University of Reading

EPSRC Data Management guidance

Introduction

This document provides guidance on:

- Why applicants for EPSRC funding must be aware of and prepared to comply with EPSRC research data policy;
- Integrating data management into your research planning at an early stage, so that if your project is funded you will be able to comply with the EPSRC expectations on research data;
- Planning for data storage and archiving, and including any relevant costs in an application.

General guidance on data management planning is available on the Research Data Management website. Contact the Research Data Manager if you have any questions about data management planning for your project or require guidance on use and costing of resources.

Contact: Research Data Manager: researchdata@reading.ac.uk / 0118 378 6161.

The EPSRC expectations, the University and EPSRC applicants

The EPSRC Policy Framework on Research Data specifies nine expectations for research data which apply to both individual researchers and organisations in receipt of EPSRC funding.

Researchers funded by EPSRC are required to preserve and enable access by third parties to data collected or generated by them that directly support published research findings. Data must be openly accessible to others, except where valid legal, ethical and commercial constraints exist.

Because the University is directly accountable to EPSRC for compliance with these expectations, applicants for EPSRC funding must be aware of EPSRC policy and must plan for the preservation and sharing of any relevant data their research will generate. The failure of researchers to comply with EPSRC policy may be reflected in individual performance assessments.

What is required in your application?

You do not need to complete a data management plan attachment as part of your EPSRC application.

But your application should clearly demonstrate that management of data for long-term preservation and sharing are integrated into your research plan. You should make it clear

in the Case for Support that consideration has been given to the resources required for data collection, management, preservation and sharing and any associated costs, which should be included in the budget.

Where the project is expected to generate data of long-term value, you should highlight this in your consideration of project outputs and academic or other impact.

Storage, computing and archiving of data

When developing your proposal you should identify any requirements for resources related to the collection and use of research data, and ensure all costs are included in your budget. In particular you will need to consider:

- how much data you will need to store during the project, where data will be stored, and any associated costs;
- whether any dedicated computing resource is required for computing-intensive proposals, and if so at what specification and cost;
- what data you expect to preserve and share over the long term and what service or services you will use.

Where costs are identified these will have to be itemised in the costs for your application and justified in the Justification of Resources section.

Storage

Data collected/held at the University should be stored using University-managed infrastructure, which will provide data security, replication in separate data centres, automated backup and file recovery. For the different options available, and information about costs, please read the guidance here.

Storage costs should be based on the volume of data to be generated/collected in the project, and should be identified on the application as a Directly Incurred cost.

Computing

If you have computing-intensive requirements, custom specifications of CPU, memory, storage and GPU can be purchased from the University on a pro rata basis. For information about Research Cloud computing visit the Academic Computing Team portal.

Archiving

To preserve and enable long-term access to research data supporting project outputs you should use managed services. These may be either:

- Data repositories, which provide long-term storage and curation, licence and manage access to data, and publish online metadata records with unique identifiers (typically Digital Object Identifiers or DOIs), so that data can be easily discovered, cited and linked to;
- Publishers, which may publish some data as Supplementary Information alongside associated articles on their websites. It is standard for Supplementary information to be freely available, even where the related article is behind a paywall.

Data repositories

As a general rule you should always use a data repository to preserve and enable access to your primary data. This is an option open to all University researchers, who are eligible to deposit data in the University of Reading Research Data Archive.

This service will preserve and enable access to data in the long-term (10 years minimum). Up to 20 GB of data per project can deposited at no charge. Deposits greater than 20 GB may be subject to a charge and must be agreed in advance. If you intend to deposit more than 20 GB of data in the Archive, contact researchdata@reading.ac.uk to discuss.

Note that the Archive should be used to deposit a defined, curated, publicly-accessible dataset that supports published project findings. It is not a private post-project storage area for all digital materials.

There may be external services that are more suitable for your data if they service particular subject communities or manage specific types of data. Guidance on choosing a suitable repository can be found here.

Publishers' websites

In some areas of research (e.g. characterisation of new chemical compounds with associated experimental data) it may be customary to supply evidence for validation of published findings as Supplementary Information alongside the article on the publisher's website. If this is the evidentiary standard in your field, you can plan for sharing of data by this means.

In general, however, it is not advisable to share substantive primary data as Supplementary Information (SI). Publishers often make SI files available in PDF format, which is not a usable data format, and there is not necessarily a guarantee that the data will be preserved and remain accessible in the long term.

Drafting a data management plan

Although EPSRC does not require you to submit a data management plan as part of your application, it publishes a list of Research Data Principles that state 'project specific data management policies and plans […] should exist for all data'. This means that if you are awarded EPSRC funding, you and your team should develop a full data management plan for your project.

It is a good idea when you are preparing your application and to start sketching out a full plan in collaboration with your team and/or partners. This can help you:

- plan the practicalities of data collection and management in your project, and anticipate any special requirements, issues or particular challenges that will have to be dealt with;
- ensure all relevant costs are identified and included in your application.

You can use a tool called DMPonline to draft a plan. Using this tool you can generate a Data Management Plan template, which includes detailed prompts and guidance to help

you complete each section of the plan. Even if you do not draft your plan using the tool, the guidance can help you with your research planning. Plans can be saved, shared with co-applicants, and exported.