GCRF Internal Application Data Management Plan Guidance

Introduction

Primary data of significant value collected or created in support of project findings must be preserved and made freely accessible to others by deposit in a suitable public data repository, unless there is a valid legal, ethical or commercial reason for restricting access to them. This is a requirement under both the UKRI's Common Principles on Data Policy and the University's Research Data Management Policy.

This requirement may not apply where data do not have significant value (for example, test data) or where the cost of archiving is in excess of any likely benefit (as may be the case with certain model output data, for example).

Applicants for GCRF Substantial Research funding must complete a Data Management Plan (DMP) of no more than 200 words as part of their application. Where the intention is not to make data freely accessible, the reason for restricting or not providing access must be specified in the DMP.

All submitted DMPs will be reviewed by the Research Data Manager.

Guidance on completing the DMP is provided in this document.

The Research Data Manager can be contacted for advice and comments on draft DMPs prior to submission. General guidance on data management planning is available on the Research Data Management website.

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

What is required?

The Data Management Plan should briefly cover the following points:

- What data outputs will be generated by the research? Identify any outputs that will not be suitable for archiving.
- Where will data be archived for long-term preservation and access?
- When will data be made accessible?
- Will there be any restrictions placed on access to the data for legal, ethical or commercial reasons?

Data outputs

Data outputs that will be generated by the project should be succinctly described. Each data type should be identified, and details about data formats and likely volume or

quantity of data included where known. If the data volume is likely to exceed 100 GB, this should be stated.

If any of the data outputs will not be suitable for archiving, these should be identified and the reason they are not suitable for archiving stated.

Archiving

Researchers should where possible use UKRI-funded or other data type-specific repositories to preserve and enable access to data. UKRI data repositories include the NERC data centres, the ESRC's UK Data Service ReShare repository, and the Archaeology Data Service, funded by AHRC and NERC. BBSRC contributes to a number of international bioscience data sharing resources, including the molecular biology databases of the European Bioinformatics Institute. The Wellcome Trust also maintains a useful list of approved data repositories. For a more general search, consult the data repository registries FAIRsharing or re3data.org.

Where no suitable external service exists, researchers can deposit data in the University's Research Data Archive. Research data in non-digital formats and digital data that cannot be made accessible or require controlled access should also be registered in the University Archive. The Archive can provide a mechanism to regulate access to controlled data under data sharing agreement where this is necessary.

Guidance on selecting a suitable data repository can be found here.

Access to data

Data should be made accessible no later than publication of the main findings. If release of data may be delayed, for example, pending confirmation of IP protection, this should be stated in the DMP.

Access restrictions

If any legal, ethical or commercial reasons for restricting access to data are anticipated, this should be stated.

Where data are collected from human subjects, in most cases these can be anonymised for sharing. A valid reason for restricting access to such data would obtain only if it will not be possible to anonymise the data (e.g. biometric data may be intrinsically identifying) or if the risk of causing harm or distress by disclosure is significant. See the section on Research ethics and data protection below for further guidance.

Even data containing personal or confidential information may be shared under certain conditions, with appropriate consent. Some data repositories, e.g. UK Data Service ReShare repository and the European Genome-phenome Archive, can manage controlled access to sensitive/confidential data. The University Archive can also offer a restricted access option. Contact Robert Darby for advice.

It is acceptable to restrict access to data if the data are commercially confidential or there is a commercial pathway for the research, for example involving an identified industrial

partner. If IP protection may be sought, it should be possible to release data once protection has been confirmed.

Further guidance

Storage and computing requirements

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.

Data collected in the field should be stored securely and backed up using local devices and transferred at the earliest opportunity to the primary storage location.

Sensitive/confidential data can also be stored in these locations; if such data are stored or shared using other devices or cloud services, this should be in accordance with the University Data Protection, Remote Working and Encryption Policies accessible here.

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. Information is available in the Academic Computing Team website.

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.

Research ethics and data protection

You have an ethical obligation to protect the confidentiality of personal information provided to you by research participants, and you must also comply with data protection law if you collect and process personal data. Where personal data are processed in jurisdictions outside the European Economic Area, they should be handled to the standards prescribed by UK data protection law.

Any research involving human subjects will need to receive approval from your School's or the University's Research Ethics Committee. Guidance can be found here.

Your application for ethical approval and information sheet/consent form should not make any commitment to destroy confidential data by a given time or not to share (anonymised) data collected from research participants. In most cases data can be shared openly if they are anonymised. It is good practice to secure consent for data sharing when you recruit participants, e.g. by including in your consent form a statement such as: 'I understand that the data collected from me in this study will be preserved and made available in anonymised form, so that they can be consulted and re-used by others'. The UK Data Service provides excellent guidance on consent and anonymization, and has sample information sheets and consent forms.

If you will be processing personal data in your research, you are advised to consult University guidance on Data Protection and Research. The Data Protection Checklist for researchers is a good starting point. There are also sample information sheets and consent forms.

Personal data is any information relating to an identified or identifiable natural person. These data enjoy statutory protection under the General Data Protection Regulation 2016 and the Data Protection Act 2018. Under this legislation any personal data collected by you must be processed fairly and lawfully. Among other things you will be required to issue a Privacy Notice to your research participants, which explains the purpose(s) for which the data are being collected, your lawful basis for processing the data, who the data will be disclosed to, and the rights of the individuals in respect of their personal data. For certain kinds of research, for example involving the processing of sensitive data or human genetic data, you will need to complete a Data Protection Impact Assessment under the advice of the University Information Management & Policy Services Officer.

You must ensure that personal data are kept secure and are not disclosed to unauthorised persons. You should use a locked storage container such as a filing cabinet in a locked office for paper-based personal data; for digital data, password-protected or, preferably, encrypted storage. This particularly applies in the case of special category sensitive personal data, which include information about an individual's: race; ethnic origin; politics; religion; trade union membership; genetics; biometrics (where used for ID purposes); health; sex life; or sexual orientation. Such personal data should be encrypted, and not stored or shared by means of cloud services other than a University OneDrive account, or transferred via unencrypted channels (e.g. via email). You can transfer data to a location on the University network using VPN, which provides an encrypted channel.

Consent forms must be retained by you and/or your School for a minimum period of five years from the completion of the research, and for as long as the personal data are held.

Working procedures should be designed to minimise the risk of inappropriate disclosure. Data can often be pseudonymised for purposes of processing and analysis, with the personally-identifying information and their linked IDs stored separately from the working dataset. When the study is complete and if there is no further need to link individuals to data, the linking key can be destroyed, so that the data become fully anonymised.

You can retain personal data on a continued basis for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes. You do not need to commit to destroy personal data at a set time, but they should be managed under a retention schedule that specifies periodic reviews, so that they can be securely destroyed when no longer needed.