

## Datasets

### Information for UKRI (RCUK) Grant Holders:

The Research Councils wish to promote open, transparent and robust research by encouraging wider data analysis, more data re-use, and the combination of datasets from multiple sources. UKRI are committed to transparency and to a coherent approach across the research base. The 'common principles on data' policy provides an overarching framework for individual Research Council policies on data policy. The seven principles and further guidance can be found at <https://www.ukri.org/funding/information-for-award-holders/data-policy/common-principles-on-data-policy/>

### Information for EPSRC Grant Holders:

From 1 May 2015 the Engineering and Physical Research Council (EPSRC) requires all grant holders to make datasets, arising from EPSRC funded research, publicly available. Information about the EPSRC Policy can be found at <https://www.epsrc.ac.uk/about/standards/researchdata/>

The EPSRC expects that all research data is made publicly available within 12 months of creation. If you have a data management plan, then the plan will guide you through the decision making process. If you do not have a data management plan, you will need to give thought whether and how your data can be made publicly available.

Guidance on how to create a data management plan, along with an online tool to create an EPSRC specific plan, is available at <https://dmponline.dcc.ac.uk>

The EPSRC require grant holding institutions to maintain a catalogue of all datasets arising from EPSRC grants, regardless of public accessibility. Pure will be used for this purpose. The University needs to be able to provide to EPSRC a list of grant numbers with linked datasets, so it is important to link your dataset to the relevant grants in Pure.

### How to add datasets in Pure:

Log in to Pure [www.abdn.ac.uk/Pure](http://www.abdn.ac.uk/Pure) with your usual log in details: you will be presented with your personal account page.

In the left hand menu, click on the + icon next to Datasets. The + sign will appear when you hover the mouse over the Datasets section. Using the + button will open up a new Dataset template. All fields marked with a red asterisk are compulsory fields and need to be completed before you can save the record.



### Filling in the template:

- **Title and Description:** The description should enable other researchers to understand the dataset and allow them to validate your outcomes or re-use the data. If you have published a journal article that describes the methodology you could cut and paste or refer to it in the description section. The description field can take at least 1,000 words (6.5k characters)

Identification ⓘ

Title \*

Description

- Date of data production: This is not a compulsory field. However it is useful to complete these fields so that other researchers can see if the time period is relevant to their research.

Date of data production

Specific date

Period of time

- People: Pure will add your name automatically, and the default role displayed is 'creator'. You can change this by clicking on 'edit' and choosing from the drop down menu. Alternative roles include, creator, owner, contributor, editor, funder, producer, rights holder, sponsor, supervisor, and other. Once edits are complete click update to return to the main template. When adding people, it is recommended that you add PI of the project that generated the data, so the dataset can be linked to the project. Add other authors/contributors etc. by clicking on 'Add person'. In the pop up box that appears, type a person's name and Pure will display all name matches underneath. If a person is on the system, click on their name to add them to the record, if there is no matching name you can use the 'create external person' option to add a new person. You can also add the names of administrators, this will enable them to access the PURE record.

People ⓘ

People \*

No persons or organisational units associated

Search and add Person – or create External Person

*Start entering the person's name, job title or organisation.  
Search results will be updated while you type.*

- Dataset Managed By: Managing organisational unit, the available units that can be selected depends on the affiliated groups and schools of the people who are added to the dataset. However you can add other units (e.g. research groups, themes etc.) or change the managing organisation.

Dataset managed by ⓘ

Managing organisational unit \*



Dataset managed by ⓘ

Managing organisational unit \*

**Biological Sciences**

Organisational unit: Discipline

Change organisation...

- Publisher: this defaults to the University of Aberdeen. If you are recording a dataset published elsewhere, e.g. the UK Data Archive or Figshare, you will need to change the publisher.

Data availability ⓘ

Publisher \*

**University of Aberdeen**

Publisher

Change publisher...

- DOI: If you have a DOI for your dataset, click on add existing DOI to enter the information.

DOI

Add existing DOI

-or-

If you do not have a DOI already one will be issued during validation where appropriate

- Electronic Data: You can upload and store your dataset directly into Pure. Pure accepts a wide variety of file formats, up to 6GB. Please get in touch with [n.strong@abdn.ac.uk](mailto:n.strong@abdn.ac.uk) if you wish to upload a dataset that exceeds 6GB. You can upload multiple files, e.g. a number of csv files, code and a text file explaining how to access the data. Preferred file formats for long term preservation are:

Data type	preferred format	not preferred
Tabular data	CSV, TSV, SPSS portable	Excel
Text	Plain text, HTML, RTF PDF/A only if layout matters	Word
Media	Container: MP4, Ogg Codec: Theora, Dirac, FLAC	Quicktime H264
Images	TIFF, JPEG2000, PNG	GIF, JPG
Structured data	XML, RDF	RDBMS

Further examples are available at <http://www.data-archive.ac.uk/create-manage/format/formats-Table>

- To upload a file, click on drag a file into the Electronic data space or click on Browse to browse your computer for the file or files.

Electronic data



Drag files or **browse** your computer.

Max filesize: 6144 MB.

- Once the file is uploaded a pop up box will prompt you to select, Visibility, licence and type for your dataset(s).

