

A comprehensive, step-by-step guide to help libraries release open data



BACKGROUND

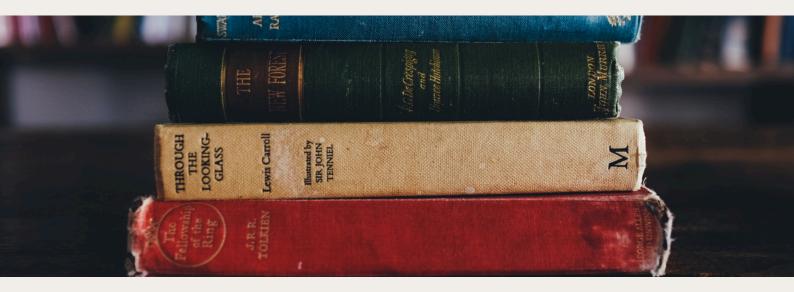




This project was made possible in part by the Institute of Museum and Library Services, IMLS National Leadership Grant for Libraries, LG-254933-OLS-23.

Toolkit creators Jennifer Ashby and Mary Neuman of Asotin County Library, Clarkston, Washington, would like to thank the <u>project</u> <u>participants</u> for their contributions to this guide.

BACKEROUND, CONT.



The Open Data Toolkit is designed to help small and rural libraries:

- Understand the basic concepts of open data
- Plan and implement an open data publishing program
- Offer open data education to their communities

This toolkit came about as the result of grants obtained over an eight-year period. The first was a Knight Foundation grant received in 2017 by the State of Washington Office of the Chief Information Officer and the California State Library. Asotin County Library was asked to participate in this project, which developed the Data Equity for Main Street curriculum. While teaching the curriculum, we discovered that there was little to no open data for the agencies within Asotin County.

We applied for and received an Accelerating Promising Practices for Small Libraries grant through the Institute of Museum and Library Services (IMLS), <u>LG-28-19-0229-19</u>, Asotin County Library Open Data Project. The focus of this grant was to publish datasets relevant to Asotin County. We developed a process that worked well for our small library and wanted to explore how it could work for other libraries. We followed up with an IMLS National Leadership Grant for Libraries, <u>LG-254933-OLS</u>, Expanding Open Data Opportunities for Small Libraries, which enabled us to create this toolkit.

Getting started with open data can be easier than many people think. If you have data and an online, publicly accessible place to put it, then you are on the right path.

Click the link below to view the Open Data Pitch video, prepared by University of Washington MSIM Students.





CONTENTS

O6 Learn About Open Data

What is Open Data?
Why is Open Data Important?
Why Should Libraries Be Involved?
Teaching Open Data Literacy

12 Getting Started

Administrative Support
Open Data Policy
Sample Data Policy



Four Basic Steps of Open Data Publication
Choosing a Dataset
Focus Groups

<u>Sample Focus Group Meeting Agenda</u> <u>Attaching an Open License</u>

Formatting the Data

• <u>Data Ingestion Tool</u>

Publishing and Making Data Discoverable

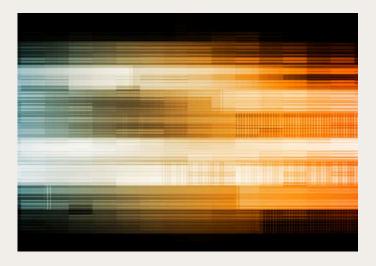
- Government Portal
- Library Website
- GItHub and Kaggle



• Visualization Tools

, Abandon per Team

- Visualization Tutorials
 - <u>Power Bl vs. Tableau Comparative</u> <u>Analysis</u>
 - Power Bl Basics Part 1
 - Power Bl Basics Part 2
 - Tableau Bar and Pie Chart Tutorial
 - Tableau Map Visualization Tutorial
- Next Steps



48 Additional Resources and Acknowledgements

- Additional Resources
 - Websites
 - Blog Posts and Articles
- <u>Glossary</u>
- Closing
- Project Participants





LEARN ABOUT OPEN DATA







WHAT IS OPEN DATA?

Open data can help people answer questions and solve problems. It generally refers to data collected by a governmental entity. To be considered open, it must meet the conditions in the definition created by opendefinition.org. "Open data and content can be freely used, modified, and shared by anyone for any purpose."

O1 Can Be Freely Used

...can be freely used means that there is no monetary cost associated with use. It also means that the data must be in a machine-readable format, such as CSV (comma separated values) which can be read by a computer and sorted, analyzed, and displayed in various ways by the user. A PDF file is not machine readable and is therefore not open data; we can read it, but the computer sees it as an image. Freely used also means that one has the ability to use the data for commercial and non-commercial purposes. No login or personal account should be needed to access the data.

02 Can Be Modified

...can be modified means that you can use open data to create a new product by presenting the data in a new way, combining data with other information, or adding additional datasets.

Can Be Shared

03

...can be shared by anyone for any purpose means that there should be no restrictions on how and by whom open data is used. Open data publishers include a license that clearly indicates that the data is available for anyone to use. The publishers often ask users to attribute the original source and/or use the same license for any new product made. Some examples of licenses include U.S. Government works, certain Creative Commons Licenses and works in the Public Domain.

WHY IS OPEN DATA IMPORTANT?



Information

03

It provides information that people need to make decisions about their personal lives as well as the health and well-being of the places they live. In addition, open data is key to government transparency and accountability.

Beneficial

02

Easy access to open government data is beneficial for communities of any size.

Engagement and equity

Easy access to open data enables civic engagement and furthers digital equity.



Executive Order 1364 Dated May 9, 2013

This 2013 executive order mandates that governmental agencies maintain their data in an open format, "Openness in government strengthens our democracy, promotes the delivery of efficient and effective services to the public, and contributes to economic growth. As one vital benefit of open government, making information resources easy to find, accessible, and usable can fuel entrepreneurship, innovation, and scientific discovery that improves Americans' lives and contributes significantly to job creation".1

OPEN DATA IN ACTION

Personal Decision Making



Trying to decide which college offers the most value? Use <u>College</u> <u>Scorecard</u> to compare college costs and outcomes.



2

Looking for a free or cheap place to camp? You'll probably need weather information, maps, and campground info.









Which is the best **nursing home** for Aunt Edna?



Solving Public Problems

Here's an <u>example</u> of open data used as evidence in a court case that resulted in a large settlement to residents of a Black neighborhood that had been denied water service for decades.



Enables Civic Engagement

Residents of New York City help count and map trees on public property and log their continuing efforts in caring for the trees. Click <u>here</u> to see the New York City Tree Map.



WHY SHOULD LIBRARIES BE INVOLVED?

Libraries are perfectly positioned to take the lead in publishing open data for the following reasons:





Access

Access to information is core to the mission of public libraries

2

Experts

Library staff are experts at collecting, organizing and providing access to information



Trusted

Libraries are trusted by their communities



Connections

Libraries have established connections to local governments



Literacy Education

Libraries have a long-established role in teaching literacy



Digital Inclusion

Libraries value digital inclusion and are well placed to play a strategic role in helping users navigate the expanding data landscape

"We didn't build libraries for a literate citizenry. We built libraries to help citizens become literate. Today we build open data portals not because we have public policy literate citizens, we build them so that citizens may become literate in public policy." ²

- David Eaves

WHY SHOULD LIBRARIES BE INVOLVED?, CONT.

