

## Developing Public and Academic Libraries as Key Participants in Civic Open Data Ecosystems

### Abstract

The University Library System (ULS) at the University of Pittsburgh, in partnership with the Carnegie Library of Pittsburgh (CLP), the Western Pennsylvania Regional Data Center (WPRDC), and the Urban Institute, which supports and coordinates the National Neighborhood Indicators Partnership (NNIP), will develop the capacity of public and academic libraries to serve as key partners in local open civic data ecosystems. The project's primary output will be a guide and toolkit to help public and academic libraries: identify local needs and contexts around open civic data; consider roles, opportunities, practices, and governance in the civic data ecosystem; anticipate and address common challenges; measure local civic open data health and capacity; and learn from examples of successful civic data partnerships. A significant secondary objective of the project is to grow a national network of shared practice and organizational capacity. The project's outputs and findings will be widely disseminated to libraries and allied groups that support the publication and use of civic open data.

The widespread availability of open data, computing power, and accessible tools for analysis is leading to growth in the use of data for discovery and argument in civic and scholarly contexts. Libraries have a strong recent history as advocates for openness in scholarly communication and cultural heritage. Parallel to these developments, government information has also seen a turn towards openness: at local and regional levels of government, open data portals have emerged to publish civic data, with goals of improved operational efficiency, transparency, public participation, and a better-informed citizenry.

Local governments are making considerable investments in open data, spending tens of thousands of dollars per year on data repositories and making investments in staff to manage, transform, and make information accessible and usable for internal and external audiences. Despite these investments, there has been a growing awareness that the act of publishing data alone does not always result in community impact. *Data intermediaries* are entities that help the public to apply and use open data to achieve impact by building relationships among publishers, users, and other members of the ecosystem, improving data, statistical, and technological literacy, enhancing data quality, providing feedback mechanisms to publishers, and building tools that enable broader data use. Libraries and librarians should be key actors in the continuing development of civic open data portals and act as core data intermediaries; their expertise adds value to a wide range of issues that affect both data publishers and users.

Despite the clear value that libraries can bring to local civic data ecosystems, few libraries now function as key participants within them. Guidance and development, at a national level, is needed to promote successful collaboration and coordination between public libraries, academic libraries, related data intermediaries, civic open data portals, and their associated data producers and users. A national framework of guidance must also draw out how local factors can impact reproducibility elsewhere. This is the gap this project will address, by 1) working with a variety of regional civic data partnerships to develop them as participants in their local civic data ecosystems, and 2) capturing and cross-testing their experiences in a guide and toolkit that can be used by other libraries nationwide.

## Developing Public and Academic Libraries as Key Participants in Civic Open Data Ecosystems Project Narrative

The University Library System (ULS) at the University of Pittsburgh, in partnership with the Carnegie Library of Pittsburgh (CLP), the Western Pennsylvania Regional Data Center (WPRDC), and the Urban Institute, which supports and coordinates the National Neighborhood Indicators Partnership (NNIP), propose a two year project that will develop the capacity of public and academic libraries to serve as key partners in local open civic data ecosystems. The project's primary output will be a guide and toolkit to help public and academic libraries: identify local needs and contexts around open civic data; consider roles, opportunities, practices, and governance in the civic data ecosystem; anticipate and address common challenges; measure local civic open data health and capacity; and learn from examples of successful civic data partnerships. A significant secondary objective of the project is to grow a national network of shared practice and organizational capacity. The project's outputs and findings will be widely disseminated to libraries and allied groups that support the publication and use of civic open data.

### 1. Statement of National Need

The widespread availability of open data, computing power, and accessible tools for analysis is leading to growth in the use of data for discovery and argument in civic and scholarly contexts. Libraries have a strong recent history as advocates for openness in scholarly communication and cultural heritage. Parallel to these developments, government information has also seen a turn towards openness; the 2013 White House Executive Order *Making Open and Machine Readable the New Default for Government Information* is a well-known example.<sup>1</sup> At local and regional levels of government, open data portals have emerged to publish civic data, with goals of improved operational efficiency, transparency, public participation, and a better-informed citizenry. In many cases, it is at the local level that the most innovative and forward-looking ideas around open data are developing. The Sunlight Foundation has written, "cities are now setting the leading examples for public data, from policy, culture, collection to engagement and reuse... they are increasingly providing best practices and learning from experimentation that federal governments should learn from and adopt<sup>2</sup>." According to the Sunlight Foundation, 73 local governments have adopted open data policies in the U.S., with many more communities operating open data programs independent of a policy framework.<sup>3</sup> The transformative potential of such data is significant. The Pew Research Center writes that open

---

<sup>1</sup> The White House. (2013, May 9). [Executive Order - Making Open and Machine Readable the New Default for Government Information](#).

<sup>2</sup> Wunderlich, J., & Sunlight Foundation. (2017). [Trump, Threats to Data, and US Cities](#).

