



LAURA BUSH 21st CENTURY LIBRARIAN PROGRAM

Sample Application RE-258986-OLS-25
Project Category: Planning

University of Alabama

Amount awarded by IMLS: \$149,938

The University of Oklahoma received this grant which was transferred to the University of Alabama in December 2025. The University of Alabama will conduct a comprehensive assessment of community needs related to advancing text analytics data services in academic libraries. The project will investigate the current state of library services relevant to text analytics and identify best practices by surveying and interviewing stakeholders, including librarians, library users, and LIS educators. The research team will also establish and solidify partnerships with data librarians. This work will lay the foundation for a subsequent project focused on developing text analytics service models for academic libraries and creating tailored training resources for data librarians.

Attached are the following components excerpted from the original application.

- Narrative
- Schedule of completion

When preparing an application for the next deadline, be sure to follow the instructions in the most recent Notice of Funding Opportunity for the grant program to which you are applying.

A Comprehensive Community Needs Assessment for Advancing Text Analytics Data Services in Academic Libraries

INTRODUCTION

The team from the University of Oklahoma and the University of Kentucky is requesting \$149,938 for a two-year **planning grant** aiming to conduct a comprehensive assessment of community needs related to advancing text analytics data services in academic libraries. This project aims to gather substantial evidence and insights to support a subsequent, larger project focused on developing text analytics service models for academic libraries and creating tailored training resources for data librarians. This funding will enable a thorough survey of the current state of text analytics services in academic libraries, and an in-depth evaluation of the needs of various stakeholders from different types of institutions, including library users, librarians, and LIS (Library and Information Studies) educators. The findings from this project will guide the follow-up effort aiming at developing text analytics service models and customized training materials for academic librarians involved in data services. The project is aligned with **Program Goal 2** of the LB21 program, specifically **Objective 2.1**, which focuses on supporting the training and professional development of current data librarians. The team has assembled an advisory board of four experienced data librarians, each with expertise in text analytics services in academic libraries. Throughout this planning phase, the research team will continue to establish and solidify partnerships with data librarians, laying the foundation for the follow-up project. Ultimately, the results of this investigation will serve as the evidence base for designing innovative data service models that incorporate text analytics and for developing practical training tools for librarians.

PROJECT JUSTIFICATION

Strategic Timing for Text Analytics

This is a pivotal moment for the broad adoption of text analytics across disciplines, driven by the increasing availability of digital text, lowered technical barriers, and the growing capabilities of text analytic models. Writing has long served as the primary medium for recording human knowledge, resulting in a vast repository of information stored in textual form. However, the unstructured nature of textual data presents challenges for researchers and educators seeking to analyze it effectively on a large scale. Text analytics uses computational methods to analyze a large amount of text to derive meaningful insights, patterns, and trends. Common techniques include natural language processing (NLP), machine learning, and statistical analysis. Text analytics has great potential for applications across domains, enabling researchers to extract relevant information, model behavior and policy, analyze writing styles and sentiment, annotate key concepts etc. [1-5]. Historically, the widespread adoption of text analytics has been impeded by constrained access to digital text, technical complexities, and the capabilities of text analytic models, leaving a wealth of textual information untapped. Currently, several developing factors in society are catalyzing the increasing use of text analytics as a method to tap into textual information across disciplines. First, an increasing amount of digital text is becoming available thanks to the growing use of digital technologies, the advance of the open access movement, and growing support from major publishers on text mining [6-8]. In addition, the advancement of generative AI offers assistance in computer programming, which lowers the technical barrier to text analytic projects. Since the advent of ChatGPT in late 2022, interest in Large Language Models (LLMs), the latest breakthrough in text analytics, has surged across disciplines due to their remarkable ability to process and understand human languages. This trend is evident in the growing number of publications exploring LLM applications in many fields [e.g. 15-17]. For example, in materials science, text analytics is used to automatically extract materials' composition, structure, and property from scientific literature, aiding the development of machine learning models to unveil their relationship [18]. In medicine, LLMs are used to analyze clinical notes to provide clinical support, such as matching patients with relevant clinical trials [19]. Humanities scholars use text analytics for storytelling and culture analytics [20], while psychologists apply it to measure cognition [21], affection [22], and child development [23]. As these advancements continue, text analytics is poised to become an increasingly valuable tool across disciplines, unlocking new insights and transforming research methodologies.

