

# Registering a dataset in Pure

Pure infosheet 10, version 13-12-2024

Part of the datasets are automatically registered in Pure (see Extra information). Unpublished datasets and datasets without a DOI must be registered in Pure manually. Also, if a datasets needs to be visible on a publication list immediately, it needs to be registered manually. To register datasets in Pure you need special rights which you can apply for via <a href="mailto:pure.library@wur.nl">pure.library@wur.nl</a>. Usually the registration of datasets will be the responsibility of the data steward, but an (experienced) local Pure administrator is also possible.

#### Log in:

- Go to <a href="https://ris.wur.nl">https://ris.wur.nl</a>
- You'll be automatically logged in. Confirm via your phone if required
- When you work from home, you don't need a VPN
- Click on Datasets in de left hand column

#### Check:

Always check first if a dataset is already registered in Pure. If so, you can add extra information if needed. If you cannot make changes to a dataset, send an email to <a href="mailto:pure.library@wur.nl">pure.library@wur.nl</a>. The Pure team will make the changes for you.

## Finding a dataset (and making changes):

- Go to the editor tab and click Datasets in the left column (not on the plus)
- Remove the filter 'My content' in the top of the screen (click on the cross next to it)
- Search in the search box that states 'Search for datasets...' using the title or DOI
- Click on the dataset and make the necessary changes
- · Click save at the bottom of the window
- If you cannot edit the dataset yourself, send an email to <a href="mailto:pure.library@wur.nl">pure.library@wur.nl</a>

#### Registration:

- Click on the plus next to Datasets in the left column, or click Add content in the right top corner and select Datasets
- Click Dataset types and then Dataset
- Enter all the information you have. The fields with a \* are mandatory.
- Title: the title of the dataset. If the dataset doesn't have a title of its own, you can use 'Data from:' or 'Data underlying the publication:' with the title of the publication.
- Description: You can enter a short description of the dataset here
- Date of data production: date when the dataset was compiled, or or date when the research was completed (not temporal coverage). This field is usually left empty.

## Extra information:

The data librarians register most published datasets with a DOI in Pure. DataCite is used to find the datasets. Additionally, some genetic repositories are searched. The datasets are imported in Pure once every three months. Therefore, a published dataset will not immediately appear in Pure.



- People: use the button 'Add person'. Search the person and click on the name. Select the
  correct affiliation. When entering external persons (not WUR), you usually have to search the
  affiliation yourself. If an external person or external affiliation is not in the system, you can
  create a new one. Usually most authors get the role 'Creator', but it can be specified if you
  wish to.
- Managing organisational unit: is filled automatically. This field determines who is allowed to make changed to the record. If it doesn't fill automatically, select one of the affiliations of the authors.
- Publisher: the organisation that deposits/archives the datasets (not the repository). This is often Wageningen University & Research.
- DOI: the DOI of the dataset (if it is published with a DOI). Do <u>not</u> enter the DOI of the publication.
- Physical data: location of an unpublished dataset. For a dataset on the W-drive or in Yoda:
  - o Title/description: Name of the file or folder
  - o Storage location: location on the W-drive or in Yoda (path or description)
  - Media: Local server Wageningen University & Research, or Yoda
- Link: link to the dataset (if published without a DOI)
- Date made available: publication date of the dataset, or the date the dataset was put in the storage
- Access: select open or closed (restricted with a free login)
- License: select the license of the dataset (for published datasets only)
- Contact person: only mandatory for unpublished datasets
- Temporal coverage: period in which the data is gathered
- Geographical coverage: place or area where the data is gathered or the samples were taken
- Dataset file types: for example csv, xlsx, ... (each in a separate box)
- Keywords: keywords improve the findability of your dataset. Copy the keywords as they are listed in the repository
- Relations to other content: click Research outputs, search for the related publication and click on it. If the publication is not yet registered in Pure, register it in Pure or <u>ask the Pure</u> team to do so.
- Relations to other content: the dataset can also be linked to a project (LVVN, EU, NWO or PhD), or to another dataset.
- Visibility: if the information of the dataset may not be made public, set the visibility to confidential. The dataset will not be visible on Research@WUR, Staff Publications and the WUR website.
- Status: For validation
- Click save

### Alternative searching:

Unpublished datasets don't always have an actual title and therefore can be harder to find in Pure. Alternatively you can search the accompanying publication in Pure and check if a dataset is linked to it. Another option is to search for the creator of the dataset and check if the dataset is on his/her publication list.

## Data policy WUR:

The data policy of WUR states that all datasets underlying a publication should be registered in Pure. This does not only apply to published datasets but also datasets that are archived on the W-drive or in Yoda.

