Research Data Management Policy

Contents

Research Data Management Policy

Summary

This University is committed to the principles of Open Research.[1] We subscribe to the principles of the UK Concordat on Open Research Data,[2] and affirm that publicly-funded research data are a public good produced in the public interest, which should be made available with as few restrictions as possible in a timely and responsible manner.[3] We recognise the fundamental importance of research data management in establishing the quality and integrity of research and enabling the verification of research findings. Datasets can also be valuable research outputs in their own right, and should be managed in such a way as to maximise their potential for use and impact.

University researchers are required to plan for data management, to collect and manage data responsibly during the research process, and to preserve and make accessible to the fullest extent possible data that substantiate research findings, considering any legal, ethical, commercial or practical restrictions that may apply. Wherever possible, data must be preserved and made accessible by means of an appropriate data repository, and in accordance with the FAIR Data Principles for Findable, Accessible, Interoperable and Re-usable research data.[4]

Scope and application

This policy covers all data collected or created under the authority of the University for the purposes of answering research questions. It includes any computer code created in research that is integral to the reproducibility of research results.

The policy applies to all researchers involved in collecting research data for or on behalf of the University, including researchers employed by the University, research students and visiting researchers. Data preservation and sharing requirements apply to undergraduate or taught postgraduate students if they are involved in research that results in the publication of peer-reviewed research outputs.

Responsibilities of researchers

It is the responsibility of researchers to ensure that research data are managed and shared in accordance with this policy and other relevant policy requirements.[5]

---

[1] University of Reading Statement on Open Research. https://www.reading.ac.uk/research/research-environment/open-research.

[2] See Concordats UKRI has signed at https://www.ukri.org/manage-your-award/good-research-resource-hub/open-research/. The Concordat on Open Research Data is also signed by Universities UK.

[3] UKRI Common principles on research data. https://www.ukri.org/manage-your-award/publishing-your-research-findings/making-your-research-data-open/.

[4] FAIR Principles. https://www.go-fair.org/fair-principles/.

[5] Information about related policies and requirements can be found in the Guidance section.

Researchers are responsible for completing relevant training, including mandatory training for University employees in Information Security, Data Protection and Freedom of Information.[6]

All researchers involved in the handling of research data in a project should understand their roles and responsibilities, and should seek clarification of these from the Principal Investigator where necessary.

Principal Investigators (including doctoral supervisors) are responsible for the management and sharing of research data collected in projects under their authority. They must ensure that the project complies with this policy and any other relevant policies, laws and contracts.

Data management planning

A data management plan (DMP) should be created at the outset of any project that involves the collection of data, and maintained for the duration of the project. The DMP should describe the data that will be collected, how the data will be collected and managed during the project, and measures that will be taken to preserve and share data in line with FAIR Principles on completion of the research and publication of findings.

Grant applications and costs

Where a DMP is required as part of a grant application, it should be reviewed by the University's Research Data Service prior to submission of the application.

Equipment and infrastructure requirements for the collection, storage, processing and archiving of research data must be considered when developing proposals, and eligible costs should be included in grant applications.

Participant-based research

Where research involves participants, researchers must inform participants during recruitment of the intention to preserve and provide access to research data collected from them, unless for any reason it will not be possible to make data accessible to others. They must wherever possible plan to prepare data for sharing, usually by anonymisation, and identify an appropriate repository to deposit the data in, which may be an open data repository or one that can provide access to higher-risk data on a controlled basis.

Collaborations and research agreements

Where research is undertaken in collaboration with third parties, intellectual property rights in data, plans for data preservation and sharing, and legal responsibilities, for example in respect of personal data, must be agreed at the outset.

Collaborative research agreements should recognise the broad objectives of this policy, and allow for the timely sharing of research data in support of project findings, while accommodating the legitimate interests of parties to the agreement.

---

[6] Mandatory online training. https://www.reading.ac.uk/human-resources/people-development/new-staff-central-induction/mandatory-training.

