



How to structure a dataverse for your project

SUMMARY

.....	
BEFORE YOU START.....	1
DATA TYPES	2
GENERAL STRUCTURE	2
DESCRIPTION AND METADATA	3
ACCESS AND SUBMISSION.....	4
FURTHER HELP	5

BEFORE YOU START

- Firstly, consult the [Checklist Prepare your dataset](#) each time you intend to create a dataset and/or a dataverse.
- **Check the quality of your data** using some [useful tools](#) and **provide the maximum level of FAIRness in your data to allow its reproducibility and reuse.**
- Select data that has not been published yet or received a DOI.
- If you have compiled it, please be sure that you are following the same data publication enlisted in your project's [Data Management Plan](#).

DATA TYPES

Initially, you and your research team should think carefully about the **data types** produced during your research project that **you are planning to share and publish in Data@UNIMI**. **Make sure you first check the [Data@UNIMI vs Drive guidelines](#).**

Preferably, raw data – that is to say, all the working data (i.e. data production or elaboration is still in progress and/or data which is not useful for the reuse and interoperability of the research's outputs) should be preserved and shared with your research team via other storage platforms and clouds. Indeed, repositories like Data@UNIMI are devoted to sharing data that have been processed and/or represent the final stages of a research process and are thus ready for interoperability and reuse, including:

- **final data resulting from the research process;**
- **work data necessary for the reproduction of the research process;**
- **data referred to scientific publications produced by the research project.**

Deliverables, such as newsletters, dissemination plans, workflows and project's standards, communication activities, proceedings of meetings, interviews, videos and presentations, management documents and data management plans can be shared via Data@UNIMI. They **can be considered as outputs of the research project only if they are granted open access**. If you want to share these types of documents only with members of your research team, then do not share them via Data@UNIMI.

GENERAL STRUCTURE

Clearly, the structure of dataverses referred to different research projects can vary because each research project is unique. When designing your research project's dataverse structure, follow the general scheme illustrated below.

After having contacted Data@UNIMI support staff for opening a dataverse, you will have a dataverse entitled to your research project inside your departmental dataverse and you and the project members will be the dataverse's administrators. Inside the project's dataverse you will be able to create as many dataverses and datasets as necessary: remember that **Data@UNIMI is built on a hierarchical tree-like structure**. Importantly, we suggest keeping a **clear and logically organized structure corresponding to the project's Data Management Plan**. Corresponding to the specificities of your project you should convey your workspace by considering the outputs of the different work packages and the diversity of data types.

Generally, consider creating:

- **a dataverse for the data produced by each work package (WP) or by each project unit (PU),** each entitled '*WP 1/2/3/etc.: Title of the WP*' or '*Title of the PU*'. For each WP/PU dataverse

you can create multiple dataverses and datasets, according to the research activities and outcomes of the work package.

According to the organization of your research project, you can then think about creating:

- **Inside each of the WP/PU dataverse:**
 - o a dataverse entitled '*Replication data for publications of WP/PU 1/2/3/etc*', where you can share data related to all the scientific publications produced by the referred work package.
 - o a dataverse entitled '*Work data of WP/PU 1/2/3/etc*', where you can share the final and work data which is not related to any scientific publication but which results from the research process and is necessary for its reproduction.
 - o a dataverse entitled '*Deliverables of WP/PU 1/2/3/etc*', where you share the deliverables and outputs of each work package, openly accessible and publicly shareable.
 - o other necessary dataverses, depending on the structure of each work package.

OR consider creating at the main level of the project's dataverse, and not inside the WP/PU dataverses:

- a dataverse for the data related to all the scientific publications produced by the research project, entitled '*Replication data for publications of "Title of the project"*'.
- a dataverse for the deliverables and outputs of the project openly accessible and publicly shareable, entitled '*Deliverables of "Title of the project"*'.

As previously mentioned, it all depends on the data organization, the data accessibility and the project structure of your research project and of each work package/project unit. Anyhow, always remember to **adequately and uniformly name all the dataverses and datasets, also showing the relationships between them**. Moreover, you should keep in mind that a good practice to follow is to store a [README file](#) with each distinct dataset you are uploading in your workspace: by so doing, you will be sure that the datasets' structures are clear and that all the details about data collection, processing, analysis, and naming convention are enlisted.