Gaps Exist in Offering Text Analytics Services

Academic libraries, positioned at the center of their institutions, have a strategic role in supporting text analytics across disciplines. As data services and digital scholarship initiatives continue to expand, some libraries have begun to offer text-analytics-related services. However, these efforts often remain limited to providing reference materials through LibGuides or offering consultations and guidance while lacking direct, hands-on involvement. Direct participation of libraries in text analytics projects is uncommon, although successful examples exist [24, 26]. For example, librarians at Marquette University Libraries teamed up with psychology researchers to develop a customized text analytics method to understand in what context social class, socioeconomic status, and associated terminology were examined in counseling publications [26]. The University of Florida Libraries hired an NLP specialist to collaborate with researchers on relevant

projects and teach introductory AI courses and workshops to students [28]. Yale University Library has a “Robots Reading Vogue” project (<http://dh.library.yale.edu/projects/vogue/>) to use text analytics to analyze the fashion magazine Vogue. The Digital Scholarship Lab at the University of Richmond used text analytics to mine the text in a digital archive called “Daily Dispatch” (<https://dsl.richmond.edu/dispatch/>). Despite the growing use of text analytics across disciplines, librarians are not yet commonly integrated into these research teams. This not only results in missed opportunities for academic libraries to better serve their users but also hinders broader participation in text analytics projects, particularly among non-technical disciplines. There is a growing and urgent need for academic libraries to take action in this area, as evidenced by recent publications, professional meetings, job postings, efforts to upskill librarians, and discussions in online communities [9-12]. However, there is a lack of guidelines and reference models for offering text analytics services in academic libraries [26]. To effectively meet users’ needs in text analytics, academic libraries must address three key gaps:

Gap 1: Lack of Comprehensive Understanding of the Current State

Academic libraries are increasingly expanding their data services and digital scholarship initiatives, with some also introducing text analytics services. However, these services vary widely among institutions. While some libraries employ dedicated NLP specialists for text analytics support, others focus solely on curating reference materials and offering project guidance. Despite its growing importance, research on the current state of text analytics data services in academic libraries remains limited.

Gap 2: Lack of In-depth Understanding of Different Stakeholders

Library users, librarians, and LIS educators are key stakeholders in text analytics services in academic libraries. However, their perspectives remain largely unexplored.

- **Gap 2.1: Lack of Comprehensive Understanding of Library Users’ Needs in Text Analytics**

While anecdotal evidence exists, a comprehensive understanding of academic library users’ text analytics needs, across various backgrounds and institution types, remains lacking. There is a dearth of literature on this topic, particularly in the context of the evolving landscape of data availability and generative AI. A thorough assessment of library users’ needs in text analytics is urgently needed to bridge this gap and inform future library services.

- **Gap 2.2: Lack of Comprehensive Understanding of Librarians’ Needs**

A significant barrier to offering these advanced services is the lack of expertise among librarians [13-14], particularly in technical skills [25]. While some ALA-accredited programs have developed text mining courses, they are often limited to elective, advanced-level offerings. Consequently, only a small number of graduates have received this training, while many current academic librarians have limited opportunities for relevant professional development in this area. While online tutorials for text analytics skills exist, they are often not tailored to the specific needs of academic librarians and do not align with the contexts of library services [26]. Library Carpentry [27] offers lessons and workshops on general computational and data skills needed for library-related roles, but they have yet to cover text analytics specifically. As a result, there is an unmet need among academic librarians for ongoing professional development to advance text analytics services, particularly given the increasing demand across academic fields.

- **Gap 2.3: Lack of Understanding in the Skills Required and Pedagogical Approaches**

Another gap in developing customized text analytics training resources for academic librarians is a lack of understanding on what curriculum design is most relevant to academic librarians and what pedagogical approaches are most effective. LIS educators have abundant experience in training future librarians. Input from LIS educators will be needed to fill in this gap.

Gap 3: Lack of Reference Models for Text Analytics Services in Academic Libraries

Lastly, there is a significant lack of established reference models that academic libraries can adopt to develop and expand their text analytics services. Although there has been a recent trend toward hiring relevant experts to support these services, most libraries currently provide limited support. These efforts are often confined to offering text analytics resources through LibGuides or providing minimal consultation, with little direct involvement in text analytics projects. While many libraries are beginning to recognize the importance and potential of text analytics services, few frameworks exist to guide key aspects of service development. These include determining the types of textual data to support, identifying appropriate analysis methods and tools, defining the extent of librarian involvement, allocating necessary resources, and fostering collaborations with other campus units and departments. This gap illustrates the urgent need for well-defined frameworks and practical models that academic libraries can adopt to build and sustain robust text analytics services.

Call to Action: Community Needs Assessment to Fill the Gaps

A comprehensive community needs assessment is required to gain a clear understanding of the perspectives of library users, librarians, and LIS educators on text analytics data services in academic libraries, particularly in light of the evolving landscape in this field. This project aims to begin to bridge these gaps by gathering evidence from

library users, academic librarians, and LIS educators to develop a comprehensive understanding of the need for text analytics data services in academic libraries. It also seeks to inform the creation of tailored training resources and pedagogical strategies for academic librarians in delivering these services. This work will lay the foundation for a future project focused on developing text analytics service models and specialized training resources for academic libraries.

The following table summarizes identified gaps and call to action to bridge the gaps to advance text analytics data services in academic libraries:

Table 1: Identified Gaps and Call to Action to Advance Text Analytics Services in Academic Libraries.