<sup>3</sup> Sunlight Foundation. (2017). Open Data Policy Collection: [List of all open data policies](#).

government data "hold the potential to change how people engage with governments at all levels," and the availability of data and tools "creates new service-delivery possibilities for government through use of data that government agencies themselves collect and generate."<sup>4</sup> Government services are not the only audience; the Pew report notes the availability of open data "as a feature to the general public. The kinds of entrepreneurial activity new government data sources can spur range from home energy management to analytics for investment decisions."

Data publishers are making considerable investments in open data. In addition to spending tens of thousands of dollars per year<sup>5</sup> on data repositories, local governments are also making investments in staff to manage, transform, and make information accessible and usable for internal and external audiences. Despite these investments, there has been a growing awareness that the act of publishing data alone does not always result in community impact.<sup>6</sup> *Data intermediaries* are entities that help the public to apply and use open data to achieve impact by building relationships among publishers, users, and other members of the ecosystem, improving data, statistical, and technological literacy, enhancing data quality, providing feedback mechanisms to publishers, and building tools that enable broader data use.<sup>7</sup>

Libraries and librarians should be key actors in the continuing development of civic open data portals and act as core data intermediaries; their expertise adds value to a wide range of issues that affect both data publishers and users. Writing on this subject, Shannon Mattern, Associate Professor at The New School, captures many of the ways that librarians and allied information professionals can -- and should -- contribute to civic open data in their communities:

We must push our civic leaders to bolster their planning teams with experts in the ethical collection, organization, preservation, and dissemination of information resources. Urban data programs should be counseled by professionals who understand the complex issues of equity, privacy, and security. Librarians on the planning commission! Archivists in the police academy! ... They are guardians of a critical, contextual approach to *information*, which is a public resource every bit as necessary as streets and sewer lines.<sup>8</sup>

There are notable examples of libraries participating in the civic data space as intermediaries, including the recent nationwide events under the DataRefuge umbrella, the emerging Libraries + Network, and Temple Libraries' Knight Foundation award on future-proofing civic data.<sup>9</sup>

Members of the project team have prior experience in some of these: as co-organizers of a Data Rescue Pittsburgh event, and as participants in the Libraries+ Network, a "nascent consortium of

<sup>4</sup> Pew Research Center. (2015). [Americans' Views on Open Government Data](#).

<sup>5</sup> Benison, M. (2016, March 17). [Open Data - How Much Does It Cost?](#). GovEx.

<sup>6</sup> Living Cities. (2012). [Field Scan of Civic Technology](#).

<sup>7</sup> Van Schalkwyk, F., Chattapadhyay, S., Cañares, M., & Andrason, A. (2015). [Open data intermediaries in developing countries](#).

<sup>8</sup> Mattern, S. (2016). [Public In/Formation](#). Places Journal.

<sup>9</sup> See [Data Refuge](#). (2017); [Libraries + Network](#). (2017).; Krotzer, S. (2017).

[Temple University Libraries wins a 2016 Knight News Challenge on Libraries award](#). Temple Now.

research libraries, library organizations, and open data communities with a shared interest in saving, preserving, and making accessible born-digital federal government information".<sup>10</sup>

While some libraries are beginning to take on greater roles around civic open data<sup>11</sup>, individual efforts usually focus on selected aspects of the local data ecosystem. For example, *public libraries* are developing programs supporting data literacy and community-centered advocacy using open data. *Academic libraries* are growing programs for research data management, data stewardship, and advocacy for data sharing. And, to be most effective, these library entities must also recognize and collaborate with both data publishers and other *non-library data intermediaries*, including members of the NNIP network of practice for neighborhood data intermediaries. Many of these non-library data intermediaries often pre-date the open data movement in providing valuable services to data users, democratizing data, and supporting equitable access to information. All of these actors form parts of a data ecosystem with overlapping and often complementary roles, expertise, and audiences.

Despite the value that libraries can bring to local civic data ecosystems, few libraries now function as key participants within them. This is, in part, because of the complexity of roles and relationships described above, and the lack of well-defined tools to help libraries understand and navigate them. Supporting use of civic data at scale requires the coordinated efforts of a complex network of data intermediaries who skillfully bridge communities of data providers and users. In practice, coordination amongst these intermediaries is a significant and ongoing challenge.<sup>12</sup>

Guidance and development, at a national level, is needed to promote successful collaboration and coordination between public libraries, academic libraries, related data intermediaries, civic open data portals, and their associated data producers and users. A national framework of guidance must also draw out how local factors can impact reproducibility elsewhere. This is the gap this project will address, by 1) working with a variety of regional civic data partnerships to develop them as participants in their local civic data ecosystems, and 2) capturing and cross-testing their experiences in a guide and toolkit that can be used by other libraries nationwide. This effort aligns with IMLS's stated goal for the National Digital Platform of "addressing intersections between digital issues facing libraries and cutting edge work in other fields" -- in this case, existing and emerging civic open data ecosystems -- "to yield broadly applicable modeling... and tools."

---

<sup>10</sup> [Libraries + Network](#). (2017).

<sup>11</sup> Further examples include work done in [Boston](#), [Seattle](#), and [Chattanooga](#).

