Open Data

We encourage staff to share their data where appropriate to do so.

What is Open Data?

Open data supports the inferences and conclusions from our studies and enables others to a) use the data and maximise utility, b) verify our conclusions and promote reproducibility. Sharing data can be facilitated through the PURE repository (preferred). Data may also be shared through the Open Science Framework based at osf.io. Your ethics approvals for your research should guide data sharing, and there is some guidance to support your engagement with Open Science in your ethics application elsewhere in these pages.

What are the benefits of open data?

- Maximizing transparency to promote reproducibility
- Maximizing usefulness of recorded data for the public good
- Reducing the file-drawer phenomenon: increasing efficiency
- In combination with preregistration it avoids suggestions of benefiting from undisclosed

How to deposit and curate your data

1. **Decide on your licence**

Researchers should apply an open licence to datasets to reduce issues around ambiguity and what can and cannot be done with the shared data . This licence is most often a specified Creative Commons licence (click on the link to identify the right licence for you). Please note without a licence it can be used without credit or acknowledgement of the source. PURE colleagues are also very helpful if you need advice about any of this.

2. **Observe the guidelines on data deposit**

The guidelines on data deposit at Queen's University Belfast (available at Standard Operating Procedures). Further support on Research Data Management is available from the library / Open Research Team.

Key publications include:

Setting Up, Maintaining and Archiving Research Files and

Data Management, Collection, Validation and Storage

3. **Make sure the data is suitable for deposit using FAIR principles**

FAIR principles should be applied. Data should be

Findable – data should have a Digital Object Identifier (DOI) which allows it to be easily found using search tools.

Accessible – data should be accessible to others and deposited on a credible repository.

Interoperable – data formats should be adaptable to multiple analytical platforms for future use. We recommend the use of .csv format to support use by others.

Reusable – data should be clearly labelled with an accompanying licence (often CC BY 4.0, but for more information see the creative commons licences links above).

FAIR principles should motivate us to organize the data logically and accompany it by documentation. Importantly, with data accessibility in mind, many of these aspects can be planned for in the data management plan to minimize the time needed, while maximizing the benefits.

4. **Link manuscript with your data**

Use the DOI in your manuscript as you submit to a journal to illustrate where your data is sourced. You may also want to do this with your Open Materials.

5. **Set access rules and encourage others to engage with your work**

You can put an embargo on the date at which data can be accessed such as 12 months from deposit to allow you to publish the papers you wish from the data, and still leave it available for others to access. You may also wish to put out a statement regarding re-use of the data:

Data may be used without express permission from the authors if the following conditions are met.

Please cite the source of the data as:

Jones, A.B. (2022). Name of the dataset from the PURE deposit [Data set]. Queen's University Belfast. https://doi.org/01234.5678910

Jones, A.B. (2022). First paper describing the results of the name of the dataset from the PURE deposit. Journal of Open Science Research. 1(2), 101-102. https://doi.org/10.01234.5678910

Please contact the lead author, Professor Jones (ab.jones@qub.ac.uk),) with any publications arising from the data.

Where to get help?

Within the university, Open data is supported by the Open Research Team. They can help with producing Data Management Plans, which are required by most funders, obtaining DOIs, and with getting access to the Active Data Storage (ADS) service for large datasets. If you have any questions about these aspects, get in touch with rdm@qub.ac.uk.

Where to share data?

There are many ways to make data open. QUB offers researchers to share their research data through PURE (https://pure.qub.ac.uk/en/datasets/), or through the aforementioned Active Data Storage service. Some QUB Psychology data is being shared through the following portals:

Open Science Framework

- Platform dedicated to Open Science.
- Multiple server locations
- Examples: Reid and Dessing (2018), Schultze, Gerlach and Rittich (2017)]

Zenodo

- Servers based at CERN

Beyond that, the following online data repositories are frequently used in our field:

Research Box

- Mirroring AsPredicted.org
- Data and materials archive

For an overview of different online data repositories, see https://www.re3data.org/

Common Questions

Should all data be shared?

No. A commonly used phrasing is "Be as open as possible, as closed as necessary". If you cannot anonymize your data, sharing may not be possible (unless explicitly consented by the participants). Moreover, sharing of data (or materials) may be constrained by considerations of intellectual property.

Is it extra work?

Yes. Many of the repositories require creating an account. Moreover, while uploading data nowadays typically is a drag-and-drop exercise, proper organization of the data using FAIR principles will take some time.