OPEN DATA GUIDEBOOKS

GETTING STARTED WITH OPEN DATA

A City's introduction and guide to setting up and launching a successful open data program.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1.1</td>
</tr>
<tr>
<td>What is Open Data</td>
<td>1.2</td>
</tr>
<tr>
<td>Launching an Open Data Program</td>
<td>1.3</td>
</tr>
<tr>
<td>Gathering a Team</td>
<td>1.4</td>
</tr>
<tr>
<td>Developing an Open Data Policy</td>
<td>1.5</td>
</tr>
<tr>
<td>Governance Committee</td>
<td>1.6</td>
</tr>
<tr>
<td>Identifying and Preparing Data for Release</td>
<td>1.7</td>
</tr>
<tr>
<td>Selecting the Right Technology</td>
<td>1.8</td>
</tr>
<tr>
<td>Launching and Marketing the Open Data Portal</td>
<td>1.9</td>
</tr>
<tr>
<td>Engaging the Community</td>
<td>1.10</td>
</tr>
<tr>
<td>Continuous Improvement</td>
<td>1.11</td>
</tr>
<tr>
<td>Conclusion</td>
<td>1.12</td>
</tr>
<tr>
<td>Additional Resources</td>
<td>1.13</td>
</tr>
<tr>
<td>Appendix A: Governance Committee Sample Meeting Agenda</td>
<td>1.14</td>
</tr>
<tr>
<td>Appendix B: Governance Committee Sample Follow Up Memo</td>
<td>1.15</td>
</tr>
</tbody>
</table>
Open Data Getting Started Guide

Cities collect and store large amounts of data that is valuable to the public. Until recently, most data was locked away in filing cabinets and only released in response to information requests. An open data program makes city data easily accessible to anyone, at any time, through the internet.

Businesses, residents, visitors, and civil servants expect to have readily available data to aid decision making. Good open data programs meet these expectations while also delivering increased transparency and strong accountability. Open data programs help governments improve work processes, enable collaboration, and optimize service delivery by harnessing the power of data to create greater efficiency and better results for constituents. Cities aren’t simply providers of open data, they are consumers as well.

- Albuquerque publicly released 311 call data, which ultimately reduced call center volume and provided demonstrable benefit to residents.
- Pittsburgh shares real time traffic flow data to engage with community partners on transportation and environmental issues.
- San Francisco made maps from state highway patrol and health open data to illustrate high traffic injury areas, which in turn informed policy decisions.

By collecting and sharing data, a city can empower constituents to use data to address community concerns and help its own civil servants find innovative solutions to key challenges. Consistently sharing and using data to act on community interests can also lead to a culture of evidence-based decision-making in which data becomes essential for good governance.

To assist cities in implementing open data programs in their own communities, the Center for Government Excellence at Johns Hopkins University, a partner in the What Works Cities initiative, has created this getting started guide. By learning from the experiences of other cities and following best practices developed by the Center, cities will gain a greater understanding of open data and be well on their way to launching their own open data program.
What Is Open Data?

Data are electronic records stored in computer systems. In the simplest terms, data are lists of things such as requests for service, inventories, or incidents, which include helpful details such as dates, locations, images, video, and more.

Open data makes these electronic records accessible in whole or in part to the public. This practice is considered proactive disclosure - making information available without it being requested. While industry experts often refer to the Eight Principles of Open Data, it is helpful to initially focus on the following elements:

- Open data is **online**. Cities proactively provide open data through the internet, giving the public the ability to find and use open data without waiting for a response or approval.
- Open data is **free**. Cities do not require payment from anyone to obtain open data.
- Open data is **non-proprietary**. Cities do not require data consumers to have specific software programs in order to use open data.
- Open data is **unrestricted**. Cities do not restrict the use, interpretation, or redistribution of open data through copyright or other terms of use.
- Open data is **machine-processable**. Cities do not require data consumers to scrape data out of "locked-up" formats like PDFs in order to use it, instead using data formats like CSV, JSON or XML.