<sup>12</sup> Petit, K., Hendey, L., Losoya, B., Kinglsey, G. (2014). [Putting Open Data to Work for Community](#).

## 2. Project Design

### Project Objectives

The project design is based on two strong convictions. First, that cultivating a healthy local civic data ecosystem depends upon the coordinated efforts of a variety of data intermediaries. In other words, no single entity can effectively cover all of the necessary roles alone.<sup>13</sup> To that end, our project team includes representation from a public library system, an academic library system, a regional open data portal, and a national membership organization of non-library civic data intermediaries. In Pittsburgh, Pennsylvania, both public and academic libraries are active participants in our regional open data ecosystem. Cultivating and strengthening similar partnerships in other regions is a core objective of this project. Secondly, we believe that the importance of local context and the variety of local civic data landscapes mean that no single model can be made to fit every city or region. This can be seen, for example, across the network of NNIP data intermediaries; because each community's civic data ecosystem is unique the structure of local data intermediaries have evolved to take different shapes.<sup>14</sup> Modelling at the national level must be done by capturing a wide variety of successful local practices. Although the prior work of the project team members serves as an example of successful library participation in *our* local civic data ecosystem<sup>15</sup>, this project is deliberately designed to involve a broad and diverse cohort of regional participants whose experiences will inform its outputs.

The project's primary output will be a guide and toolkit to help public and academic libraries become more engaged in their civic data ecosystem. The guide will help libraries do the following:

- identify local needs and contexts around open civic data
- consider roles, practices, and governance in the civic data ecosystem
- anticipate and address common challenges data intermediaries may encounter as they help the public to apply and use open data
- measure local civic open data health and capacity
- promote use of data in alignment with community-based concerns
- learn from examples of successful civic data partnerships

A significant secondary objective of the project will be to grow a national network of shared practice and organizational capacity.

---

<sup>13</sup> Van Schalkwyk, F., Chattapadhyay, S., Cañares, M., & Andrason, A. (2015). [Open data intermediaries in developing countries](#).

<sup>14</sup> See Chapter 3, "Deciding What's Right for Your Community" in Hendley, L., Cowan, J., Kingsley, G., Pettit, K. (2016). [NNIP's Guide to Starting a Local Data Intermediary](#).

<sup>15</sup> E.g., Brenner, A., Gradeck, R., Mattern, E., Bolam, M. (2015). [Libraries Will be An Asset for Us: Emerging Roles for Libraries in Civic Data Partnerships](#); Carnegie Library of Pittsburgh's [work on data in the community](#); Gradeck, R. (2017). [Role of Data Intermediaries in Civic Data Ecosystems](#).

## Project Staff and Tracking

The project team is comprised of:

- Aaron Brenner, Coordinator of Digital Scholarship at the University Library System, University of Pittsburgh, who will serve as Principal Investigator and will oversee project activities and progress;
- Bob Gradeck, Project Manager of the Western Pennsylvania Regional Data Center at the University of Pittsburgh, who will focus on ways libraries can fit into other aspects of civic data ecosystems, how roles may be structured in other cities, and will help participants make connections with national organizations such as Code for America and the Sunlight Foundation;
- Toby Greenwalt, Director of Digital Strategy and Technology Integration at the Carnegie Library of Pittsburgh, who will focus on engagement with the public library community;
- Eleanor Mattern, Postdoctoral Researcher at the University Library System, University of Pittsburgh, who will focus on the development of the guide;
- Liz Monk, currently Research Specialist at the University Center for Social and Urban Research at the University of Pittsburgh and Project Manager of the Southwestern Pennsylvania Community Profiles, who will serve as a .5 FTE Project Manager for this work (see Project Manager position description in support documents);
- Kathryn Pettit, Director of the National Neighborhood Indicators Partnership who will focus on building connections between this project and the NNIP network of data intermediaries;
- Eleanor Tutt, Open Data and Knowledge Manager at the Carnegie Library of Pittsburgh, who will focus on the development of the guide.

The project team will meet monthly as a full group to track overall project activities against the schedule of completion. The project design, detailed below, is based on an iterative model of development with explicit components that focus on revising and adjusting our primary product. The Principal Investigator and Project Manager will meet at least monthly to track project costs and budget. Smaller team meetings will be scheduled as needed for work related to specific tasks. The entire project team will make use of web-based task tracking tools and file sharing platforms with which they have existing experience. The Principal Investigator and Project Manager will regularly assess the frequency of full team meetings and will adjust if necessary.

## Project Components

The project design has the following main components that will take place over its two year timeline. More specific information about the timing and sequence of these components can be found in the separate *Schedule of Completion*.