Lacking Areas	Description	Call to Action
GAP 1 Understanding of the Current State	- Academic libraries are expanding data services, including text analytics, but implementation varies widely. - A lack of comprehensive understanding hinders the strategic deployment of relevant services.	- Conduct a nationwide investigation of current text analytics services in academic libraries. - Document best practices and challenges.
GAP 2-1 Understanding of Library Users' Needs	- Limited understanding of user needs for text analytics across disciplines and institution types makes targeted support challenging. - Existing literature is scarce, especially amid evolving AI and data trends. - This gap hinders the development of services that align with user demands, limiting user-centered library support.	- Conduct comprehensive needs assessment with potential users, including faculty, researchers, and students. - Develop a report to inform future user-centered library services.
GAP 2-2 Understanding of Librarians' Needs	- Lack of expertise among librarians, particularly in technical skills. - MLIS coursework on text mining is often limited to elective, advanced-level offerings. - Existing online tutorials for text analytics are not customized to librarians' needs or library services. Many current academic librarians lack opportunities for professional development in text analytics.	- Develop specialized training materials tailored to librarians' roles and needs. - Expand professional development opportunities for librarians through workshops, webinars, and certifications. - Encourage LIS programs to integrate text analytics into curricula.
GAP 2-3 Understanding of the Skills Required and Pedagogical Approaches	- Lack of understanding of the most relevant curriculum design for academic librarians in text analytics. - Effective pedagogical approaches for teaching librarians text analytics remain unclear. - LIS educators, who have experience in training future librarians, need to be involved in addressing this gap.	- Develop open-access teaching materials to support scalable training. - Collaborate with LIS educators to design curriculum guidelines and pedagogical frameworks. - Conduct research on best practices in text analytics instruction for librarians.
GAP 3 Reference Models for Text Analytics Services	- Few established reference frameworks exist for text analytics services in academic libraries. - Current support is largely limited to LibGuides and basic consultations. - Clear guidance is needed for service development, resource allocation, and collaboration strategies.	- Develop adaptable reference models for different institution types. - Define core service components, technical tools, and collaboration strategies.

Target Group

The project will gather comprehensive evidence for developing text analytics service models in academic libraries and creating training resources for academic librarians. The groups that will be most immediately and positively affected by the project are academic librarians. Academic librarians will benefit from a more comprehensive understanding of library users' needs in text analytics and identify the gaps in current library services. The findings from this project will provide a comprehensive picture of the need for text analytics data services in academic libraries. The ACRL (Association of College and Research Libraries) 2023 survey reported an average of 36.2 full-time equivalent (FTE) library staff from 1,414 responding institutions out of 3,125 institutions in the sample [29]. This indicates that there may be up to 113,245 FTE library staff in academic libraries, although this is an upper estimate due to many non-responding institutions are smaller institutions. The other group that will be positively impacted is the LIS educators. LIS educators can use this information to help design or update their courses to prepare future data librarians. U.S. Bureau of Labor Statistics estimates that 4,220 LIS educators were employed in the U.S. in 2023 [30].

Beneficiaries

Library users will be the ultimate beneficiaries of this project as the findings from this project will support a future project to develop text analytics service models and customized training resources for academic libraries. Once implemented, library users will benefit from these services to support their research and education. Certainly, academic librarians will benefit from more tailored and updated training materials for offering text analytics services. This offers opportunities for professional development to catch up with trends in society. Academic libraries will also benefit from offering these services to support their missions, strengthening their positions in their institutions. LIS educators can also consider using or adapting these training resources in their courses.

Alignment with Project Goal and Objective of the LB21 Program

The project is most aligned with Project Goal 2 of the LB21 program on “Train and retain current library and archives professionals, faculty, and staff,” objective 2.1, focusing on supporting the training and professional development of current data librarians. The project is both timely and urgent in the evolving landscape of text analytics. Academic libraries have a strategic opportunity to become involved in the increasing adoption of text analytics across disciplines. This project aims to remove the barriers for academic libraries to be involved in this development by assessing the community needs to plan for a future, larger project to develop service models and training resources in text analytics.

PROJECT WORK PLAN

Project Personnel

Kun Lu (Principal Investigator), Associate Professor at the School of Library and Information Studies, University of Oklahoma, brings extensive expertise in applied NLP, text mining, and information retrieval. With many years of experience teaching text mining courses in an ALA accredited MLIS program, Dr. Lu has a strong background in both research and pedagogy in this field. **Soohyung Joo** (co-PI), Associate Professor at the School of Information Science, University of Kentucky, also specialized in applied text mining and currently teaches an applied text analytics course for library science students. Both investigators have a record of numerous scholarly contributions, including journal articles and conference presentations on NLP and text analytics. This proposal aligns with their shared interest in advancing text analytics services in academic libraries. In addition, the project team will hire three hourly student assistants to support data collection and content analysis.

The advisory board includes four external experts who will provide valuable insights and guidance to support the project's goals. **Tara Baillargeon** (Ed.D., MLIS), Dean of Raynor Library at Marquette University, has conducted research on digital scholarship. **Borui Zhang** (Ph.D. in Linguistics), an NLP specialist at George A. Smathers Libraries, University of Florida, focuses on making text analytics and AI resources accessible across disciplines in her libraries. **Kara Handren** (Master of Information), a Data Librarian at the University of Toronto Libraries, specializes in metadata and data librarianship. **Mark Lauferweiler** (Ph.D. in Meteorology), a Research Data Specialist at the University of Oklahoma Libraries, has extensive experience in data analytics and served as the Head of Data Analysis, Visualization, Informatics Syndicate. The advisory board members will contribute to the project in several ways. They will participate in two regular meetings each year, provide feedback on survey design and research instruments, offer insights to help interpret findings, facilitate connections with library communities and professionals in the field, and assist with the distribution of project results.

