

CT Open Data

Metadata Guidelines

Introduction

This document provides guidance on metadata for publishers of open data in Connecticut state government, including through the CT Open Data Portal and the CT Geo Data Portal. Metadata is the data that describes other data. It helps answer the question “what is this data about?” High quality metadata provides helpful context about the data’s creation, quality, and uses and is key to improving data discovery.

Metadata Checklist

When publishing data as open data use the checklist below to make sure your metadata is complete.

- Does the metadata describe **what information is contained** in the data?
- Does it include a **clear title** that non-technical users will be able to understand?
- Does it use **plain language** and avoid use of overly technical language, jargon, and acronyms?
- Does it include the **source of the data** and describe how the dataset was created?
- Does it list the **data owner**?
- Does it include information about **how frequently the data is updated**?
- Does it provide guidance on **how the data should be used and any limitations**?
- Does it align with the **metadata standard** of the platform where the data is being published?

CT Open Data Portal Metadata Schema

The table below summarizes the metadata element on the CT Open Data Portal. When creating a new dataset, data publishers should provide as much detail as possible in the metadata, providing the following elements where applicable. A more detailed version of the table below is available as an Excel spreadsheet [here](#).

Dataset Metadata Elements			
	Field	Definition	Permitted values
Basic Descriptive	Title	Human-readable name of the asset. Should be in plain language and include sufficient detail to facilitate search and discovery. Avoid acronyms.	Text
	Description	What the dataset describes. Provide a longer description of the data that can be readily understood by non-technical users.	Text
	Category	The category of the data set identified by the list of possible values. If a data set can fall into multiple categories, select the one which is most significant.	Drop down menu: Business, Education, Environment and Natural Resources, Government, Health and Human Services, Housing and Development, Local Government, Public Safety, Tax and Revenue, Transportation
	Agency	The agency that collects and manages the data as the authoritative source.	Drop down menu

Open Data Metadata Guidelines

Detailed Descriptive	Number of Rows	Number of rows in the dataset.	Auto-generated
	Row Label	What each row in the dataset represents.	Text
	Tags/Keywords	Tags (or keywords) help users discover your dataset; include terms that would be used by technical and non-technical users.	Text
	Source Link	The URL to the program area web pages	URL
	Geographic Unit	At what geographic unit is the data collected? For example, if the data is collected by address, it would be Street Address.	Drop down menu: Latitude/longitude, Street address, Parcel (block/lot), Census block, Census block group, Census tract, Planning District, Zip code, Municipality, County or county-equivalent, State, Other, Not applicable
	Temporal Coverage	This field should contain an interval of time defined by the start and end dates for which the dataset is applicable.	Text
	Related Documents/Attachments	Related documents such as technical information about a dataset, developer documentation, etc.	Attachments
Internal Management	Data Provided By	The name of the data source, for example the name of the publishing agency, organization, or individual.	Text
	Dataset Owner	The name associated with the account publishing a dataset. The name will be automatically associated with the dataset.	Auto-generated
	Contact Email	This address will not be displayed publicly, but inquiries submitted via the "Contact Dataset Owner" button will be routed to this email. If left blank, it will default to the dataset owner email address	Auto-generated
Publishing Details	Update Frequency	Frequency with which dataset is updated.	Drop down menu: Not updated (historical only), As needed, Annually, Bi-annually, Quarterly, Bi-monthly, Monthly, Bi-weekly, Weekly, Every weekday, Daily, Other
	Last Updated	Most recent date and time when the dataset was changed, updated or modified.	Auto-generated
	Public License Type	The license with which the dataset is published.	Drop down menu

Open Data Metadata Guidelines

Column Metadata

In addition to the dataset metadata (as detailed in the table above), data publishers should also provide “column metadata,” or descriptions of the columns in the dataset. Especially when column names are not easily understood by users not familiar with the dataset, data publishers should provide short column descriptions in the metadata. Additional technical documentation can be provided as an attachment accompanying the dataset.

CT Geo Data Portal Metadata Schema

The table below summarizes the metadata element on the CT Geo Data Portal, adapted from the [ArcGIS Hub documentation](#). Data publishers should provide as much detail as possible in the metadata, providing the following elements where applicable.

Dataset Metadata Elements			
	Field	Definition	Permitted values
Basic Descriptive	Thumbnail	The thumbnail image is displayed on layout cards and content views.	Image
	Title	Use a title that is succinct and informative. Avoid underscores as they will be removed.	Text
	Summary	Give an overview in several sentences that cover the key elements of the item.	Text
	Description	Write a description that is clear and informative, as it is shown on content views and in search results.	Text
Detailed Descriptive	Number of Records	Number of records in the data.	Auto-generated
	Tags	These help users discover your datasets through searches. For individual items, add tags in ArcGIS Online. All layers in map and feature services will have the same tags as those set for the entire service in ArcGIS Online.	Text
	Categories	Category of the data. These help organize items and facilitate their discovery and use.	URL
Publishing Details	Data Publisher	The name associated with the account publishing a dataset. The name will be automatically associated with the dataset.	Auto-generated
	Published Date	Date item was created.	Auto-generated
	Data Updated	Date item was last modified	Auto-generated
	Info Updated	Date metadata was last modified.	Auto-generated
	License	Choose a structured license or enter a custom one. The license is shown on the full details page and on the information side panel if the item has an explore view. It also displays (and you can add it here as well) in ArcGIS Online, under Terms of Use on the item details page.	Drop down menu

Marking Data as Authoritative

In addition to providing complete metadata, publishers of data on the CT Geospatial Data Portal can also improve the usability of the data on the Portal by marking their data as authoritative. Designating items as authoritative makes it easier for users to find authoritative items. The [ArcGIS Online documentation](#) explains:

Open Data Metadata Guidelines

“Organization administrators and those with administrative privileges to update content can specify that an item is authoritative using the Mark as Authoritative button. Items designated as authoritative are identified with an Authoritative badge on the Overview tab. If your organization is verified, items that are shared with everyone (the public) and marked as authoritative display the organization name as the item owner on the Overview tab.”