### *Create initial draft of guide and toolkit*

The project manager will be hired and will work with the Pittsburgh-based project partners to create an initial draft and structure for the guide and toolkit. Project partners will draw on both local partnerships with data publishers and user groups, as well as their involvement in national communities of practice, including the NNIP Network, the American Library Association (including the ACRL and PLA divisions), DLF, and DataRefuge / Libraries+Network. This guide and toolkit will be drafted in an open platform that supports versioning, comments, contributions, and re-use (e.g. GitBook, or a GitHub repository with static site overlay). The guide will draw inspiration from similar "playbooks" and field guides that are increasingly common within open data and civic tech communities.<sup>16</sup> The content will be structured to provide guidance and models around core aspects of civic open data, with a focus on opportunities for libraries' participation. The core aspects will include: motivations and drivers; roles and stakeholders; institutional contexts; the value of libraries and librarians' expertise; surveying and mapping the local landscape; pathways of data publishing, access and use; and the importance of relationships and trust. Case studies will be added during later phases of the project. The associated toolkit will contain example activities, assessment tools, communications, and data collection instruments. Project partner The Urban Institute and supporters such as The Sunlight Foundation, Code for America, and the Digital Library Federation will promote the initial work and help identify other civic data partnerships that the team will recruit to participate in the workshops and further contributions to the guide.

### *Design and conduct workshops; establish participant communications*

To build relationships between libraries and other non-library data intermediaries, and to gather additional information for the guide, the project will conduct two workshops that bring together regional teams consisting of at least one library partner and one NNIP or comparable civic data partner. One workshop will take place at a national NNIP meeting in spring 2018 and another at the Digital Library Federation (DLF) Forum in fall 2018. At NNIP and DLF, the project will provide stipend awards to subsidize the workshop participation of regional teams (consisting of at least one library partner and civic data partner). We will make awards to support the attendance of up to ten pairs representing different cities or regions. In addition to the workshop

---

<sup>16</sup> E.g., What Works Cities, Sunlight Foundation. (2017). [Tactical Data Engagement Guide](#); OpenNews. (2017). [Field Guide to Open Source in the Newsroom](#); and Headd, M. (2016). [How to Talk to Civic Hackers](#)

stipends, the project will encourage cross-pollination of library and non-library data intermediaries through *Cross-Pollinator Registration Awards*. Through a mix of cost shared contributions and requested funds, up to five of the workshop participants will receive complimentary registration to stay on to attend the conference of their non-affiliated community (either NNIP or DLF), further strengthening partnerships and cross-pollinating the communities. We will strongly encourage the participating regional teams to attend both workshops in order to advance their work and to grow the connections between the library and NNIP communities. The recipients of the workshop stipend awards and the cross-pollinator registration awards will be identified through a process asking applicants to describe their local data ecosystem, articulate their prior or current engagement with regional civic data initiatives, share their perspective of the role of libraries in civic data initiatives, and express their interest in attending a workshop that would bring together librarian and non-librarian participants to discuss models for collaboration.

Two to three members of the project team will conduct and facilitate each workshop, and all members of the project team will contribute to their design. The workshops will introduce the themes of the guide, solicit input and discussion from participants, and give regional teams concrete actions that they can take when they return to their local contexts. The workshops will provide a forum for participants to learn about one another's work, areas of expertise, and needs; to identify opportunities for the regional partners to collaborate; and to demonstrate the structural forms that local partnerships may take. The project team will use these workshops to test the draft guide and gather additional examples of civic data partnerships.

The project team will establish means of communication for workshop participants and any other interested parties; workshop attendees will be encouraged to continue communicating and sharing after the events.

#### *Create revised version of guide and toolkit*

The project team will use notes and artifacts generated during the workshops along with feedback provided by workshop participants, partners, and supporting organizations including the Urban Institute, Living Cities, Code for America, and the Sunlight Foundation to make revisions to the guide and toolkit. At this stage, specific examples from other regions will begin to be incorporated.

#### *Field awards for testing and activating the guide, contributions of case studies*

Beginning in year two, the project team will make contract "field awards" to five libraries with existing or strong potential roles in their civic data ecosystem. These awards are modeled in part after the IMLS Sparks! Ignite program, in which libraries are encouraged "to test and evaluate specific innovations in the ways they operate and the services they provide;"<sup>17</sup> in this

---

<sup>17</sup>IMLS (2016). [Sparks! Ignition Grants for Libraries](#).

case the project will support libraries with ideas for testing and further developing aspects of work represented in the guide. Awards will be determined through an separate application process that will favor libraries working in close partnership on local civic open data with a mix of academic, public, and non-library data intermediaries. Applicants will be provided with a set of guiding principles for the kinds of work the project desires to support, and then will be asked to provide a budget and brief narrative specifying how support would be used. It is likely, though not required, that the awardees will come from among the libraries who have already participated in the workshops in year one; the project team will strive to make awards that represent a diversity of geographic, population, and institutional makeups. Each awardee will agree to contribute a case study, documenting their local experience, as well as participate in regular communications of this cohort, described below. The case studies will be featured in the guide to illustrate the variety of library roles in civic data partnerships. The project team will work with applicants to ensure that proposed activities align with the guiding principles.

