

**Title:** Curating Data to Enhance Public Library Effectiveness

**Summary:** The University of Michigan’s Inter-university Consortium for Political and Social Research (ICPSR) and the University of Missouri’s School of Information Science & Learning Technologies (SISLT), in collaboration with the Public Library Association (PLA) and the Association of Bookmobile and Outreach Services (ABOS), requests **\$225,148** for an Institute of Museum and Library Services (IMLS) **Implementation Grant**. We propose supplementing the IMLS Public Library Survey (PLS) with data from ABOS’ Bookmobile and Outreach Information Repository (BOIR) and surveys of public library program attendees collected through PLA’s Project Outcome initiative. This data will be aggregated, curated, and archived at ICPSR, then provided to practitioners. The result will be a one-of-a-kind dataset to augment the PLS and document the impact of library programming and outreach on communities. Our project will also serve as a prototype for future PLS data integration efforts.

**Project Justification:** The PLS is an annual survey of American public libraries. The PLS collects data used by policymakers, practitioners, and academics to make decisions about funding and managing libraries, as well as to study libraries’ effectiveness. Data in the PLS comes from over 17,000 sources, and it represents the “gold standard” for national data about public libraries. While the PLS is an invaluable resource, other organizations collect data that complement the PLS. The Association of Bookmobile Outreach Services BOIR tool, for instance, is gathering data about bookmobile and outreach services. PLA’s Project Outcome (PO) initiative collects patron evaluations of library programming.

Measures That Matter, an IMLS-funded effort, states that “**there is no over-arching national plan for the collection, storage, use, and dissemination of public library data and service outcomes.**” As a result, data collected by libraries, consortia, and professional associations documenting library programming and outreach outcomes are often inaccessible, fragmented, or in formats that are difficult to use. We propose combining ABOS and PO datasets in a user-friendly format for practitioners and researchers. Curating ABOS and PO datasets, then making enhancements by linking variables to the PLS, or computing new ones, will let librarians, researchers, and other library stakeholders evaluate library outreach and programming decisions. The PLS contains scant information about library outreach and programming outcomes. Complementing the PLS with ABOS and PO data will produce a one-of-a-kind dataset at a low cost by leveraging the resources of four complementary organizations: ICPSR, SISLT, ABOS, and PLA.

This project meets six National Leadership Grants for Libraries program goals and objectives. First, it will strengthen libraries’ capacity to work collaboratively to benefit the communities they serve (**IMLS Goal 5**). We will achieve this goal by developing a model to aggregate library outcome data and creating interoperable data structures and standards. We will also establish workflows to curate and archive data, then provide infrastructure for long-term data access that users can leverage to increase library outreach and programming effectiveness (**Objective 5.1**). Second, this project will improve the capacity of libraries to serve the public (**IMLS Goal 2**) by enhancing librarians’ and academics’ ability to use library outreach and programming data to establish evidence-based “best practices” (**Objective 2.3**). Third, our work will help create an educated, national workforce (**IMLS Goal 1**) by advertising our dataset to interested parties and conducting workshops that encourage dataset reuse. We will also provide summary statistics describing curated data to practitioners and create open-source learning modules for graduate courses (**Objective 1.2**).

Our primary **target audience** for this project is public library administrators and researchers. Combining BOIR and PO data with the PLS will let our target audience predict which programming and service decisions have the most significant community impacts. A second audience is library students and practitioners, whom we will train to use our data and replicate our model for augmenting the PLS. The **ultimate beneficiaries** of our work will be library patrons whose needs will be better served by programs and services designed to maximize beneficial outcomes.

**Project Work Plan:** The activities, partners, and schedule for this project are:

- 1) Collecting and curating third-party library outreach and programming data from ABOS and PLA to supplement the PLS.** Project lead A.J. Million and ICPSR staff will work with ABOS and PLA to obtain their data, curate it, and create a new dataset for the library community. This dataset will be a one-time extract of available data that may be updated annually in the future. ABOS will advertise their BOIR tool during Year 1 to expedite data collection and provide an extract to ICPSR during Year 2. PLA will export their data and provide support to ICPSR in Years 1 and 2. After the project ends, ICPSR will preserve data in their CoreTrustSeal repository and offer perpetual support to data users.
- 2) Creating a data module to support and document curation and guide future PLS enhancements.** Denice Adkins, a SISLT faculty member, will lead data module creation in Years 1 and 2. This module will be a document to support data users. The module will map BOIR and PO variables to the PLS by establishing pre-defined units of analysis, identify and document new variables of interest that we compute, and carefully describe any transformations we carry out. The module will provide transparency for our work and serve as a “how-to” guide for integrating other datasets with the PLS.
- 3) The creation of educational and outreach modules to promote the use of data to multiple stakeholder groups.** In Year 2, Denice Adkins will supervise a SISLT Graduate Assistant. Together, they will create open-access curricular modules for use in graduate courses. These modules will showcase data in the scope of real-world projects and teach data analysis and management skills. PLA, ICPSR, and SISLT will partner to offer free online workshops introducing the dataset and provide hands-on “hack time” for in-service librarians and administrators.

A.J. Million and staff at ICPSR will spend **1,583 hours** keeping the project on track, curating data, documenting datasets by creating codebooks and metadata, supporting outreach, and tracking citations of project data to measure long-term usage. Denice Adkins and the SISLT Graduate Assistant will spend **600 hours** creating our data module, finding and suggesting data enhancements, and developing educational materials for classroom use. PLA staff will spend **416 hours** extracting PO data, advising on creating educational materials and workshops, and supporting data curation. ABOS requests financial support to operate their BOIR tool and expedite data collection.

**Diversity Plan:** We will form an advisory committee during Year 1 to support Dr. Adkins in creating our data module. Committee members will represent various associations and public libraries in multiple contexts (rural, urban, suburban, majority-minority libraries, and others). Members will advise us on demographic and other datasets to provide contextual detail for our project.

**Project Results:** Intended results are 1) the creation of a novel dataset that provides information about the outcomes of public library outreach and programming, 2) a template for future data integration projects with the PLS, and 3) education for library practitioners and students. Curating this dataset, and training people in its use, will help library administrators make informed service decisions with limited resources, provide evidence for researchers to determine the value of services public libraries provide, and support library leader outreach to stakeholders like municipal leaders, regional policymakers, and potential partner organizations. Our data module will provide structure and support for future data integrations to increase the utility of PLS data and help address the lack of a national plan for data collection, storage, and use as noted by Measures That Matter. Finally, while this project seeks support for the one-time extraction and curation of data, we anticipate future efforts will be more efficient requiring only data curation.

**Budget:** We request **\$225,148** for this project. Requested funds include **\$94,045** for salaries, wages, and benefits; **\$94,405** for indirect costs, such as facility operations; **\$12,249** for ICPSR systems infrastructure to perpetually store and provide data; and **\$24,449** for ABOS to expedite ongoing data collection.