

Open Data Retirement Policy

Purpose

The City of Cambridge is committed to maintaining a robust and reliable Open Data Portal. As a part of this commitment, it is necessary to manage the life cycle of the data assets available on the portal. Over time, some datasets may become inactive or obsolete. This policy outlines the criteria and processes for identifying, marking and removing such assets, ensuring that users can navigate and utilize the portal effectively without being encumbered by outdated or irrelevant data.

All decisions regarding the retirement of assets will be made by the Open Data Manager in collaboration with the department responsible for generating or maintaining the asset.

1. Definitions

- **Open Data Assets:** Resources made publicly available by the City of Cambridge on the Open Data Portal, including datasets, visualizations, and other relevant materials.
- **Obsolete Asset:** A data asset may be deemed obsolete if it is no longer relevant, either due to its replacement by a more accurate source, its misalignment with the City's current administrative, legal, or operational standards, or because it is no longer being updated and lacks historical interest.
- **Historical Asset:** A data asset that is no longer being actively updated but retains sufficient historical interest to be remain on the Open Data Portal for the time being.

2. Criteria for Classifying an Asset as Obsolete

An asset may be marked as obsolete if:

- It has been superseded by a newer version or a superior source.
- It no longer aligns with the City's current administrative, legal, or operational norms.
- There are identified issues pertaining to the asset's integrity or accuracy.

3. Criteria for Classifying an Asset as Historical

An asset may be classified as historical if:

- It is no longer being actively updated.
- It still holds public or research interest.
- It does not meet the criteria for being marked as obsolete.

4. Process for Classifying an Asset as Obsolete

Once an asset is deemed obsolete:

- It will be tagged as “obsolete” in the Socrata system. This tag is visible on both the public dataset primer page and in the “tags” field of the “classification” object using the [discovery API](#).
- Its title will be prefixed with “OBSOLETE”
- Its status will be changed to “hidden,” removing it from public browsing and search results. As a hidden asset, it will remain accessible via direct URL and API, to prevent disruptions to existing applications or automated systems.
- A note will be prepended to the asset description in the format, “This dataset is OBSOLETE as of m/d/yyyy and will be removed from the Open Data Portal on m/d/yyyy”
- If the asset has been replaced by another, a link to the new asset will be provided in the description. Append after the note above in the format: “An updated version of this dataset is available at [url]”
- After a specified period (by default, one year), the obsolete asset will be removed from the system and will no longer be accessible by direct URL or API.
- Dataset owners/contributors will be notified at each step of this process.

5. Process for Classifying an Asset as Historical

Once an asset is determined to be historical:

- Its estimated update frequency will be set to “Historical Data.” This field is visible on both the public dataset primer page and in the “domain metadata” field of the “classification” object using the [discovery API](#).
- The maintenance plan details will be set to “This dataset is no longer being updated. It is being retained on the Open Data Portal for its potential historical interest.”
- The time period of the data in the dataset will be appended to the title “m/d/yyyy - m/d/yyyy” (or a single date if a snapshot)
- Its status will remain unchanged, allowing it to remain visible in public browse and search.
- The asset will undergo periodic review to determine whether it should remain historical or be reclassified as obsolete.
- Dataset owners/contributors will be notified at each step of this process.