*Grow a national network of shared practice and organizational capacity*

The establishment of a connected network of libraries participating in their local data ecosystems is a key outcome of the project. Through such a network, librarians participating in the project can provide encouragement and mentorship to others looking to become involved in their civic data ecosystems. The design of the workshops, in particular their connection to established professional communities of practice, as well as the distribution of field awards are intended to directly support this growth. And, while the project will emphasize ongoing connections between its direct participants, it also seeks to reach a wider audience and provide broader opportunities for participation. To do this, the project will establish a communications infrastructure that will include a mailing list / group, monthly open conference calls, and a project blog, featuring contributions from the project team as well as from the project's regional participants. To effectively cultivate these communities team members will take active roles such as regularly monitoring and responding to online discussions, soliciting reports and agenda items for monthly calls, and soliciting and working with participant contributions to the project blog. Participants from NNIP will also serve as ambassadors with their networks of practice informing other members of collaborations with libraries.

*Disseminate findings, arrange transfer of guide responsibility*

Although the project's guide and toolkit will be open and available during the course of the project, and a great deal of communication and outreach is built into the project plan already described, in year two project team members will make special efforts to publicize a summary form of the project work and findings. Project members are targeting four conferences for this purpose; these may include the Coalition for Networked Information meeting in winter 2018, the NNIP Partnership Meeting in 2019, DLF Forum 2019, Code for America Summit, What Works Cities Summit, and Research Data Access and Preservation (RDAP) Summit. The project team will

also make use of its network of partners and supporters, including the Urban Institute, NNIP, DLF, Code for America, Living Cities, Libraries+Network, and The Sunlight Foundation to spread word of the project's work.

Also in year two, the project team will work with its supporters to seek an organizational or institutional host to take responsibility for ongoing promotion and maintenance of the guide after the conclusion of the project. Examples of possible homes are the DLF's Government Records Transparency and Accountability Group, or some part of the emerging Libraries+Network, with support from the Association of Research Libraries. Project team members will be participating in both of these groups during the course of the project term.

### **Milestones and Performance Measures**

The project will use the following milestones and measures of success (Year 1 or 2 identified):

1. Creation and initial publication of a guide and toolkit in open format (Y1)
2. Use of the guide and toolkit, with positive user assessment, to deliver two workshops, at national conferences, to at least ten civic open data partnerships (Y1)
3. Award of five field awards promoting use and development of the guide and toolkit, with positive user assessment, over several months by a variety of regional civic data partnerships (Y2)
4. Revision and expansion of guide and toolkit based on user testing and feedback (Y1,Y2)
5. Identifiable use of the guide and toolkit by libraries in regions not participating directly in the project (Y2)
6. Establishment and growth of communications between libraries participating in this project, demonstrating a sustaining community of practice (Y2)
7. Promotion of project work by team members speaking and writing at national forums (Y2).

### **Project Sustainability**

Sustainability is addressed in several ways. First, through the deliberately distributed activities in the project design, which aim to grow a national network and emphasize local capacity-building among a variety of regional participants. At the successful conclusion of the funding period, each participating region should be well-positioned to carry on activities beyond the term of the project, and moreover should have a network of allies performing similar work. Second, the project's primary output, the guide and toolkit, will be published in an open format supporting copying, reuse, and modification. Third, as described above, the project team will seek

an organizational or institutional host to take responsibility for ongoing promotion and maintenance of the guide and toolkit.

### 3. National Impact

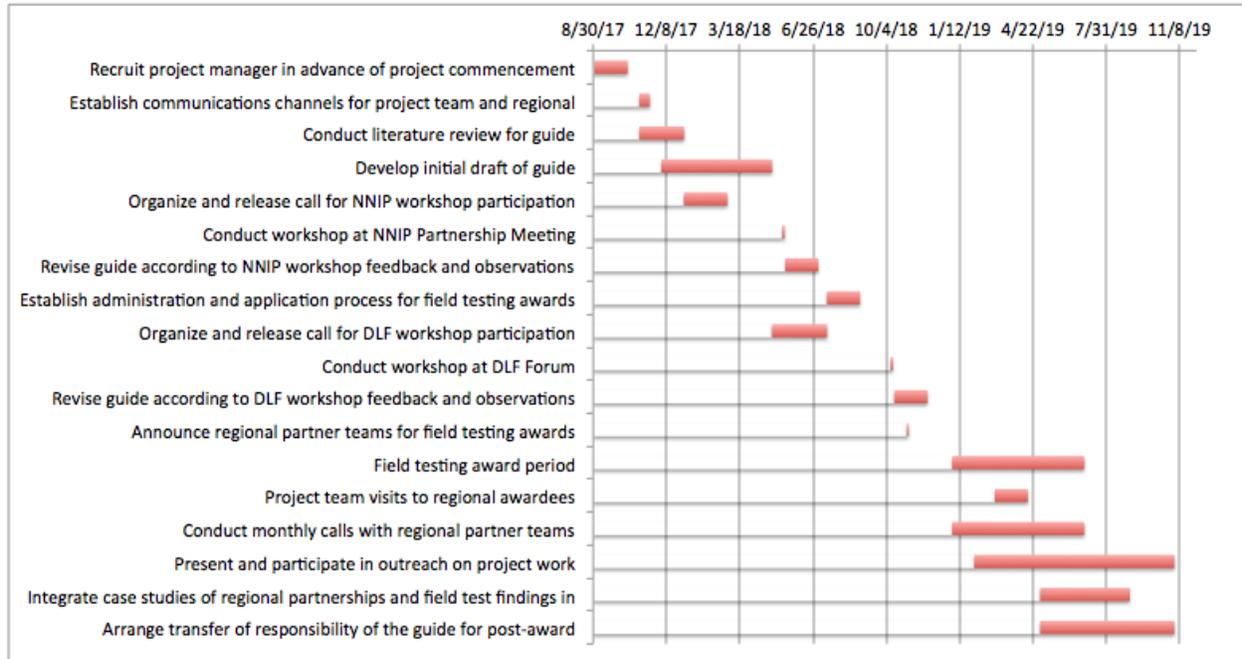
As the amount and availability of available local civic data increases, the need for such data to be accessible, usable, and stewarded grows as well. Much of this work falls to data intermediaries, entities that help the public to apply and use open data to achieve impact by building relationships among publishers, users, and other members of the data ecosystem, improving data, statistical, and technological literacy, enhancing data quality, providing feedback mechanisms to publishers, and building tools that enable broader data use. Libraries, with their history of supporting the information needs of their communities, are poised to contribute tremendous value as local data intermediaries, connecting civic data publishers and users. As local institutions, libraries have significant reach because of their distribution nationwide; the most recent IMLS Public Library Survey indicates that "more than 306 million Americans lived within a public library service area in 2014."<sup>18</sup> This project will meet the selected IMLS program goal, to strengthen libraries as essential partners in addressing the needs of their communities, by:

