Indiana University Indianapolis (IUPUI)

Project Title: Data Reuse for Local Community

Summary: The research team from the Department of Library and Information Science (DLIS) at Indiana University in Indianapolis (IUPUI), will explore community data reusers' practices to offer insights into how public libraries' services can be adapted to meet the needs of community reusers. Project outcomes will include a prototype library data service and workshops for librarians to help them work with community members.

Statement of Need

The potential for data to help address current societal problems (e.g., education, health, economic development, environment) is significant not only at the federal and state levels but also in smaller communities, neighborhoods, and individual lives. While the proposition for this potential is that data are and will be shared with and reused by and for communities at different levels, many data are not systematically or routinely shared for reuse with communities, particularly data collected by university and research institutes.[1] Government data (e.g., census) are traditionally shared with citizens, but research data are usually shared within scientists' own communities of practices, although they share data interpretation to support community development. The ultimate goal of this project is to support community members' data reuse[2] *by facilitating their connection with existing resources and data infrastructure in order to create equitable access to the vast stores of data being collected about individuals and their communities*; therefore, our focus is public libraries. We broadly define community members as citizens who utilize data to make an impact on the community (e.g., educational consultants, city planners, social activists, community foundations), as well as those who reuse data for community-related decision making (e.g., parents looking for educational recommendation to schools for their children). To achieve the goal, we will (1) characterize community members' data reuse practices and strategies, and (2) map these practices and strategies to the existing infrastructure focusing public libraries, to better support communities.

RQ1. What are the processes, strategies, and challenges in community members' data reuse?

Little is known about community members' data practices, such as how they utilize data for what purpose through what kind of support. Community members contribute to data collection in citizen science projects (e.g., eBird), but in many cases, they are not positioned as "active reusers" who can read and interpret data for their own purpose. To build more ground, the PI conducted a preliminary study in 2016, funded by the RDA/US Data Share program through the Alfred P. Sloan Foundation, in a small city in the Midwest. The findings suggest that community members' data needs and reuse practices are different from scientists, and they have limited interactions with local data infrastructure. Instead, they relied on regional community research centers (CRCs) that connect them with data sources, partnerships (e.g., between academic and community), and consultations.[3] The findings informed the overall study design, and we will continue to explore community's data reuse practices in a different context, such as in small, medium, and large-sized cities. The findings from the proposed study will help us develop a theoretical model of community-inclusive data exchange with practical recommendation for public librarians to create community-inclusive data workflow.

RQ2. How do public librarians see their role in supporting data reuse for communities? How can the current data infrastructure become more inclusive to support communities for data work?

One striking finding from the preliminary study was the lack of notable interaction of community members with public libraries. While the role of CRCs is critical to support community's data work, because the CRCs' priority is not on individuals (e.g., parents in the example above), there are needs for those remaining population to be served. Public libraries serve communities by providing access to some of the existing data resources (e.g., census, school statistics), and have the potential to facilitate citizen engagement with data and data professionals. Boston Public Library's recent initiative to develop open data collection for the City of Boston demonstrates this potential. We will explore the views of public librarians on their roles, perceived barriers to those roles, and the ways in which public libraries can more actively engage with data. Using the

---

[1] *Data for Development Interest Group* at Research Data Alliance: https://www.rd-alliance.org/groups/data-development.html

[2] Data reuse is defined as the secondary use of data beyond the original purpose of data collection.

[3] In northeast Indiana, for example, the Community Research Institute (CRI) in Fort Wane serves to promote the growth and vitality of the regional community and provides contract-based research and analytical services. We identified about 50 CRCs in US, which also exist as nonprofit organizations and are often affiliated with universities.

findings from R1 and R2, we will develop a prototype library data service, and design workshops for librarians to enhance their awareness, inform significance of their roles, and help them work with community members.