Managing research data

Research data should be managed by researchers in accordance with a data management plan, following defined procedures for data storage and backup, file naming and organisation, version control, quality control, documentation and metadata creation.

Researchers must store and process data securely using University-approved infrastructure in accordance with the University's Information Compliance policies.

Preservation, sharing and disposal of data

Data that substantiate published research findings (including doctoral theses) or that have long-term value must be preserved and made accessible using a suitable data repository wherever possible in accordance with the FAIR Data Principles for Findable, Accessible, Interoperable and Re-usable research data, considering any legal, ethical, commercial or practical restrictions that may apply.

Data that are shared must be accompanied by sufficient documentation and relevant materials (such as code written to process data) to enable verification or reproduction of published results and re-use for research or other purposes.

Data collected from research participants should be shared wherever possible in anonymised form as open data or, in the case of higher-risk anonymised or identifiable data, by using a data repository that can hold and provide access to datasets under a controlled access procedure.

Where high volumes of data have been generated in the research, and it is not viable to deposit them in a data repository, they should where possible be held in University infrastructure. A metadata record describing the data should be published in the University's Research Data Archive, so that they can be cited in related publications and interested parties can request access.

When data should be shared

Data supporting research findings must be made available no later than the date of first publication of research results, unless there is a legal, ethical, commercial or practical reason why data cannot be released at that time.

Disclosure of data associated with commercial exploitation of research may be delayed while IP protection is sought. Once protection is in place, data should be shared.

Use of data repositories

Researchers must use a data repository to preserve and share data wherever possible. The University recommends the use of disciplinary and data type-specific repositories where appropriate; but all University researchers have the option of using the University's Research Data Archive.

When a dataset has been deposited in a data repository external to the University or otherwise stored for preservation, the researcher should register the dataset with the University by creating a metadata record in the University's Research Data Archive.

Formats and standards

Data should be preserved in open or widely-used file formats wherever possible, so that access is not dependent on purchase of a software licence. Data and associated metadata should be structured and described using relevant standards where these exist.

Rights and licensing

Data that are deposited in a data repository or disclosed to parties other than the members of the University and research collaborators must be accompanied by a rights statement identifying the owner(s) of intellectual property rights in the data, and a licence specifying the terms on which the data can be used.

Wherever possible, data should be made available under an open licence, such as the Creative Commons Attribution (CC BY) licence, in order to provide the broadest possible scope for re-use, unless legal obligations, third party rights, intellectual property rights or privacy rights preclude this.

More restrictive licences should be used where there is a well-founded reason to do so, for example, to protect commercial interests. Wherever possible in such cases a standard licence should be used, such as the Creative Commons Attribution-NonCommercial (CC BY-NC) licence, which grants permission for use in research and teaching and other non-commercial activities.

Data availability statement and data citation

Any publication based on the use of data must include a data availability statement citing the supporting dataset in the holding repository by means of its digital object identifier (DOI) or other unique identifier, or referencing any alternative storage location, and stating on what terms the data may be accessed. The full citation for the dataset should be provided in the publication's list of references.

Preservation term

Relevant data should be preserved for at least ten years after completion of a project, or for any longer minimum term specified in a funder's policy.

Disposal of data

Data that do not substantiate published research findings and do not have long-term value should be disposed of when they no longer serve a purpose. Data that contain personal or confidential information must be disposed of securely in accordance with University guidelines.[7]

---

[7] Guidelines for record retention and disposal. https://www.reading.ac.uk/imps/records-management/record-retention-and-disposal-schedules.

Intellectual Property Rights

The University has ownership of intellectual property rights (IPR) in research datasets created in the course of research undertaken by researchers in its employment, unless rights have been contractually assigned to another party.

Ownership of IPR in a research dataset created by a student rests with the student, unless it has been assigned to another party under the terms of a contract of sponsorship or other agreement.

Where a dataset is the creation of more than one hand, IPR may be jointly held by more than one party.

