Guide – How to publish your research data in Yoda

This guide outlines the basic steps for creating an (electron) microscopy-related Earth science data publication using Yoda at Utrecht University. An overview of the terminology used in this guide can be found at the bottom of this page.

Yoda (short for 'your data') is a research data management service developed by Utrecht University that enables researchers to deposit, publish, and preserve their research data. Once data is published in the Yoda repository, the associated metadata (including a DOI) can be harvested by research data catalogs, such as the EPOS Multi-scale laboratories data catalog, making your data publication *Findable*, *A*ccessible, *I*nteroperable, and *R*e-usable (*FAIR*). The steps to go from locally storing research data to a formal data publication that can be accessed and cited by others, are outlined in the guide below.

The process to publish your data in Yoda is outlined by the 7 steps below:

1) Obtain access to Yoda.

Contact the Yoda Geosciences data manager (Maisam Mohammadi Dadkan - m.mohammadidadkan@uu.nl or Aristotelis Kandylas - a.kandylas@uu.nl) for

a) access to the Yoda data repository to store and publish your research data, and
b) any questions that you may have after browsing through this tutorial.

2) Create the dataset that you would like to publish.

Decide on the data that you want to include in the dataset for publication. According to the EXCITE TNA data policy (attached for reference), all collected data that can potentially be of use to other researchers should be included. If you are unsure what that means for your data, you are encouraged to contact the Facility Manager of the installation where you obtained the data, to discuss what data to include in the data publication. Subsequently,

a) Create a folder structure in which (sub-)folders are hierarchical and follow each other logically, and organize your data. Make sure to assign appropriate names to the (sub)folders and files, making them easy to interpret for others. Examples of folder structures can be found here (1, 2).
b) Create the necessary README file that contains essential information for others to use the dataset, notably how the data was collected, how the data publication is structured, and how it should be read. Examples of README files can be found here (1, 2).
c) Create the StraboMicro project following the attached tutorial. Include the resulting .smz file in the root directory of your data publication.

3) Upload your dataset in Yoda.

Once the Yoda data manager has granted you access, you can upload your data with the desired folder structure. The process to add data using the Web Portal is described here.

4) Add structured metadata to your dataset.

This step creates a so-called data package. How to add structured metadata to your dataset is described here. In the process you will be requested to;

a) Add keywords. Keywords helps other researchers in finding your data publication. Please use the EPOS MSL keyword selection tool (here) to select keywords from the controlled vocabulary.
b) Choose a license. We advise to use a CC BY 4.0 license. This license allows anyone to share (copy and redistribute the data in any medium or format) and adapt (remix, transform, and build upon the data) the dataset, under conditions of i) correct attribution

(you receive credit through citation), ii) duly indicate if and how the dataset is adapted, iii) linking to the license information.

5) **Archive the data package in the Yoda Vault.**

This step transfers the data package for sustainable storage in the Yoda archive (called the 'Vault'). Your data package will be retained unchanged during its retention period in the vault. The steps to archive your data are found here.

6) **Obtain feedback from Yoda data manager.**

After submission to the Vault, the Yoda data manager will review the file types, folder structure, file formats, and metadata. Apply the corrections that the data manager proposes. After you have implemented the suggested changes, the process can be repeated if necessary, or the dataset can be published.

7) **Submit your data package for publication.**

After archiving the data package in the vault, it is ready to be published. Your data package will obtain a DOI and the structured metadata will be published so it can be harvested by data catalogs, making your data publication findable, accessible, interoperable, and re-usable! The process of publishing your data package is described here.