Project Tasks

To bridge the gaps and prepare for the follow-up project, several tasks will be completed during the project period. Detailed activities for each task, along with the specific sequence of activities, are outlined in the Schedule of Completion document. The PIs conducted a thorough assessment of the feasibility of each task, leveraging their prior experience,

preliminary questions for each data collection method, literature reviews, checking with IRB processes, and careful estimation of time for sub-tasks and individual activities. For each specific activity, we calculated the estimated time and workload, securing student assistant hours and the PIs' workload to ensure all tasks are achievable within the scheduled timeframe. Special attention was given to managing multiple simultaneous surveys, with a systematic distribution of workload to maintain manageability.

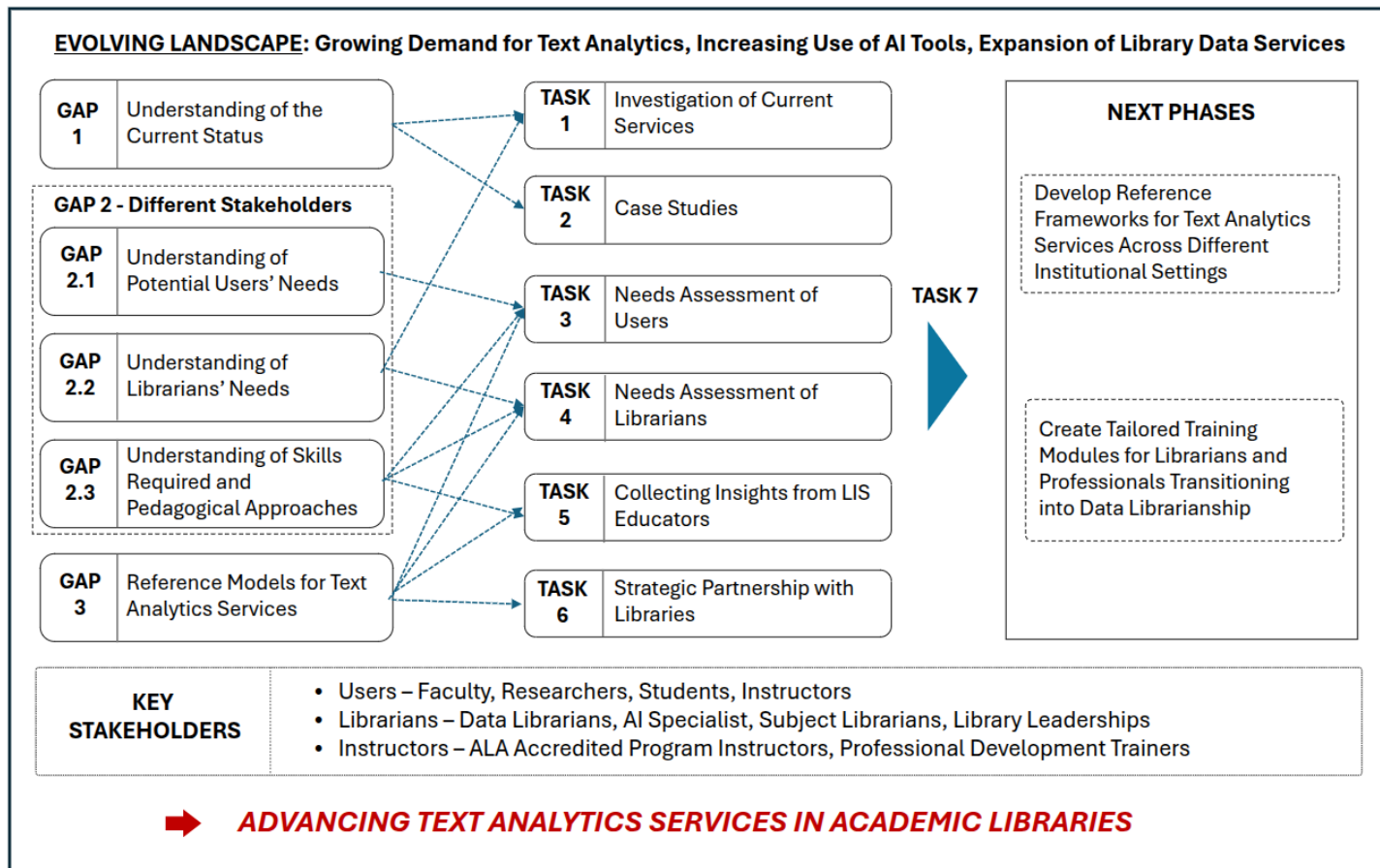


Figure 1: Addressing Gaps Through Project Tasks.