Disclosure of data, for example by deposit in a data repository, must be with the permission of all rights-holders and subject to any relevant authorisation, such as by a project's Principal Investigator. Where the University is the rights-holder, disclosure of research data must be in compliance with the Code of Practice on Intellectual Property.[8] Subject to the above conditions, University employees are empowered under this policy to disclose or authorise the disclosure of data collected by them for research purposes, unless a reason for restriction exists.

Support for research data management

The University provides infrastructure and services to support the effective storage, management, preservation and sharing of research data collected under its authority, in compliance with this policy and related policies and external requirements.

The University's Research Data Service promotes and supports good practice in research data management and compliance with this policy, and manages the University's data repository, the University of Reading Research Data Archive.

The Research Data Service is based in the Research Engagement team in the Library. Contact details can be found on research data management website.[9]

Monitoring and compliance

The University will monitor compliance with this policy on the part of its researchers and may take appropriate action where cases of non-compliance are identified.

Responsibility and review

Responsibility for the Research Data Management Policy sits with the Pro-Vice-Chancellor (Research and Innovation) and the University Board for Research and Innovation.

The Policy will be reviewed every three years.

---

[8] Intellectual property. https://www.reading.ac.uk/essentials/the-important-stuff/rules-and-regulations/intellectual-property.
[9] Research data management. https://www.reading.ac.uk/research-services/research-data-management.

VERSION	KEEPER	APPROVED	APPROVED BY
1.0	Research Data Manager	2 March 2015	UBRI
1.1	Research Data Manager	26 September 2017	Research Dean for Environment
2.0	Research Data Manager	23 January 2024	UBRI

Guidance

Related policies and other requirements

Research data management is subject to some or all of the following policies and requirements:

- the University's Information Compliance policies, including *inter alia* the Data Protection, Encryption, Bring Your Own Device and Classification policies;[10]

- the University's Code of Practice on Intellectual Property;[11]

- the University's Code of Good Practice in Research;[12]

- any legal obligations, including those specified in the Freedom of Information Act 2000, the Environmental Information Regulations 2004, and the Data Protection Act 2018;[13]

- any contractual obligations, such as the provisions of research agreements relating to the use and disclosure of intellectual property and publication of results;

- the data sharing policies of research funders.

Data management planning

A data management plan (DMP) describes the data that will be collected in the course of a research project, how the data will be collected and managed, how relevant ethical and legal requirements will be met, and measures that will be taken to ensure the preservation and sharing of data on completion of the research and publication of findings. A DMP also defines data management roles and responsibilities within a research project.

The University provides guidance on data management planning, and can advise on and review DMPs on request.[14]

Grant applications and costs

Some funders require researchers to submit a DMP or discuss how they will manage, preserve and share research data as part of the grant application. Where this is required, the Research Development Manager supporting the application will ensure that the DMP is reviewed by the Research Data Service before the application is submitted. The University provides information about the requirements of funders and guidance on preparing DMPs.[15]

---

[10] Information Compliance policies. https://www.reading.ac.uk/imps/information-compliance-policies.
[11] Intellectual property. https://www.reading.ac.uk/essentials/The-Important-Stuff/Rules-and-regulations/Intellectual-property.
[12] Quality assurance in research. https://www.reading.ac.uk/academic-governance-services/quality-assurance-in-research.
[13] Information Management and Policy Services. https://www.reading.ac.uk/imps/.
[14] Data management planning. https://www.reading.ac.uk/research-services/research-data-management/data-management-planning.
[15] Writing a DMP for a grant application. https://www.reading.ac.uk/research-services/research-data-management/data-management-planning/writing-a-dmp-for-a-grant-application.

