

2023-03-24

Library Learning Analytics Project White Paper

Kabo, Felichism; Varnum, Kenneth J.; Alexander, Laurie; Korner, Sebastien; Bradley, Doreen

<https://dx.doi.org/10.7302/7034>

<https://hdl.handle.net/2027.42/175994>

<http://creativecommons.org/publicdomain/zero/1.0/>

Downloaded from Deep Blue, University of Michigan's institutional repository

Library Learning Analytics Project White Paper

24 March 2023

Felichism Kabo, Kenneth J. Varnum, Laurie Alexander,
Sebastien Korner, and Doreen Bradley

<https://libraryanalytics.org/>

<https://doi.org/10.7302/7034>

The Library Learning Analytics Project was funded by the
[Institute of Library and Museum Services \(LG-96-18-0040-18\)](#).

1. About the Library Learning Analytics Project

1a. Library Learning Analytics Project Overview

The Library Learning Analytics Project (LLAP) was a collaborative partnership between the [University of Michigan](#) (U-M) and the institutions that comprise the [project advisory group](#) (PAG). It was funded for three years (2018-2021) by the Institute of Museums and Library Services (IMLS LG-96-18-0040-18).

The project used U-M campus and library data in order to prototype and pilot data collection, storage, and analysis with the expectation that the tools developed would be made available to other libraries and institutions. A proximate goal was to provide guidance to PAG institutions and other libraries on how to best design and implement empirical, holistic library learning analytics (LLA) studies of the links from library usage, to learning outcomes such as research, course instruction, and publications. The project produced a set of tools, scripts, and protocols that are freely available to all libraries.

LLAP was designed to serve as a template for other libraries with respect to:

1. collecting and storing library use data with individual identifiers while maintaining the privacy of individuals;
2. designing and implementing a holistic LLA study of the link from library use to multiple learning outcomes; and
3. creating a secure cyber infrastructure for LLA research.

We refer to the Society for Learning Analytics Research (SoLAR) [definition of learning analytics](#), which is “the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs.”

We also aligned with the University of Michigan learning analytics guiding principles which were developed by the University Registrar, University Privacy Officer, and the Office of Academic Innovation. These guiding principles ensure that the creation, collection, analysis, and reporting of learning analytics data remains transparent and secure, and places privacy and ethical considerations at the forefront: <https://ai.umich.edu/learning-analytics-guiding-principles/>.

1b. Learning through Partnership

At the time the [Library Learning Analytics Project: Demonstrating How Libraries Impact Learning](#) proposal was submitted, the University of Michigan Library was in the very early days of learning analytics. Several faculty members and librarians were thought leaders, examining the role and impact of analytics on teaching, learning, training and development. The focus was on

supporting initiatives to understand and improve teaching and learning through projects such as [E²Coach](#) (a tailored educational support system) and [REBUILD](#) (use of evidence-based methods in introductory STEM courses).

It is interesting to note that increased interest and movement around learning analytics was neither central nor distributed, but came from multiple directions, driven by a desire to collect, analyze and use student data to improve learning. It arose out of a broad aspiration to assist instructors in achieving a wide range of teaching goals, to explore possibilities around emerging technology, to fulfill a commitment to service enhancements and, most importantly, to provide student tools to take charge of their own learning in new ways.

There were also many questions around ethics, data sources, intervention processes, and impacts. Within the University Library, we were exploring how and when to engage with learning analytics and to articulate the potential correlations between library use and both student and research success. Emergent scholarly practices, dynamic user needs, and new expectations about when and where learning happens called our attention to learning analytics. The University Library was called by a desire to balance our engagement, explore possibilities, live our values, connect with campus initiatives, and contribute to student learning. We wanted data to be part of our service design tools and to embrace data-informed practices to enable emergent scholarship.

Dr. Kabo from the Institute for Social Research invited the University of Michigan Library to collaborate on an IMLS grant to study how academic libraries impact learning. Primary goals included understanding library impacts on learning, especially in the area of course instruction, research, and publication, and developing sharable tools, scripts, and protocols based on principled engagement and professional agency. We were enthusiastic partners in the grant because it gave us a way and a framework to further understand how to create secure infrastructures (such as a data repository and virtual enclave) for this type of research and highlight principled, ethical treatment of data and user privacy in alignment and integrity with library professional values.

At the time of that invitation, the Library had more questions than answers. What library data would enable us to make different decisions? What data would we need to improve library services? How does the library partner with campus learning analytics efforts? Are we prepared and ready to be that campus partner? How do ethics show up in our learning analytic efforts? What role might our library play in institutional learning analytics activities? What should our librarians know about learning analytics in order to participate in this endeavor in ways that maximize student support and honor professional values? How do we create the possibilities for this future state?

Most of our focus to date at that time had been on descriptive analytics (understanding what is happening in our learning environments and activities of our users) and we had yet to strategically engage with diagnostic (what is facilitating the success of our users) or predictive (future outcomes) analytics. The shared thread through our questions was drawing connections

among existing and potential partners. The emerging goal was to build our expertise, focus on impact, turn research into practice, and use library data ethically to improve services, enhance student success and personalize services. With those questions in mind, we partnered with Dr. Kabo and in doing so, made a commitment to iterate and bring together possibility with strategy and experimentation to develop emergent processes in order to learn, share and approach analytics through the library profession lens.

