



Depositing and Managing Data

Supported File Formats

Any kind of file can be uploaded to Dataverse, but extra functionality is supported for some file types:

Tabular data

Tabular data (Stata, SPSS, Excel, R, & CSV) is **normalized** to **.tab** format on upload—a non-proprietary archival format that a variety of programs can read. See bit.ly/2VmKTGF for more about this process. Normalization is important for long-term preservation of digital data; however, the deposited files can be downloaded in multiple formats, always including the original. You will choose the format from a list when you click **Download**.

In Dataverse, tabular normalization also allows you to perform statistical data exploration and visualization right in your browser. Click the **Explore** button in the **Files** tab to see what it can do! Learn more about the TwoRavens Data Explorer application at bit.ly/2H2QQOC.

Compressed Files (.zip, .tar)

Compressed files are the preferred method for uploading large datasets or many files to Dataverse.

Zip files are automatically extracted on upload, and the contents will appear as a list under the Files tab. Folder structure and file hierarchy within the zip file are maintained on extraction (as of DV version 4.11).

Geospatial shapefiles and Astronomy data (FITS) also have special functionality in Dataverse. Contact Data Services (dataservices@carleton.ca), for more information if your data involves these file types.

File Size

Dataverse is not intended to handle Big Data. Current **file size limits** are:

- For upload: **2.5 GB** per file (unlimited number of files in a dataset)
- For tabular normalization: **500 MB** per file (tabular files over this size will remain in their original format, e.g. Excel (.xlsx))
- While there is no cap on the overall number of files that you can upload. **If your data exceeds 10 GB**, please contact dataservices@carleton.ca to discuss the best repository options.

Access Control

Sensitive files can be **Restricted** so they are not freely available to download. Files can either be fully restricted, or set up so users can send access requests to the contact email for review.

To restrict a file:

Under the dataset **Files** tab, click the checkbox beside the file(s) you want to restrict, then select **Edit Files** -> **Restrict**. In the popup window, you may describe the conditions of use or license for this file in the Terms of Access box. Check the **Request Access** box to allow users to send the contact an email asking for permission to download a restricted file. Click **Continue** to finish.

Files can also be **Unrestricted** if the terms change — for example, once an embargo period has passed.

To Unrestrict files click the checkbox for the restricted file under the **Files** tab, then **Edit Files** -> **Unrestrict**.

Licenses and Terms

You have control over how your data can be used. Dataverse allows for a variety of licenses and terms of use:

- **Built-in License Templates** — can be selected at dataset creation, and changed at any time. These automatically apply the right information about the license to the dataset's metadata.

Apply a template to your dataset by selecting one from the drop-down menu under Dataset Template, at the top of the New Dataset creation form.

- **Custom:** One-size-fits-all licenses don't suit every dataset, you can create customized terms and conditions of use.

To customize terms:

Any dataset saved in Draft or Published form will have a Terms tab on the dataset page. Click **Edit Terms Requirements** and fill in the form with your Terms of Use and/or Terms of Access for Restricted files. If you need help creating a custom license, contact dataservices@carleton.ca.

Version control

Every change made to a dataset—adding files, editing metadata, etc. — is saved as a new version of the dataset. This allows you to track the change history of the project, which can be viewed under the **Versions** tab. This is useful if you need to roll back to a previous version, or find out who made what changes, when.