Task 1: Investigation of Current Services

To address **Gap 1**, the project team will start out with a content analysis of existing library services related to text analytics, focusing particularly on LibGuides. This investigation will involve visiting approximately 250 academic library websites representing a broad range of institution types, including R1 and R2 universities, liberal arts colleges, and Historically Black Colleges and Universities (HBCUs), across various regions of the United States. Libraries will be selected from five geographic regions (Northeast, Midwest, South, West, and Southwest), with approximately 50 libraries sampled from each region. The sample will include approximately 80 R1 institutions, approximately 80 R2 institutions, approximately 80 liberal arts colleges, and approximately 10 HBCUs.

Our initial exploration indicates advanced text analytics services are primarily centered in top R1 institutions. Given that many smaller institutions offer few or no data services related to text analytics, the proposed sample size (250) is considered manageable. To assess the feasibility of this investigation, the PIs conducted a preliminary review of 33 institutions, including 10 R1, 10 R2, 10 liberal arts colleges, and 3 HBCUs. The results indicated that 15 out of these 33 libraries provide at least a LibGuide or resource list on text analytics. Among the R1 libraries, 9 out of 10 offered text analytics services, with some providing extensive services supported by dedicated professionals. However, among R2 and liberal arts colleges, only 6 out of 20 offered even minimal resources, typically in the form of LibGuides. Given that many R2 institutions, liberal arts colleges, and HBCUs do not offer dedicated text analytics services in their libraries, we determined that investigating 250 libraries is feasible based on the estimated time required to assess each library.

For the content analysis, the project team will develop a detailed template to examine the following aspects: (a) whether relevant text analytics services are provided, (b) the types of services offered (e.g., data sources & licensing, data

acquisition, data cleaning, data analysis, use of AI and LLMs, etc.) (c) the nature of the services (e.g., resource listing, consultation, direct involvement, etc.), (d) the number and titles of librarians or professionals involved, (e) the name of the department or unit providing the services, (f) tools or software made available, and (g) any other pertinent information. Using this template, student assistants will review the selected library websites, explore available text analytics services, and complete the template for each institution. To ensure accuracy and consistency, student assistants will undergo thorough training with the PIs. As part of the training, they will analyze approximately 10% of the sample under the direct supervision and guidance from the PIs.

The collected data will be analyzed using both quantitative and qualitative methods. First, the project team will explore basic statistical measures such as the percentage of libraries providing relevant services, counts of different types of services, the average number of personnel involved, and others. Second, the project team will examine the specific nature and characteristics of the services offered, as well as the resources and tools utilized. In addition, comparative analyses will be conducted to explore differences among institution types (e.g., R1 and R2 universities, liberal arts colleges, and HBCUs) using Chi-Square tests and ANOVA (Analysis of Variance). Moreover, relationships between service characteristics and factors such as library size and institution size will be analyzed through correlation and regression analyses. These analyses will provide a comprehensive understanding of the current state of text analytics services in academic libraries and the factors influencing their provision, such as institution type and library size.

Estimated Time Needed for Task 1: Sample selection (1 month), Student assistant training (1 month), Data collection (4 months), Data analysis (4 months).

Task 2: Case Studies

Task 2 builds on Task 1 to address **Gap 1** by conducting a supplementary analysis of approximately 10 institutions with unique approaches or well-defined service structures in text analytics. The project team will select a small subset of cases from Task 1 for an in-depth examination to identify best practices that can serve as benchmarks for the academic library community. The selection process will involve a collaborative effort between the investigators and student assistants, who will meet to identify institutions that stand out based on the data collected in Task 1. These selected cases will be analyzed in-depth to uncover insights and strategies that can inform text analytics services in academic library communities, ultimately serving as benchmark models for the community. The case study results will be widely distributed across academic library communities via multiple channels such as professional journal articles and conference presentations.

Estimated Time Needed for Task 2: Case selection (1 month), In-depth analysis (1 month).

Task 3: Needs Assessment of Users

To address **Gap 2-1**, Task 3 focuses on the needs assessment of library users. It will start after student assistants have been trained for Task 1 and have started collecting data on current services. An online survey will be administered to potential user groups, such as faculty, researchers, and students from selected institutions. The survey aims to investigate the needs of potential users on text analytics services offered by academic libraries, focusing on identifying the situations where text analytics would be useful, and the types of support researchers and educators require for their work.

Reaching out to potential library users requires a significant amount of manual work. To balance the representation from various categories and the practical constraints, the project team will select seven institutions including two R1 universities, two R2 universities, two liberal arts colleges, and one HBCU. The selected institutions will be chosen to reflect different geographic regions across the United States. To gather contact information, student assistants will manually visit the college and department websites of each selected institution to collect publicly available email addresses of faculty, researchers, scientists, and students. The co-PI has prior experience with this method, having successfully collected a substantial number of email addresses in a previous study targeting potential library users [31]. While it may not be possible to gather every potential user's email address, this approach is the most practical and effective way to reach out to potential users using publicly available information. The project team acknowledges the limitations of this method, including its inability to capture the entirety of the target population. However, this remains the only viable and realistic option to connect with potential users and investigate their needs regarding library services.