A commitment to becoming "open by default" means that as long as data are not specifically prohibited from being published, they will eventually become open and freely accessed, used, modified, and shared by anyone, for any purpose. Cities can use open data to deliver results, increase resident engagement, and promote transparency.
Launching an Open Data Program

The following list is an overview of the steps a city should take to implement a successful open data program:

1. Gather a team
2. Develop an open data policy
3. Identify data for release
4. Launch and market the open data program
5. Engage communities of users
6. Review and improve the open data program (continuously)
Gathering a Team

Planning a successful open data program requires choosing the right leaders to help. Communicators, lawyers, line managers, and technologists must collaborate effectively with the public to develop the best policies and practices. Like any other initiative, civil servants need to understand how open data can make them more successful and what they can contribute to the process. Buy in and participation in the opening of data is critical, so it is important to gather key players and engage them from the outset. A city needs certain people to lead with a bird’s eye view of the entire initiative, while others will get into the day-to-day details of releasing and managing data.

The following list provides a description of each key player. Depending on a city’s resources, a single person may take on multiple roles. Many of these team members will form the program’s governance committee, once it is formalized.

Program Sponsorship

- **Mayor or City Manager**: Executive level support and encouragement of the open data process is critical to gaining buy in citywide. The mayor need not be engaged all the time, however, it is extremely valuable for he or she to publicly announce a new open data program or policy, as well as provide encouragement and support to the city’s open data governance committee.

- **Deputy Mayor or Assistant City Manager or Other Senior Executive**: Responsible for overall program leadership; ensures alignment with overall city and administration goals; creates accountability for agency heads and delegated agency staff; helps set and approve priorities for data releases; reports to the Mayor or City Manager on program activities.

- **Chief Counsel**: Assists in developing policy and guidance materials for releasing datasets; develops terms of service and licensing for data access and usage; helps resolve legal considerations for complex datasets; connects dataset publishing to freedom of information requests.

- **Press or Communications Manager**: Ensures public datasets have messaging consistent with other city goals and programs; assists other senior executives with potential responses to public input on published datasets; oversees public engagement activities and events.

- **Chief Information Officer, Chief Technology Officer, or Chief Data Officer**: Oversees the technical aspects of open data, such as maintaining the public data catalog, extracting data from internal IT systems, and ensuring public data is kept up to
Program Execution

- **Project Manager**: Coordinates the open data program; plans and oversees data releases; plans daily activities for the team members; reports on progress to program sponsors.
- **Data Evangelist**: Works with the public and important partners to foster the consumption of data; handles questions about data on behalf of data owners; presents open data at appropriate community events; provides technical assistance to data consumers.
- **Data Owner(s)**: Usually the person or team in an agency who uses the data to carry out their daily business, data owners work with the project manager, data analysts (see below,) and others to ensure the publicly released data is accurate, well documented, and up to date.
- **Data Analyst(s)**: Reviews source data systems; recommends specific data elements for release approval; helps develop public documentation; handles manual data publishing if needed.
- **Programming or Database Engineers**: Build and configure the tools needed to routinely copy data from city technology systems to the open data website.
- **Internal Users**: Agency staff who are important consumers of open data and use data to make decisions, streamline processes, and improve operations; offer perspectives on using data for cross-agency and cross-organizational collaboration.
- **Community**: External stakeholders validate the utility of datasets being opened; the community offers perspective to the city on how the open data program is perceived and employed. Involvement from community members can be formalized through creation of a public advisory group to the open data program, if so desired.
Developing an Open Data Policy

Open data policies are a critical foundation for effective open data programs. Generally, policies are developed on two levels:

- An **executive order** or **city ordinance** (or "law") creates a broad legal framework including open data definitions, scope of participation, high-level timelines, roles, a governance plan, and funding allocations;
- **Governance documents** detail how each aspect of the law will be accomplished, including prioritizing data for release, keeping public data up to date, and engaging with the public. Note that this is not the same as a project plan, which may also be needed.