In many parts of the country, large libraries have taken a role in publishing open data in partnership with their cities' technology departments. An example of this is Chapel Hill Public Library, which initiated an open data project with the Town of Chapel Hill in 2016 to create Chapel Hill Open Data.

Small libraries do not usually have those kinds of resources. It is vitally important that they not be left behind in the open data arena.

Libraries publish open data in order to encourage greater public participation, government transparency, and further expand on the library's role as a collector of materials to spark greater civic engagement.

To remain relevant to their communities, libraries cannot merely be repositories of collections, but must instead offer information and materials with which their communities can create something of value.

TEACHING OPEN DATA LITERACY



More information about open data and its value can be found in the <u>Data Equity for Main Street</u> project, which has developed a curriculum to promote open data literacy by training librarians and community members to find, use, and give feedback about open data.



Examples of how libraries can foster open data knowledge in their communities can be found in <u>The Roles of Libraries in Open Data Literacy</u>, a presentation by <u>Open Data Literacy</u>, a collaborative project led by the Information School at the University of Washington.



GETTING STARTED





Now that you have a basic understanding of open data and the role of libraries in its publication, it's time to determine whether or not you have the budget, time and capacity to do an open data project. Here is a blueprint that will guide you through the process of publishing, beginning with determining if your library has the capacity for this work.

Administrative Support

Before beginning an open data project, it is essential to ensure that the library has adequate capacity in the following areas:



Ayoung Yoon and Andrea Copeland of Indiana University developed a rubric to use as a guide in evaluating a library's capacity for an open data project as part of the Institute of Museum and Library Services grant "Data Reuse for Local Community" <u>LG-96-17-0184-17</u>. Click <u>here</u> to begin the assessment process.

Open Data Policy

Once it has been determined that your library is ready to publish open data, the next step is to develop an open data policy. This will be similar to, but more specific than, a collection development policy. A library's open data policy outlines its commitment to making data publicly accessible, usable, and reusable, often with the goal of increasing transparency and community engagement. This policy typically covers how the library collects, stores, and shares data, as well as the data's format and licensing.

Here is a more detailed look at the elements which should be included in an open data policy:

Open Data Definition and Principles (S): Privacy and Security Data Publication

Data Format and Metadata

Data Quality Licensing

Exemptions

Data Stewardship

Community Engagement

Open Data Definition and Principles

Include the standard definition of open data: Open data and content can be "freely used as """ content can be "<u>freely used, modified, and shared by anyone for</u>

An open data policy ensures that government-held data is made publicly available for unrestricted use and reuse, promoting transparency and accountability. It requires data to be published in a machine-readable format, accessible without charge or licensing restrictions, and accessible to anyone without registration or other barriers.

Data Publication

Specify where the data will be published, e.g., an open data portal, directly on library website.

Data Format and Metadata

Data should be presented in a standard, machine-readable format, along with relevant metadata and documentation to aid in understanding and use. Each published dataset should be associated with contact information for the responsible party, as well as a data dictionary explaining the column headings, units of measure, etc.

Data Quality

An open data policy should include provisions for ensuring the accuracy and quality of the published data. Focus on formatting accuracy and quality of the published data. Focus on formatting, sourcing, and timeliness, keeping in mind the credibility, accuracy, completeness, consistency, and uniformity of the data.

Open licenses should be used to confirm that the data can be reused and redistributed. Open data policies often include a clear statement that the data is in the public domain or is licensed under a permissive license like Creative Commons CCO.

Privacy and Security

An open data policy should address the handling of personal information and ensure compliance with privacy laws and regulations.

Exemptions

An open data policy should clearly define which types of data are exempt from open access, such as confidential financial records or personally identifiable information.

Data Stewardship

The policy should outline procedures for data curation, quality control, and long-term preservation. A template to create a Catalog of Published Datasets is included in this toolkit to assist in keeping a complete record of each published dataset.

Community Engagement

Offering training and support to patrons and staff on how to use and interpret open data will help further the use of the open data collection. Libraries can use open data to engage with the public, demonstrate their value, and foster collaboration with local government and other organizations. Seeking public perspectives helps ensure that the published datasets are relevant to the community.

SAMPLE OPEN DATA POLICY

Click the document below to access a sample open data policy.





Open Data Policy

Scope

This policy applies to the Asotin County Library, Asotin County, Washington.

Purpose

To provide guiding principles by which Asotin County Library (ACL) publishes local open data.

Policy

Open data is defined as data that can be freely used, modified, and shared by anyone for any purpose" https://opendefinition.org/ Asotin County Library is committed to making library data publicly available for unrestricted use and reuse, thereby ensuring transparency and accountability. In addition, Asotin County Library cooperates with local government agencies and community organizations to acquire and openly publish other locally relevant data. The data is published in a machine-readable format, accessible without charge or licensing restrictions, and accessible to anyone without registration or other barriers. Datasets are published on the Washington State Open Data Portal at https://data.wa.gov and also on the Asotin County Library website at https://asotincountylibrary.org/open-data



BASIC STEPS OF OPEN DATA PUBLICATION





FOUR BASIC STEPS OF OPEN DATA PUBLICATION

The four basic steps in publishing open data are:

Choosing a dataset.



Formatting the data.



Publishing and making discoverable.



1: CHOOSING A DATA SET



The key to a successful project is getting high quality, relevant datasets. Working with governmental agencies and community organizations, as well as local residents, will lead to obtaining data that matters.

Libraries may choose to start by publishing data about their operations, collections, programming, and budget. Beginning with this easily obtainable data allows the opportunity to get familiar with the Data Ingestion Tool and other resources in this toolkit. The next step would be to connect with organizations within your community. For some ideas on partners and the kinds of data they could contribute, please see the List of Community Partners. Involving the general public by holding a focus group is a good way to assess community needs and interests.

COMMUNITY PARTNERS

Expanding upon existing relationships, as well as building new ones, is the key to a successful open data project. We found that multiple conversations were needed with some governmental agencies to reassure them of the following:

- Their data was not at risk of being altered
- (💽 The data could be delivered in any format (Excel, pdf, handwritten)
- Context could be provided to explain any anomalies
- Data could be anonymized to protect the privacy of individuals

To start the conversations with potential partners, the tools below can be helpful:

- List of Community Partners
- Sample Phone Call
- Open Data Talking Points
- Follow-up Email
- Data Inventory Form (Click on "Use Template" to make your own copy that can be customized and sent to partners).
- Open Data Memorandum of Understanding

MMUNITY PARTNERS, CONT.



CASE SSSS GLADWIN COUNTY DISTRICT LIBRARY STUDY

GLADWIN COUNTY

John Clexton, Library Director President of the Association of Rural and Small Libraries

When the Open Toolkit project was introduced, I saw it as a timely and collaborative effort to gather our data, analyze it, and share our story. The opportunity to include the library in this initiative was particularly valuable, as it allowed our county to explore partnerships with other entities, enabling us to collectively share data and narratives.

The first step was to make sense of our numbers. After figuring out how to input the data we already had, with the work of the "Open Data Toolkit" it created a simple system to organize and interpret it. I presented this idea to our library board, which had many questions regarding access and responsibilities. However, they were very receptive to the concept. Once things were explained and so forth, they were all for it.

