City of Chattanooga Data Inventory Guide

The following steps are meant to guide the identification and release of City of Chattanooga datasets in a strategic, transparent, and meaningful way. This guide was developed in coordination with the City's Open Data Advisory Committee, the City's Office of Performance Management and Open Data and Bloomberg philanthropies What Works Cities initiative. This guide is a living and breathing document and will be updated as the City's open data program matures. The **five** steps included in this guide are as follows:

Step 1: Identify potential datasets

Step 2: Provide details about the dataset

Step 3: Determine internal dataset publication priority

Step 4: Determine external publication priority

Step 5: Create release schedule based on all the above steps

Step 1: Identify Potential Datasets from Each City Department

Each city department collects data as part of their daily operation. These datasets have great potential value for increasing transparency, operational efficiency and actionable insights. For all these reasons it is important to understand the data assets currently housed within the city. As the Open Data Coordinator of your department you know best the datasets found within your department. The very first step for the inventory is to start thinking about the data assets maintained by your department.

The following questions may help the you identify data assets in your department:

- What data populates your monthly or quarterly reports?
- What data is currently publicly available online?
- What data does your organization use for internal performance and trend analysis?
- What information is published as a performance metric?
- What data is included in reports to federal, state or local departments/agencies?
- What are the trends in Public Information Act requests?
- What data do other departments/agencies ask for?
- What kinds of open data are similar departments across the country publishing?
- Which datasets are related to citywide goals and priorities, or your department's mission?
- What are the metrics used in the city's budgeting process?
- What datasets that you regularly create, maintain, or manage are used to record a measurement or transaction and/or provide information on government services, initiatives, or resources?
- What is the data that drives lists, tables, graphs, charts, maps, or other visualizations that your department uses?

Dataset Title
Human-readable name of the asset. Should be in plain English and include sufficient detail to facilitate search and discovery. Avoid acronyms.

As you start answering these questions you can keep track of the potential datasets in the City of Chattanooga Data Inventory Spreadsheet. At this point, there is no need to fill out all the fields in the spreadsheet, just fill in the "Dataset Title" column as you identify as you ask yourself the questions above. It is a good best practice to share this initial list with others in your department to see if they can think of any other datasets. Remember, at this point the goal at this point is breadth not depth. Potential datasets could be anything from information keep solely in paper files, to spreadsheets in personal computers or on a network drive to databases for the third party systems you use.

To access your department's inventory spreadsheet and begin inventorying datasets click on your department's name below.

City Attorney's Office	CARTA	CDOT	City Courts	Economic and Community Development
Finance	Fire		Human Resources	Information Technology
	Library	Mayor's Office	Office of Performance Management	Office of Multicultural Affairs
Police	Public Works	Purchasing	Regional Planning Agency	Youth and Family Development

Step 2: Provide details about the dataset

Once your department has identified any and all potential datasets the next step is providing some basic information about those datasets. This is where we shift to depth and start to provide the details that will be needed to evaluate, prioritize, schedule and ultimately begin releasing your department's datasets. Below are the basic fields you need to fill in for the inventory. You should have already filled out the "Dataset Title" column of the inventory spreadsheet in step 1 of this guide. The next step is to determine if the data has sensitive data and if so, how best to manage this data to still make the data sharable.

A note about privacy: Do not exclude any datasets based on privacy or confidentiality concerns! To make the data inventory as useful as possible, it should include data that may be sensitive, private, or unlikely to be released. Always include a description of the sensitivity concerns. In addition to setting up the open data program on a strong and transparent foundation, an open data inventory can provide a roadmap for creating release timelines and help identify where new data may need to be collected.

Division	Data Steward's Email Address	Contents	System of Record	Contains Sensitive Data	Sensitive Data Comments
Division where the data is maintained. If more than one department or division is responsible, list the primary owner of the dataset.	*Data Steward's email address. Who manages the data and / or is responsible for granting permission to access the data? Who understands what the dataset includes and can answer questions about it?*	*What the dataset describes. Provide a longer description of the data that can be readily understood by non-technical users.*	*What is the original system/application/file that houses it? Please provide additional comments as necessary.*	*Does the dataset contain sensitive data, such as personal phone numbers, Social Security Numbers, etc.?*	*What could be done to make the data shareable?*

Step 3: Determine Internal Dataset Publication Priority

After identifying datasets and their technical characteristics, the next step is to determine the relative value in publishing a given dataset. To assist with that process a scoring rubric show below was created to prioritize the datasets.