Many funders of research accept that data management costs are legitimate research expenses. Requirements for the collection, storage, processing and archiving of research data must be considered when developing proposals, and eligible costs should be included in grant applications. Any third-party suppliers used for the processing of data must be approved by the University.[16]

Participant-based research

Applications for ethical review of research proposals submitted to the University Research Ethics Committee (UREC) and to some School Research Ethics Committees must be accompanied by a DMP describing the data to be collected, measures that will be taken to protect participants' data, and plans for sharing of research data. The DMP will be reviewed by the relevant Research Ethics Committee, which may specify requirements for the proposed research to receive a favourable opinion.[17]

Planning for sharing of data collected from research participants

Wording that can be used to notify an intention to preserve and share data, either openly or under safeguards, is provided in the sample consent form provided by Information Management and Policy Services (IMPS). The IMPS Data Protection and Research guide provides guidance on anonymisation of research data.[18]

If data collected from participants are considered to be sensitive or higher-risk and not suitable for public sharing, they can be archived in some repositories under controlled access procedures. The UK Data Service ReShare can hold anonymised safeguarded data and the University's Research Data Archive can accept restricted datasets, which may include identifiable and confidential information.[19]

Collaborations and research agreements

Standard research agreements include provisions that support the objectives of this policy. Research Contracts will ensure the appropriate implementation of research agreements.[20]

Managing research data

Effective management of research data according to a data management plan will increase the efficiency of the research process, safeguard against corruption or loss of data, and make it easier to prepare data for preservation and sharing at the end of the project.

---

[16] Information about University storage and computing services and guidance on information security is provided here: https://www.reading.ac.uk/research-services/research-data-management/managing-your-data/data-storage.

[17] Guidance on the UREC procedure can be found here: https://www.reading.ac.uk/academic-governance-services/research-ethics/what-do-i-need-to-do. The REC DMP template, with guidance for applicants and reviewers, can be found here: https://www.reading.ac.uk/research-services/research-data-management/data-management-planning/research-ethics-and-data-protection.

[18] Data protection and research. https://www.reading.ac.uk/imps/data-protection/data-protection-and-research.

[19] See guidance at https://www.reading.ac.uk/research-services/research-data-management/data-management-planning/research-ethics-and-data-protection.

[20] Research Contracts. https://www.reading.ac.uk/research-services/research-contracts.

The University's research data management website provides guidance on managing research data, including data storage, data organisation, file formats, quality control, documentation and metadata, and online survey tools available through the University.[21]

Preservation, sharing and disposal of data

Preservation and sharing of FAIR data entails:

- the use of appropriate repository services to preserve, publish information about and provide access to the data;

- the use of open and widely-used standards to encode and describe the data;

- the use of open or other standard licences which define re-use permissions;

- the assignment to datasets of globally unique and persistent identifiers (such as a Digital Object Identifier or DOI), in order to enable citation, discovery and re-use.

The FAIR Principles are described in detail on the Go FAIR website.[22]

What data should be preserved and shared

Sufficient data and supporting information to enable verification or reproduction of the research results and meaningful re-use should be preserved. Documentation should include details of the methods and instruments used to collect and process the data, and interpretive metadata, which may be in the form of a data dictionary or codebook.

While there may be legitimate legal and ethical reasons why some data cannot be shared, these can often be addressed by redaction of a dataset, for example to remove personal data, and/or by depositing data in a repository that provides access to data on a controlled basis (for example, subject to authorisation and under a data access agreement). Most data collected from participants can be preserved and shared on some basis. Researchers must inform participants of the intention to share data, either openly or under safeguards, during recruitment. Failure to do this is poor research practice.

Sharing of data obtained from and about Indigenous Peoples should respect Indigenous Peoples' rights and interests in their data and should be undertaken in accordance with the CARE Principles for Indigenous Data Governance.[23]

Some areas of research, for example computational modelling, can generate large volumes of synthetic data. In such cases it may be more important - and more viable - to preserve the reproducibility of a method rather than a specific set of outputs, which might have little long-term value and be prohibitively expensive to maintain. It may be sufficient to preserve the code used to execute a model along with information about the input data used and the computing environment in which the model was run.

---

[21] Managing your data. https://www.reading.ac.uk/research-services/research-data-management/managing-your-data.
[22] FAIR Principles. https://www.go-fair.org/fair-principles/.
[23] CARE Principles for Indigenous Data Governance. https://www.gida-global.org/care.