- Directly supporting, through in-person workshops, at least ten regional partnerships made up of librarians and other civic data intermediaries to become closer and more effective participants in their local civic data ecosystems;
- Directly supporting, through field awards, at least five regional partnerships experimenting with expanding library roles in civic data and capturing these in case studies;
- Producing a guide and toolkit, incorporating a variety of regional experiences, that will enable libraries nationwide to better understand their local civic data landscape, to contribute libraries' expertise and resources to local civic data activities; and to more effectively partner with other libraries and data intermediaries while doing so. Ultimately, this project output will guide libraries towards specific actions that will benefit both data publishers and users in their local context.
- Growing a nationwide network of shared practice for libraries working with local civic open data that can be sustained by the participants after the project's funded term.

---

<sup>18</sup> IMLS (2014). [Public Libraries Survey](#).

## Developing Public and Academic Libraries as Key Participants in Civic Open Data Ecosystems Schedule of Completion



| Task Description  | Start Date | End Date |
|---|------------|----------|
| Recruit project manager in advance of project commencement                | 9/1/17     | 10/15/17 |
| Establish communications channels for project team and regional partners  | 11/1/17    | 11/15/17 |
| Conduct literature review for guide                                       | 11/1/17    | 1/1/18   |
| Develop initial draft of guide  | 12/1/17    | 5/1/18   |
| Organize and release call for NNIP workshop participation                 | 1/1/18     | 3/1/18   |
| Conduct workshop at NNIP Partnership Meeting                              | 5/15/18    | 5/18/18  |
| Revise guide according to NNIP workshop feedback and observations         | 5/19/18    | 7/3/18   |
| Establish administration and application process for field testing awards | 7/15/18    | 8/29/18  |
| Organize and release call for DLF workshop participation                  | 5/1/18     | 7/15/18  |

|  |          |          |
|--|----------|----------|
| Conduct workshop at DLF Forum                                    | 10/10/18 | 10/13/18 |
| Revise guide according to DLF workshop feedback and observations | 10/15/18 | 11/29/18 |
| Announce regional partner teams for field testing awards         | 11/1/18  | 11/2/18  |
| Field testing award period                                       | 1/2/19   | 7/1/19   |
| Project team visits to regional awardees                         | 3/1/19   | 4/15/19  |
| Conduct monthly calls with regional partner teams                | 1/2/19   | 7/1/19   |
| Present and participate in outreach on project work              | 2/1/19   | 11/1/19  |

## DIGITAL PRODUCT FORM

### Introduction

The Institute of Museum and Library Services (IMLS) is committed to expanding public access to federally funded digital products (i.e., digital content, resources, assets, software, and datasets). The products you create with IMLS funding require careful stewardship to protect and enhance their value, and they should be freely and readily available for use and re-use by libraries, archives, museums, and the public. However, applying these principles to the development and management of digital products can be challenging. Because technology is dynamic and because we do not want to inhibit innovation, we do not want to prescribe set standards and practices that could become quickly outdated. Instead, we ask that you answer questions that address specific aspects of creating and managing digital products. Like all components of your IMLS application, your answers will be used by IMLS staff and by expert peer reviewers to evaluate your application, and they will be important in determining whether your project will be funded.

### Instructions

You must provide answers to the questions in Part I. In addition, you must also complete at least one of the subsequent sections. If you intend to create or collect digital content, resources, or assets, complete Part II. If you intend to develop software, complete Part III. If you intend to create a dataset, complete Part IV.