The survey will examine several aspects of potential users' needs related to text analytics services. Specifically, it will investigate whether respondents have textual data for analysis, potentially require text analytics, the scenarios in which they find text analytics useful, the types of analysis they need, the challenges they face, and their awareness and perceptions of library-provided text analytics services, and others. The survey will include logic to identify respondents with potential needs for text analytics services, prompting them with more in-depth questions about their experiences with text analysis tasks, specific needs, and expectations for library-provided services. For respondents without such needs, the survey will explore their reasons for their non-use. This design will ensure a more targeted understanding of potential users. Data collected will be analyzed primarily using quantitative methods, supplemented by qualitative analysis for open-ended responses. The analysis will focus on comparing various needs and expectations across different user groups (e.g., faculty,

graduate students, etc.), academic disciplines (e.g., humanities, social sciences, sciences, engineering, medical sciences, etc.), and institution types (e.g., R1, R2, liberal arts colleges, etc.). Statistical methods such as Chi-square analysis, t-tests, and ANOVA will be primarily employed to examine differences in needs between research-oriented and education-oriented tasks, as well as variations among user groups and institution types. The findings will provide insights into the types of services users need, the extent of those needs, their willingness to use library-provided text analytics services if offered, the relevant challenges, and the factors associated with non-use.

Estimated Time Needed for Task 3: IRB (2 months), Survey instrument development (2 months), Email address collection (3 months), Survey distribution (2 months), Data analysis (2 months).

Task 4: Needs Assessment of Librarians

To address **Gaps 1, 2-2, and 3**, Task 4 focuses on assessing the needs of librarians. It will begin in parallel with Task 3. A nationwide online survey will target librarians involved in data services, aiming to examine areas where text analytics can be applied, the training needs of librarians, and strategies for integrating and implementing text-analytics-related services. The survey instrument will include questions to collect the types of services offered, resources and tools utilized, and the challenges librarians face in providing text analytics services. More importantly, the survey will also assess librarians' specific needs for training resources and professional development opportunities to enhance their skills in this area. Moreover, it will investigate the factors that influence librarians' involvement in text analytics initiatives and solicit their suggestions for advancing these services.

The project team will construct a database of email addresses for relevant librarians and professionals identified during Task 1. While student assistants conduct website content analysis, they will also gather publicly accessible email addresses for the survey. Where feasible, the research team will also utilize email discussion lists (listservs), such as those provided by ACRL, Code4Lib, and JESSE, to distribute survey invitations. The advisory board will help identify further channels for distribution. The collected data will be analyzed using both quantitative and qualitative methods, focusing on three key areas: (a) librarians' needs for training resources and opportunities to enhance their skills, (b) factors that motivate or hinder librarians' involvement in text analytics services, and (c) their suggestions for improving and advancing these services across various types of institutions. Quantitative analysis will include descriptive statistics to assess the current service landscape and identify key opportunities and challenges from librarians' perspectives. Inferential statistical methods, particularly ANOVA and regression analysis, will be used to explore differences in needs across institution types and examine factors influencing librarians' engagement in text analytics services. Qualitative analysis will focus on open-ended survey responses, providing deeper insights and recommendations for enhancing and effectively implementing text analytics services.

To supplement the survey, we will conduct in-depth, semi-structured interviews with approximately 5 to 10 librarians selected from the survey respondents. The PIs will administer these online interviews, and the collected data will be transcribed for content analysis. These interviews will provide deeper insights into the underlying reasons and motivations behind their survey responses, as well as gather their perspectives on improving text analytics services and shaping future directions. The qualitative findings will complement the survey results, providing a richer understanding of librarians' needs and suggestions. Furthermore, these interviews will serve as an opportunity to build professional connections with librarians, aligning with one of the project's key goals: fostering future collaborations on service models and training resources.

Estimated Time Needed for Task 4: IRB (2 months), Survey instrument development (2 months), Survey distribution (1 month), Survey data analysis (2 months), Interview (2 months), Interview data analysis (2 months).

Task 5: Input from LIS instructors

To address **Gap 2-3**, Task 5 focuses on gathering input from LIS instructors. It will begin after the completion of Tasks 3 and 4, as their findings will inform the discussion.

This project will examine the curriculum and pedagogical approaches for text analytics courses in LIS programs. An online focus group will be conducted to gather information on several key areas, including (a) current topics and courses related to text analytics, (b) the tools and resources used for instruction, (c) how library-specific contexts are incorporated into these courses, (d) instructors' perspectives on students' needs, and (e) the challenges faced in teaching text analytics for library science students. More importantly, the focus group will seek suggestions for improving teaching pedagogies in this area in LIS programs. To identify potential participants, student assistants will visit around 68 ALA-accredited programs and compile a list of instructors who teach text-mining-related courses. The decision to conduct a focus group rather than a survey was based on our preliminary research, which showed that only a small number of ALA-accredited programs offer text mining courses. With a limited pool of potential participants, a survey might not generate enough responses to be useful. A focus group, on the other hand, offers the advantage of collecting more in-depth input, particularly through open discussions that can yield detailed suggestions for course development strategies. Given the relatively small sample size,

approximately five instructors will be invited to participate. The focus group will follow a semi-structured format, with the first half dedicated to addressing specific questions about the topics identified above and the second half allowing for a more open discussion. This open discussion will focus on strategies for creating courses that align closely with library-specific contexts.