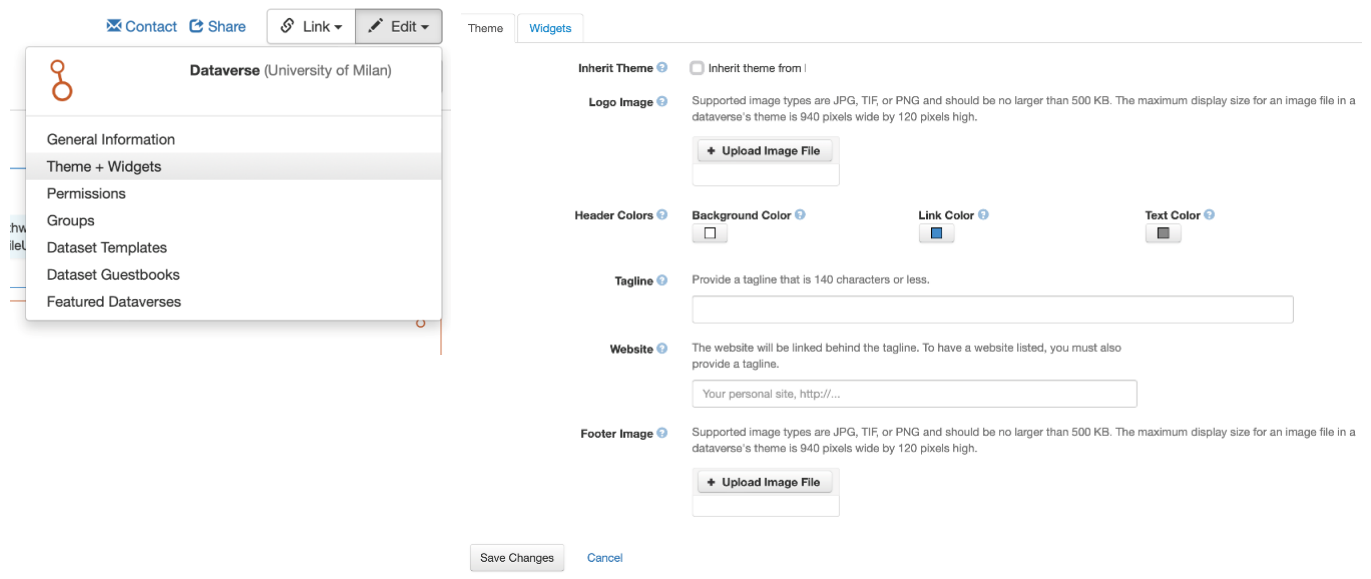
Importantly, when sharing data related to scientific publications, check the guidelines on [How to share data related to a scientific publication](#).

DESCRIPTION AND METADATA

Remember that well-structured and well-described dataverses and datasets are key for a [FAIR](#) use of research data, for this allows others to access, understand and use your research data more easily.

Importantly, at the level of the main dataverse of your research project you should fill in the '**Description**' field with the abstract of your research project: just click on the 'Edit' drop-down menu in the sidebar of your dataverse, then choose 'General information' and compile the referred field. In addition, you can add the logo image of your research project and the direct link to its website:

again, click on the 'Edit' drop-down menu in the sidebar of your dataverse, then choose 'Theme + Widget' and fill in the necessary fields, as shown in the images below. **Remember to always save the changes you make.**



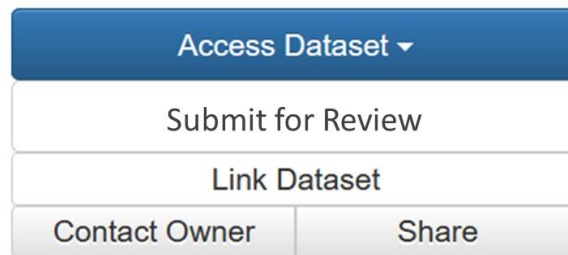
Similarly, each WP/PU dataverse's 'Description' field should contain the abstract and activities of the WP/PU. For each dataverse generated inside the main dataverse of the project you can choose to apply the same **theme** of the "parent" dataverse, so that each dataverse will have the same logo image and website. As previously described, to obtain it click on 'Inherit Theme' in the 'Theme + Widget' option of the 'Edit' drop-down menu.

Finally, keep in mind that **the metadata of a single dataset is related only to that specific dataset**; therefore, you should ensure that **your dataset's metadata is extensively filled in regardless of the general structure of the project's dataverse**. Further, **make sure that you always include the reference to the research project**: the 'Description' field should contain a description of the content and goal of the dataset, whilst the 'Notes' field should specify how data was produced during the research project, indicating the project title. Similarly, fill in also the '**Funding information**' field specifying the funder of your project and, if available, the grant identifier. **Ensure that you provide this information for all of the datasets related to the research project**. You can also fill the 'Related dataset' field to link datasets produced within the project to each other.

ACCESS AND SUBMISSION

Once you have defined the structure of your dataverse, created your dataset(s), exhaustively filled in metadata sections, uploaded data files and documentation, defined license and terms of access to your dataset(s) and files, and checked the quality of your data, you should submit your dataset(s) for publication.

Just click on 'Submit for review' button in the left menù of your dataset, so that Data@UNIMI support staff will automatically receive a notification, review your dataset and, eventually, publish it.



Indeed, for projects that are still in progress, it is preferable to wait before submitting for publication dataverses or datasets: you can decide to keep them unpublished and in draft form and submit everything for review after the conclusion of the project, paying attention to applying the appropriate license.

Otherwise, close to the end of the project, you can submit your dataverses and datasets for review, having them published by Data@UNIMI support staff but keep the access to the uploaded data closed. In this case, you should fill in the 'Terms' field of each dataset with the statement that the files are restricted because the research project they are related to is still on going, and that they will be accessible, accordingly to the chosen license, soon after its conclusion. During the embargo, your metadata will be publicly available. Once the project is finished, make sure that you readily update the terms of access to your research data. Importantly, the key point is not that the project's data will be open, but that it will be accessible: research data should be as open as possible and as closed as necessary.

For further information on licenses, check [Data@UNIMI License Management](#).

FURTHER HELP

- Didn't find what you are looking for? Check the [guide](#) or the UNIMI research data management [website](#).
- Still have some questions? Contact [Data@UNIMI support](#).