## PART I: Intellectual Property Rights and Permissions

**A.1** What will be the intellectual property status of the digital products (content, resources, assets, software, or datasets) you intend to create? Who will hold the copyright(s)? How will you explain property rights and permissions to potential users (for example, by assigning a non-restrictive license such as BSD, GNU, MIT, or Creative Commons to the product)? Explain and justify your licensing selections.

This output of this project, a guide to assist libraries in building successful civic data partnerships, is designed to be freely used. We will assign the guide a Creative Commons license, specifically a "CC By" license, which is described as the "most accommodating" license and would support "maximum dissemination and use" of the guide.

**A.2** What ownership rights will your organization assert over the new digital products and what conditions will you impose on access and use? Explain and justify any terms of access and conditions of use and detail how you will notify potential users about relevant terms or conditions.

The CC By license will encourage free sharing and use of the guide. The guide will be marked by the CC By license, which will notify users that they have permission to share and adapt the guide freely on the condition that they provide attribution to the original. No limitations will be placed on the guide based on the copyright of the project team members.

**A.3** If you will create any products that may involve privacy concerns, require obtaining permissions or rights, or raise any cultural sensitivities, describe the issues and how you plan to address them.

There are no privacy concerns associated with this guide. We will inform and request permission to feature from regional libraries and civic data intermediaries in the guide as case studies of models for civic data partnerships.

## Part II: Projects Creating or Collecting Digital Content, Resources, or Assets

### A. Creating or Collecting New Digital Content, Resources, or Assets

**A.1** Describe the digital content, resources, or assets you will create or collect, the quantities of each type, and format you will use.

The central digital output that we will produce is an open, versionable guide. One guide, continually updated and accessible during the life of the project, will be created. The project team will keep digital notes from workshops and team meetings.

**A.2** List the equipment, software, and supplies that you will use to create the content, resources, or assets, or the name of the service provider that will perform the work.

We plan to use an open and versionable platform like GitBook or a GitHub repository with a static site overlay for the creation of the digital guide. The project team will use a collaborative file sharing tool like Google Docs to record project team notes, including observations from the NNIP and DLF workshops.

**A.3** List all the digital file formats (e.g., XML, TIFF, MPEG) you plan to use, along with the relevant information about the appropriate quality standards (e.g., resolution, sampling rate, or pixel dimensions).

GitBook, a tool we are assessing for use, uses Git to manage text files written with Markdown syntax.

## **B. Workflow and Asset Maintenance/Preservation**

**B.1** Describe your quality control plan (i.e., how you will monitor and evaluate your workflow and products).

The project team will meet monthly as a full group to track overall project activities against the schedule of completion. The project design is based on an iterative model of development with explicit components that focus on revising and adjusting our primary product. The entire project team will make use of web-based task tracking tools and file sharing platforms with which they have existing experience.

**B.2** Describe your plan for preserving and maintaining digital assets during and after the award period of performance. Your plan may address storage systems, shared repositories, technical documentation, migration planning, and commitment of organizational funding for these purposes. Please note: You may charge the federal award before closeout for the costs of publication or sharing of research results if the costs are not incurred during the period of performance of the federal award (see 2 C.F.R. § 200.461).

The open platform selected for the guide will accommodate copying and web harvesting. Additionally, Gitbook, a tool that we may select as a platform for the guide, allows the digital product to be exported as a pdf. For preservation purposes, we will deposit a PDF version of the guide, with a clear capture date, in D-Scholarship@Pitt, the University of Pittsburgh's institutional repository. Also in year two, the project team will work with its supporters to seek an organizational or institutional host to take responsibility for ongoing promotion and maintenance of the guide after the conclusion of the project.

## **C. Metadata**

**C.1** Describe how you will produce any and all technical, descriptive, administrative, or preservation metadata. Specify which standards you will use for the metadata structure (e.g., MARC, Dublin Core, Encoded Archival Description, PBCore, PREMIS) and metadata content (e.g., thesauri).

The metadata associated with the version deposited in D-Scholarship@Pitt will clearly indicate the date for the version. Metadata for the guide that lives the open platform (likely GitBook) will include elements such as title, date of last update and date of first publication, authors and contributors, license, and description.

**C.2** Explain your strategy for preserving and maintaining metadata created or collected during and after the award period of performance.

The metadata that we will maintain is basic descriptive metadata that will live beside the guide. The project team will be developing a sustainability plan for the stewardship of the guide, which will include the maintenance of associated metadata.

**C.3** Explain what metadata sharing and/or other strategies you will use to facilitate widespread discovery and use of the digital content, resources, or assets created during your project (e.g., an API [Application Programming Interface], contributions to a digital platform, or other ways you might enable batch queries and retrieval of metadata).

The primary digital content is the guide itself; project partner The Urban Institute and supporters such as The Sunlight Foundation, Code for America, and the Digital Library Federation will promote the initial work and help identify other civic data partnerships that the team will recruit to participate in the workshops and further contributions to the guide.