The next challenge was to reach out to selected county partners, such as the Gladwin Fire Department, Gladwin County Offices, and the Gladwin Sheriff's Department. This project was part of a larger oral presentation to major county officials, including city managers and emergency personnel during various public speaking engagements. There was little hesitation or resistance in obtaining data; in fact, much of the information we sought was already available in their files, online, or in some obscure computer systems... (continued on next page).

When the Open Toolkit project was introduced, I saw it as a timely and collaborative effort to gather our data, analyze it, and share our story

CASE STUDY, CONT.



We are currently collecting this information and incorporating it into our databases so we can provide the public with a centralized source of information. Initially, I was apprehensive, thinking there would be resistance. However, when I started making phone calls and meeting with officials, they turned out to be more than willing to share information. Most of the data is required by law in Michigan, and the officials were generally appreciative and helpful in providing it.

The most significant outcome of this project has been the establishment of previously unexplored partnerships. Not only were they willing to provide information, but there were also discussions about how the library and their departments could collaborate more effectively. For instance, I recently met with our county emergency manager, who expressed a desire to provide more data than I could handle and indicated interest in conducting other collaborative programming.

The process of obtaining buy-in for this amazing Toolkit project was much easier than I had anticipated. I now feel a remarkable sense of cohesiveness among the various departments in the county as we move toward a collaborative approach. Information is a powerful tool and having the correct information and all in one spot, is gold.



WASHINGTON STATE LIBRARY

Kathleen Sullivan, Open Data Librarian

Open data helps a community understand itself and connects government to the public it serves. Like most valuable information resources, though, open data needs time and care to do well – along with someone to get it started. This toolkit makes it possible for any library to play that role, with clear steps, a just-right selection of resources and examples, and tips for heading off common snags.

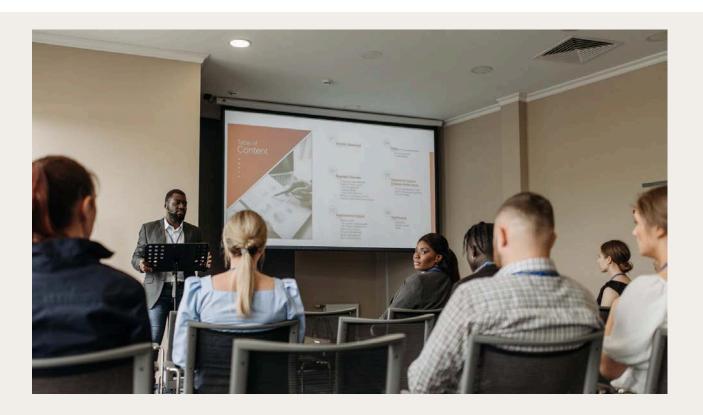
Yohn Clextor

FOCUS GROUPS



Have you ever wondered what your community really wants, and how they might be able to solve those problems collaboratively? Focus groups enable the library to determine the issues and datasets that are most important to stakeholders. The basic components of a focus group are:

- Scheduling
- 💽 Logistics: Zoom or in-person
- identify the sample: aim for a broad representation of community sectors (education, health, business, civic organizations); include as many non-library users as possible; strive for diversity
- Recruit participants: include non-library users. We asked some of our library supporters to invite a non-library user friend.
- 💓 Identify and train a facilitator and an assistant who will take notes
- Develop a discussion guide with open-ended questions (example below)
- Convene the focus group
- 💓 Analyze the results







Below is the general outline, including Discussion Guide, that Asotin County Library followed when we held a focus group to determine the topics in which our community was interested.

- Opening
 - ·Welcome and introductions: facilitator, assistant and participants
 - ·Review purpose of discussion
 - ·Go over ground rules
 - ·Show YouTube video <u>Hans Rosling's 200 Countries</u>, <u>200 Years</u>, <u>4 Minutes</u> <u>The Joy of Stats</u> (we showed this because it is a compelling example of the power of open data)
- O Discussion
 - ·Define open data, show <u>data.wa.gov</u> as an example of what published open data looks like
 - ·Ask questions in discussion guide, making sure all participants have the opportunity to speak (see discussion guide below)
 - ·Summarize comments after each question
 - ·Explore unexpected comments if they add to the discussion
 - ·Ask participants to rank their top three choices of datasets they'd like to see published
- Wrap Up
 - Provide one last chance for comments
 - ·Summarize the major points
 - ·Thank the participants (we gave each person a small gift card donated by our Friends of the Library group)





Discussion Guide

- ·Have you ever tried to find government data online?
- ·Any local data? (911, traffic data, restaurant inspections)
- ·What issues most concern you about Asotin County and/or the Lewis Clark Valley?
- ·What would you like to see get better in the next five years?
- ·What kind of data could we use to bring about this change/answer that question? AND/OR Who might have information (data) that could help us answer that question?
- ·What's one question, even if it seems silly, that you would like to answer about our area?
- · Can you think of any types of data that you would be interested in combining to inform your understanding of a topic?



2: ATTACHING AN OPEN LICENSE





WHAT IS AN OPEN LICENSE?

An open license outlines what can and cannot be done with a work (whether dataset, sound, text, image or multimedia). In broad terms, an open license grants permission to access, re-use and redistribute a work with few or no restrictions. The most common licenses used for open data include: Public Domain, Attribution, and Attribution and Share-Alike.

PUBLIC DOMAIN

The organization completely waives all rights to the particular data that it publishes. People can reuse the data without needing to say where it is from, and they can create their own applications or services. A subset of public domain is government public domain; the term refers to works created by the United States federal government, which are automatically in the public domain.





ATTRIBUTION

The organization still waives all rights, and people can reuse the data in any way. The only requirement is that they must attribute the publishing organization as the source of the data.

ATTRIBUTION AND SHARE-ALIKE

As well as attributing the publishing organization as the source, the user of the data must share any derived content or data under the same license.









CREATIVE COMMONS AND OPEN DATA COMMONS LICENSING

Typically, the platform (open data publishing software) being used will give license choices in a drop-down menu. As an example, the Socrata platform (available through Tyler Technologies), used by data.wa.gov gives the option of public domain or government public domain. If there are no prescribed choices, choosing a Creative Commons license is a good option. CC0 (CC Zero) is the designation for public domain. More information on Creative Commons licensing and Open Data Commons licensing can be found by clicking on the links.

<u>Open Data License Metadata Guidance</u>, published by <u>data.wa.gov</u>, is a good overview of the considerations that need to be taken in account when choosing an open license.







3: FORMATTING THE DATA

There are several aspects to readying the data for publication. Before publishing, ensure that you have met the five characteristics of quality data:

- Credibility -reliability of the data source
- Accuracy exact values (or as close to exact as possible. You may need to reduce the number of decimal points displayed for ease of consumption)
- Completeness the degree to which all required data is known Consistency the degree to which data is consistent within the same dataset and/or across multiple datasets
- Uniformity the degree to which the data is specified using the same unit of measure

To simplify formatting the data and creating metadata, the toolkit includes a <u>Data Ingestion Tool</u> developed by Lillian Curanzy as her Capstone Project when obtaining her MLIS degree from the University of Washington iSchool. It has been slightly modified as part of the Expanding Open Data for Small Libraries grant project. It is designed to help you prepare and organize the data prior to uploading it into a data platform. Elements common to all data platforms are represented in the Data Ingestion tool. Completing the tool gathers all metadata and other information, as well as a spreadsheet of the data, in one location so that when you are ready to publish, everything needed is in one document.

