

# Methodological Improvements to the Public Libraries Survey

JUNE 2024

The Institute of Museum and Library Services (IMLS) has undertaken a series of integrated initiatives over the past five years to improve the methodological rigor and utility of the annual [Public Libraries Survey \(PLS\)](#). In this research brief, we highlight these efforts and preview the plans for the next phase of this initiative. The end goal is to enable library staff, researchers, and the public to more effectively use library data to document the ongoing value of libraries as important social institutions in their communities.

## Challenges with Linking PLS Data to Other Data

IMLS directly and indirectly sponsored two research projects that sought to link PLS data with other federal datasets to enhance the narrative about the contributions of public libraries in the United States. Despite these aims, both research endeavors met significant challenges, largely due to the structure of the PLS data.

First, in 2020, Lisa Frehill and Melissa Cidade presented a [study](#) to the Federal Committee on Statistical Methods that blended data from the PLS and the [American Housing Survey \(AHS\)](#). This study asked research questions about the relationship between the civic engagement and social capital within a community and the levels of public activity in the library that serves that community. Second, in 2022, the American Institutes for Research (AIR), on behalf of IMLS, conducted preliminary analyses by linking the [National Household Education Surveys \(NHES\)](#) with PLS data. This study asked research questions about the relationship between two attributes: whether households with children visited a library, and the level of children's services at the nearest library. In both studies, the researchers attributed common PLS indicators from the Administrative Entity (AE) file to the associated outlets for each AE. They then used a spatial join<sup>1</sup> to merge these library data to household

records in the AHS and NHES datasets. Researchers involved in both studies acknowledged the potential for bias and inaccuracies due to two related factors:

- Linking a household to the nearest library outlet may not reflect the actual outlet that the household visits or would consider visiting. Additionally, joining household data to the nearest library outlet does not guarantee that the household is within the legal service area of that outlet, which could affect the services available to household members at that outlet.
- Attributing AE-level data to each library outlet in larger library systems, often found in urban and suburban locales, tends to mask neighborhood-level variation among those outlets.

To mitigate these issues and enhance the quality of data analyses, IMLS is exploring options to improve and extend the PLS dataset. The first factor above is being addressed by updating the geographic identifiers included in the AE file. The second could be improved by collecting and publishing data on a more granular level, such as information on visits, circulation, and Internet computer use. Below, we further describe each of these efforts.

## Updating the Geographic Identifiers

To enable data users to link PLS data with other geographical data more easily (e.g., the American Community Survey), IMLS revised the categories for the Census Bureau geography type that best aligns with a library's legal service area, starting with the FY 2022 PLS data to be released in late spring 2024. Historically, the published PLS data files—for both AEs and outlets—have included geographic identifiers for the county, census tract, block group, and congressional district in which each AE and outlet is located. While some AEs do serve an entire county (and some serve multiple counties), others cover areas

<sup>1</sup>Like a traditional database that uses a common identifier to join two datasets, a "spatial join" merges two datasets based on location. In both research studies discussed, the join was based on the closest library outlet to each household record in the other dataset.

smaller than a county but larger than a census tract—often encompassing a municipality, township, school district, or combination thereof. By revising the categories for this geography type indicator in the PLS data to match the categories currently used by the Census Bureau for most of its data products—and by including the geographic identifier (GEOID) for those varied geographic areas in the PLS data files—IMLS provides researchers with more geographically accurate public library metrics that align with the demographic data about the communities served by those libraries. For example, for a public library that serves a single municipality, the Census Bureau GEOID for that municipality will be included in that library's PLS record to make it easy to join demographic data for that municipality to the library data. Data users could also merge these GEOIDs onto older PLS datasets, as long as the library has not undergone any changes to its legal basis or structure in the meantime.

Furthermore, IMLS has begun to collect geographic data files, such as shapefiles, of library service area boundaries from several states. While some states lack these files and have no plans to develop them, IMLS will evaluate the utility of existing files on a trial basis. This assessment will help determine the potential use cases for these data. If deemed suitable, IMLS may share them with the public as a provisional dataset to motivate public library research.

## Exploring the Collection of Outlet-Level Data

To address the challenges outlined above and make it easier for researchers to connect library activity to other community-related data, IMLS will again partner with AIR over the next two years to explore the feasibility of collecting certain PLS data items at the outlet level. Historically, almost all PLS data items are collected at the AE level. In some cases, data items occur at the outlet level but are aggregated to the AE level before libraries or states report them via the PLS. For example, while libraries already track visits at the outlet level, these figures are summed to the AE level, resulting in a lack of geographic granularity for this fundamental library use metric within the PLS. However, capturing and reporting certain data items, such as physical circulation, at the outlet level may pose challenges in developing a standardized method. The new study will examine the feasibility and burden associated with collecting data items at the outlet level.

IMLS will engage with key stakeholders knowledgeable about library data using the framework described in the Policy and Procedures for making changes to the PLS. Working with chief officers of state library agencies, state data coordinators who collect the public library data in each state, public library staff, and PLS data users, IMLS and AIR will determine the feasibility and value of collecting outlet-level data. This group will collaborate to develop a plan to test and implement any approved data collection and reporting changes.

## Continuous Improvement: Annual Process for Reviewing PLS Data Elements

In 2022, after a two-year pilot, IMLS enacted new policies and procedures for making changes to PLS data elements with the goal of improving survey development transparency and to advance the relevancy, accuracy, and timeliness of the PLS.

The policy lays out requirements for making change requests, developing data elements, and reviewing suggestions. Discussions with defined stakeholders, such as the Library Statistics Working Group, state data coordinators, and chief officers of state library agencies, lead to methodological support through cognitive testing with library-level respondents to rigorously develop and test items. Starting with the FY 2024 cycle, changes to the policy's timeline resulted in states having more time to add the changes to their state-level surveys, with the aim of improving data response rates and quality in the first year of data collection.

Changes to the PLS are ratified when a minimum of two-thirds of at least 45 states and outlying territories agree to implement them. This process exemplifies the cooperative federal-state partnership of the PLS, resulting in annual response rates of 95% or higher from the over 9,000 public libraries in the United States. Over the past five years, the new collaborative process has enabled a robust update to program and program attendance data items (FY 2021), the rapid addition of COVID-19 variables (FY 2020–FY 2022), the addition of a new indicator of overdue fine policy (FY 2022), updates to the GEOCODE item (FY 2022), and revisions to electronic content, circulation, and funding sources (FY 2024) to increase data quality and reliability.

## About the Institute of Museum and Library Services

The Institute of Museum and Library Services (IMLS) is the primary source of federal support for the nation's libraries and museums. We advance, support, and empower America's museums, libraries, and related organizations through grant-making, research, and policy development. Our vision is a nation in which museums and libraries work together to transform the lives of individuals and communities.

### Publication Information

This publication is available online at <https://www.ims.gov/publications>. IMLS will provide an audio recording of this publication upon request.

For questions or comments, contact [research@ims.gov](mailto:research@ims.gov).

All information contained herein is in the public domain and may be used and reprinted without special permission. Citation of this source is required.

**Suggested Citation:** Nielsen, E. & Pelczar, M. (2024). *Methodological Improvements to the Public Libraries Survey*. Washington, DC: Institute of Museum and Library Services.

Institute of Museum and Library  
Services

[www.ims.gov](http://www.ims.gov)

[imsinfo@ims.gov](mailto:imsinfo@ims.gov)

202-653-IMLS (4657)