National Impact and Intended Results

The project will have major impacts in two areas. Overwhelmingly, data reuse research focuses on the scientific and academic contexts. This project will contribute to broadening the scope of data reuse research and begin to bridge the connection between scientific and local community data reuse. The project will also help build a collaborative, community-inclusive culture of data reuse among different stakeholders. Second, our work will offer insights into how current data services can be adapted to meet the needs of community reusers. In particular, it will raise awareness of the role of public librarians in community data reuse through the inclusion of their expertise as well as community stakeholders' in the design/extension of library data reuse services.

Project Design

Year 1: Study of community data reuse practice – Interview and survey. We will interview community data reusers and relevant stakeholders. We will partner with 10 CRCs in small, mid, and large cities in U.S. (e.g., Allendale, MI, Kansas City, KS, Chicago, IL) and recruit about 30 data reusers from their network to interview them. We will also interview staff in regional data infrastructure (e.g., director of CRCs and data repositories) to understand their role in supporting data reuse. An online survey will be designed and distributed to community members through 50 CRCs we identify. From qualitative and quantitative analysis, a theoretical model of and practical recommendations for community-inclusive data exchange/workflow will be produced.

Year 2: Study with public libraries – Environmental scan follow-up interview, partnership building. We will conduct an environmental scan to identify librarians working with data for their communities' use, followed by interviews with librarians identified through the scan as well as help from our advisory board, to learn about librarians' roles, current efforts to support communities. From the qualitative analysis of interview data and examination of services introduced during the interviews, we will update the theoretical model and practical recommendations developed in phase 1 to reflect the roles of public libraries. The outcomes will be distributed online to seek input from members of PLA and ALA. With that, we will start recruiting participating libraries for phase 3, library service design. The environmental scan will help us identify partner libraries that have implemented some efforts to support community data needs or that are interested in building data services for their community. We will also attend the PLA conference to advertise our project, and recruit partners.

Year 3: Public library service prototype design and workshop. In year 3, we will utilize evidence from our findings to improve or develop library services. We will work with our partner libraries to design a prototype service. We will also run two half-day workshops at the PLA conference: the first session will focus on opportunities for public libraries to support the community through data (conceptual grounding); the second session will delve into the role of public libraries in building community data infrastructure and service design. We cannot specify what types of service we will be designing; our goal is to let the data guide us with the project progress. All workshop materials will be freely available through the project website. Our outcome will also be presented at the PLA and other relevant conferences to be broadly distributed.

Project Evaluation. We will routinely evaluate the progress of the project to ensure the project success. We will adopt a participatory evaluation framework, and invite our partners to evaluate our project in each phase. An advisory board (**Nate Hill**, Executive Director at Metropolitan New York Library Council, **David Leonard**, President at Boston Public Library, **Ellen Cutter**, former Director at CRI Fort Wayne, 3 to be determined) will help with study recruitments, participate in the ongoing evaluation (reviewing instrument development, implementation processes, the team's collaborative work with partners) and be invited to Indianapolis for half-day annual evaluation. We will also collect post-workshop evaluation from our workshop participants.

Project Personnel: **Ayoung Yoon, Ph.D**.(PI) is an assistant professor in DLIS at IUPUI. She is a RDA/US data share fellow. Her dissertation, *Data Reuse and Users' Trust Judgments*, received the Eugene Garfield Doctoral Dissertation Award. **Andrea Copeland, Ph.D**.(Co-PI) is an associate professor in the DLIS at IUPUI. She has researched the relationship between communities and public libraries.

Budget Requested: $362,405 (Total direct costs: $233,459, Indirect costs (F&A @57.5%): $128,946)
Project personal $172,559 (PI 20%, Co-PI 10%, 1 GA half time tuition with stipend), travel $18,000, CRCs consulting fee $3,000, participants $2,500, Public library partners $3,000, Advisory board (travel, hospitality) $ 24,000, Workshops $5,000, Hardware $500, Printing/publication $900, Transcription $4,000