

Introduction

The Pro Vice-Chancellor Research is the owner of the Research Data Storage Policy. The Research Data Storage and Management Executive will modify this policy in the light of developments and guidance from funders and other bodies. It is intended that this policy should align with the requirements of UK Research and Innovation (UKRI) and the University of Bristol's Research data management and open data policy. The Research Data Storage and Management Executive reports through the HPC Executive, to the Advanced IT Board as set out in its terms of reference, https://www.acrc.bris.ac.uk/acrc/RDSME_TOR.pdf

Funder policy

The Research Data Service has put together a guide for the major UK and international funders outlining each funder's policies and requirements and detailed guides to writing a Data Management Plan (DMP) for each funder. <http://www.bristol.ac.uk/staff/researchers/data/writing-a-data-management-plan/>

RDSF pricing policy

Information about the current costs of storing data in the RDSF is provided at <http://www.bristol.ac.uk/acrc/research-data-storage-facility/rdsf-costs/>. The principle underlying the costs is modelled on the Princeton 'Pay Once, Store Forever' (POSF) model, which calculates a one-off charge for data storage which will cover future refresh of the media by applying Moore's Law to storage costs.

We have defined 'forever' in this context to be 20 years.

The University of Bristol Research data management and open data policy

This is a University-wide policy that provides guidance and support on the responsibilities of the University and its staff in managing and preserving current or future research data.

<http://www.bristol.ac.uk/research/environment/governance/research-data-policy/>

University of Bristol Research Data Service

The University's Research Data Service offers advice, support and training in all areas of research data management and is responsible for the University's data.bris Research Data Repository - <https://data.bris.ac.uk/data/>.

The Repository enables researchers to upload subsets of data from the RDSF and make them publicly available, linking them where appropriate to a publication through creation of a Digital Object Identifier (DOI).

Policy for the use of the Research Data Storage Facility V 2.0 January 2021

- This document should be read in conjunction with the Research Data Storage Facility Terms of Use.
- For clarification of terminology, refer to the Glossary in Appendix 1 of this document.
- This policy will be reviewed by the Research Data Storage and Management Executive on a regular basis.

1. Management of the Research Data Storage Facility (the Facility)

- 1.1. The scope of this policy concerns only the digital research data assets held in the Facility.
- 1.2. The Facility must not be used for the storage of non-research-related data. Other University facilities exist for such data.
- 1.3. The Facility not only provides storage for the research data assets, but allows them to be processed, manipulated and mined.
- 1.4. The Facility will store both live data and archived data. Live data will be active with changes being made to it, whilst archived data will only be accessible in its current form without the possibility of making changes.
- 1.5. The Advanced Computing Research Centre (ACRC) will assume responsibility for maintaining the Facility infrastructure, that is, the storage hardware, storage file system and storage media. Maintaining the integrity and validity of the data is the responsibility of the Data Steward.
- 1.6. Day to day management of the Research Data Storage Facility is undertaken by the ACRC under the direction of the Research Data Storage and Management Executive.
- 1.7. Assuming appropriate capital expenditure is available, the data held will be migrated as appropriate on to new media as required and as technologies change. If appropriate capital expenditure is not available, the data will be held on low-cost media such as tape until the media decays, at which point the data will no longer be available.
- 1.8. The ACRC will support researchers to comply with the **University of Bristol Research data management and open data policy** and with relevant funder policy. This will be achieved by endeavouring to provide storage for the research data that meets both the legal and regulatory framework for particular types of research and the terms and conditions imposed by external research funders.
- 1.9. All data stored within the Facility will have an owner, who must be an employee of the University and who must be able to take long-term responsibility for the data. This person will be known as the Data Steward. A Data Steward will normally be the PI of the project in the first instance, but in every case, the Data Steward must be a University of Bristol employee. Students and honorary staff members cannot be Data Stewards.

If a Data Steward is leaving the University of Bristol, they must inform the ACRC and provide the name of an alternative Data Steward for any data they hold in the RDSF before they leave the University.

If the Data Steward subsequently leaves the University or is absent for a prolonged period of time (e.g. on sabbatical) without nominating an alternative Data Steward, their line manager or the Head of School, as appropriate, will be contacted by the ACRC and asked to nominate a successor. If no successor is nominated, the line manager or Head of School will assume the responsibilities of Data Steward until an appropriate alternative can be identified. In exceptional circumstances, the Research Data Storage and Management Executive may agree that an existing Data Steward can take over as Data Steward of the departing Data Steward's project(s) as well, even if this increases the total amount allocated to the existing Data Steward above the current 'free' allocation.

Where personal data (as defined by Data Protection Legislation) is processed for research purposes in the course of a researcher's employment at the University of Bristol, the University will be the Data Controller under the Act for that Data. This will include where personal data is stored in the Facility.