3: FORMATTING THE DATA, CONT...



Introducing the Data Ingestion Tool

On the next page, you can launch the tool to create your own copy and watch a short video explaining how to use the tool. Inside the tool, you will see the following worksheet tabs at the bottom. The function of each tab is briefly explained below and in greater detail in the Data Ingestion Tool with Embedded Guidelines.

Column Headings

This tab gives a record of the column names and a description of the information in each column. This information will be included in the published data set.

Record Level

This tab is where the metadata for the dataset is recorded. The information recorded here will be transferred to the open data platform and provides the context needed to understand the data.

Context

This tab allows users to record information that will aid others in updating, data interpretation, and reuse.

Primary Data

This tab holds the raw data as received from the community agency.

Clean Data

This tab holds the data after it has been cleaned. For example, ensuring consistency in capitalizations and abbreviations, checking spelling, removing spaces from column headings, specifying the data type for columns, etc.

Cleaning Protocol

This tab allows users to record the steps required for cleaning in order to streamline updates.

Sustainability Plan

This tab is for in-house use in evaluating the costs (time and money) for preparing the dataset.

Documentation

This tab provides a glossary of terms used in the ingestion tool along with considerations for their use.

For Further Reference, see:

Guidelines for Data Ingestion Tool

Completed Data Ingestion Tool Sample

<u>Data Ingestion Tool with Embedded</u> <u>Guidelines</u> Column Description Worksheet

General Metadata Guidance

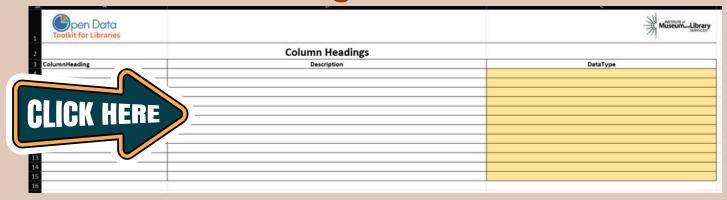
<u>Open Data License Metadata</u> Guideli<u>nes</u>

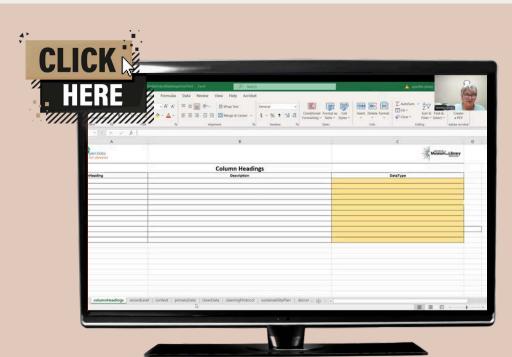




Introducing the Data Ingestion Tool

Data Ingestion Tool





Using the Data Ingestion Tool

4: PUBLISHING AND MAKING DISCOVERABLE



Your dataset has been gathered from a community partner, cleaned and formatted, and metadata carefully chosen. Now it's time to share it with the world!

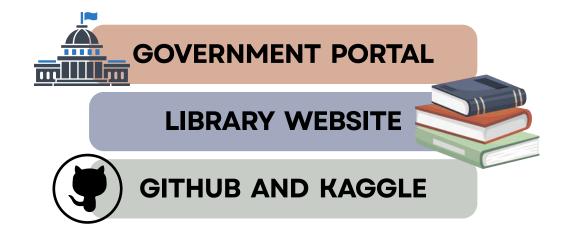
Publishing datasets will be a different process for each library, but certain aspects apply to all. The two key goals are to make the data discoverable and to make it harvestable.

Making open data discoverable is essential for maximizing its value and usability. It involves organizing and presenting data in ways that make it easy to find, access, and understand by a wide range of users, from researchers to policymakers and the general public.

Data harvesting is an automated process "that copies datasets and their metadata between two or more data catalogs—a critical step in making data useful." The harvested data is then organized into a structured format, such as a spreadsheet. An example of a site that utilizes data harvesting from a multitude of federal agencies, as well as state, county and city, is <u>Data.gov</u>.

Key strategies to ensure both discoverability and harvestability include using standardized metadata, adopting common data formats, and ensuring datasets are catalogued in open data portals. By following these dictums, open data initiatives can drive transparency, innovation, and informed decision-making.

The first step is to find a home (portal) for your datasets. The following pages will discuss some options to consider when choosing where to publish your data.









Ideally, there is a state government open data portal that will welcome your datasets. When the Asotin County Library team embarked on our publishing venture, we were given permission to house our datasets on <u>data.wa.gov</u>. It was a surprising setback when we began working with our first partner library, Auburn Public Library in Georgia, and discovered that it wasn't going to be so easy to find an open data portal in every state.

To aid in locating a portal in your state, Kathleen Sullivan, Open Data Librarian at the Washington State Library put together some <u>steps to take</u>. Another option is to contact your state data officer. If you do not know who that is, the state library probably can help you make contact.

If there is access to a state open data portal, check their requirements for metadata and add fields to the Data Ingestion Tool Record Level tab if necessary. Below is a short video that explains how to take the dataset you prepared and publish it with Socrata, which is a common platform used by many open data sites.



Using Socrata to Publish Your Dataset





LIBRARY WEBSITE



If there is no state or regional portal, then what? The datasets can be parked on the library's website and will be easily available to library users and the local community. Some steps can be taken to make the datasets discoverable and harvestable to a certain extent.

Data platforms are expensive and not something a small library would invest in on their own, but when publishing datasets directly onto your library's website, a platform isn't necessary.

However, since the commonly used platforms such as Socrata and OpenDataSoft are programmed to display the metadata associated with each dataset, it is imperative that the same information be included on your library website's open data pages. Without consistent descriptions, categories, column definitions, and other core documentation (metadata), the content will be harder to find, understand, and use.



4: PUBLISH AND MAKE DISCOVERABLE



LIBRARY WEBSITE



A standard spreadsheet program like Excel or Google Sheets can be used on your webpage to display the dataset. This format allows you to include one tab for the dataset itself and another tab for the necessary metadata, ensuring that all required information is published together. The Record Level tab in the Data Ingestion Tool provided in this toolkit can be used for this purpose.

A good practice to follow when publishing datasets to your library website is to develop an Open Data page. The datasets can be set up by category; this structure will aid in search engine optimization. A feedback and suggestion form, such as this <u>example</u> (which you can customize to meet your needs), should also be provided.

Lastly, creating a MARC record for the library's ILS for each dataset creates another avenue for a site like <u>Data.gov</u> to harvest the data and add it to their extensive library of datasets.









GitHub and Kaggle are two possible options to house datasets. Both websites help data look professional, and provide benefits for organizing a catalog, making data easier for people to find and use.

<u>GitHub</u> is a website and cloud-based service that helps developers store and manage their code, as well as track and control changes to their code, but is increasingly being used as a repository for open data.