Although a small team of city staff will likely develop these documents, feedback should be obtained from a broad set of stakeholders both within and outside city government.

The Sunlight Foundation provides guidelines and assistance for developing open data policies.

Executive orders and city ordinances generally:

- establish broad program goals
- define terms (e.g., “data,” “open data,” etc.)
- specify program scope (e.g., government branches, departments, etc.)
- assemble a governance committee and define member roles (see Governance Committee)
- identify governance documents to be developed (see below)
- outline a high-level schedule (for developing the governance plan, launching a site, and reaching data publishing goals)
- create of a central, online catalog to help the public find data
- require public input for each of the above, as appropriate

Governance documents generally include:

- goals for measuring success
- rules and guidance on identifying and prioritizing data for publication
- requirements to ensure the protection of sensitive information
- policies which will be embedded into other government functions, such as procurement or IT project management
- processes for publishing and refreshing data
- quality standards for the data being published
- communications strategy, including engaging external stakeholders and getting internal
buy in

- liability limitations for published data
Governance Committee

A governance committee gives life to the policy and oversees its proper implementation. This team makes policy decisions, prioritizes data publishing, resolves systemic issues raised by city staff or the public, monitors and mitigates risks, and communicates success to the public. During the early stages of the program, this group meets frequently; once the program becomes stable, the meeting frequency can be reduced, and the membership may be reconfigured to reflect the specific requirements of the program. Although establishing a formal charter and committee procedures are usually unnecessary, larger committees may find them helpful.

Members

The governance committee is usually composed of leaders and managers who will be held accountable for the success of the open data program, such as:

- Deputy Mayor, Assistant City Manager, or Senior Executive, who serves as the committee chair
- Council Members or designees (if a law or ordinance was passed and sponsors expect to be engaged on an ongoing basis)
- Chief Counsel
- Press or Communications Manager
- Chief Information Officer, Chief Technology Officer, or Chief Data Officer
- Project Manager
- Agency or department heads or their representatives
- Data managers from all relevant agencies and departments
- Community members, such as open data advocates, nonprofits or service providers, or businesses that rely on open data
- As needed presenters or other guests

Community members are key in providing an external perspective; checking the priorities of the governmental committee members; and communicating the committee's work back to the public at large. Community members must be empowered to speak candidly and confidently, and know that their input is valued with equal weight.

Meetings
During the formative stages of the open data program, the governance committee meets bi-weekly, as frequent reviews of the developing governance plans and implementation progress are needed. As the program matures, the committee can elect to reduce the meeting frequency to review ongoing progress and address outstanding concerns, and may designate subcommittees to achieve particular regular tasks (e.g., reviewing proposed datasets for sensitivity.)

Agendas for Open Data Governance Committee meetings should bolster the Committee's charge to promote the city's open data policy, which may include:

- providing guidance and policy recommendations and, when appropriate, recommending additional legislation or regulations for:
  - procedures, standards, and other deliverables for open data, including for open data portals
  - promotion, advertising, and marketing of open data
    - best practices for sharing open data while taking into account privacy and security concerns
- coordinating the appropriate staff at each agency or department for the development, maintenance, and use of open data and open data portals
- identifying the collective cost of operating and investing in open data and funding mechanisms to support open data; and
- advising the Mayor, City Manager, or City Council on budget matters related to open data
- inviting and encouraging local entities and the legislative and judicial branches to:
  - use open data portals established by city entities;
  - create their own open data portals; and
  - adopt policies consistent with the city's open data policy
- establishing a plan for providing all open data to the public at no cost
- advocating for sound records management and data preservation practices
- making recommendations to ensure that the purchase of new data processing devices, systems, and software by the city includes a review of compliance with the open data policy

Refer to Appendix A for a sample meeting agenda and Appendix B for a sample follow up memo.
Identifying and Preparing Data for Release

One of the first steps in an open data program is developing a plan to release data to the public. This plan is often broken into two stages: collecting a specific selection of data in preparation for launching the portal, and releasing additional data periodically after launch.