Beyond gathering insights, this focus group will also serve as an opportunity to build professional connections with LIS instructors. These connections could be valuable for future projects, particularly in developing training resources and continuing to improve pedagogy in this evolving field.

Estimated Time Needed for Task 5: IRB (2 months), Compiling a list of instructors (1 month), Contacting and inviting participants (1 month), Conducting the focus group (1 month), Data analysis (2 months)

Task 6: Partnerships with libraries

The project team will establish partnerships with library professionals to support the next phase of the project, which will focus on developing text analytics service models and training modules tailored for librarians. To achieve this, the team will utilize multiple engagement strategies, including attending relevant professional conferences and meetings, as well as directly reaching out to librarians involved in text analytics services. To generate interest among librarians, the project team will present preliminary findings from earlier tasks at professional conferences such as the ACRL Conference, the Research Data Alliance (RDA) Meeting, and the Code4Lib Conference. In addition, the team will use its project blog to share progress updates and preliminary findings, expanding its outreach to academic library communities and connecting with professionals interested in text analytics services in academic libraries. Moreover, during planned interviews (Task 4), participants will be asked if they would be open to maintaining connections and providing input on the follow-up initiatives. These efforts will help raise awareness and build relationships with potential partners, which will be essential for developing the services and resources in the next stage of the project. Task 6 will be throughout the project period, with year 1 focusing on identifying partners and year 2 focusing on developing and solidifying partnerships.

Estimated Time Needed for Task 6: Throughout the project.

Task 7: Planning next stages

Building on the evidence collected from Task 1 to Task 6, the project team will develop a plan for integrating text analytics into library services and creating relevant training resources for librarians, tailored to the specific needs of these services. The follow-up project will focus on the practical implications of the findings from this planning project, particularly in developing comprehensive reference models for text analytics services in academic libraries and designing targeted training resources for librarians.

The reference models for text analytics services will be proposed for institutions of different sizes and types, taking into account their resources and user needs, as these factors vary based on institutional purpose and scale. These models will encompass (a) specific services and their alignment with user needs, (b) resources and tools associated with each service type, (c) optimized allocation of library resources, (d) guidelines for effective communication with users, (e) strategies for promoting and marketing these services, (f) collaboration approaches with other campus units, and (g) professional development and upskilling opportunities for existing staff and librarians. Considerations will be incorporated to ensure adaptability across various library settings.

More importantly, the follow-up project will produce practical training resources designed for (a) current data service librarians, (b) liaison librarians seeking to upskill their expertise, (c) LIS program students, and (d) professionals transitioning into library text analytics services. These training materials will be structured into modular levels, customized for different library service contexts and professional needs. The development of these resources will be informed by the findings of this project to address the needs of both librarians and users. Moreover, insights and recommendations from LIS educators will be incorporated into the training design to ensure pedagogical effectiveness.

Ultimately, the follow-up project will contribute to the adoption and advancement of text analytics in academic libraries, expanding the territory of library services and reinforcing their role in academia and knowledge creation. By supporting librarians with well-structured reference models and comprehensive training resources, this initiative will help position libraries as key contributors to data-driven research and scholarly innovation. Task 7 will be carried out during the final four months of this project.

Estimated Time Needed for Task 7: Final four months of the project.

Table 2: Project Design - Tasks, Objectives, Methods, Outcomes and Preparations for Next Phases.

Tasks	Objectives	Methods	Outcomes/ Deliverables	Connection to Next Project
Task 1: Investigation of Current Services	Assess the current state of text analytics services in academic libraries.	Content analysis of ~250 library websites; Literature review.	Report on existing services, trends, and gaps.	Identified gaps will inform new service models and training programs.
Task 2: Case Studies	Identify best practices and innovative text analytics services.	In-depth analysis of selected libraries.	Benchmarking report of exemplar cases.	Best practices will guide the development of scalable service frameworks.
Task 3: Needs Assessment of Users	Identify user needs for library support in text analytics.	Online survey of faculty, researchers, and students.	Data on user needs and preferred services.	Findings will shape user-centered service models that fully reflect specific user needs.
Task 4: Needs Assessment of Librarians	Identify existing and potential text analytics services. Assess librarians' current skills and training needs.	Nationwide survey of librarians; Semi-structured interviews.	Report on current and potential service areas. Inventory of required skills and knowledge.	Insights will guide targeted training programs and strategic service deployment.
Task 5 Input from LIS Educators	Determine knowledge and skills required for text analytics services. Identify pedagogical approaches for librarian training.	Focus groups	Detailed list of subjects and topics for training resources. Pedagogical recommendations.	Findings will inform curriculum development for librarian training.
Task 6 Library Partnerships	Establish and strengthen collaborations with libraries and professionals.	Outreach to librarians; Participation in professional meetings.	Agreements for collaboration on developing service models and training resources.	Partnerships will ensure service models and training programs align with real-world library needs.
Task 7 Planning Next Project	Plan the next phase, which aims to design text analytics service models and develop tailored training resources.	Synthesis of evidence and takeaways from Tasks 1-5.	Actionable framework for implementation.	Evidence-based approach will ensure practical, tailored solutions for libraries.

