EPSRC Frequently Asked Questions:

How can I get help with research data management?
Contact the Research Data Management team in the UofG Library.
Email: research-datamanagement@glasgow.ac.uk
Webpages: www.gla.ac.uk/services/datamanagement/

What is research data / a dataset?
Research data include data, records, files and other evidence on which research conclusions are based. This can include (but is not restricted to):
  • results of experiments or simulations
  • statistics or measurements
  • models and software
  • fieldwork observations
  • survey results
  • interview recordings and transcripts
  • images and videos
  • textual source materials and annotations

A dataset is all the data (and supporting info such as readme files and reuse metadata) that together support a finding or project.

What is metadata?
Metadata is ‘data about data’. The EPSRC place an emphasis on publishing metadata for discovery (metadata which allows relevant datasets to be found). The expectation is that published metadata ‘should be sufficient to allow others to understand what research data exists, why, when and how it was generated, and how to access it’. Where the research data referred to in the metadata is a digital object, it is expected that the metadata will include use of a robust digital object identifier (DOI).

What is a Data Management Plan?
A data management plan (DMP) is a formal document that details how you will handle your data both during your research and after the project is completed. Plans typically state what data will be created and how, and outline the plans for sharing and preservation, noting what is appropriate given the nature of the data and any restrictions that may need to be applied.
The Digital Curation Centre (DCC) has developed useful guidance on DMPs including a summary of the expectations of different funding bodies:
www.dcc.ac.uk/resources/data-management-plans
There is also an online tool for creating DMPs that incorporates the requirements of various funders and some University of Glasgow-specific guidance:
https://dmponline.dcc.ac.uk/
What training does the UofG provide on research data management?
The Staff Development Service provides a range of training courses for researchers. The following courses contain elements of RDM training that are appropriate to the course context:
Winning Research Income
Managing Successful Research Projects
Research Integrity
Principle Investigators Programme
There is also a specific course available via SDS on Managing Your Research Data, and a new course planned for this year on Research Data Management Plans.
Within the Colleges, data management training is being integrated into Supervisor Training Workshops and new College-specific Research Data Management courses are being developed for research students. It is intended that the new student course will be available in the Grad School Training Brochures from April 2015.
In addition to the formal training courses available, the RDM team based within the Library are happy to visit Colleges and Schools to give more tailored talks on aspects of RDM. Contact the RDM team on research-datamanagement@glasgow.ac.uk

My data includes information that is sensitive (commercially, legally or ethically). Am I obliged to make this publically available?
The EPSRC recognises that while it is desirable to make data as openly available as possible, there are situations in which this is not possible. If your data are sensitive and can only be made available under certain conditions (or not at all), publications which refer to this data should contain a statement detailing the conditions on which the data will be available. NB A statement to ‘contact the author’ is not considered to be sufficient.
In addition, although the dataset itself might not be openly available, researchers should, when possible make metadata about the data openly available. This metadata should give the reasons why access to the data is restricted and on what terms it might be available.

How quickly am I required to deposit my data?
Metadata describing the research data is expected to be published and made freely available within 12 months of the data being generated. However, it is often difficult to determine the date of generation for data which is acquired over time or which required extensive processing before it is usable. The following guidance has been issued by EPSRC:
If the data directly supports research findings published after 1st May 2015, this expectation will be met if the publication contains an appropriate data citation. The cited data and supporting documentation should be accessible no later than the date of first online publication of the article. For data which is clearly generated after 1st May 2015 and where there is no intention to publish findings which rely on that data, metadata about these datasets should be available online within one year of the date of generation.
For data which has no clear date of generation and the relevant EPSRC grant ends on or after 1st May 2015, and where there is no intention to publish findings which rely on that data, metadata about these datasets should be available online within one year of the end of the grant.
Where an intention to publish is used as a justification to withhold research data, a publication plan must be in place.

Do I just need to deposit the image I published (which is representative of my data), or do I need to deposit all the images that went into making the figure?
The EPSRC expectation is that all data relied on in published findings will, by default, be available for scrutiny by others. Researchers are strongly encouraged to deposit all the data which went into the making of a figure or experimental outcome.

I generate large amounts of data from simulations or computer modelling. Do I need to deposit / keep this data?
The EPSRC recommend that in this situation, it may be more effective to preserve the means to recreate the data – the code, input and details of the environment – rather than preserving the outputs themselves.

The data I produce is not digital. How do I go about making it openly available?
Data which is not digital should be stored securely for the same time-frames expected of digital data (the EPSRC require storage for 10 years from generation or last access). Digital metadata about the data should be published on the UoG registry (Enlighten: Research Data). The EPSRC do not consider it unreasonable to expect those requesting access to non-digital data to travel to the place the data is held in physical form, if it is not practical to digitise it.

How should I consider RDM issues at the award application stage?
Yes. Unlike most major funders, EPSRC do not require a Data Management Plan (DMP) to be submitted as part of a funding application, however, they do require a DMP to be prepared and held locally. Advice on preparing a DMP can be found on the DCC website:

http://www.dcc.ac.uk/resources/data-management-plans

Additionally, funders consider that costs associated with RDM (for example to pay for repository space or to buy in assistance prior to depositing datasets) are an acceptable cost in a funding application. EPSRC have also indicated that in the future, funding applications in which RDM is not a costed expense might not be considered to be complete. Advice on RDM costing can be sought from the RDM team: research-datamanagement@glasgow.ac.uk

What do I need to do when I publish research?
When you have a paper ready for publication, the dataset underpinning the research should be passed to the RDM team for deposit in the repository (if the data is not already in a subject-specific repository). You will be issued with a Digital Object Identifier (DOI) for the dataset which can be included in the manuscript as part of the data citation.
What counts as an appropriate data citation?
Data citation refers to the practice of providing a reference to data in the same way as researchers provide a bibliographic reference to printed resources. An example of a basic data citation would be...

‘Creator (Publication Year): Title. Publisher. Identifier.’

In this citation, the Identifier field would ideally be a robust, persistent URL such as a digital object identifier (DOI). The Research Data Management Service at the University of Glasgow can mint a DOI for your dataset.

Data citation facilitates the easy verification and reuse of data and can allow the impact of data to be tracked. More information about data citations can be found on the DataCite webpage:

https://www.datacite.org/