



## ODC SCI Version Control Policy

### Editing Published Datasets (Public with DOI)

PIs can request permission to make changes/updates to **published** datasets from their lab. In order to maintain proper data provenance, if a published dataset needs to be edited, **please contact the ODC SCI Data Team ([data@odc-sci.org](mailto:data@odc-sci.org)) with an explanation of the changes you wish to make.** The Data Team will assess the proposed changes on a case-by-case basis to inform you of how the changes will be applied and guide you through the process.

For published datasets, the ODC SCI provides a version control policy to maintain data provenance. Minor corrections may not require the creation of a new version and can be applied to the current DOI and data file. Larger changes will lead to the generation of a new DOI; each DOI and landing page associated with the dataset will include new sections providing:

- Links to the other versions (including linking previous versions to the most recent version)
- An explanation of changes from previous versions

### What are minor corrections?

Minor corrections can be applied to a specific DOI and landing page and do not result in the creation of a new DOI and landing page. The ODC-SCI will keep a change log of all minor corrections. Minor corrections include:

- Correcting typos in the dataset, data dictionary, or metadata
- Editing existing author/contributor information
  - Correcting typos
  - Updating affiliation
  - Adding ORCID
- Minor update to the title/metadata abstract
- Changing the contact author to another existing author
- Adding new originating publication

### What changes involve creating a new DOI?

If the Data Team concludes that a change should result in the creation of a new DOI, they will guide you through an expedited data reupload and publication process. The resulting published DOI landing page will also be linked with previous versions.

The following changes will likely require new DOI generation:

- Major update to the title/metadata abstract
- Adding/Removing Authors, Contributors, or Acknowledgements



- Adding new columns (variables) or rows (samples) to the dataset
- Removing columns or rows from the dataset
- Updating definitions in the data dictionary
- Adding new rows to/Removing existing rows from the data dictionary
- Changing the supplementary files

## Deleting Published Datasets (Public with DOI)

Once a DOI and landing page have been published to the public on the ODC-SCI, we will not delete the information. However, PIs can ask for the dataset and associated data dictionary and supplementary files to be retracted.

**To initiate the process, please contact the ODC SCI Data Team ([data@odc-sci.org](mailto:data@odc-sci.org)) with an explanation of why the dataset needs to be retracted.** The Data Team will assess the request. If approved:

- The dataset and associated files will be removed from the public DOI landing page.
- A notice and explanation of retraction will be attached to the public DOI landing page.

## Editing/Deleting Datasets in the Personal and Lab Spaces

Unpublished datasets in the Personal or Lab data spaces do not have the same data versioning restrictions as published datasets. Datasets in the Personal and Lab Spaces can be altered and deleted at the discretion of the data uploader and respective PI. No version control is applied to these datasets at the moment. We are developing version control features in the Personal and Lab space to be implemented in the near future.

## Editing/Deleting Datasets in the Community Space

Unpublished datasets in the Community (i.e. Commons) data space do not have the same data versioning restrictions as published datasets. Currently, no version control is applied to data released into the ODC Community Space. Editing or deleting such datasets is within the purview of the respective PI. We are developing a version control feature in the Community space to be implemented in the near future.

Note that if work (e.g. a paper) is produced with a dataset in the Community data space, we require that the authors of the work contact and collaborate with the PI of the dataset to publish the dataset in some form. In this way, data reused from the Community data space will be maintained with proper data provenance.