Before the portal is launched, the focus is usually on low-hanging fruit that can be quickly gathered and organized. A quick analysis of what other cities have published can help identify data for publishing. In many cases, cities already publish some data, though it is rarely catalogued in a single place. Using a search engine to find data file types on the city website can yield some excellent results. Though they should be reviewed, they likely do not need evaluation to determine if they should be made public, because they already are. Finally, data that drives, informs, or tracks administration priorities are strong candidates for quick release. This data often already has a lot of attention both within and outside the government; making it open reinforces commitments to solving civic challenges.

Once the portal launches, a more strategic approach to releasing public data is adopted. This approach is usually documented in the governance plan, and includes considering public requests for data. The opportunities to open data relevant to administration priorities increases, especially if the executive sponsor(s) of the open data program are closely involved in planning and managing those priorities and their related public engagement activities. Data can also be periodically released in themes such as education, recreation, or transportation. This approach can be excellent for fostering data consumption, but can be challenged by limited resources at a few specific agencies which represent the theme.

Data preparation, another important part of the governance plan, involves extracting information from its sources, transforming it as needed to suit the purposes of the city and the public, and reviewing it for compliance and other issues before publishing it. This process ensures that sensitive information isn’t inappropriately released, data values are understandable to consumers, and quality is preserved. During the early stages of an open data program, data preparation is often done by hand, but this is unsustainable as the program grows. Information technology tools are frequently used to automate data publishing at scale, and are critical to ensuring pre-existing public information is kept up-to-date.
Selecting the Right Technology

At their core, open data programs leverage technology to make data easy to obtain. There are a wide array of options, and significant technology changes may be implemented as the program evolves. These choices should be made by the governance committee, and strongly influenced by the Chief Information, Technology, or Data Officer. In general, cities tend to start simple by leveraging their existing websites, and then move towards advanced platforms offered by third-party companies. Consider migrating to a more complex site if your data providers or customers are asking you for capabilities which are not found through a simple open data site.

- **Simple open data sites** usually consist of a single webpage that contains a list of datasets. At a minimum the list includes dataset titles, descriptions, and links to downloadable data files stored elsewhere. They are often published as pages or sections of pre-existing city websites. Sites like these are easy to maintain and are highly suitable for a quick launch, but they lack features that become important as the volume of public data increases.

- **Complex open data portals** usually provide a variety of ways to interact with both the catalog of available data and the data itself. They are often separate technology solutions, provided by companies who specialize in open data management. While some companies or products can be deployed using the city’s own technology infrastructure, nearly all of them have cloud-based solutions designed to grow as the program does, and they will automatically deliver new features to cities and their data consumers as the open data industry evolves.
Launching and Marketing the Open Data Portal

Getting an open data portal off the ground can involve several steps. It’s worth noting that members of the public open data community generally don’t expect perfection on the first try, and are quite tolerant of problems in both data catalogs as well as the data itself. City leaders earn public trust and praise by acknowledging these imperfections and working to correct them.

Pre-Launch and Beta-Testing

Although it isn’t mandatory, launching a publicly accessible open data portal before formally announcing it can be valuable. It provides the opportunity to work with a limited community to identify and correct any major problems. During this pre-launch period, it is important to ensure all site visitors understand that it may be unstable, difficult to use, and that the available data may be temporary and subject to change. During this stage, it is common for websites to include a “beta” label somewhere prominent or require visitors to acknowledge a message indicating the site is “pre-launch.” Cities may opt to include a beta-testing period for public feedback as part of their executive order or governance plan.

Official Site Launch

Formally launching an open data portal requires the voice of the Mayor. This step is largely a communications and marketing effort. Although an open data launch announcement is powerful by itself, its reach can be dramatically magnified when tied to other policies, initiatives, or even local events. Cities often use a wide array of media channels to get the word out, including a press conference, press release, email outreach, website homepage banner, social media messaging, and more. An in-person announcement by the Mayor is often organized through local technology sector events.

