

Open Data Retirement Guidelines

The following guidelines dictate the appropriate circumstances by which an asset may be removed from the Connecticut Open Data Portal (data.ct.gov), as well as the process for retiring the asset. Assets may include datasets, filtered views, maps or other visualizations, data stories, or other files published to the Open Data Portal.

Connecticut state agencies should use the “Criteria” and “Process for assessing assets” sections below to determine whether their data is eligible for retirement. If your agency has data that you would like to retire (or consolidate with more recent data), contact the Open Data Portal Administrators at OPM via dapa@ct.gov.

Criteria

To retire an asset, one or more of the following criteria must be met:

- The data or metadata is factually inaccurate, misleading, or incomplete;
- The asset contains protected data;
- The data has been discontinued, consolidated, or replaced by the owner agency;
- The asset was mistakenly published (e.g. a draft was published instead of the final version);
- The asset is rarely used;
- The data has become outdated; or
- The data does not belong to the agency that published it, especially if the data is repeated in assets published by the owner agency.

Process for Assessing Assets

To determine if an asset should be retired, the following steps will be followed by the data publisher in consultation with Open Data Portal administrators:

- **In the case of assets being retired solely because of their age, consider the value of such data in larger context.** Could this data be consolidated with more recent data so users can track trends over time? If so, this data should be consolidated with similar data from other time periods rather than retiring it if possible. If the consolidated dataset would be so large that it would become unmanageable, then it is acceptable to retire the oldest portions, preferably in an archive, such as the Connecticut State Library Digital Archive.¹
- **Determine whether there are similar or associated assets that are also eligible for retirement.** For example, if the dataset in question is being retired because it is ten years old, and there is another dataset with the same data but that is nine years old, both may be eligible for retirement or consolidation. If that is the case, Open Data Portal administrators should consolidate the datasets before proceeding with the data retirement process.
- **Consult with the agency that owns the data.** If the agency that owns the data and Open Data Portal administrators agree that an asset should be retired, Open Data Portal administrators will proceed with the removal process below. If site administrators believe that an asset should be retired, but the owner agency is actively maintaining the data and requests that it remain, Open Data Portal administrators should not retire the asset.

¹ Accessible via: <https://ctstatelibrary.org/>

If they disagree and the owner agency is not actively maintaining the data, then the agencies must continue to converse until a successful compromise is reached.

- If the dataset was created or owned by someone no longer with the agency, Open Data Portal Administrators will discuss it with the Agency Data Officer. Open Portal data administrators can change ownership or disable old accounts, if needed. Once the owner agency has approved retirement, consider whether the asset could or should be archived. To make this determination, consider whether:
 - The asset could have historical value;
 - The data and metadata are complete enough that the asset can be understood in context; and
 - The owner agency still has the data, or whether the asset on the Open Data Portal is the only remaining source of the data.
- If the asset is eligible to be archived, the Open Data Portal administrators will consult with the Connecticut State Library to make a final decision. This decision should be finalized before proceeding with retirement.

Process for Retiring Assets

To retire an asset, the process outlined below should be followed by an Open Data Portal administrator:

- Add the word “ARCHIVE” to the title of the asset and add a tag “archive” in the metadata and add it to the dataset of retired assets² from the portal with ‘mid-retirement’ status. In the description, add a note indicating that the asset will be removed in 30 days and that users should download the asset or contact site administrators if they will need it in the future. The following information should be included in the dataset of retired assets:
 - The metadata included in the Asset Inventory for the asset being retired;
 - The date the asset was made private in the retirement process;
 - An explanation of why the asset was removed (including a link to the asset that supersedes it, if applicable); and
 - A link to the archived asset, if applicable.
- After 30 days, set the asset to private and update the status in the dataset of retired assets.
- If the asset is going to be permanently archived, send it to the State Archives at the Connecticut State Library. Confirm with the archive staff that they have received the data. Once the State Archive staff has approved the deletion of the asset from the portal, delete the asset.
- If the asset is not going to be archived, wait 90 days, then delete the asset.
- If the asset has only just been published (within the last 7 days), it may be removed without adhering to the waiting periods, since it is unlikely that anyone would have had time to use it in their research or for other substantive purposes.

Acknowledgments

This policy was developed based on similar policies in other open data programs, including the examples listed below. Guidance from the staff of the Connecticut State Library, and research on data deprecation policies in open data programs from Kathleen Sullivan at the Washington State Library and Andrew McKenna-Foster from the University of Washington also greatly informed the development of this policy.

² Accessible via: <https://data.ct.gov/dataset/Retired-Assets/i9r5-65tv>

1. NYC Open Data, "Removing data assets from the Open Data Portal – Process," <https://opendata.cityofnewyork.us/wp-content/uploads/2018/02/Open-Data-Removals-Process-and-Guidelines.pdf>.
2. NYC Open Data, "Dataset Removals," <https://data.cityofnewyork.us/dataset/Dataset-Removals/tm5c-buy3>.
3. Minnesota Geospatial Commons, "Maintaining and Updating your Resources," https://gisdata.mn.gov/content/?q=help/maintain_resources.
4. Arkansas Office of Information Technology, "Policy Statement: Arkansas Spatial Data Infrastructure (ASDI) Data Loading and Retirement Procedures," http://www.gis.arkansas.gov/Docs/LAW/AGIO%20Policy%20Statement%20PS-01%20ASDI%20Data%20Loading%20Procedures_rev20190313_FINAL.pdf.
5. Delaware FirstMap, "Information Service Change," <https://firstmap.delaware.gov/contentFolder/pdfs/public/FirstMapServiceChangePolicy.pdf>.
6. Utah Automated Geographic Reference Center, "SGID Database Policies," <https://gis.utah.gov/about/policy/sgid/>.
7. Data.texas.gov, "Dataset update schedule," <https://data.texas.gov/stories/s/TDA-Data-Overview-Food-and-Nutrition-Programs-Open/nk79-w2cs/>.