The features that make it attractive for open data use include:

- **Hosting and Sharing:** GitHub provides a platform for individuals and organizations to store, manage, and share open datasets publicly.
- **Collaboration:** The platform enables collaborative efforts on open data projects. Users can suggest changes, identify issues, and contribute to the development and refinement of datasets.
- **Discoverability:** GitHub's search functionality and topic tagging ("opendata") help users find and access relevant datasets.
- **Version Control:** Git's version control system, integrated within GitHub, allows tracking of changes made to datasets, promoting transparency and reproducibility.
- Government Open Data Initiatives: Government bodies, like the Department of Homeland Security, leverage GitHub to share data and encourage public collaboration.
 - **Specialized Open Data Tools:** Projects like Open Data Editor (ODE), available on GitHub, provide tools to manage and publish open data easily.

To learn more about publishing open data to Github, take a look at these <u>step-by-step instructions</u>.

4: PUBLISH AND MAKE DISCOVERABLE





<u>Kaggle</u> is similar to GitHub in that it allows you to host datasets and collaborate with others. It is user-friendly and provides features such as data documentation and community forums.



Kaggle provides a vast repository of public datasets spanning various domains. These datasets are often cleaned and organized, making them suitable for practicing data science skills and exploring different problem types. Kaggle's commitment to open data aligns with the principles of transparency and accessibility, promoting collaborative learning and innovation.

<u>Github Open Dataset</u> <u>Example</u> <u>Kaggle Open dataset</u> <u>Example</u>



USING DATA AND VISUALIZATIONS TO TELL A STORY





VISUALIZATION TOOLS

OPEN DATA TELLS A STORY



Open data can be used to tell the story of the agency or organization providing the data. Whether it be to show a decrease in funding and subsequent need for additional funds or to show the impact the organization has on the community, data provides factual information which adds credibility to the story being told.

For example, your library can use revenue and expenditure data to show the need for increased funding, and data about circulation, library visits, programs and program attendance to show community impact.

When published openly, data can be used by anyone to tell a story, and, by combining two or more datasets, users can create new meaning.

An example of this would be <u>Walk</u>
<u>Score</u>, a website that measures
pedestrian friendliness by analyzing
population density and road metrics
such as block length and intersection
density. Walk Score is a commercial
endeavor whose data sources
include Google, Factual, Great
Schools, Open Street Map, the U.S.
Census, Localeze, and places added
by the Walk Score user community.



MULTIPLE OPEN DATASETS CAN CREATE NEW MEANING

VISUALIZATION TOOLS, CONT.

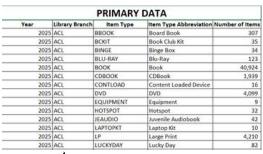
SELECTING THE RIGHT VISUALIZATION



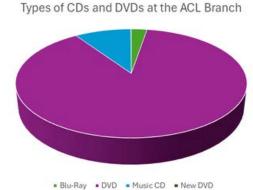
Selecting the right visualization for the information is essential to telling an impactful story. See the Visualization section of the **Data Equity for Main**Street curriculum for discussion of choosing the right visualization and avoiding misleading visualizations.

0 1 1 1 0 0

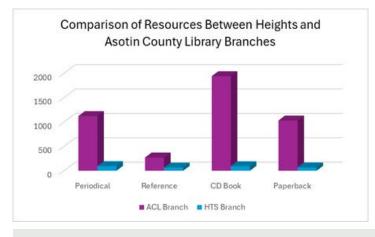
VISUALIZATION EXAMPLES

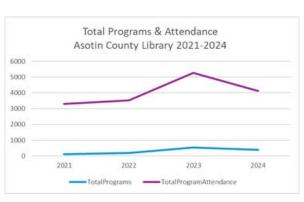












Open data platforms typically have visualization components, as do Microsoft Excel and Google Sheets. If you do not find what you need within the platform, try using external visualization software such as PowerBI or Tableau. Below are some tutorials created by students from the University of Washington's iSchool to give you an overview of these tools.

USING LIBRARY DATA AND VISUALIZATIONS FOR ADVOCACY AND DECISION-MAKING

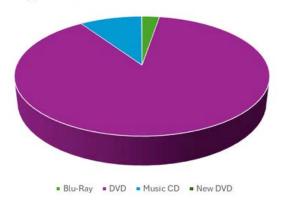


A well-chosen chart, graph or map can bring data to life and help people understand and interact with information more effectively. Instead of scanning hundreds of rows and columns trying to spot trends, a good visualization can bring instant meaning to all those numbers. Visualizations can be used for both advocacy and decision-making.

USING VISUALIZATIONS FOR ADVOCACY

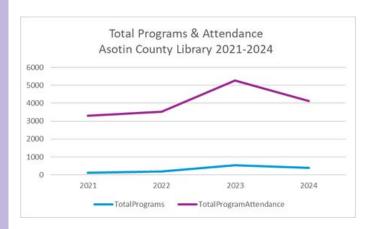
Using visualizations can help stakeholders appreciate the value the library brings to the community. Two common visualizations that can be used include pie charts and line charts.

Types of CDs and DVDs at the ACL Branch



Pie charts show parts of a whole and can illustrate the types of materials being borrowed, thus showcasing the breadth of resources offered and their popularity.

Line charts are good for displaying trends, such as an increase in attendance for programs, resulting in an easy way for people to see the library's role in community engagement and learning. The line graph on the right shows a rebound in attendance after Covid. The drop between 2023 - 2024 was caused due to a vacancy in the Youth Services Librarian position.



USING LIBRARY DATA AND VISUALIZATIONS FOR ADVOCACY AND DECISION-MAKING

USING VISUALIZATIONS FOR DECISION MAKING

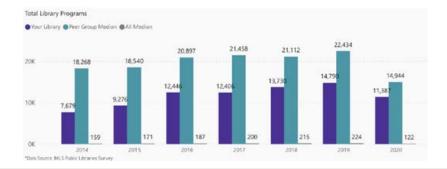
Visualizations are also useful for decision making, particularly in resource allocation and program development.

limeslot	Mon	Tue	weds	Thu	FII	Sat	Sun
00:00-01:00	3849	4003	6905	3730	4366	4936	2401
01:00-02:00	1637	1749	3938	1673	1882	1457	1245
02:00-03:00	950	897	2146	874	990	865	774
03:00-04:00	652	712	769	1024	560	808	507
04:00-05:00	1187	1107	1215	1037	1206	917	425
05:00-06:00	3170	3578	2612	2760	2340	1729	1090
06:00-07:00	11077	10264	8715	9244	7104	5468	2974
07:00-08:00	22808	23407	17770	17914	14317	13440	7058
08:00-09:00	37845	38268	30708	28293	23260	21004	10457
09:00-10:00	104190	96024	91963	85761	82301	99003	11410
10:00-11:00	168007	164731	164485	140371	149563	193962	12098
11:00-12:00	189041	186322	180085	159610	163393	240862	12533
12:00-13:00	171801	156837	154166	143693	158581	217561	20318
13:00-14:00	172687	160413	160838	153598	164897	195309	59680
14:00-15:00	196650	173971	175158	162763	175993	177469	67058
15:00-16:00	215985	201732	196098	189687	204813	190780	57397
16:00-17:00	240877	233032	224738	215796	240400	188727	57405
17:00-18:00	210998	211144	206110	194060	150598	17918	11253
18:00-19:00	181838	177355	169650	152661	61021	7036	11101
19:00-20:00	154001	148367	139938	128032	54193	7038	14057
20:00-21:00	76867	73288	71771	61636	31393	7347	16295
21:00-22:00	18306	18722	14513	13316	8850	7615	15469
22:00-23:00	12258	15535	10285	9660	6857	5702	11470
23:00-24:00	6199	9504	5032	5114	4326	3902	6077

Heat maps help show patterns, trends, or outliers in a data set through color shading that indicates where data falls along a scale as well as its location within a table or map. The heat map shown here is a report of a library's circulation by time of day, useful for making staffing decisions.⁴

Another use of a heat map would be to visualize website traffic patterns, revealing which pages are most popular and which pages may need improvement. A website heatmap tool like <u>Hotjar</u> helps you see what people do on your website pages: where they click, how far they scroll, what they look at or ignore. Hotjar is a subscription service but click on "Pricing" to see the free tools available.