2. Project Advisory Group

2a. Purpose

The Project Advisory Group (PAG) served multiple purposes within the grant. Both as a collective and as individuals, PAG members provided feedback and guidance to the researchers on perspectives regarding learning analytics capabilities and priorities at multiple types of institutions of higher education, including community colleges, four-year public and private universities, and research universities. PAG institutions also provided guidance on the types of data and methods of access to data, as well as insight into the types of expertise and technologies that are available to our partner institutions. We shared and discussed findings, methods, and protocols with PAG institutions, both electronically and face-to-face, over the entire project timeline. PAG members' professional knowledge and experience aided in shaping the tools, methods, and protocols that researchers developed to share with the PAG and broader library learning analytics community.

PAG members, accomplished library faculty and administrators, represented the following institutions: Northwestern University, Association of College & Research Libraries (ACRL), Big Ten Academic Alliance (BTAA), New England College, Saint Mary's College of California, University of Illinois at Urbana-Champaign, Wayne State University, Ferris State University, Western Michigan University, Ohio State University, University of Detroit Mercy, Oakland Community College, Washtenaw Community College, Los Angeles Public Library, and the Midwest Collaborative for Library Services (MCLS).

2b. Surveying PAG Member Needs

As mentioned, we sought a diverse group of institutions including community college, four-year public, four-year private, and large research universities to ensure we could integrate varying perspectives into our work and that diverse needs were represented in the project output. An initial [survey](#) of all PAG institutions was conducted in April 2016 to establish a baseline of discussion points, values, and needs (see Appendix A). Important questions from the survey included:

- Q.1 As the higher education environment changes, with online degree programs, changing demographics, and return on investment concerns, how would you describe the role of the academic research library?
- Q2. We also asked a series of questions around the top issues important to faculty, researchers, and students, focusing on the most important issues, selected from the list below, that we need to better understand with respect to academic libraries and research universities and how learning analytics may play into these issues.
 - Access to scholarly literature and resources
 - Access to archives, primary sources and special collections
 - Expertise in academic integrity
 - Expertise in copyright and intellectual property
 - Support for publishing and open access publishing
 - Discipline-specific research consultations
 - Expertise in digital scholarship tools
 - Data management planning
 - Information literacy expertise
 - Campus engagement programming
 - Numeric and spatial data and data visualization
 - Space for research and scholarship to take place, quiet study space
 - Partnerships on research projects
 - Partnerships on course development
 - Preservation of knowledge in all forms

2c. Privacy Guide

As the early work on this IMLS grant began, the issue of patron privacy repeatedly surfaced in conversations with our PAG institution members, in library literature, and at conferences across the country. The right to privacy and the obligation for libraries to protect the privacy of their patrons are long held values in librarianship. This principle is evident in the American Library Association's (ALA) [Library Bill of Rights](#), many statements from the ALA on privacy and confidentiality, as well as in numerous state laws related to libraries and privacy.

As an institution, the University of Michigan Library had worked through the issue of privacy and learning analytics early on in our journey into analytics, asking questions such as *How does learning analytics fit with our professional values? Who should we involve on campus? What campus policies or state laws do we need to be aware of? How does FERPA fit in?* Our discussions involved many stakeholders over the course of many conversations, some of them difficult ones, and offered opportunities for all library staff to engage in this journey.

It became clear that PAG institutions were each at various stages of discussions around privacy on their respective campuses. The U-M Library sought to share our own experiences in hopes of enriching their conversations. As several extensive learning analytics toolkits were already available, we sought to make a concise, practical guide. To do this, we created a *privacy guide* as a means to offer guidance on necessary steps to having thoughtful and productive

conversations. Our goals were twofold: first, to facilitate discussions focused on privacy for libraries engaging with learning analytics, and second, to make existing resources readily accessible for these discussions.

Development of the privacy guide began by outlining the milestones and essential conversations that we had on our own campus. These started with conversations with library staff, and continued with discussions with campus learning analytics experts including representatives from the Office of the General Counsel and the campus Chief Privacy Officer.

Based on feedback from the survey of PAG members, a subset of the LLAP team delved into deeper discussions with privacy and legal experts on the U-M campus then created a draft outline of a privacy guide. The privacy guide went through several iterations as we sought feedback from the PAG members, campus experts, and our grant team. Sections include: *Introduction; Working Definitions; Laws & Guidelines; Potential Stakeholders; Stakeholder Engagement; Videos & Presentations; Templates & Privacy Statements; and Further Readings*. The [Library Learning Analytics Privacy Guide](#) was released in 2019 and is available in the Resources section of the grant website. We hope the guide is instrumental in starting critical conversations about privacy for our PAG members as well as librarians at other institutions.

