

Preparing your data for deposit / Checklist

Use the following checklist to prepare your data for deposit in any data repository, including the St George's Data Repository

- I have the required permissions to share this dataset

I own the copyright to this data. I have the permission from any other rights holders, including rights holders of any third party material included in the dataset, to share this data

- All datasets have been appropriately anonymised/pseudonymised

- My files are accessible, interoperable and reusable

My files are in open, commonly-used file formats. I have practiced good records management - my files and folders have meaningful, logical filenames and my folders follow a logical structure

- I have prepared a 'readme.txt' file with study and data documentation. This will be included in the deposit to assist others in using my data. In this file I have included:

- Structural information: a list of the files included in the dataset along with a brief description. Information on the folder structure and any relationships between the files. Information about any requirements needed to make use of the files – e.g. software, hardware, original source code – which should be shared either as part of the dataset, or independently and linked to the dataset with its DOI
- Methodological information: describe methods used to collect and process the data and maintain data quality. May include instrument guides and survey questionnaires
- Data-level documentation: variable lists, definitions of abbreviations - including abbreviations used in column headings, units of measurement, codes or symbols used to record missing data

A template for creating readme style metadata can be downloaded from Cornell University's [Research Data Management Service](#)

- All owners of any third party material included in my dataset have been appropriately credited

- I have decided on a suitable embargo period for the dataset – if required

You can embargo a dataset until your research publication is public or under guidance from your funder

- I have familiarised myself with the copyright licenses available on SGUL's Data Repository and have selected an appropriate license for my data

- I have checked that the title of my dataset is not identical to the title of any of my other research outputs

For data supporting a publication it is acceptable to use "Data supporting" or "Data used in" along with the title of the publication as the title of your dataset