PROJECT RESULTS

The primary outcome of this project will be **a series of research findings and reports**, covering various aspects: (a) the areas of text analytics services in academic libraries (from Task 1-2 and 4); (b) the specific needs of text analytics support among users (from Task 3); (c) case studies and best practices of text analytics services in libraries (from Task 1-2); (d) the topics and content of text analytics tailored to librarian training (from Task 3-5); (e) pedagogical approaches to training librarians (from Task 5); (f) partnership agreement on developing text analytics services and training resources (from Task 6), among others. These findings and reports will paint a comprehensive picture of the current state of text analytics services and the perspectives of different stakeholders, including librarians, library users, and LIS educators, on advancing text analytics services in academic libraries.

The findings will be disseminated through multiple channels, specifically targeting library practitioners. The project team will share outcomes and discussions across various professional networks, ensuring broad reach to and engagement with librarians involved in data scholarship. Key dissemination avenues will include major library conferences focused on data services, such as the ACRL Conference, the Research Data Alliance (RDA) Meeting, the Code4Lib Conference, and the Midwest Data Librarian Symposium. Findings will be published in professional journals with librarians as the primary audience, such as *The Journal of Academic Librarianship*, *The Journal of Librarianship and Information Science*, and *College & Research Libraries*. For findings related to education and pedagogical approaches, the project team will present at the Association for Library and Information Science Education (ALISE) meetings and publish in *The Journal of Education for Library and Information Science (JELIS)*. To enhance accessibility, all sharable research outputs and collected data will be deposited in institutional repositories at the PI's institution. To further expand engagement, the project team will create a dedicated blog to share preliminary findings and project updates, serving as a central hub for collaboration and discussion among academic library communities and LIS schools. The advisory board will be consulted for recommendations on additional dissemination venues and will assist in distributing project outcomes and follow-up plans through their libraries and professional networks.

ASSESSMENT OF THE PROJECT

The insights and evidence gathered from this planning project will be used for a follow-up initiative focused on developing text analytics service models and creating specialized training resources tailored to the needs of users and libraries. This project will also foster collaborative relationships with data librarians, who will be invited to participate in the next project. Strengthening these partnerships will be crucial for ensuring the long-term success and sustainability of this project.

Project success and sustainability will be evaluated through a combination of qualitative and quantitative measures. **First**, the number of survey participants and respondents will be tracked to ensure the reliability and representativeness of the data. **Second**, engagement with the project blog will serve as an indicator of successful dissemination and interaction with library communities and library educators, providing a platform for comments, questions, and suggestions from relevant stakeholder groups. **Third**, the number of partnerships established with academic libraries and professional organizations will be another key measurement. These collaborations will not only validate the project's relevance but also provide opportunities for continued engagement in text analytics service development. **Fourth**, the number of conference presentations, workshops, and publications resulting from the project will also serve as an indicator of its influence in library communities. In addition, tracking the number of attendees in presentations or workshops will help measure the project's dissemination impact. **Fifth**, institutional repositories will provide another avenue for assessing the project's reach and engagement. Metrics such as the number of views, downloads, and shares of project outcomes will indicate broader interest in the findings. These repositories will ensure that the project's outputs remain accessible for future research and implementation efforts, further extending its impact.

PROJECT SUSTAINABILITY

The findings from this project will offer a comprehensive picture of the current state of text analytics services in academic libraries and perspectives from different stakeholders. This will enhance the community's understanding of the landscape of text analytics needs and services across various institutions. The presentations and publications from this project will be deposited into the institutional repository at the PI's institute to support sustainable open access. The raw data from the current services investigation will be shared through the institutional Dropbox account provided by the PI's institute. More details are provided in the Digital Products Plan document.

The project team plans a follow-up project built on the foundation of this project to develop text analytics service models for academic libraries and create tailored training resources for academic librarians. These service models and training resources will have a long-term impact on advancing text analytics data services in academic libraries.

Year 1:

[illegible]

Year 2:

		Months											
Tasks	Activities	1	2	3	4	5	6	7	8	9	10	11	12
Task 3 Needs Assessment of Librarians	3.6 Interview data analysis												
Task 4 Needs Assessment of Users	4.4 Survey distribution												
	4.5 Data analysis												
Task 5 Input from LIS instructors	5.2 Compile instructors list												
	5.3 Invite participants												
	5.4 Online focus group												
	5.5 Focus group data analysis												
Task 6 Partnership with libraries	6.2 Develop and solidify partnership												
Task 7 Planning next stages	7.1 Compiling evidence												
	7.2 Plan the next phase												