Public Engagement

Open data signals the city's openness to collaboration on important city problems and can represent a new mode of public participation in city problem identification and decision-making. Beginning an open data program with visible and authentic public involvement
improves the momentum for open data both inside and outside of government. For these reasons, it's useful to begin thinking about how to involve members of the public in the open data program from the earliest stages of your open data release. See Engaging the Community for more.
Engaging the Community

Simply launching an open data portal is not valuable if sustainable communities are not built around it. It is important to engage with these communities to help them understand the data the city makes public, listen to their changing needs, and encourage them to help solve civic challenges. The city’s open data program needs a person, often called a data evangelist, who helps build the connections between internal data suppliers and external data consumers. The following section describes how to accomplish this.

Raising Awareness

After launching a portal, there will undoubtedly be some awareness that the city has removed barriers to obtaining valuable government data, but the program will be far from a household name.

- Attending local events can be an excellent opportunity to connect with existing communities who may be able to take advantage of the city’s open data. Many cities are home to vibrant civic technology communities with members who actively participate in open data activities. Presenting open data at events not only raises awareness, but also helps to identify new opportunities for collaboration and question answering.
- Convening roundtables connects community members with each other and with the city’s program. Roundtables can be a particularly effective tool for learning about how individuals, businesses, and nonprofit organizations are using the city’s data to help stay informed, and they always have ideas for future growth.
- Social media outreach helps spread the word about the city’s open data. Use existing services, identities, and group memberships to spread the word about data relevant to each audience.
- Media stories highlight and celebrate valuable use of the city open data. Whether it is a local blog, a regional TV news program, a nationally circulated newspaper, or an international internet media platform, getting the word out about the data and how it has been used invites more people to find other value from it. It also helps other cities note successes and learn from them, improving their own programs in the process.

Listening to Feedback
Accepting and responding to feedback from people who use the city’s data is important to cultivating a culture of openness. Even if the city cannot agree to every request, the city will quickly earn a reputation of being trustworthy and committed to the program. Feedback from people who use the data can be invaluable for helping the city decide to release more data, improve quality, and understand how it is being used. This type of feedback generally comes in any of three forms:

- **Suggestions.** At the simplest level, this means allowing anyone to suggest unpublished data which they would like to have. At a deeper level, this can mean allowing the public to propose changes to the published data, which, if accepted, can help improve internal data quality.

- **Questions.** Published data can be confusing, even if it is extremely well-documented. Providing a way for data consumers to improve their understanding of the city’s data results in it being used more accurately and effectively. A firm understanding of the data enables better media reporting, fewer research errors, and more useful applications.

- **Technical Assistance.** Software programmers may need specialized assistance to access the data, especially if the portal supports advanced features like record querying. If a city is using a vendor to provide its open data portal, the vendor usually provides resources to address these questions. It is important to know where they can be found and how to direct programmers to them when necessary.

Some vendor-provided open data platforms provide different tools for all of the above types of feedback. However, the simplest approach is to provide an email address and ensure a city employee is responsible for monitoring it and responding.
Continuous Improvement

Congratulations, the open data program has been successfully launched! After the launch party and media outreach, it is time for the city to measure and evolve its program.

Measuring Success

Just as measurement is the foundation of performance management, it is equally important to track the success of a city’s open data program. Standards do exist to provide guidance in publishing certain datasets, but this is not necessarily the best way to determine the value of datasets. A successful open data program is one that engages the public, and measures and tracks outputs and outcomes. Monitoring the public’s engagement with a city’s open data program is critical to maintain and grow the program. Below are two ways to think about measuring success of an open data program:

- **Outputs** - Metrics collected around a city’s open data program will most likely reflect outputs related to usage. This can include the number of datasets made public, how often data is viewed or downloaded, how many people visit the data catalog, and the number of assistance requests. Although setting goals is an inherent part of measurement in performance management, goal setting for these metrics will not necessarily help a city provide greater value or increase engagement. However, establishing baseline data can help a city monitor program growth or decline. Note that these metrics are different than the ones the city uses to measure progress toward overarching citywide priorities.

