 What documentation (e.g., data documentation, codebooks) will you capture or create along with the dataset(s)? Where will the documentation be stored and in what format(s)? How will you permanently associate and manage the documentation with the dataset(s) it describes?

A.7 What is your plan for archiving, managing, and disseminating data after the completion of the award-funded project?

A.8 Identify where you will deposit the dataset(s):

Name of repository:

URL:

A.9 When and how frequently will you review this data management plan? How will the implementation be monitored?