

Restricted Dataset Deposit Instructions

University of Reading Research Data Archive

Contents

Introduction	2
Contact	2
What are restricted data?	2
Is the Research Data Archive the right place for the data?	3
Where is the data stored?	4
How long will it take?	4
Ethical and legal basis for archiving and sharing restricted datasets	5
Intellectual property rights and permissions	5
How is access to the data provided and managed?	6
Deposit instructions	6
Citing the dataset	10
Access requests	11
Ongoing management	11

Introduction

These instructions describe the process to be followed for depositing a restricted dataset in the [University of Reading Research Data Archive](#).

Contact

To discuss archiving a restricted dataset, contact the Research Data Service at researchdata@reading.ac.uk. You can book a consultation via the appointment booking link on the [research data management website](#).

What are restricted data?

The University's [Classification Policy](#) identifies three categories of information:

- **Unrestricted:** information that can be made freely available to the public;
- **Restricted:** information that shall not be made public, and shall only be handled by members of well-defined groups of individuals that have a reasonably foreseeable need. It might, for example, include personal information, commercially sensitive information or information originating elsewhere that has been made available to the University on a restricted or confidential basis;
- **Highly Restricted:** information that shall be restricted to a small number of named individuals, each with a clearly defined need. It might, for example, include sensitive personal information, such as details of health status, or information with the potential to cause harm to individuals or to the University if compromised lost or disclosed.

In terms of research data **Unrestricted datasets** are research datasets that can be made openly accessible via the Research Data Archive (or any other suitable data repository).

Restricted datasets are research datasets that meet the definition of Restricted or Highly Restricted information in the Classification Policy. They cannot be made openly accessible via the data repository, but providing certain conditions are met, they can be deposited with the University Archive to be preserved and shared with authorised users subject to a Data Access Agreement. This will involve publication of an Archive metadata record, so that a dataset is discoverable and citable, secure storage of the related data files in internal University storage, and a process by means of which access to the dataset can be requested, reviewed by a Data Access Committee, authorised where appropriate, and provided for use by the applicant subject to a Data Access Agreement between the University and the recipient organisation.

The category of restricted data may encompass datasets containing identifiable information and data that have been de-identified where a higher risk of disclosure is considered to exist, for example through linkage to private databases, and/or where a there is deemed to be a higher risk of harm in the event of disclosure, for example because of the sensitivity of the subject matter.

The following are examples of datasets that may fall within the definition of restricted data:

- A database of video files showing research participants producing utterances of various lengths in spoken and sung conditions of various emotions. The recordings show the faces of participants. Although the participants are not identified by name in the dataset or its documentation, the recorded facial images make the participants identifiable and constitute personal data.¹
- Transcripts of interview with students at UoR about their perceptions and experiences of racism. Although direct identifiers such as names have been removed, the transcripts contain information about students' courses, experiences, places of residence, social and family relationships, which in combination could be sufficient to identify individuals. Full anonymisation of the transcripts would have been difficult to achieve, and removal of possible identifying information would have reduced the value of the data as a research resource.
- A research student has collected experimental data as part of a PhD funded by a commercial sponsor. Under the sponsorship agreement, ownership and disposal of IP generated by the student belongs to the industrial sponsor. The sponsor has a commercial interest in the data and is unwilling for them to be made available as open data, but is willing for the data to be archived at the University on a restricted access, with any access requests being subject to authorisation by sponsor before data can be supplied to an applicant.

Is the Research Data Archive the right place for the data?

The University's Research Data Archive is not the only data repository that can manage research data under a controlled access procedure, and you should consider whether it is the best place for such data.

The process of depositing data and of managing them on an ongoing basis involves considerable administration. Legal Services will have to draft a data access agreement as part of the deposit process, so that it is ready for execution in the case of a successful request to access the dataset. The depositor will need to identify a University PI for the dataset, who would consider any request for access to the data as part of a Data Access Committee convened whenever a request is received. This person must be empowered to authorise or refuse data access requests. Additional contacts for a dataset may need to be identified if organisations other than the University hold IP rights in the data.

The Research Data Archive restricted datasets option is best suited for high-risk data containing identifiable or other confidential information where a highly-managed approach is justified.

¹ This is a description of Ong, Jia Hoong, Leung, Florence and Liu, Fang (2021): The Reading Everyday Emotion Database (REED). University of Reading. Dataset. <https://doi.org/10.17864/1947.000336>.

In the case of anonymised data where there is considered to be some risk of identification through linkage with other existing data or where access only by authorised researchers is required or preferred we advise deposit with the UK Data Service as 'safeguarded data'.² This is a suitable option for data in a broad social science scope including biomedical data. Data deposited under the 'safeguarded' policy would only be made available to users registered with the UK Data Service under the terms of its safeguarded data end user licence, which requires users to maintain the confidentiality of the information provided to them.³ This is still a regulated process, but it is managed according to a standard protocol, and does not require authorisation by the PI on every occasion or the involvement of an applicant's organisation in signing a data access agreement.

The UK Data Service is not the only data repository that can manage datasets under a controlled access procedure. For example, the European Genome-Phenome Archive can be used to archive personally identifiable genetic, phenotypic, and clinical data generated in biomedical research projects.⁴ There may be other services catering to specific disciplines or types of data that can offer archiving for controlled access. You can search for a suitable data repository via re3data.org and [FAIRsharing](https://fairsharing.org).