2d. Site Visits

As part of the grant, site visits to each Program Advisory Group institution were planned for the first year. Site visits were an opportunity for the library staff at each member institution to learn more about the grant as well as about learning analytics work at the University of Michigan. Likewise, these visits provided the grant researchers an opportunity for in-depth understanding of the expertise, resources, and priorities of each PAG institution so that we could address their needs in years two and three of the grant. A typical schedule for the site visits included a presentation about the grant followed by focused conversations about each institution's climate and priorities, data sources and infrastructure, and research and statistical expertise available on each campus. At least two members of the U-M grant team participated in each site visit, in addition to as many members of the PAG institutions as they chose to involve. Site visits took place for 10 of the 15 PAG institutions before needing to end visits due to COVID-19. For the remaining five PAG institutions, there were some additional in-depth conversations about their individual campus priorities and needs, but no actual campus visits. In all, these site visits were helpful to everyone in developing a shared understanding of the work of the grant, the needs of participating institutions, and furthering our relationships as library and research colleagues.

2e. PAG Activities

One of the first engagements for the PAG was to help us create a Library Learning Analytics Privacy Guide that could help other libraries work through issues around privacy and library learning analytics. At the beginning of this grant, national conversations were threaded with questions about professional ethics and how library data and policies fit into broader learning analytics efforts as well as library professional ethics. The U-M grant team used this as an

opportunity to have discussions with PAG members about various campus priorities and climate regarding the use of library data. We explored issues such as national and state laws, FERPA, library and institutional data policies, and professional ethics. We spoke with university privacy officers and general counsel officials. From these discussions, we created a [Library Learning Analytics Privacy Guide](#) to facilitate conversations about library learning analytics and made it available to the library community through our grant website. (See more about the privacy guide in section 3.)

Our quarterly *Research Meetings* were the locus for most of the engagement with the PAG members. These meetings brought together PAG members and the entire U-M grant team. Research meetings would highlight progress on grant activities such as site visits, establishment of new data pipelines, data sharing and analysis from queries to test our scripts and tools, guest speakers, opportunities to share analytics milestones from PAG institutions, and the chance to address any issues identified by PAG members or the U-M research team. Additional meetings with individual PAG institutions took place as needed, including meetings to provide consultations about data issues and strategies unique to those individual institutions. These individual meetings provided additional mentorship opportunities to advance expertise and address challenges for our PAG members.

2f. Annual Events

In addition to quarterly meetings, the work of the grant also called for annual meetings of all members including U-M researchers and PAG members. These annual events were designed to provide extended time for knowledge sharing and expertise building for everyone involved. The first, and as it turned out only, in-person event was held in May 2019. This day-and-a-half-long event was held at the University of Michigan Library and provided hands-on workshops in a variety of tools including SQL, R, and Python. The workshop curriculum was developed in response to feedback the LLAP team received from PAG participants in the form of a pre-workshop survey (see Appendix B). It also included discussion of learning analytics priorities at our respective campuses and tours of U-M libraries.

Our second onsite event, originally scheduled for May 2020, was to focus on conceptual approaches to library learning analytics and focus on theory and practice rather than technical topics as we had covered those in the year prior. However, this event was canceled due to COVID-19.

For our third, end-of-grant, event, we are hosting a virtual event May 27-28. The first day will feature a review of the grant work from the past three years, highlights from each of our PAG member institutions, and guest speakers including Megan Oakleaf. The second day of this event will be dedicated to individual consultations for PAG institutions to work in partnership with the U-M grant team on topics unique to their institutions to help them plan for future advancements in library analytics.

3. Sharing Code

3a. Overview

One of the lessons we learned through the course of this project overall is that the data from distinct source systems was formatted slightly differently from the others. This had two effects. First, we needed to develop log file parsing software for each source, so that we could correctly attribute each action (by an identified or anonymous user) and action appropriately in the database structure we developed. Second, while the parsing code we developed worked well for our environment, it is likely that it would need small adjustments to work in another library's because of differences in configurations at each institution.

Code is described in this section for each of two main areas of activity:

1. Preparing log files and extracting data from raw log files to be added to the SQL Server database;
2. R scripts to analyze data from the SQL Server database

This computer code has been added to the project's GitHub repository at <https://github.com/Learning-Library-Analytics-Project>, where it is available to anyone for use.

3b. Description of GitHub Code Repository

The GitHub repository currently contains four folders. Two of them are associated with data processing and data analysis. The other two are related to the 2019 and 2020 training workshops. However, the 2020 folder is currently empty as the workshop was canceled due to the COVID-19 pandemic.

3c. Data Preparation and Transfer

The University of Michigan Library has a complex web hosting environment. Many of our web services (the library web site, the library discovery interface, etc.) are supported by multiple web servers spread across two campus data centers. Other services (such as the Integrated Library System [ILS] that manages the library's books and other items), the proxy server, and others, are supported by single servers. The library manages many systems that generate access and transaction logs that are not within LLAP's scope, as well.

As part of the University of Michigan Library's work on the grant, we developed and implemented a process to centralize log file storage in one location within the library's server environment and routinely transfer those logs securely to a secure enclave staging area managed by the Institute for Social Research (ISR). ISR developed scripts to read the log files and add them to the project database for later analysis.

Some log files are automatically generated by the Apache web server. Other log files were essentially custom database reports generated periodically by the library's ILS or interlibrary loan software specifically to provide input for LLAP.