Guidance on preserving and providing access to restricted and high-volume data is provided in the research data management website.[24]

When data should be shared

Researchers are entitled to a limited period of privileged use of the data they have collected to enable them to publish the results of their research. Once the research findings are placed on public record, the evidence to support those findings must be made accessible unless a justifiable exemption applies.

The University provides assistance with the protection of intellectual property.[25]

Use of data repositories

A data repository is a service that exists to preserve and provide access to research data. It supports the long-term management of and access to data in accordance with the FAIR Data Principles. Among other things a repository will:

- apply data and metadata standards to support interoperability and re-use;

- publish machine-readable metadata about a dataset to enable online discovery;

- assign a persistent unique identifier, such as a DOI, to a dataset to make it citable and discoverable;

- publish licence information to facilitate re-use.

A data repository should be the preferred solution for data preservation and sharing. Sharing of raw data as supplementary information alongside the article on the publisher's website should be avoided. It is not acceptable to retain data in closed storage, to be made available on request, except where there is a well-founded legal, ethical, commercial or practical reason for doing so.

Most data collected from research participants can be deposited in a data repository either as anonymised open data, or, in the case of higher risk anonymised or identifiable data, by using a data repository that can hold and provide access to datasets under a controlled access procedure. For example, the UK Data Service can hold safeguarded data, and the University's Research Data Archive can manage restricted datasets.

Data centres and repositories serving disciplinary communities and catering to specialised types of research data should be used where they are available and suitable. Requirements or preferences of funders should be taken into account when considering the most suitable place of deposit for research data. The University of Reading Research Data Archive can be used in most cases in the absence of a suitable external service.

---

[24] Where to archive data. https://www.reading.ac.uk/research-services/research-data-management/preserving-and-sharing-data/where-to-archive-data.
[25] Intellectual property and commercialisation. https://www.reading.ac.uk/research-services/intellectual-property-and-patents.

The research data management website provides guidance on where to archive data[26] and on depositing data in the University's Research Data Archive.[27]

Formats and standards

If data have been stored in a proprietary format, researchers should convert the data into an open or widely-accessible file format for preservation purposes. The research data management website provides guidance on preparing data for archiving.[28]

Standards are community-defined sets of requirements, specifications, or guidelines that can be used to encode, structure and describe data. Examples of standards are:

- data models and associated formats for encoding specific types of data, e.g. the Network Common Data Form (NetCDF) standard for the numerical representation of scientific data, widely used in climate modelling, and the Brain Imaging Data Structure (BIDS) standard for organising neuroimaging and behavioural data;

- metadata standards for the structured description of data objects;[29]

- controlled vocabularies;

- persistent identifiers for entities, such as DOIs and ORCIDs.

Standards are essential to the machine-actionability of data and metadata, and support discovery, access, interoperability and re-use in accordance with the FAIR Principles. Repositories may require submission of data and metadata in conformity with specified standards.

Rights and licensing

The University does not prescribe the use of a specific licence, but recommends the Creative Commons Attribution (CC BY) licence or one with similar permissions for open data. An Open Source licence such as Apache or GNU GPL can be used to share software code. Where there is a need to restrict permissions, standard licences should be used where possible. The Creative Commons suite of licences includes versions with non-commercial, no derivatives and share alike terms.

The research data management website provides guidance on licensing data[30] and on publishing and licensing software code.[31]

---

[26] Where to archive data. https://www.reading.ac.uk/research-services/research-data-management/preserving-and-sharing-data/where-to-archive-data.

[27] UoR Research Data Archive. https://www.reading.ac.uk/research-services/research-data-management/preserving-and-sharing-data/uor-research-data-archive.

[28] Preparing for data archiving. https://www.reading.ac.uk/research-services/research-data-management/preserving-and-sharing-data/preparing-for-data-archiving.