Bar and column charts compare values side-by-side. An effective example of using a bar chart to aid in resource allocation comes from a scenario developed by Hayley Park and Izak Hosmer-Dillard for their MLIS capstone project, "Data-Driven Decision Making in Libraries" at the University of Washington Information School in spring 2023. In this scenario, librarians used a column chart as part of their strategy to convince administration to allocate more resources for children's programming. The chart clearly shows that their library lags behind peer libraries in the number of children's programs.



^{4.} Barniskis, Greg. Circulation by Time of Day Heat Map. Techbits, https://scls.typepad.com/techbits/2013/12/turning-library-reports-into-heat-maps.html. Accessed August 5, 2025.

USING LIBRARY DATA AND VISUALIZATIONS FOR DECISION-MAKING AND ADVOCACY

12 MONTHS TO BETTER LIBRARY DATA WEBINAR SERIES

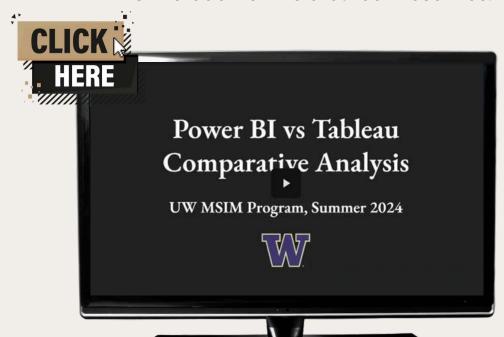
The Research Institute for Public Libraries (RIPL) hosted a monthly webinar series from March 2024-February 2025 called 12 Months to Better Library Data. Several of the webinars focused on how to use data for management, strategic planning, program improvement, and advocacy. The entire series is well worth exploring, particularly the webinars listed below:

- Begin With a Strong Foundation: Developing a Data Plan
- Building Bridges: Collecting Culturally Relevant Data
- Elements of Data Visualization and Design
- Inspire, Advocate, Communicate: Library Data Storytelling
- <u>From Insight to Action: Using Public Library Data for Decision-Making, Strategic Planning, and Advocacy</u>



POWER BI AND TABLEAU TUTORIALS

A special thank you to the University of Washington MSIM students who created these files.



Power BI vs Tableau Comparative Analysis



Power BI Basics Part I

VISUALIZATION TUTORIALS, CONT.

A special thank you to the University of Washington MSIM students who created these files.



Power BI Basics Part II

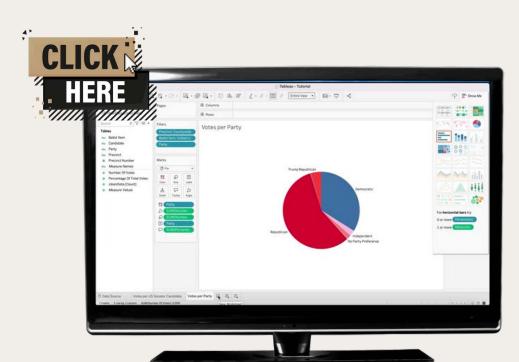


Tableau Bar and Pie Chart Tutorial

VISUALIZATION TUTORIALS, CONT.

A special thank you to the University of Washington MSIM students who created these files.



Tableau Map Visualization Tutorial



WHAT'S NEXT...

So, you've opened up some data - what comes next?

The concept of open data, and the possibilities of what can be done with it, are new to many people. The following strategies can engage stakeholders and encourage them to use the data to its full potential.

Promote Your Open Data Collection

- Feature it on your website and social media
- Issue a press release
- Contact organizations who are likely to be interested in the data, such as journalists and community groups

2 Engage With the Community

- Provide a way for users to provide feedback and suggest improvement by providing an online form on your website.
- Offer data literacy classes to help potential users understand and utilize data effectively.

A good avenue for this is the <u>Data Equity for Main Street</u> curriculum. A complete set of materials to provide the classes in person, as well as a self-directed online curriculum, can be found here.

- Provide examples of how the data can be analyzed and visualized to help users understand its potential by offering classes in using the charts and graphs built into Microsoft Excel, Google Sheets, and more sophisticated tools such as Power BI and Tableau.
- Consider hosting or participating in civic tech events to showcase the data and connect with users.

WHAT'S NEXT...

What is Civic Tech?

"Civic tech, in short, is technology that positively impacts society – but that's not the whole story. It's about creating tools that help make more democratic, transparent and people-centered public services, but it isn't only about the tools themselves – it's about the process through which these tools are created. One important element of the process, for example, is that the people who will be using the tool must be involved in its creation."

Jennifer Pahlka's TEDTalk, "<u>Coding A Better Government</u>" is a great introduction to civic tech.



Coding a Better Government Video

For more information, take a look at these web pages:

"About Civic Tech", Code for All

"Exploring Civic Tech and Gov Tech", WeSolve.app

For a board game about using open data for the public good, play <u>Datopolis</u>.

WHAT'S NEXT...

Monitor Usage and Impact

- Determine the intended goals of publishing the data (e.g., increasing transparency, improving public service, enabling economic growth)
- Track downloads and usage

Note: While most data repositories provide analytics on dataset downloads, views, and other usage metrics, the accessible nature of data published under an open license means that the publisher of the data is unlikely to know how, or by whom, the data is being used. However, some repositories offer forums, discussion boards, or other ways to connect with users.

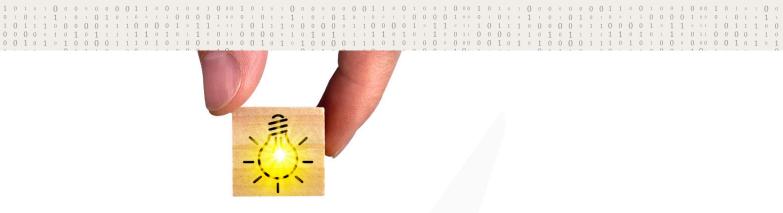
Using an online feedback form, gather user feedback, suggestions, and reuse cases to help inform further publishing efforts. If using a portal, such as Socrata, there should be a feedback form built in. If not, a <u>Google Form</u> can be created and posted on the library website.