The first step in the ISR log ingest process was to develop a mechanism to make daily copies of the previous day's log files from the library's common log storage space. This storage area is managed by the central campus Information Technology Services unit. The centralized log files were produced in real time using log forwarding and aggregation tools on library servers. The library's intention was to preserve these log files in the consolidated server area only long enough to support transfer to the ISR data enclave and to meet U-M data retention guidelines.

ISR scripts copied log files from the library server space to the ISR data enclave for processing. Any files added or changed in the previous 24 hours were copied anew for processing or reprocessing. This redundancy ensured that a clean copy of the log files would be available to the library, and that ISR could reprocess any log files from the original source as needed. Only a small number of specifically-authorized ISR staff had access to the raw log files in the ISR data enclave.

3d. Data Ingest

Library data sources came in several formats, each needing its own process to read the individual log file lines and import them into the SQL Server database. Where the log files were generated in the same way -- for example, the first row in Figure 1 below represents logs generated by the Apache web server, all of which followed the same format. One script could process all of these files. Other formats were idiosyncratic to the source, such as checkout history from the Aleph ILS or Special Collections Library access requests. These logs needed their own individual script to read and process the log files.

Once in the ISR SQL Server database, however, information about users and transactions was normalized so that similar queries could be run across all data sources.

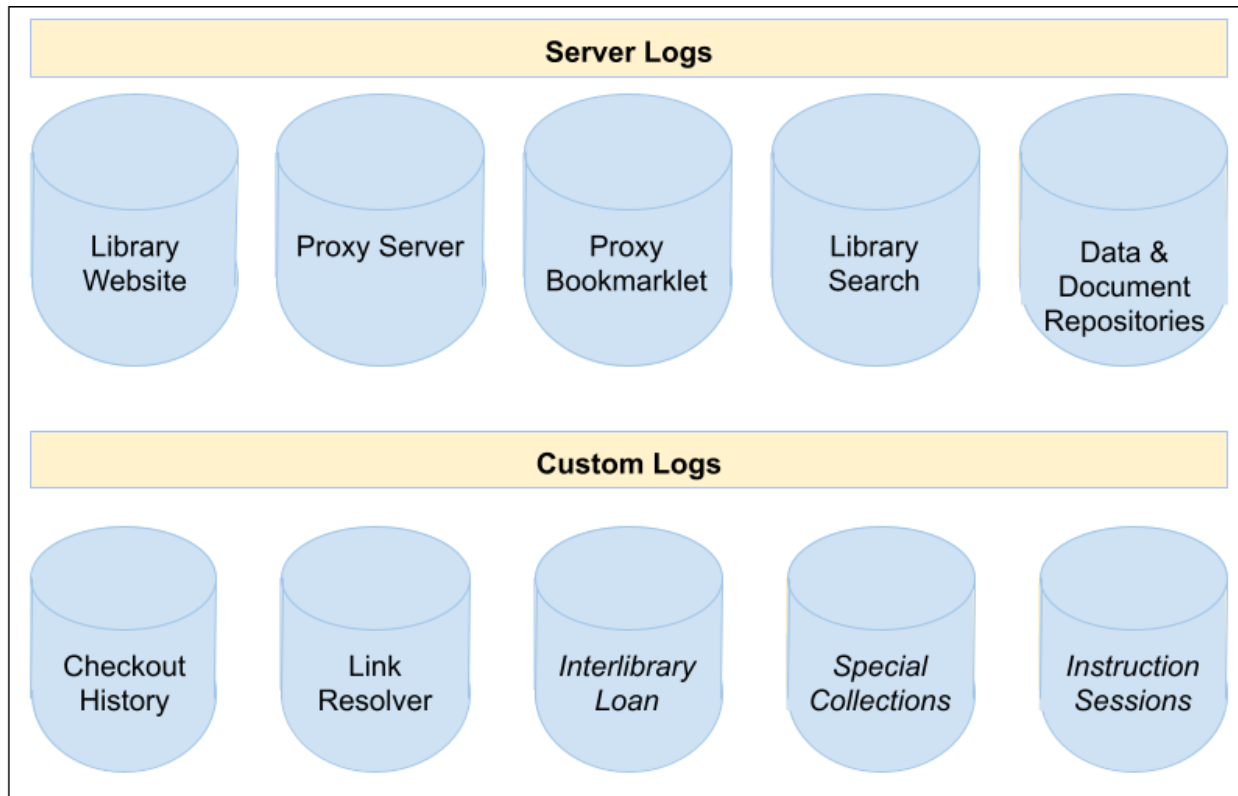


Figure 1. Data sources. Library search interactions were in the original plan but were ultimately not included because the custom coding needed to capture Library Search user interactions was not completed.

3e. Statistical Analysis

LLAP’s studies employ years of large-scale administrative data (library usage and institutional) that provide sufficient statistical power for longitudinal analyses of, for example, the links between library usage and student outcomes. These data also enable us to perform analyses for subgroups within the student population, such as by gender, racial/ethnic status, and academic units. We have run different types of descriptive analyses and regression models contingent on the outcomes and potential causal relationships that we are investigating. For example, longitudinal analyses of the links between library usage and student performance must account for factors that are common to all students, such as race/ethnicity, and factors that vary and thus cluster student behaviors and outcomes, such as academic units. To do that, we have run panel linear mixed effects regressions that account for “fixed” (common to all students) and “random” (vary across students) effects. We have also run analyses that account for the fact that disparities in academic performance are not randomly across the student population as they may be a reflection of what happened prior to the students matriculating at the university. To account for that, we have sample selection models that enable us to correct for the bias in our non-randomly selected sample of students.