Data Quality	Data Value	Data Accessibility	Internal Priority Score
What is the quality of the dataset?	How valuable is the dataset to the city internally? External value will be gathered by engaging the public.	How easy is it to access and publish the data on the open data portal?	Sum of the data quality, data value and data accessibility scores.

Internal Prioritization Scoring Rubric

	Value	Value Description
Data Quality	10	Excellent Data Quality: The data is complete, accurate and updated regularly.
	5	Adequate Data Quality: The data is somewhat complete, accurate and updated fairly regularly.
	0	Poor Data Quality: The data is not complete, accurate or updated regularly.
Data Value	10	Excellent Data Value: The data is relevant to city's goals and or department core functions and would greatly increase transparency.
	5	Adequate Data Value: The data is somewhat relevant to city's goals and or department core functions and would greatly increase transparency.
	0	Poor Data Value: The data is not relevant to city's goals and or department core functions and would not increase transparency.

Data Accessibility	10	Excellent Data Accessibility: The data is machine readable, easily accessed, and publishable in current format.
	5	Adequate Data Accessibility: The data is easily convertible to be machine readable, somewhat difficult to access, and somewhat publishable in current format.
	0	Poor Data Accessibility: The data is not easily convertible to be machine readable, very difficult to access, and not publishable in current format.

Step 4: Determine external publication priority

Now that the data has been identified, inventoried and the internal prioritization set, the next step is to engage the public in the decisions about what to release, to meet the needs of residents. The process on how to gather community input has yet to be worked out and will be crafted by the Open Data Advisory Group.

Screenshot of [Philadelphia's Data Inventory](#)

Citywide Data Inventory

| Fire | Licenses & Inspections | Streets | Parks and Recreation | Commerce | City Commissioners | Ethics | Emergency Management | Energy Office |

search by keyword   search

Showing 1-100 of 168   export   100 per page   Page 1 of 2   < >

Title	Contents	Comments	Details
Fire Incident Reports	Any time there is a fire, information is gathered relating to the event and the people involved, including the incident number, when the incident occurred, if any died in the fire, extent of the fire, location of the property, if any injuries occurred.	10	add comment
Emergency 911 Calls	This data set contains informaiton related to the calls received by the Fire Communications 911 unit. These entries are timestamped with when the call was received, when a unit was dispatched, when the unit arrived on the scene, and has additional information regarding what happened while the incident was being mitigated.	5	add comment
Budget	Budget reported monthly, quarterly, and annually	4	add comment
Jury Duty Paper Work	Employees that have been called for jury duty, when they were out and additional details	3	add comment
FCC Phone Database	The FCC Phone Database contains information on all incoming calls for a given day. This information is broken down into specific time frames as well as type of call which could be 911EMS, 911 FIRE, or NON-911 calls. Cumulative totals for time frames and type of calls are also recorded.	3	add comment
FCC Dispatcher Database	This database contains the total number of dispatches for a given day. Data contained within each record is broken down by time frames, and type of incidents (fire, medical, etc.)	3	add comment
Daily number of calls received	The number of 911 calls received on a given day.	3	add comment

In addition to asking the public for feedback, Philadelphia hosts and curates a discussion board focused on open data and transparency. Topics include mislabeled datasets, 404 errors, posts about new datasets with requests for feedback,

questions and requests for new datasets, research questions, and notifications about new data initiatives. An example of an interaction between Philadelphia's Chief Data Officer and the community is below.

Screenshot of Philadelphia's Open Data Discussion Board

Tapping the networks of stakeholders, such as a local civic tech group or Code For America brigade, nonprofits that may be interested in the subject matter of the datasets, and businesses that may have an economic use for the data can be a valuable way to leverage these groups to get the word out about a city's open data and transparency efforts.

Step 5: Create Release Schedule

The final creation of the release schedule will be done by the Office of Performance Management and Open Data in coordination with the Open Data Advisory Group.