However, in all circumstances, the Data Steward will be personally responsible to the University for ensuring the proper administration and oversight of any data they have caused to be stored in the Facility. This will include the Data Steward providing such information as is reasonably required by the Research Data Storage and Management Executive to make an adequate risk assessment for data storage, and to meet the University's legal and ethical obligations.

- 1.10. Ensuring that the research project data is readable and accessible in future is the responsibility of the Data Steward. Readability and accessibility could be affected by, for example, changing document formats. It may not be possible to exactly reproduce computational data on a long-term basis if the computer on which the data was originally run no longer exists or is no longer functional. Data will be returned to the Data Steward in its original form. Data will be treated as raw data, not interpreted data.
- 1.11. It is the responsibility of the Data Steward to ensure that the data is stored in as resilient a state as is appropriate to the data type and research. The University will use its reasonable endeavours to recover any data which may be lost.
- 1.12. It is the responsibility of the Data Steward to choose appropriate data storage classes for their data.
- 1.13. It is the responsibility of the Data Steward to validate any data to be uploaded into the University's *data.bris* repository for publication and to authorise the upload. As part of the validation process, the Data Steward will be asked to reconfirm compliance with current and relevant legislation (such as Data Protection Legislation and the Freedom of Information Act) and the University's ethical requirements. Generally data in the repository is published as 'open data', however if the data has ethical or commercial restrictions the repository can be used to deposit materials under controlled access agreements, see <https://www.bristol.ac.uk/staff/researchers/data/publishing-research-data/publishing-data-under-access-restrictions/> for more information.
- 1.14. If the capacity of the Facility nears saturation, the Research Data Storage and Management Executive will review the data held and may contact a Data Steward to see whether part or all the data held on their behalf is still required. The ACRC will aim to work with the Data Steward to move data to another infrastructure where appropriate.
- 1.15. Each project must have an exit strategy in line with the requirements of the relevant funder. The RDSF project application form will ask how long the data needs to be held for after the end of the project. Data will be kept for 12 months if a Data Steward leaves without appointing a replacement, and then the Research Data Storage and Management Executive will consider at the next review whether the data remains within the Facility or is transferred elsewhere.
- 1.16. In the event of a security breach, the ACRC will investigate what has occurred and then make a report to the University's Information Security Manager. If appropriate, the system will be quarantined and appropriate computer forensic techniques applied. Any investigation will be in line with the University's Investigation of Computer Use policy - <http://www.bristol.ac.uk/infosec/policies/docs/>.

2. Application and review process

- 2.1. Data Stewards will apply to use the storage facility using a web-based application form detailing the amount of storage required, giving a brief description of the data to be stored and for how long it should be stored, and providing a name for the project. Requirements for levels of security will also be requested (see University of Bristol policies relating to information security and management - <http://www.bristol.ac.uk/infosec/policies/docs/>)

Each application will be given the actual or notional cost of the storage that they are applying for. The minimum project size is 100GB.

The Data Steward must ensure that no sensitive, confidential or personal information is included in the project name.

A Data Steward can nominate one or two senior RAs in the research group to be Deputy Data Stewards. A Deputy Data Steward can add users to the project and remove users from the project. A Deputy Data Steward can also prepare data for publication, although the Data Steward still authorises publication, as set out in the RDSF Terms of Use.

A Data Management Plan or a similar set of guidelines may be submitted with the application. This plan or set of guidelines can be more or less detailed depending on the amount and type of storage requested and the complexity of the data to be stored.

On application, the Data Steward will be asked to confirm that all machines connected to the storage have adequate virus protection. It is the Data Steward's responsibility to ensure that all data is virus scanned before it is stored on the RDSF.

If a Data Steward subsequently finds that more data storage is required than originally requested, an amended request will need to be submitted to the Research Data Storage and Management Executive.

- 2.2. If a Data Steward wishes to share data with a researcher at another institution, the Data Steward can apply for a Collaboration Project and arrange for the External Data User to register to join the Collaboration Project. The Data Steward then authorises the External Data User to become a member of the Collaboration Project.
- 2.3. The Data Steward will be responsible for the production and storage of metadata to address the semantic issues of what the data means.
- 2.4. Where the project generating the data has required an ethical review by an Ethics Committee, such as a Department or Faculty Ethics Committee, the Data Steward will be responsible for providing a copy of the ethics application and documented evidence of the Ethics Committee's approval and any conditions.

Proposed storage of unanonymised sensitive and/or personal data (as defined by Data Protection Legislation) will require the Data Steward to provide a copy of the relevant ethical review application and documented evidence of the Ethics Committee approval. Please also refer to section 2 of the Research Data Storage Facility Terms of Use.