We have also done exploratory machine learning using random forests, a type of ensemble learning method for classification, and k-means clustering with the elbow method, an unsupervised learning algorithm. This work has been useful in helping us identify areas that need improvement if we are to use the library usage data to identify broad categories of users based on search behaviors.

4. What Have We Learned So Far

LLAP has significantly impacted library learning analytics. One of the strongest contributions that LLAP has made to the field of library learning analytics is in terms of lessons learned, primarily in the four areas related to academic outcomes, intra- and inter-organizational collaboration, infrastructural development, and policy and technical outcomes.

4a. Academic Outcomes

We have learned that library usage has positive short- and long-term associations with student academic performance. We merged fine-grained library usage data with student outcomes data from the central university data warehouse and used longitudinal regression modeling. These steps increase our confidence in the robustness of our findings. Our analysis also revealed that not only is library curriculum-integrated instruction (CII) positively associated with academic performance, but the results also suggest that library CII could be used to mitigate academic disparities. The most important lesson we have learned, however, is the integral nature of the academic library within its parent college or university. This calls for a change in perspective with respect to important issues like privacy, data collection and management, and learning analytics. With respect to the last issue, learning analytics is of particular interest and importance to the broader university community and it is imperative that the library be an active participant in this process and dialogue.

4b. Collaboration

The LLAP team learned invaluable lessons on the nature of collaboration between different units and groups at the University of Michigan, and among a diverse group of organizations (PAG members) that had different goals and missions. LLAP activities necessitated two different types of collaboration: *internal* (tightly coupled), between several units at U-M (e.g. Institute for Social Research, the University Library), and *external* (loosely coupled), among the PAG institutions. These two forms of collaboration are best understood in the context of the two main goals of the LLAP project.

The first goal of LLAP was to pilot methods for modeling the relationship between library use and learning outcomes. This was accomplished through the analysis of datasets that were created by merging library usage data and university data such as on student performance. This effort involved the collaborative activities of individuals from several units at U-M, primarily the U-M Library, the Institute for Social Research (ISR), and the School of Information (SI). Support

was also provided by the Office of Enrollment Management and Information and Technology Services (ITS) which is the university's central IT provider. The primary challenges, especially for the individuals from the library, ISR, and IT were centered on developing shared frames of reference such as on the best practices for working with library usage data. In that instance, the consensus was to accord the library usage data the highest sensitivity rating as the LLAP team realized that it was important to do everything possible to safeguard the data since they were dealing with identifiable library usage data. As of when LLAP was launched, it was typical for libraries to purge identifiable data as soon as the transaction was completed; until 2016, the U-M Library was no different. The overwhelming majority of libraries did not have anything like the revised privacy policy of 2016 that allowed the U-M Library to collect and retain library usage data with identifiers.

The second LLAP goal was to develop tools, scripts, and protocols to serve the library analytics needs of the community of PAG institutions. At the outset of the project, the LLAP team had assumptions about the capacity of learning analytics across the PAG institutions that proved to be inaccurate as the project progressed. PAG member institutions turned out to have such a diversity of institutional goals and outcomes that made empirical library learning analytics an impractical means of demonstrating value for some PAG members. Some of the lessons learned by the LLAP team are as follows:

- I. Libraries need to clarify and resolve issues related to their internal collaboration with other units at their parent institutions before they should plan to launch external collaborations with other libraries. For example, most academic libraries cannot do library learning analytics without collaborating with their institutional research (IR) offices. Unfortunately IR offices tend to be under-resourced and under-staffed, and IR staff are apt to prioritize other units' projects (e.g., the admissions office) over library projects. Libraries may also need to work with other units that are reluctant collaborators e.g. central IT for library and institutional data, enrollment office for student data, etc.
- II. For learning analytics projects that entail rapid prototyping of ideas and processes, it is more efficacious when external collaborations are with similar institutions. In other words, the more homogenous the libraries in an exploratory project the more likely it is that tools produced are usable to all, and that resources and knowhow can be pooled for the greater good. This is less likely in the case of collaborations among a set of diverse institutions.
- III. Collaboration is much harder across library types. For example, the missions of public and academic libraries can be rather divergent, as they have different user bases and their users demand and expect different things from them. This significantly complicates any collaboration between academic and public libraries as the two library types may be unlikely to benefit from shared tools and resources given the divergences in their missions and goals. There may even be important differences within a library type, such as academic libraries, that complicate external collaboration. For example, the LLAP team found that four-year colleges and universities were more likely to have a higher interest in learning analytics than were two-year community colleges. Further, four-year institutions were more likely to have access to the computational resources and human capacity for library learning analytics than were their two-year counterparts.