[29] Information about metadata standards can be found at https://rdamsc.bath.ac.uk/ and https://fairsharing.org/search?fairsharingRegistry=Standard.

[30] Licensing data. https://www.reading.ac.uk/research-services/research-data-management/preserving-and-sharing-data/licensing-data.

[31] Research software and code. https://www.reading.ac.uk/research-services/research-data-management/about-research-data-management/research-software-and-code.

Data availability statement and data citation

The research data management website provides guidance on data availability statements and data citation.[32]

The inclusion of a data availability statement in a publication is a requirement of the UKRI Open Access Policy.[33]

Preservation term

Most data repositories undertake to hold datasets indefinitely. A minimum retention or preservation period may be specified in the repository's policy.

The University's Research Data Archive guarantees a minimum retention term of ten years, but is committed to the long-term preservation of research data. A dataset would only ordinarily be in scope for deletion if it had not been accessed for ten years. Datasets may be removed for other reasons, for example, if they are found to be in breach of the law or the rights of other parties. If a dataset is withdrawn from a repository, a 'tombstone' metadata record will be retained to preserve a citable record of the dataset.

Disposal of data

IMPS provides guidelines for the retention and disposal of records.[34]

Intellectual Property Rights

Further information is provided in the Code of Practice on Intellectual Property.[35]

Researchers who leave the University should ensure that any data belonging to the University are handed over to an appropriate data steward, such as their line manager, or, in default of any other person, the Head of School.

Support for research data management

The Research Data Service promotes good research data management practice through the provision of information, support and training; supports data management planning for research funding applications; manages the University's data repository, the Research Data Archive; and monitors and supports compliance with this policy.

The Research Data Manager works with the Committee on Open Research and Research Integrity, the Committee for Research Infrastructure, relevant professional services and members of the research community to ensure that research data management is integrated into University policy and infrastructure.

---

[32] Data access statements. https://www.reading.ac.uk/research-services/research-data-management/preserving-and-sharing-data/data-access-statements.

[33] UKRI Open Access Policy. https://www.ukri.org/publications/ukri-open-access-policy/.

[34] Guidelines for record retention and disposal. https://www.reading.ac.uk/imps/records-management/record-retention-and-disposal-schedules.

[35] Intellectual property. https://www.reading.ac.uk/essentials/the-important-stuff/rules-and-regulations/intellectual-property.

Glossary

Data repository: a service that provides long-term preservation of and access to research data in accordance with the FAIR Data Principles. Among other things a repository will:

- apply data and metadata standards to support interoperability and re-use;

- publish machine-readable metadata about a dataset to enable online discovery;

- assign a persistent unique identifier, such as a DOI, to a dataset to make it citable;

- publish licence information to facilitate re-use.

Data repositories may serve defined disciplinary research communities or members of a research-performing organisation or operate as public data sharing services.

FAIR Data Principles: a widely-accepted set of principles for the management of research data in ways that facilitate discovery, access and re-use. FAIR stands for Findable, Accessible, Interoperable and Re-usable.[36]

Principal Investigator: The individual at the University with overall responsibility for the conduct of a research project. Doctoral supervisors are included under this definition.

Research data: the information collected or created, processed and studied for the purpose of answering a research question, and which constitute the evidential basis that substantiates research findings. They may be primary data generated or collected by the researcher or existing sources consulted and processed as part of the research activity. Research data may be digital or non-digital, and may include, but are not limited to: results of experiments or simulations, statistics and measurements, computational models, software and code, fieldwork observations, survey results, interview recordings and transcripts, images from cameras and scientific equipment, databases, textual or linguistic corpora, lab books, and physical objects, such as samples and specimens.

Research data management (or RDM): the sum of activities undertaken by researchers and research organisations in relation to the collection, processing, preservation, sharing and disposal of research data.

Researcher: anyone undertaking research or involved in collecting research data for or on behalf of the University, which includes but is not limited to research staff, research students and visiting researchers.

---

[36] FAIR Principles. https://www.go-fair.org/.