Adding descriptions of research datasets to PURE

Overview
PURE can be used to store descriptive data on research datasets including access location and conditions. Effective research data management (RDM) is a requirement of research funders such as EPSRC. The current document provides instructions on how to record information on research datasets within PURE.

Guidance on effective RDM has been provided by the Information Directorate (this includes an expanded version of the current instructions).

NB. PURE is not being used to physically store research datasets, only descriptive data related to them. For details on storage options, please see the Information Directorate RDM pages.

How to create a datasets record in PURE
To add a new dataset, click Add new on the right hand side of the Personal overview page.

This will open up a new window with options for adding new content. Click Dataset.
This will open the Metadata window, a new record where you describe your dataset. Fill in as much information as possible, using the next section to help you.

- Fields marked with a red asterisk must be completed
- Help is available by pressing the ‘I’ information icon at the top of each section.

**What information do I need to add?**

This section gives guidance on the most important information you need to enter into the form. You can also enter more detailed information (e.g., geospatial coordinates and temporal coverage).

Note: Some information entered in the form will become visible to the public via the University’s Data Catalogue (the publicly accessible web interface for datasets) if the dataset record visibility is set to Public (see Visibility guidance below).

Other information will only be visible to University of York PURE users, even if visibility is set to Public. This is indicated in brackets in the sub-headings below.

There is an option to make the whole dataset record completely confidential (see Visibility guidance below).

**Title (public)**

Enter a title for your dataset. Keep the title short but descriptive, bearing in mind that it will need to be understood by people who may not be familiar with your project.

You may need to include a date or a version number in the title, to distinguish the current version of the dataset from other versions.

**Description (public)**

Please give as full a description of the dataset as you can, including how people can access the dataset and what form the dataset takes.

If there are reasons why the dataset can’t be made openly available, a brief outline of those reasons should be noted here. It is a requirement for EPSRC-funded projects to publicly state reasons why datasets are not openly available, and also good practice for all projects.

More detailed and confidential statements about the issues relating to a dataset can be entered in the Ethics/legal section (see guidance below)

**People (public)**

Use this field to record all contributors to the dataset. Click Add person and enter the name of the person that you wish to add into the search bar. Click on the relevant name to confirm.

Once you have done this, you will be able to assign that person a role from a predefined list and add any organisational affiliations. Fill in as much information as possible and then click Create.

All University of York staff with a research or academic contract (grade 6 or above) should be in the system. If you cannot find the name of the person you want in the listing, you’ll need to add the individual as an external person.

To do this click on the Create external person option in the Add person window.

You can add external organisations in the same way using the Add organisational unit button.
DOIs/URLs (public)

If you have deposited your data with an external repository and you have a DOI for it, enter it in the DOIs field in the Data availability section.

If you have deposited your data with an external online repository, and you do not have a DOI for it, click on the Add link... button at the bottom of the Data availability section and enter the URL for the repository into the Web address field.

Note: The digital object identifier [DOI®] system provides an infrastructure for persistent unique identification of objects of any type. Therefore, if URLs or services change over time (e.g. the resource moves) the DOI will continue to resolve to the correct resource or service at its new location. www.doi.org provides further information on the DOI® system and the benefits.

Data location (not public)

Use the Data location field in the Data availability section to describe the precise location and form of the data.

If this information is already contained in a Data Management Plan (DMP) and the DMP is accessible via a URL, then you can enter the URL in the Title/description field of the Data location window, in order to save duplicating the information in PURE.

Note: This field will not be visible in the University's Data Catalogue, so it can be used to create an internal administrative record of the location and the form of the datasets.

Access to the dataset (not public)

Please use the Access options field to record the access conditions for the dataset.

- Use Open if access to your data is openly available to anyone, with no restrictions
- Use Closed if access to your data is closed and unavailable to all
- Use Restricted if access to your data is allowed if certain conditions are met (e.g. with a letter of recommendation, adherence to and signing of a Data Transfer Agreement)
- Use Embargoed if your dataset is subject to an embargo period.

If you select Embargoed, an additional field will appear for you to record the embargo period.

If you select Closed, please include a brief explanation in the Description field in the Identification section (please note, the explanation will be public). This is a requirement for EPSRC-funded projects, but also good practice for others.

Access contact details (not public)

Use this field for the primary contact for access to the data.
Note: The contact details here will not be displayed publicly in the University's Data Catalogue, but will be used by University Library staff to arrange access to the data.

**Legal/ethical (not public)**

Use this section to record any legal/ethical issues with the dataset. If there are issues of this nature, please carefully consider the access conditions you apply to your dataset and the visibility you set for the dataset record (see **Visibility** guidance below).

If you enter issues here, you may be contacted by the University Library to confirm that the access conditions and visibility are correct.

**Relations (public)**

You can link your dataset record to other content in PURE (eg persons, projects, publications, activities and impact) thereby building a dynamic and interconnected picture of York’s research.

You should link your dataset record to its associated project, where appropriate. To meet many funder requirements, you should also link your dataset record to all associated publication records. Clicking an **add** button will add one of these associations.

**Visibility (not public – but determines the fields that are made visible)**

Use this field to set the visibility of the dataset record (as distinct from the dataset itself).

You have three options:

- Select **Public** and your dataset record will be visible in the University Data Catalogue and PURE
- Select **PURE users** and your dataset record will visible within PURE only (ie visible to University of York research staff and administrators)
- Select **Confidential** and your dataset record will be visible to authorised users only (ie contributors listed in the **People** section, editors of the record and PURE administrators)
Do not use the **Campus** option, as we are not implementing this option at York.

Only use the **PURE users** and **Confidential** options if there are good reasons why the dataset record itself may not be publicly visible in the University Data Catalogue.

If your dataset is not openly available but you still want the dataset record to be visible on the University Data Catalogue you should select **Closed** in the **Access options** field and **Public** in the **Visibility** field. It is still of value to show that a dataset exists, and is a requirement of some funders (eg EPSRC), even if no access to the dataset can be granted (especially if accompanied by an explanation).  
Note: Dataset records will be visible in the University Data Catalogue by default, unless you change the setting.

*Saving the record*

Set the status bar at the bottom of the page to **Entry completed by User** if the entry is completed, or **Entry in progress** if you have not fully completed the entry.

**Make sure you click Save after completing the record.** If you close the window without doing this, there is no automatic save prompt and the information you entered will be lost!

Note: It is not possible to import information about datasets from an external source, unlike with publications and other research outputs.

*What happens next?*

After your record has been saved with the status ‘Entry completed by User’, it will be checked by University Library staff.

If there is not an appropriate external repository to deposit your dataset in, the University will store it on secure filestore and manage access to it (in accordance with any conditions/embargo noted in the record). In this case, Library staff will be in touch to arrange for a copy of your digital dataset to be transferred to secure storage. Library staff will also check whether a DOI is required and create one if necessary.

Once all of this has taken place, University Library staff will change the record’s status to Entry checked by Library and it will then be visible in the University's Data Catalogue, if the record's visibility is set to Public.

*Who can I contact for advice?*

Questions on RDM should be directed to the Library Research Support team ([lib-research-support@york.ac.uk](mailto:lib-research-support@york.ac.uk)), Questions on PURE should be directed to PURE support ([reo-pure-support@york.ac.uk](mailto:reo-pure-support@york.ac.uk)).