

How to Inventory Data

[GovEx](#)

September 2, 2017

+ Share

The datasets worth inventorying are those which are considered *assets* to employees, departments, executive leadership, and the general public.

The data inventory process is carried out in the following steps:

Step 1: Establish an Oversight Authority

Conducting a data inventory across departments requires coordination, oversight, and leadership. The first step to conducting an inventory is establishing who will manage the inventory process. Oversight authorities can come in a variety of shapes and sizes and are often defined in a city's Open Data Policy. Some are led by a Chief Data Officer (or similar role), others leverage existing enterprise data management bodies, and others are working groups that include public representation. While the breadth and depth of data governance authorities can range to best suit your city's needs, establishing a clear authority body to oversee the data inventory process is key to success. In the absence of a data governance committee, you may wish to identify a lead liaison, preferably within the Mayor or City Manager's office, to interact with departments and facilitate this process.

- **Note:** Establishing a data governance committee, or repurposing an existing committee, is an optional, but highly recommended step in successfully completing the inventory process.

Step 2: Determine the Data Inventory Scope and Plan

The oversight authority, such as a data governance committee, should

manage the inventory process by providing an unambiguous scope, deadlines, performance metrics, and guidelines.

Scope: If the scope is not already defined in your city's Open Data Policy, the oversight authority should determine the scope of the data inventory at hand. If your city does not already have a data inventory in place, creating a city-wide comprehensive data inventory can range in difficulty depending on how many data assets your city manages, how siloed those assets are managed, and your available capacity to conduct the inventory. When defining the scope of the data inventory, the oversight authority should consider the following:

- Any relevant data definitions or inventory requirements that are included in your city's Open Data Policy
- Any government records definitions outlined in your local Records Management policies (i.e. Distinguish government datasets from non-record data and personal data notes)
- All data assets
- Strategic-priority-specific assets
- Individual departmental assets

Plan: The Data Inventory plan

- Required [Metadata](#)
- Deadlines
- Guidance
- Performance Metrics

Step 3: Catalog Data Assets in Accordance with Inventory Plan

Liaisons in each city department or agency catalogue and describe the data assets within their departments. Liaisons are employees who are responsible for managing the inventory process at the departmental/agency level. The lead manager of the data inventory compiles the individual department

inventories into a larger citywide data inventory. Inventories should be structured in machine-readable format (Spreadsheet, CSV, JSON, etc.)

Step 4: Data Inventory Quality Checks

The data governance committee establishes the extent to which the inventory is made public.

Data governance committees can publish high level inventory summaries that specify which datasets are to be published, or remain unpublished.

Philadelphia Example

- Philadelphia’s inventory lists all datasets that it uncovers, even those that will never be released due to sensitive content such as Personally Identifiable Information (PII) or security concerns. Philadelphia lists the following notes with datasets that contain sensitive information: *“Some data sets in this inventory cannot be published as open data. Others could be published after sensitive data is removed (such as personal information).”*

Step 5: Initiate Data Prioritization Efforts

The data governance committee establishes the priority and scheduling of the publication of datasets described in the inventory.

Publication priority can be established in a number of different ways:

- Strategic-priority specific
- Departmental priorities
- Public demand