4c. Infrastructure

LLAP entailed working with data from diverse sources at U-M, and at multiple scales including very large logs and datasets. A critical challenge was creating the infrastructure that enabled the LLAP team to collect, clean, structure, merge, and analyze data that were held by different units at the university. For example, resource usage logs were obtained from the library, while data on student academic outcomes were obtained from a research-focused learning analytics dataset that was created by the university's Office of Enrollment Management, and maintained by the university's central IT office at the U-M Data Warehouse. The library usage data were a mix of smaller, structured datasets (e.g., interlibrary loan transactions), and larger, unstructured logs (e.g., proxy server usage). The LLAP team spent roughly six months working on a secure data pipeline that would be used to move the library usage data from a secure repository at the library, to a secure enclave at ISR. A primary reason for taking these measures is that the library usage data were assigned the "Restricted" sensitivity or classification level, which is the highest level for all institutional data at U-M.

The LLAP team also had to secure access to servers powerful enough to run models using some of the very large datasets. For example, some of the logs had nearly a billion records each at the official conclusion of LLAP. Thus, the LLAP PI (Dr. Kabo) had to work closely with ISR's computing staff to ensure that LLAP had access to a cyberinfrastructure that could cope with high data volumes, and the analysis of very large datasets. This was made all the more challenging by the fact that, due to its "Restricted" classification, the analysis of the library usage data had to be done in the secure environment. The LLAP team developed Python scripts for ingesting data into the ISR enclave, cleaning and structuring the data, and then feeding the data to a SQL database that sat in the enclave. This automated the process of retrieving data from the library to the ISR along the secure data pipeline, and significantly accelerated the availability of the research datasets that the LLAP team used for analysis and modeling. The last component of the secure cyberinfrastructure for LLAP was a secure virtual enclave. Coincidentally, this process was completed in early 2020 and only a couple of weeks before the national and university-wide shutdowns due to the COVID-19 pandemic. The secure virtual enclave made it possible for the LLAP team to continue its work even at the height of the pandemic-induced shutdowns.

4d. Policy & Technical Outcomes

Before LLAP, the Library did not have comprehensive policy and technical solutions in place for managing application and system logs. This resulted in *ad hoc* log management and data sharing decisions, an inability to easily perform rich analysis of log data across applications, and lack of compliance with emerging university log management requirements.

From a technical perspective, LLAP motivated the library to implement a technical solution for comprehensive aggregation of the majority of our system and application logs in real time. This solution allowed the library to meet the data sharing terms of the grant and to come into compliance with the new U-M standard policy, [Security Log Collection, Analysis, and Retention](#).

covering data retention. The solution also allowed the library to make use of the archive for technical troubleshooting of application and system problems by our engineers.

From a policy perspective, LLAP motivated the library to think critically about what data it collects, how the data is managed, how the data is shared, and how we communicate to users about library data collection and use. This led to an updated privacy policy for the library, a statement to the university community about library data collection and use, and an ongoing effort to maintain an accurate inventory of library data.

In summary, it is accurate to assert that LLAP supported the library's creation of a data analytics program. A foundation now exists that the library will build on to sustain a mature analytics program that meets both the operational analytics needs of the library and the research needs of those who wish to leverage library data.

5. Looking Forward

Our engagement with learning analytics over the past years is best described as one of iteration, emergence and increased nuance. We started with many more questions than answers. We conclude the grant with much learned, an increased understanding of practices and a deeper appreciation for collaboration. Within the library community, we have seen an increase in the number and types of libraries engaging with learning analytics. Topics include privacy and ethics, technical capabilities, student engagement and consortial collaborations. As we look forward, our work is informed by core values intended to be aspirational while also guiding our day-to-day engagement: excellence, engagement, diversity and anti-racism, interdependence and humanity. We aspire to use analytics with thoughtful engagement to both potential benefits and risks to the communities of users; to work to be aware of these biases with a focus on equity, including through incorporating reflective bias checks as we move through the work; to engage with diverse groups to ensure we are gathering feedback from a diverse range of voices; to be transparent about the data we hold and ways we are and may use it; to involve welcoming expertise from across the Library, campus, and beyond; and to commit to treating all people in an inclusive, respectful, and caring manner.

5a. Perspectives

Throughout our partnership, we worked to build models that characterize the relationship between library use and institutional learning outcomes. As we worked, two perspectives emerged as connected yet sometimes distinct. The library service perspective emphasized application at the front end of processes. It focused on how we can take action on the analytics to understand the users we serve, and how the data patterns inform decisions about outreach, promotion and service delivery. The research perspective also engaged with application at the back end, forefronting emphasis on theoretical and data science. These orientations made for rich discussion, back and forths about models and purposes, and resulted in a richer outcome.

They also indicated a need for an infrastructure strategy that can sustainably support inquiries from both perspectives.

5b. Expanded Partnerships

The Library's work with the LLAP grant has sparked other and different types of partnerships. We have created an internal Learning Analytics Task Force, increased our engagement with students and focused our attention on the need for sustainable and cooperative solutions.