## **D. Access and Use**

**D.1** Describe how you will make the digital content, resources, or assets available to the public. Include details such as the delivery strategy (e.g., openly available online, available to specified audiences) and underlying hardware/software platforms and infrastructure (e.g., specific digital repository software or leased services, accessibility via standard web browsers, requirements for special software tools in order to use the content).

The guide and toolkit will be openly available online, using GitBook or a GitHub repository with a static site overlay.

**D.2** Provide the name(s) and URL(s) (Uniform Resource Locator) for any examples of previous digital content, resources, or assets your organization has created.

WPRDC Data User Guides: <http://www.wprdc.org/data-user-guides/>; WPRDC Metadata Guidelines: <http://www.wprdc.org/documenting-data/>; WPRDC Tutorials: <http://www.wprdc.org/tutorials-2/>; ULS Data Management: <http://www.library.pitt.edu/data-management> and <http://pitt.libguides.com/managedata>; NNIP Publications: <http://www.neighborhoodindicators.org/library/catalog>; NNIP Training Resources: <http://www.neighborhoodindicators.org/data-tech/training>

## **Part III. Projects Developing Software**

### **A. General Information**

**A.1** Describe the software you intend to create, including a summary of the major functions it will perform and the intended primary audience(s) it will serve.

Part III is not applicable as no software will be created during the course of this project.

**A.2** List other existing software that wholly or partially performs the same functions, and explain how the software you intend to create is different, and justify why those differences are significant and necessary.

### **B. Technical Information**

**B.1** List the programming languages, platforms, software, or other applications you will use to create your software and explain why you chose them.

**B.2** Describe how the software you intend to create will extend or interoperate with relevant existing software.

**B.3** Describe any underlying additional software or system dependencies necessary to run the software you intend to create.

**B.4** Describe the processes you will use for development, documentation, and for maintaining and updating documentation for users of the software.

**B.5** Provide the name(s) and URL(s) for examples of any previous software your organization has created.

### **C. Access and Use**

**C.1** We expect applicants seeking federal funds for software to develop and release these products under open-source licenses to maximize access and promote reuse. What ownership rights will your organization assert over the software you intend to create, and what conditions will you impose on its access and use? Identify and explain the license under which you will release source code for the software you develop (e.g., BSD, GNU, or MIT software licenses). Explain and justify any prohibitive terms or conditions of use or access and detail how you will notify potential users about relevant terms and conditions.

**C.2** Describe how you will make the software and source code available to the public and/or its intended users.

**C.3** Identify where you will deposit the source code for the software you intend to develop:

Name of publicly accessible source code repository:

URL:

### **Part IV: Projects Creating Datasets**

**A.1** Identify the type of data you plan to collect or generate, and the purpose or intended use to which you expect it to be put. Describe the method(s) you will use and the approximate dates or intervals at which you will collect or generate it.

The workshops will generate qualitative data artifacts. Findings from those data will be incorporated into the guide. The data themselves are not considered a product of this project.

**A.2** Does the proposed data collection or research activity require approval by any internal review panel or institutional review board (IRB)? If so, has the proposed research activity been approved? If not, what is your plan for securing approval?

This project will not be studying human subjects but rather organizations and organizational structures.

**A.3** Will you collect any personally identifiable information (PII), confidential information (e.g., trade secrets), or proprietary information? If so, detail the specific steps you will take to protect such information while you prepare the data files for public release (e.g., data anonymization, data suppression PII, or synthetic data).

No, this project is focused on organizations and models for civic data partnerships rather than on individuals. We will not collect any trade secrets or proprietary information during the course of the project. The field award recipients will be asked to write a case study on their regional civic partnership and the project team will clearly communicate that the case studies will be integrated into the guide. We will obtain permission to disseminate these case studies in writing and will ask awardees to avoid recording information that may be sensitive.

**A.4** If you will collect additional documentation, such as consent agreements, along with the data, describe plans for preserving the documentation and ensuring that its relationship to the collected data is maintained.

The project team will ask field awardees to complete the University of Pittsburgh paperwork that is necessary for disbursement of funds. The agreement between the University and the awardee will be submitted to and maintained by the University of Pittsburgh's Purchasing Services Department.

**A.5** What methods will you use to collect or generate the data? Provide details about any technical requirements or dependencies that would be necessary for understanding, retrieving, displaying, or processing the dataset(s).

The workshops will produce observational data, recorded by a project team member.

**A.6** What documentation (e.g., data documentation, codebooks) will you capture or create along with the dataset(s)? Where will the documentation be stored and in what format(s)? How will you permanently associate and manage the documentation with the dataset(s) it describes?

The guide will report on observational qualitative data from workshops at NNIP and DLF. We will document the structure of the workshop and corresponding observations using Google Docs.

**A.7** What is your plan for archiving, managing, and disseminating data after the completion of the award-funded project?

The qualitative observations will be incorporated as findings in the digital guide and toolkit.

**A.8** Identify where you will deposit the dataset(s):

Name of repository: N/A

URL: N/A

**A.9** When and how frequently will you review this data management plan? How will the implementation be monitored?

The file management approach for this project will be reviewed on a biannual basis to ensure that all project team members are comfortable and aware of its contents.