- 2.5. Under the Freedom of Information Act 2000, third parties may request access to information held by Public Authorities, subject to certain exemptions. Such exemptions are interpreted strictly. Universities are defined as Public Authorities under the Act and research data may thus be requested under FOI legislation. An amendment to Section 22 of the Act in 2014 provides an exemption which should protect the majority of ongoing research where there is an intention or requirement to publish the data at some future date - <http://www.legislation.gov.uk/ukpga/2000/36/section/22>.

The Data Steward must first consult with the University Secretary's Office, and then inform the Research Data Storage and Management Executive, at the time of application, if they believe an exemption to third party access under the Freedom of Information Act 2000 should apply to the data they wish to store in the Facility. They must provide details to the Secretary's Office of why the exemption applies to the data, and how long the exemption should last. If a FOI request is received, whether the exemption applies will be assessed by the Secretary's Office on receipt of the request.

- 2.6. The Data Steward must consult members of the ACRC when completing a grant application and Data Management Plan if a significant and long-term storage need is identified.
- 2.7. All applications for new RDSF projects will be reviewed by the members of the Research Data Storage and Management Executive. Once approved, the Data Steward's storage account will be created.
- 2.8. Data Stewards will be asked annually if the data is still required, and if so, whether the work that produced it has been cited. Details will be requested of further research projects using the data and any grants related to the research data held. This annual review and impact analysis will help in ensuring sustainability of the Facility.
- 2.9. Storage holdings and assets will be reviewed by the Research Data Storage and Management Executive. If there has been no demonstrable activity with the data within that 12 month period, the Research Data Storage and Management Executive will consider the need to migrate the data down the storage hierarchy, for example from disk to tape media. The Research Data Storage and Management Executive reserves the right to review how data is held in the light of overall storage requirements.
- 2.10. The conflict resolution process will involve disputes being escalated by the Research Data Storage and Management Executive to the HPC Executive and then to the Advanced IT Board, who will act as the final arbiter.
- 2.11. Users are encouraged to mention use of the storage facility in publications, where the data produced by or underpinning the research is stored in the Facility. The suggested template wording is "*The work was made possible in part by using the Research Data Storage Facility of the University of Bristol - <http://www.bris.ac.uk/acrc>*".
- 2.12. If an external party is part of a joint project with a member of staff from the University of Bristol and requires access to the Facility, they may apply to join a Collaboration Project as an External Data User for a maximum period of 3 years. If the External Data User's account is still required after 3 years, the Data Steward can ask External Data User to re-register for a further 12 months. The University reserves the right to terminate external accounts if circumstances require.
- 2.13. If a linkage file relating to anonymised datasets needs to be stored in the RDSF for back up purposes, this can be stored by the Research Data Service, acting in the role of independent Data Steward and the Data Steward of the anonymised datasets will have no access to the linkage file.

Appendix 1

Glossary

Archive – a separate, static copy of the digital data held, usually on a read-only basis.

Collaboration Project – a project registered by a Data Steward, specifically for sharing data with researchers from outside the University of Bristol.

Data Controller – a person who determines the purposes for which, and the manner in which, personal information is to be processed. This may be an individual or an organisation and the processing may be carried out jointly or in common with other persons e.g. the University.

Data Management Plan (DMP) - a structured plan increasingly asked for by funding bodies, usually developed in two stages: an initial version at the grant application stage, and a fuller version which is developed at the early-project stage. It provides a summary of the data that will be created during the project and how it will be stored, curated and made accessible.

Data Protection Legislation means the **UK Data Protection Legislation** and (for so long as and to the extent that the law of the European Union has legal effect in the UK) the General Data Protection Regulations (EU) 2016/679) and any other directly applicable European Union regulation relation to data protection, where “**UK Data Protection Legislation**” means any data protection legislation in force from time to time in the UK including the Data Protection Act 1998 or 2018 or any successor legislation.

Data Steward - the person with responsibility for the data, usually the Principal Investigator (PI) of the project. The Data Steward must be an employee of the University of Bristol.

Data User – someone authorised by the Data Steward to have access to the data assets of the project. The Data User must be an employee of the University of Bristol.

Digital Object Identifier (DOI) – a unique and permanent character string which identifies an electronic document (for example a publication or a data set.)

External Data User – someone collaborating with a Data Steward, who is not an employee of the University of Bristol, but is authorised by the Data Steward to have access to the data assets of a Collaboration Project.

Research Data Storage Facility (RDSF) or “the Facility” – the University’s integrated, resilient facility for the storage of research data.

Standard Project – a project registered by a Data Steward to enable use of the RDSF and which can be accessed only by members of the University of Bristol.