5b1. Learning Analytics Task Force

One outcome of the Library's partnership with the LLAP grant was the recognition by the Library we needed to integrate coherence among various engagements with data and learning analytics. While not part of the LLAP grant, the formation of the Library's Analytics Task Force has a direct link. Our work on the grant elevated the need to reinforce and connect around the principles of transparency, consistency, accountability, and agility. The formation of a Library Analytics Task Force in 2019 was one way to provide a framework to enhance consistency and our ability to develop a shared understanding of our current engagement (e.g. library, campus and national) with library analytics, and articulate strategies to contribute to the strategic objectives focused on data and instruction. Focused engagement needed to be done in the following areas to guide our directions:

- Describing the library's participation in recent IMLS analytics projects, outlining specific deliverables and resource commitments from the Library.
- Building a common understanding and definitions for future projects and research around several large, overlapping areas: learning analytics, library analytics, and assessment.
- Recommending proposals for continued and further engagement with library analytics, informed by our values, inquiry needs, and resources

We are currently piloting a peer engagement process related to library analytics. We want to engage with students and bring their perspectives to the forefront as we use data and analytics to inform and transform our program and service developments. Our commitment is to develop diverse, equitable, inclusive and accessible services that are informed by student voices and experiences. The goal is to increase transparency and student participation in refining library practices around data collection & storage, data analysis and service development. To that end, two undergraduate students are interviewing 16 of their undergraduate peers to explore how students think about the patron data collected by the Library, and how they feel about the Library using patron data in analytics projects. Analysis of these interviews found that students are resigned to the inevitability of data collection but expressed a desire for more and clearer information about the purposes and intent of these processes and desire opt-out mechanisms.¹

¹ Jorgensen, D., Garden, Z., & Smith, C.E. (in press). Big Data at The Big House: What Do University of Michigan Patrons Think About Their Library Data? In H. Staines, & D. Bradley (Eds.), *Navigating Library Privacy: Planning, Policy, and Technology*. Against the Grain (Media), LLC.

The Task Force also recently collaborated with U-M Information Technology Services to create a public-facing web page that shares information about that patron data that the Library collects, stores, and uses. We are excited about this program and view it as a way to further our commitment to our values. <https://safecomputing.umich.edu/viziblue>

5b2. Consortial & Campus

We are focused on being partners in the broader student success engagement. With the LLAP wrapping up, the U-M Library is directing efforts towards furthering campus and/or consortial-based (e.g. Unizen, CLLASS2) analytics solutions that are sustainable and can support both our operational and research analytics needs. An expected outcome of this strategy is to identify a campus data warehouse solution or other suitable platform to try and sustainably meet the learning analytics needs of both the library and researchers.

Appendix A: Initial PAG Survey

See the following pages.



Default Question Block

As the higher education environment changes, with online degree programs, changing demographics, and return on investment concerns, how would you describe the role of the academic research library?

Please use the text boxes below to let us know what you think are the most important issues (*no more than 10*) we need to better understand with respect to academic libraries and research universities.

Issue 1	<input type="text"/>
Issue 2	<input type="text"/>
Issue 3	<input type="text"/>
Issue 4	<input type="text"/>
Issue 5	<input type="text"/>
Issue 6	<input type="text"/>
Issue 7	<input type="text"/>
Issue 8	<input type="text"/>
Issue 9	<input type="text"/>
Issue 10	<input type="text"/>

The library is involved in the success of the university in many areas, including teaching, research, and publishing. Thinking about all the aspects for which users seek the library's help, from the list below please identify up to **10** areas that are most important for **faculty and researchers**. Click on a specific area to highlight it.

- Access to scholarly literature and resources
- Access to archives, primary sources and special collections

- Expertise in academic integrity
- Expertise in copyright and intellectual property
- Support for publishing and open access publishing
- Discipline-specific research consultations
- Expertise in academic, research, and learning technology (including digital scholarship tools)
- Institutional repository
- Research data repository (comply with funding agency mandates, provide access to research data)
- Data management planning (support for research data management in general)
- Information literacy expertise (teaching research skills and concepts)
- Campus engagement programming (events, lectures, workshops)
- Numeric and spatial data; data visualization
- Partnerships on research projects
- Partnerships in course development/teaching
- Space for research and scholarship to take place, especially interdisciplinary work, both quiet study space and space for group discussion and collaboration
- Central hub for cross-disciplinary sharing of best practices among faculty
- Preservation of knowledge in all forms

The library is involved in the success of the university in many areas, including teaching, research, and publishing. Thinking about all the aspects for which users seek the library's help, from the list below please identify up to **10** areas that are most important for **research and technical staff**. Click on a specific area to highlight it.

- Access to scholarly literature and resources
- Access to archives, primary sources and special collections
- Expertise in academic integrity
- Expertise in copyright and intellectual property
- Support for publishing and open access publishing
- Discipline-specific research consultations
- Expertise in academic, research, and learning technology (including digital scholarship tools)
- Institutional repository
- Research data repository (comply with funding agency mandates, provide access to research data)
- Data management planning (support for research data management in general)
- Information literacy expertise (teaching research skills and concepts)
- Campus engagement programming (events, lectures, workshops)
- Numeric and spatial data; data visualization
- Partnerships on research projects
- Partnerships in course development/teaching
- Space for research and scholarship to take place, especially interdisciplinary work, both quiet study space and space for group discussion and collaboration

- Central hub for cross-disciplinary sharing of best practices among faculty
- Preservation of knowledge in all forms

The library is involved in the success of the university in many areas, including teaching, research, and publishing. Thinking about all the aspects for which users seek the library's help, from the list below please identify up to **10** areas that are most important for **graduate students and postdocs/fellows**. Click on a specific area to highlight it.

