College of Computing & Informatics, Drexel University (CCI, Drexel)

LIS Education And Data science-Integrated Network Group (LEADING)

Drexel University requests $905,624 for **LIS Education And Data science-Integrated Network Group (LEADING)**, a Laura Bush 21st Century Librarian (LB21) National Digital Infrastructures and Initiatives project, to **build capacity** through **recruitment, training, and development of the library workforce**. LEADING will train **40 Fellows** to be future data leaders and extend the mentor network across the national digital infrastructure. The LEADING network partners hubs include the Schomburg Center for Research in Black Culture (SC-NYPL), Montana State University (MSU), the University of California San Diego (UCSD), OCLC, and seven additional mentor partners sites (*see* Section 6). The LEADING project **scales-up** the original IMLS LEADS (RE-70-17-0094-17) pilot project through education, collaboration, and community development. LEADING will **extend the education pipeline** and build a diverse, inclusive, national, continuing education network of data-competent professionals connecting educators, researchers, library managers, mentors, and front-line librarians.

1. STATEMENT OF NEED: Libraries, archives and related institutions face an unprecedented opportunity to leverage data and transform their services and operations. Unfortunately, libraries continue to face a number of significant barriers to achieving this end, due to a nationwide skills gap due and a limited data-competent workforce (Shifting to Data Savvy: The Future of Data Science In Libraries, 2017). Indeed, there is progress toward closing the gap in libraries with exemplary initiatives (e.g., Data Curation Network, Open Data Literacy project, and Collections as Data). In parallel, LIS graduate programs have advanced data-related curricula (e.g., Ortiz-Repiso & Greenberg, 2018). All of these efforts are making important contributions in their respective domains. A critical next step is to build strong, **nationwide interconnections** among LIS programs, libraries, and librarians and **strengthen the capacity** of libraries to innovate and advance within the data science space.

The LEADS pilot program successfully made this **critical connection** training a cohort of **21 doctoral student Fellows** from across the country, who completed a formal data science curriculum at Drexel University and an immersive data science internships with a national digital infrastructure partner (OCLC, DPLA, California Digital Library, etc.). LEADS' impact is exciting and gratifying and has underscored the need for an extended LEADS-like network to frontline librarians. Key support includes: **1)** 100% advocacy from all 21 LEADS Fellows, based on their LEADS and professional experiences; **2)** LEADS advisory board recommendations to scale-up, beyond doctoral students toward sustainability; **3)** success with Temple University and Montana State University formally joining LEADS, after the project began, given the program's capacity to address their pressing, frontline library objectives; and **4)** a clear need for greater capacity, with LEADS being able to support only 40% of the applicants each year. Overall, there is a growing need to recruit, train, and develop a data science community across libraries, nationwide, as underscored in the Positioning Library and Information Science Graduate Programs for 21st Century Practice 2017 report, further motivating LEADING, as proposed here.

2. PROJECT DESIGN: We propose **LEADING**—a **national network for library professionals** and **advanced LIS graduate students** (doctoral students, or advanced LIS master's students). LEADING's project design adapts the National Science Foundation's (NSF) Big Data Innovation Regional Hub model, to scale a network of hubs and spokes. LEADING's plans draw from IMLS' resource lists, Biscuits vs. Granola, specifically Strategies to Scale Up Social Programs, with partners allowing librarians to customize their immersive LEADS experience. Drexel will serve as the chief coordinating and educational hub; and the four network hubs (SC-NYPL, MSU, UCSD, and OCLC) will serve as community spokes. **LEADING goals are to:**

1) Extend the Education Pipeline. Drexel University is one of the first LIS programs to implement data science at the doctoral and master's level. The LEADING project will leverage Drexel's curriculum and faculty expertise and extend the educational pipeline to frontline information professionals.
2) Collaborate and Develop Community. The project will bring together academics, researchers, library leaders, frontline information professionals, in a blended network to collaborate and advance data science in LIS education and practice. The Fellows will develop invaluable, enduring partnerships supporting future collaboration.
3) Promote Diversity and Inclusivity. During the original LEADS project, over half the participants represented diverse communities, and almost 30% were from underrepresented minorities. The LEADING program will carry this goal forward through our diversity plan, included below (Section 5).