Multiple file upload					
File name	File size	Visibility	Licence	Type	Embargo end date
Book1.xlsx	89 KB	Public – No restriction	No value	Audiovisual	<input type="text"/>

The system defaults to No Value. So you will need to select the most appropriate licence for you data. CC-BY, which allows any user to reproduce all or part of your data provided it is properly attributed to the original author, is the preferred licence type for the EPSRC. If you have any questions about the appropriate licence please contact the Copyright Officer, Mary Mowat, [m.mowat@abdn.ac.uk](mailto:m.mowat@abdn.ac.uk). Further information on licences is available at: <http://creativecommons.org/licenses/>

- Physical data: Most funder’s policies apply to physical data as well as electronic data. You are required to complete the metadata for all datasets, and provide reasonable access to physical data. You will need to describe the data (e.g. paper records of questionnaires) and describe the arrangements that are in place to access the records (e.g. by contacting the person named in Access contact details)

**Physical data**

- Add link: if your dataset is stored in an external repository, please click ‘add link’ and enter the URL.

**Links**

- Date made available: please enter the date. Only the year of publication is compulsory but it is helpful to be more specific if you can.

**Date made available**

Year *	Month	Day
<input type="text"/>	<input type="text"/>	<input type="text"/>

- Access to the Dataset: Options are: Open/Closed/Embargoed/Restricted. For some datasets, the data management plan will set out the appropriate access option. For all data that is not open, a reason for not allowing public access should be given in the Legal/Ethical section.

**Access to the dataset** ⓘ

**Access options**

Not set ▼

- Contact person: you can nominate a contact person for access requests –the default contact displayed is [data@abdn.ac.uk](mailto:data@abdn.ac.uk) this allows the university to monitor/manage access to dataset centrally.

#### Access contact details

##### Contact person



The generic contact email [data@abdn.ac.uk](mailto:data@abdn.ac.uk) will appear when this record is displayed online.

- Geo location: geographical coverage, this is not mandatory but it can be helpful to indicate the regional, national or international coverage of the date. For further information or geospatial references, visit: <http://mygeoposition.com/>. Further information about the metadata fields can be found by on any i-icon on the screen.

#### Geo location

##### Geographical coverage

Point

Polygon

##### Geospatial point (Coordinates in decimal format)

Test

e.g. 57.013543, 9.98934

- Legal / Ethical: If your dataset cannot be made publicly available, you will need to record the reason for that in the Legal/ ethical section. Check the relevant tick box, and enter the reason in the free text box (takes at least 2,000 words). You can draw on your data management plan if you have one, or your ethics application for the research project the data relates to.

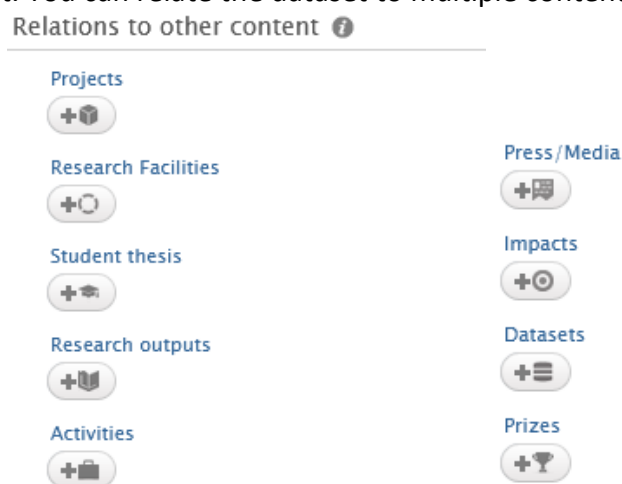
#### Legal/ethical

Is the data subject to any of the following constraints?

- Data protection
- Ethical approval
- Commercial constraints
- Sensitive (Animal, tobacco, security sensitive, GMO, stem cell)

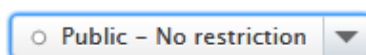
1. If there are any intellectual property issues, please seek advice from your Business Development Officer. If the dataset cannot be made public because of IP issues, then this needs to be recorded.
2. If your data is contains personal information, you will need to give thought to how the dataset can be anonymised prior to publication. If this is not possible, then the reason for not making the data accessible needs to be recorded. You can discuss this further with the University's data protection officer.
3. If your data relates to a project that required ethical approval, you may wish to seek advice from your school ethics contact.

- Relations to other content: This allows you to relate the dataset to other content held in Pure.
  1. Projects: this is the grant that supported the generation of the data. Click on the + icon, then enter either grant number or grant title (or words from the grant title,) and Pure will display the relevant grant(s). You can relate the data to multiple grants (i.e. parent grant, follow on grant etc.). For the EPSRC mandate it is essential that you relate your dataset to your EPSRC grant.
  2. Publications and other content types: You can relate the dataset to relevant publications (e.g. that refer to the dataset), activities (e.g. academic visits or collaborations which helped to generate the data), impacts (e.g. non - academic use or application of the data) or equipment (which was used to generate the data). Click on the + icon under each content type, then in the search bar that appears start entering the title, and Pure will display relevant content in a drop down menu for you to select. You can relate the dataset to multiple content types and records.



- Visibility: the Access options earlier in the template relate to the datasets and the physical data, Visibility at this point in the template relates to the metadata record or data catalogue record, if you are happy for the description and 'metadata' to be visible then the visibility should be set to 'Public - No restriction'.

#### Visibility ?



- When you're finished, remember to save using the blue Save button at the bottom of the screen



The Cataloguing Team within the library will be validating the records prior to publication, i.e. checking the metadata is complete, setting the access options and ensuring that any links to documents or webpages work