- Access to scholarly literature and resources
- Access to archives, primary sources and special collections
- Expertise in academic integrity
- Expertise in copyright and intellectual property
- Support for publishing and open access publishing
- Discipline-specific research consultations
- Expertise in academic, research, and learning technology (including digital scholarship tools)
- Institutional repository
- Research data repository (comply with funding agency mandates, provide access to research data)
- Data management planning (support for research data management in general)
- Information literacy expertise (teaching research skills and concepts)
- Campus engagement programming (events, lectures, workshops)
- Numeric and spatial data; data visualization
- Partnerships on research projects
- Partnerships in course development/teaching
- Space for research and scholarship to take place, especially interdisciplinary work, both quiet study space and space for group discussion and collaboration
- Central hub for cross-disciplinary sharing of best practices among faculty
- Preservation of knowledge in all forms

Thinking about **faculty and researchers**, how important are the following areas with respect to the library's direct contributions to the success of a university?

Access to scholarly literature and resources

Access to archives, primary sources and special collections

Expertise in academic integrity

Expertise in copyright and intellectual property

Support for publishing and open access publishing

Discipline-specific research consultations

Expertise in academic, research, and learning technology (including digital scholarship tools)

Institutional repository

Research data repository (comply with funding agency mandates, provide access to research data)

Data management planning (support for research data management in general)

Information literacy expertise (teaching research skills and concepts)

Campus engagement programming (events, lectures, workshops)

Numeric and spatial data; data visualization

Partnerships on research projects

Partnerships in course development/teaching

Space for research and scholarship to take place, especially interdisciplinary work, both quiet study space and space for group discussion and collaboration

Central hub for cross-disciplinary sharing of best practices among faculty

Preservation of knowledge in all forms

Thinking about **research and technical staff**, how important are the following areas with respect to the library's direct contributions to the success of a university?

Access to scholarly literature and resources

Access to archives, primary sources and special collections

Expertise in academic integrity

Expertise in copyright and intellectual property

Support for publishing and open access publishing

Discipline-specific research consultations

Expertise in academic, research, and learning technology (including digital scholarship tools)

Institutional repository

Research data repository (comply with funding agency mandates, provide access to research data)

Data management planning (support for research data management in general)

Information literacy expertise (teaching research skills and concepts)

Campus engagement programming (events, lectures, workshops)

Numeric and spatial data; data visualization

Partnerships on research projects

Partnerships in course development/teaching

Space for research and scholarship to take place, especially interdisciplinary work, both quiet study space and space for group discussion and collaboration

Central hub for cross-disciplinary sharing of best practices among faculty

Preservation of knowledge in all forms

Thinking about **graduate students and postdocs/fellows**, how important are the following areas with respect to the library's direct contributions to the success of a university?

Access to scholarly literature and resources

Access to archives, primary sources and special collections

Expertise in academic integrity

Expertise in copyright and intellectual property

Support for publishing and open access publishing

Discipline-specific research consultations

Expertise in academic, research, and learning technology (including digital scholarship tools)

Institutional repository

Research data repository (comply with funding agency mandates, provide access to research data)

Data management planning (support for research data management in general)

Information literacy expertise (teaching research skills and concepts)

Campus engagement programming (events, lectures, workshops)

Numeric and spatial data; data visualization

Partnerships on research projects

Partnerships in course development/teaching

Space for research and scholarship to take place, especially interdisciplinary work, both quiet study space and space for group discussion and collaboration

Central hub for cross-disciplinary sharing of best practices among faculty

Preservation of knowledge in all forms

How important is the academic library in assisting faculty and researchers from the following

disciplines in finding funding opportunities?

	Very important	Important	Neutral	Unimportant	Very unimportant
Humanities	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Social sciences	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Natural sciences	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Biomedical/health sciences	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Engineering	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

In what parts of the scholarly research lifecycle do you think that academic libraries have the most impact?

Powered by Qualtrics

Appendix B: PAG Pre-Workshop Survey

See the following pages.



Default Question Block

Please enter your name in the space below using the "Firstname Lastname" format e.g. "Jane Doe."

What is your institutional affiliation (e.g. "University of Michigan")?

What is your prior statistical and programming experience?

- SQL
- R
- Python
- Matlab
- Stata
- SAS
- SPSS
- Other

Please indicate below your other statistical and programming experience. Separate multiple entries using commas e.g. C++, Java, Julia, etc.

We would like to learn more about your proficiency in the languages below. How much would you agree with the following statements:

	I feel that I have received adequate formal and/or informal training in this language					I have a strong desire to learn more about this language					I think that I will be able to use this language on a regular basis or in my performance of library analytics					I can easily find someone in my organization that can help me with this language should the need arise				
	Strongly agree	Agree	Neutral	Disagree	Strongly Disagree	Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
SQL	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	I feel that I have received adequate formal and/or informal training in this language					I have a strong desire to learn more about this language					I think that I will be able to use this language on a regular basis or in my performance of library analytics					I can easily find someone in my organization that can help me with this language should the need arise				
	Strongly agree	Agree	Neutral	Disagree	Strongly Disagree	Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
R	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Python	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Powered by Qualtrics