- **Outcomes** - The real value of the city’s open data is not something which can be easily quantified, and most cities use anecdotes when describing successes. A more rigorous way to measure outcomes is to track program improvements and efficiencies, public perception of government, understanding of government’s activities, awareness of open data, and trust in government, before and after the creation of an open data program.

Evolving the Program

Open data programs continually evolve as public expectations, value of data, and technologies change. The program may have started with a simple list of data files, but there are a variety of ways to make it easier for the city’s data to be found, used, and valued.
There are so many ways to improve the program over time that the Center has a whole other guide for it. Explore the Johns Hopkins Center for Government Excellence’s Open Data Long-Term Guide (coming soon) for more information.
Conclusion

Cities all over the US are actively using data and evidence in decision making, investing in what works, and improving the lives of their residents. Open data programs are a fundamental and powerful practice in this important transformation of governing.

While implementing an effective open data program takes a concerted effort from cities -- establishing policies, publishing data, and engaging the public -- the results can be impressive.

Looking for further information about how to get started with open data? Please let us know!
Additional Resources

This guide is designed for mid-sized U.S. cities which aim to start an open data program as part of the What Works Cities initiative. Other guides to launching open data programs are available; they are listed here for reference.

- Open Data Policy Guidelines - Sunlight Foundation
- Open Data Playbook - Code for America
- Open Data Handbook - Open Knowledge Foundation
- Open Data Guide - Mark Headd
- Open Data Field Guide - Socrata
Appendix A: Governance Committee Sample Meeting Agenda

Open Data Governance Council Meeting

[DATE] [TIME] (90 minutes)

[LOCATION]

Welcome (Chair/Co-Chairs)-10 minutes

- What is open data
- Purpose of the council
- Structure of the council (leadership, staffing, meeting frequency, membership: state agency secretaries or designees; legislators; elected officials or employees from local entities who have knowledge of and interest in open data; members of the private, private utility, academic, or nonprofit sectors who have knowledge of and interest in open data)
- Duties of the council, including deliverables such as a plan for providing open data to the public at no cost; regular reports to the public; etc.
  1. Introductions-5 minutes
  2. Thank you and remarks from Mayor-5 minutes
  3. Review of existing Open Data policy-10 minutes
  4. Discuss authorizing legislation or Executive Order, and the charge of the Mayor
  5. Compliance requirements review (ethics, financial disclosure)-10 minutes
  6. Background and demonstration of current open data portal (if the city does not have an open data portal, review portals of other cities)-10 minutes
  7. Open discussion of open data portal-10 minutes
  8. Open Discussion: Business of the Council-30 minutes

- Choosing technology (if applicable)
- Inventory existing datasets
- Identifying criteria for datasets to be released by agencies (FOIA requests, 311, agency priorities, city priorities, trends, strategic releases, map to performance management priorities, etc.)
- Review and consolidate previous open data policies
• Linking Open Data with performance management

Please note the following dates and times for future meetings

[Date, time, location] [Date, time, location] [Date, time, location]
Appendix B: Governance Committee Sample Follow Up Memo

To: Open Data Council Members

From: [STAFF]

Subject: Open Data Council Meeting,

Date: [DATE]

As a result of our recent meeting, please provide work through the following action items before the next Open Data Council meeting.

The next several meetings are scheduled for the following dates:

- [Date, time, location]
- [Date, time, location]
- [Date, time, location]

1. As requested, please complete required [Ethics and Financial Disclosure] forms.
2. Each agency should identify its top five datasets for making publicly available on open data portal, using the criteria agreed upon by the council.
3. Before the next meeting, each agency should upload three new datasets to the open data portal. One of these three datasets should be a dataset listed in your response to question #2.
4. Please communicate the requirements of the Open Data Policy and the Mayor’s charge on open data to your agency's staff.
5. Each agency should prepare suggestions on how to organize the work and discussions of the Council. Does your agency recommend any specific working groups?