Project Phases:
■ **Phase 1: Curriculum Enhancement and Recruitment** *(Nov. 2020-April 2021)*. Project PIs prepare curriculum modules covering introduction to data science, data curation, big metadata, algorithms, analytics, and visualization. Curriculum enhancements include AI (artificial intelligence), ethics, and data preservation. Hub partners will contribute to curriculum development. Phase 1 will also include virtual meetings with Hub partners and national digital infrastructure mentors to design Fellowship projects and initiate recruitment.
■ **Phase 2: LEADING Fellows/Program Launch** *(May 2021-Aug. 2021)*. Fellows will complete a 15-hour online curriculum, attend a 3-day data science bootcamp at Drexel, and pursue a 10-week immersive data science internship with their mentor partner or at their hub. Internships will be virtual or on-site. During Phase 2 Drexel will conduct formal assessments, drawing rom [IMLS Performance Measures](#). Our approach will build on LEADS evaluation instruments, and include surveys, interviews, and focus groups with all stakeholders.
■ **Phase 3: LEADING Program 2** *(Sept. 2021-Aug. 2022).* The project team will review and revise the LEADING Fellowship program for its second iteration. Phases 1 and 2 will be repeated, and Fellows will disseminate results and learning experiences at relevant venues (e.g., CNI, DLF, RDA, ALA, ALISE, ASIST, etc.).
■ **Phase 4: Sustainability** *(Sept. 2022-August 2023).* During Year 3, stakeholder (mentors, hub partners, advisory board members, educator, researchers, and Fellows) will gather to develop a sustainable plan. Feedback from LEADS has been instrumental in developing LEADING. All LEADING digital products (curricula lesson plans, internship models, and assessment resources), will being made available in an open repository (e.g., GitHub).

4. DIVERSITY PLAN: LEADING's diversity plan covers recruitment, collection use, and community outreach.
Recruitment: We will work with ALA Office for Diversity, Literacy, and Outreach Services to target applicants from under-represented populations; and Hub partners will also work through their Offices of Equality and Diversity for professional recruiting. **Collections:** Fellows will have access to unparalleled collections via the Schomburg Center's resources covering the history of race and black diaspora, MSU's holdings covering indigenous cultures, and DCIC's Mapping Inequality project and the Densho.org project with Japanese internment camp data. **Community outreach:** Drexel will host LIS career events at the Dornsife Center for Community Partnerships in Philadelphia's Promise Zone, and the Schomburg Center in Harlem, New York City.

5. BROAD IMPACT: LEADING aligns with IMLS/LB21 goals and will advance **student** and **frontline professionals'** knowledge, skills, and capacity to work with data, and apply data science techniques in the workplace. LEADING will have broad impact in four key areas **1) Continuing Education:** LEADING will recruit and train 40 Fellows in data science, through a blended program for both frontline librarians and advanced graduate students); **2) Network partnership development:** Building on LEADS' success, LEADING extends to library partner hubs and the OCLC network, to support future collaboration; **3) Curriculum:** An enhanced continuing education curriculum available to LIS graduate programs, libraries, archives, museums, and other cultural institutions nationwide**; and 4) Model:** LEADING will serve as a proof-of-concept model that can be extended to other libraries, librarians, and graduate programs, and enhance the LEADNIG national network.

6. PROJECT TEAM: PIs: Drs. Jane Greenberg Il-Yeol Song, Weimao Ke, Jake Williams, and Erjia Yan; instructors, Drs. Mat Kelly, and Rosina Weber, will develop and deliver the curriculum; **Project hubs** (*see* Section 1) and **Mentor partners:** Academy of Natural Sciences, Drexel; California Digital Library (CDL); Digital Curation Innovation Center (DCIC), Univ. of MD; DPLA, Historical Society of Pennsylvania (HSP), Univ. of Pennsylvania Libraries, will host Fellows, and the **Advisory Board** will guide the project development.

7. BUDGET: We request: $905,624, 1:1 cost-share match less student support is: $663,560. The total budget will cover a three-year period (Sep. 1, 2020-Aug. 31, 2023). The Drexel budget of $685,624 primarily supports advanced graduate students and bootcamp travel for students and mentors, with funds supporting project management, faculty preparation, network coordination, and project dissemination. A budget of $220,000 is allocated to project partners, to support early and mid-career professionals attending the Drexel bootcamp, mentor training, and institutional outreach and sharing the LEADING experience to build LEADING network capacity toward a sustainable model.