

High Value Data Inventory Guidance

Overview

[Section 4-67p of the Connecticut General Statutes](#) requires executive branch agencies to maintain an inventory of their high value data. [Agency Data Officers](#) are responsible for updating their agency's high value data inventories on an annual basis. The inventories for all state agencies are compiled on the Connecticut Open Data Portal at the following link: [CT High Value Data Inventory](#). This document provides guidance for Agency Data Officers on how to draft and update a high value data inventory.

Why inventory data?

Data is a valuable strategic asset. As state agencies we invest public funds in the creation and maintenance of data to administer state policies, programs, and services. To ensure that we are managing the state's data appropriately, we must maintain an inventory of these assets. In the same way that we manage our physical assets such as buildings, vehicles, or other property, we must manage our data assets.

What to inventory?

[Conn. Gen. Stat. Sec 4-67p](#) requires each executive branch agency to inventory high value data. The act defines "high value data" as meeting any one or more of the following criteria:

- Critical to the operation of an executive branch agency,
- Can increase agency accountability and responsiveness,
- Can improve public knowledge of the agency and its operations,
- Can further the core mission of the agency,
- Can create economic opportunity,
- Is frequently requested by the public,
- Responds to a need or demand identified by the agency through public consultation, or
- Is used to satisfy a legislative reporting requirement.

Agencies should inventory data they possess that meets one or more of these criteria. Agencies should not inventory datasets that are generated from primary sources or data that are obtained from another agency.

Instructions

If this is your agency's first time completing the high value data inventory...

1. Request access to the [CT High Value Data Inventory](#) in Microsoft Lists if you don't already have access.
2. Begin to identify high value data sources at your agency. You can use the IT Application inventory as a starting point or contact data owners at your agency for information about the data they manage.
3. Add the high value data sources you have identified to the list in Microsoft Lists. You can work in the Microsoft Lists app, in [the Teams app](#), or in [your browser](#). To add a new item to the list, click the "+New" button in the top left corner. Fill out as much detail as possible for each row.
4. Once you have finished adding the high value data assets for your agency, notify the Chief Data Officer, Scott Gaul, by email before December 31. You can make updates to the inventory at any time.

If your agency has completed the high value data inventory in the past...

1. Review your agency's high value data inventory in Microsoft Lists for needed updates: [CT High Value Data Inventory](#). You can work in the Microsoft Lists app, in [the Teams app](#), or in [your browser](#).
2. Make any changes to the inventory. Double click on a row to make edits to an entry. To add a new item to the list, click the "+New" button in the top left corner. Fill out as much detail as possible for each row.
3. Once you have finished adding the high value data assets for your agency—or if no changes are needed—notify the Chief Data Officer, Scott Gaul, by email before December 31.

Key Terms

Data

Under [CGS Sec. 4-67o](#), means the final version of statistical or factual information that: (A) Is reflected in a list, table, graph, chart or other non-narrative form that can be digitally or nondigitally transmitted or processed; (B) is regularly created or maintained by, or on behalf of, an executive branch agency; and (C) records a measurement, transaction or determination related to the mission of the agency or is provided to the agency by third parties pursuant to law.

High Value Data

Under [CGS Sec. 4-67o](#), means any data that the department head determines (A) is critical to the operation of an executive branch agency; (B) can increase executive branch agency accountability and responsiveness; (C) can improve public knowledge of the executive branch agency and its operations; (D) can further the core mission of the executive branch agency; (E) can create economic opportunity; (F) is frequently requested by the public; (G) responds to a need and demand as identified by the agency through public consultation; or (H) is used to satisfy any legislative or other reporting requirements.

Agency

This column is a drop-down option to indicate the appropriate agency acronym.

Business Unit

This column describes the part of your agency which is the business owner of the entity.

Data Asset Name

This column is the name of the item being added or updated in the inventory. The National Institute for Standards and Technology defines a data asset as "any entity that is comprised of data. For example, a database is a data asset that is comprised of data records. A data asset may be a system or application output file, database, document, or web page. A data asset also includes a service that may be provided to access data from an application".

Data Asset Description

This column is a high-level description of the data asset being added or updated in the inventory.

GIS or non-GIS

Is this geospatial (GIS) data? is a drop-down column to indicate whether the data asset is Geospatial Information System (GIS) data or non-GIS data (for example, tabular data).

Data Asset Structure

What is the data structure (i.e. the format for organizing, processing, retrieving and storing the data)? Choose one option below from the dropdown menu.

- Non-GIS Structured (e.g. relational database or file (excel, csv, etc.). If a relational database has a table holding a photograph, for example, this would be structured data.)
- Non-GIS Unstructured (e.g. Mobile data and photographs)
- Non-GIS Mixed (contains both structured and unstructured)
- GIS Cloud/Online Database
- GIS Enterprise Database
- GIS Shapefile
- GIS Raster Image
- GIS Other Native File Format

Data Classification Impact Level

What is the data classification impact level (i.e. the level of impact that a breach of the data asset would have)? Choose high, medium, or low from the dropdown menu. The Office of Policy and Management [Data Classification Policy](#) requires that each Executive Agency Branch assign a classification to all data for which it has custodial responsibility. The [Data Classification Methodology](#) should be followed to assign a classification impact level of High, Medium, or Low.

Policy Requirements

What are the laws or policy requirements that govern the collection, use, and management of the data asset? Choose all that apply from the dropdown menu or write in the applicable laws or policies. Some examples of policy requirements include HIPAA, FERPA, or COPPA. Using this column to indicate that there is PII or PHI is also acceptable if you are unsure of the applicable policies.

Demographic Data

What demographic information is included in the data asset? Choose all that apply from the dropdown menu or write in the type of demographic information in the data.

Location Data

What location information is included in the data asset? Choose all that apply from the dropdown menu or write in the type of location information included in the data.

Data Owner

Who is the data owner (i.e. who is responsible for the data within the asset, or who has authority to make decisions about the data within their domain)? Write the name or role of the data owner.

Data Steward

Who is the data steward? Enter the email of the data steward to tag them. A data steward is responsible for the governance of data to ensure fitness for purpose of content and metadata. Stewards exercise established process, policy, guidance, compliance and rule in this effort.

Used in AI System? (Y/N)

Is the data asset used in an AI system or to train an AI system? Select yes or no from the dropdown menu.

Additional AI Details

This column is an optional text field to add details about the data asset's use within an artificial intelligence system. For example, if an application uses AI, has AI embedded, or is used to train AI systems, you can share here.