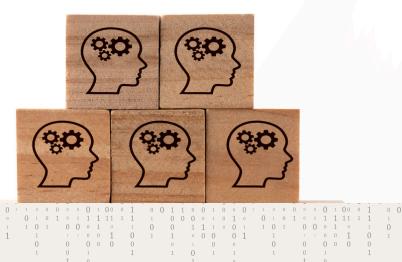
Small libraries may not have the resources to carry out some of these activities, and there may not be individuals with the skills to engage in civic tech activities. That doesn't diminish the work you've done in publishing open data. Your hard work has put the open data out there in the universe and, whether you ever are aware of it or not, wonderful things can come of it! Your efforts have helped inspire and empower citizens to use open data and maximize value to the public in practical ways.





ADDITIONAL RESOURCES AND ACKNOWLEDGEMENTS







ADDITIONAL RESOURCES

WEBSITES



Civic Hacker Network

Bringing together the heads, hands and hearts of people using data and technology to create change in their communities.



Civic Switchboard Guide

A living document designed to help libraries become more engaged in their local civic data ecosystems.



<u>Open Data Handbook</u>

Guides, case studies and resources for government & civil society on the "what, why & how" of open data



Open Data: What Is It and Why Should You Care?

An explanation of open data, why it's important and how you can do it yourself.



What Is Open Data?

A quick primer on open data concepts from the Open Knowledge Foundation.



What Is Open Data? - Practical Guide

Covers the facets of open data and links to many examples.

ARTICLES & BLOG POSTS



<u>3 Tips for Creating Data Visualizations That People Understand</u>

EPL Blogs, Feb 22, 2018 Author: Lydia Zvyagintseva



<u>8 Considerations for Libraries That Want to Host Open</u> Data

Civic Switchboard Blog Post, May 1, 2019

ADDITIONAL RESOURCES



<u>The Changing Role of Libraries in an Open Data</u> <u>World</u>

Business of Academic Publishing, July 21, 2023 Author: David Burbidge



<u>Civic Data Partnerships: by working with local</u> <u>experts on civic open data projects, libraries can</u> <u>become the heart of the smart city</u>

Library Journal, vol. 145, no. 1, Jan 2020, pp. 26+ Author: Matt Enis



Leveraging Open Data for the Public Good

Carnegie Mellon University Libraries Author: Beltran, Lencia



<u>Libraries and Open Data</u>

IFLA SpeakUp! blog, March 6, 2020



The Myth of the Obvious

Medium: Open Data Literacy, Aug 9, 2019

Author: Kaitlin Throgmorton



Open Data Literacy and the Library

Alki: The Washington Library Association Journal, 36(2), pp. 27-29,

ISBN/ISSN: 8756-4173

Authors: Kaitlin Throgmorton, Bree Norland, Carole L Palmer

ADDITIONAL RESOURCES



<u>Opening Rural Library Data: Applying Coursework</u> <u>to the Field</u>

Medium: Open Data Literacy, April 20, 2019

Author: Lillian Curanzy



<u>Persuading the Gatekeepers - Open Data Literacy</u>

Medium: Open Data Literacy, Sept 4, 2019

Author: Lillian Curanzy



<u>Public Libraries and Open Government Data:</u> <u>Partnerships for Progress</u>

Library Journal, Feb 25, 2021

Authors: Carole L Palmer, Nicholas Weber, Kaitlin Throgmorton, Bree

Norlander



Should Libraries Be the Keepers of Their Cities' Public Data?

CityLab, Feb 11 2019 Author: Linda Poon



Using Data for Civic Good: academic and public libraries and local governments are working together to make community data more accessible

Information Outlook, vol. 24, no. 1, Jan-Feb 2020, pp. 17+ Author: Stephen Sherman



What Is Data Visualization? Benefits, Types & Best Practices

Syracuse University School of Information Studies, March 28, 2025

GLOSSARY: TERMS USED IN THE TOOLKIT AS WELL AS COMMON TERMS IN THE OPEN DATA LEXICON

Anonymization: A process of removing or modifying personally identifiable information from data, making it impossible to associate it with any specific individual. This process is crucial for protecting privacy and ensuring compliance with data protection regulations. Anonymized data can still be used for research, analysis, and other purposes without the risk of exposing individual identities.

API: Application Programming Interface. For data, this is usually a way provided by the data publisher for programs or apps to read data directly over the web. The app sends the API a query asking for the specific data it needs. For example, the time of the next bus leaving a particular stop.

Attribution: Acknowledging the source of data when using or re-publishing it. A data license permitting the data to be used may include a requirement to attribute the source. Data subject to this restriction may still be considered open data according to the Open Definition.

Civic Engagement: Actively involving the public in policy and decision-making. Citizen engagement is a central aim of open government, with the aims of improving decision making and gaining or retaining citizens' consent and support. Open data is an essential tool for ensuring informed engagement.

Civic Tech: Online tools and communities that address particular civic or social problems. An important element of civic tech is that the people who will be using the tool must be involved in its creation. Examples: tools that help users meet likeminded people locally based on particular interests, report broken infrastructure to their local government, or collaborate to clear litter from their neighborhood. Locallevel open data is particularly useful for civic tech projects. People who engage with data and technology in in this way are often called civic hackers.

Clean Data: A tab in the Data Ingestion Tool in the Open Data Toolkit for Libraries to be used for cleaning primary data in preparation for publishing. See Data Cleaning.

Cleaning Protocol: A tab in the Data Ingestion Tool in the Open Data Toolkit for Libraries to be used for recording the data cleaning steps taken in order to simplify updating the dataset.

Column Headings: Column headings are the labels or descriptors found in the first row of a dataset that identify the contents of each column. They are crucial for understanding and using open data effectively. Also, a tab in the Data Ingestion Tool in the Open Data Toolkit for Libraries to be used for recording the Column Heading names, descriptions of column contents, and the data type.

Community Partners: Local government agencies and community organizations contributing data.

Context: A tab in the Data Ingestion Tool in the Open Data Toolkit for Libraries to be used for recording any context that will aid others with data interpretation and reuse. Items in this tab do not show up in the metadata, but rather, are for the use of those publishing the data currently or in the future.

Creative Commons: A non-profit organization founded in 2001 that promotes reusable content by publishing a number of standard licenses, some of them open (though others include a non-commercial clause), that can be used to release content for re-use, together with clear explanations of their meaning.

CSV: 'Comma-separated values', a standard format for spreadsheet data. Data is represented in a plain text file, with each data row on a new line and commas separating the values on each row. As a very simple open format it is easy to consume and is widely used for publishing open data.

Data Cleaning: Various processes to make a data resource easier to use. Data cleaning may involve fixing inconsistencies and errors, removing non-machine-readable formatting (such as styles and formatting), using standard labels for row and column headings, ensuring that numbers, dates, and other quantities are represented appropriately, or converting the table to a preferred file format.

Data Harvesting: Data harvesting is a process that copies datasets and their metadata between two or more data catalogs—a critical step in making data useful. It's similar to the techniques that search engines use to look for, catalog, and index content from different websites to make it searchable in a single location.

Data Ingestion Tool: The purpose of the Data Ingestion Tool is to organize the raw data, metadata, and clean data prior to publishing. It serves as an internal aid for staff and includes all of the elements needed for publishing as well as documentation of the source of the data, who worked on it, how long it took, and the resources used. A well-prepared Data Ingestion Tool ensures that all relevant information needed for publishing open data is gathered in one document. The Open Data Toolkit contains a blank ingestion tool as well as a sample completed ingestion tool and one with guidelines for use.

Data Inventory: Google form included in toolkit and given to potential community partners to guide them through the assessment of their data for possible open publication.

Data Wrangler: A person converting data into a form easily used with automated or semi-automated tools. Data wrangling may include further data cleaning.

Dataset: Any organized collection of data. Dataset is a flexible term and may refer to an entire database, a spreadsheet or other data file, or a related collection of data resources.

Discoverable: The state of being easily found by search engine queries. It is not enough for open data to be published if potential users cannot find it, or even do not know that it exists. Rather than simply publishing data on websites, governments and other data publishers can help make their datasets discoverable by indexing them in catalogs or data portals, and/or employing search engine optimization techniques.

Dataset Catalog Template: An organizational tool included in the Open Data Toolkit to be used by data publishers for tracking the status of datasets within an open data collection.

Documentation Overview: A tab in the Data Ingestion Tool in the Open Data Toolkit for Libraries used as a reference for understanding and determining what metadata to use in publishing the dataset.

GIS: Geographical Information System; any computer system designed to read, display, analyze and manipulate geodata.

License: The legal instrument that accompanies the publication of a dataset to convey how a user can use or reference the data.

Machine-readable: Information or data that is in a format that can be easily processed by a computer without human intervention. CSV, JSON, and XML among others, are formats that contain structured data that a computer can automatically read and process. Non-digital materials such as photos and handwritten documents are not machine readable even when scanned. For example, a PDF document containing tables of data is digital but is not machine-readable because the tables are still simply images.

Metadata: Information about a dataset such as its title and description, method of collection, author or publisher, area and time period covered, license, date and frequency of release, etc. It is essential to publish data with adequate metadata to aid both discoverability and usability of the data.

Open Data: Data that can be freely used, modified and shared by anyone for any purpose. In order to be open, the data must be machine-readable.

Open Data Publishing: The process of collecting, cleaning, formatting data, developing metadata, and making it publicly available in an open format.

PDF: Portable Document Format, a file format for representing the layout and appearance of documents on a page independent of the layout software, computer operating system, etc. Data in PDF files is not machine-readable.

Portal/Platform: In the context of the Open Data Toolkit, platform is used to describe the software used to publish open data. Data portal indicates the online location where open data can be found. For example, open data pertaining to the State of Washington can be found at their portal, https://www.data.gov, which is published on the Socrata platform.

Primary Data: Original, uncleaned data as received from an agency or community organization. Also, a tab in the Data Ingestion Tool in the Open Data Toolkit for Libraries to be used for recording the data the way it was received from the community partner, which may include errors and other formatting issues which will be corrected in the Clean Data sheet.

Public Domain: Content to which copyright does not apply, for example because it has expired, is free for any kind of use by anyone and is said to be in the public domain. CCO, one of the licenses of Creative Commons, is a 'public domain dedication' which attempts as far as possible to renounce all rights in the work and place it in the public domain.

Publisher: Anyone who distributes and makes available data or other content. Data publishers include libraries, government departments and agencies, research establishments, NGOs (non-governmental organizations), media organizations, commercial companies, individuals, etc.

Record Level: A tab in the Data Ingestion Tool in the Open Data Toolkit for Libraries to be used for recording the metadata for the dataset: title of dataset; brief description, row label, category, tags/keywords, public license type, provider of data, email address, attachments, notes, period of time covered by dataset, posting frequency, and source link.

Share-alike License: A license that requires users of a work to provide the content under the same or similar conditions as the original.

Spreadsheet: A table of data and calculations that can be processed interactively with a specialized program such as Microsoft Excel, Google Sheets, OpenOffice Calc, and Apple Numbers.

Sustainability Plan: A tab in the Data Ingestion Tool in the Open Data Toolkit for Libraries to be used to track the resources used in both the initial preparation and the publication of the dataset.

Transparency: Governments and other organizations are said to be transparent when their workings and decision-making processes are well-understood, properly documented and open to scrutiny. Transparency is one of the key aspects of open government. An increase in transparency is one of the benefits of open data.

Tags: A tag is a keyword or term assigned to a piece of information or a file. Tags are assigned to data assets to help organize, describe, and make them more discoverable.

Visualizations: A graphic interpretation of data often in the form of a chart, graph, map or dashboard. Many visualization tools exist such as those in Google Charts, Excel, ArcGIS, Tableau, and PowerBI. Creating a dataset's visualization requires careful attention to the meaning of the variables, the relations between them and the stories inherent in the data, to design a visual representation that clearly communicates the message of the data.

CLOSING

This project was made possible in part by the Institute of Museum and Library Services, IMLS National Leadership Grant for Libraries, <u>LG-254933-OLS-23</u>.



The <u>Institute of Museum and Library Services</u> 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 grantmaking, research, and policy development. IMLS envisions a nation where individuals and communities have access to museums and libraries to learn from and be inspired by the trusted information, ideas, and stories they contain about our diverse natural and cultural heritage. To learn more, visit <u>www.imls.gov</u> and follow us on <u>Facebook</u> and <u>X.</u>

The views, findings, conclusions or recommendations expressed in this publication do not necessarily represent those of the Institute of Museum and Library Services.



PROJECT PARTICIPANTS

GRANTEES

Mary Neuman, Project Director Jennifer Ashby, Library Director Asotin County Library Clarkston Washington

PARTNER LIBRARIES

<u>Auburn Public Library</u>

Bel Outwater, Library Manager Holly Burrell Auburn, Georgia

<u>Crawford County Federated Library</u> <u>System</u>

Jessica Hilburn, CEO Meadville, Pennsylvania

Gladwin County District Library

John Clexton, Director President, Association for Rural and Small Libraries Gladwin, Michigan

Lake County Library District

Marsha Richmond, Director Lakeview, Oregon

Lied Lincoln Township Library

Wendy Ketelsen, Director Wausa, Nebraska

Northfield Public Library

Natalie Draper, Director Tyler Gardner, Emerging Technologies Librarian Northfield, Minnesota

ADVISORY BOARD

Cindy Aden, Teaching Professor, MLIS Program Chair, Information School University of Washington

Toby Greenwalt, Principal Flywheel Strategies

Wendy Pearson, Strategic Initiatives Manager, Kansas City Public Library

Kathleen Sullivan

Open Data Librarian for the Washington State Library Office of the Secretary of State

Ayoung Yoon

Associate Professor Luddy School of Informatics, Computing, and Engineering Indiana University Indianapolis

DATA INGESTION TOOL CREATOR

Lillian Curanzy, Supervising Librarian Newport Public Library Newport, Oregon

VIDEO CREATION

Eshita Gupta Bao G. Hoang Tiff Y. Nguyen Kathleen Wong Snehel Yannam

TOOLKIT DESIGN

Anna Aldous, Technology Instructor Asotin County Library



THANK YOU

Toolkit creators Mary Neuman and Jennifer Ashby of Asotin County Library, Clarkston, Washington, would like to thank the partner libraries, advisory board members, Lillian Curanzy, University of Washington iSchool MSIM students, James Neal, and Anna Aldous.