Where is the data stored?

Restricted datasets will be held in a secure storage location on the University network under the direct management of the Research Data Service. Only staff supporting the Research Data Service will have access to the data in this location.

Deposit of a restricted dataset will involve creating a public metadata record in the Research Data Archive and uploading a publicly available readme or documentation file describing the dataset and any other files suitable for being made public. The metadata record will provide information about the dataset and the means by which access can be requested. It will also include a Digital object Identifier (DOI), which can be used to cite the dataset and link to the metadata record.

How long will it take?

The deposit process for restricted datasets can take several weeks to complete. Legal Services will need to be instructed to draft and set up a data access agreement, which may take some time. If other parties have intellectual property rights in the data, agreement for the data to be archived at the University will need to be secured. This may have to be formalised in a data storage agreement, which would also require the involvement of Legal Services.

² <https://reshare.ukdataservice.ac.uk/>.

³ <https://reshare.ukdataservice.ac.uk/legal/#Safeguarded>.

⁴ <https://ega-archive.org/>.

The actual deposit process should not take more than a few days to complete, provided that there is an established ethical basis for archiving, permissions and agreements have been finalised, and the dataset and documentation have been prepared.

Ethical and legal basis for archiving and sharing restricted datasets

The ethical basis for archiving and sharing of restricted datasets is established through approval to conduct research by a Research Ethics Committee, and consent to participate in research given by participants on the basis of information provided to them. The intention to preserve and share data collected from participants must be recorded in the ethical approval documentation, must be communicated clearly and transparently to participants through the consent process, and must be documented in auditable form.

If the ethical basis is not clearly established, it may not be possible to archive the dataset as a restricted dataset. An ethical basis for archiving as a restricted dataset may fail to be established if the participant:

- has been informed that the data will be destroyed by a certain time;
- has been informed that data will not be shared outside the project team;
- has not been informed that data will be preserved and may be shared with other researchers subject to safeguards.

The legal basis for archiving datasets containing personal data is referable to data protection laws. Data protection laws contain provisions and exemptions for uses of personal data for *processing for archiving purposes in the public interest, [and] scientific or historical research purpose[s]*, subject to safeguards. Archiving of research data containing identifiable information is lawful on a public interest basis, providing the conditions for such archiving are met, and that the data subject has been informed of the purpose of archiving their data, how long the data will be archived for, and who the data will be shared with.⁵

As part of the deposit process we will ask you supply us with copies of your information and consent documentation, including completed consent forms from any data subjects who are identified in the dataset.

Intellectual property rights and permissions

Intellectual property rights follow creators, but creators are not necessarily owners of intellectual property rights. It is usual for IP created by employees to belong to employers, so for any creators employed by the University, the University will be the owner of IP rights.

⁵ For a guide to archiving personal data see the National Archives Guide to archiving personal data: <https://www.nationalarchives.gov.uk/information-management/legislation/data-protection/>. For data subject's rights see guidance from the Information Commissioner's Office: <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/individual-rights/individual-rights/>. For guidance on use of personal data in research refer to guidance and support provided by IMPS: <https://www.reading.ac.uk/imps/data-protection/data-protection-and-research>.

If the dataset was created in whole or part by employees of other organisations, these organisations will have IP rights. We will need to establish that these organisations agree to the archiving of data at the University for purposes of providing controlled access to authorised users. The organisation may wish to exercise a veto over any requests for access to the data, in which case one or more employees who can act as members of a Data Access Committee will need to be identified. We may need to implement a data storage agreement with the relevant organisation formalising this authorisation. This is more likely to be required where commercial organisations are involved. Where third parties are universities or other research-performing organisations, permission from the lead contact for the dataset at that organisation may be sufficient.

Students own IP they create, unless IPR have been assigned under a sponsorship or IP assignment agreement. If this is the case, and IP rights have been assigned to an organisation other than the University, we may need to negotiate with this organisation as above.

It will be important in all cases for any creators and rights-holders to give permission for archiving the restricted dataset at the University of Reading.

For guidance on identifying dataset creators and rights holders, see the relevant section in the [preparing for data archiving guide](#).

How is access to the data provided and managed?

Requests to access the data will be reviewed by a Data Access Committee, which will include the UoR PI for the dataset. The applicant must be a researcher affiliated to a recognised research organisation, and must require access to the data for a legitimate and valid research purpose. The applicant's organisation must be able to demonstrate that it has appropriate information security policies and protocols.

If a decision to grant access is made, the Research Data Manager will arrange for the applicant's organisation to sign a Data Access Agreement with the University. Once this is in place the dataset will be transferred by secure means to the applicant's organisation.

The Research Data Manager will undertake post-delivery administration of the Data Access Agreement. For information about the access request process see the [Access requests](#) section.

Deposit instructions

1. Initial discussion and acceptance in principle

We recommend you contact the Research Data Service team in the first instance to discuss a proposed restricted dataset deposit, in order to obtain an acceptance in principle. This is not required, and you can if you wish proceed directly to submission of a dataset deposit request (stage 2).

In order to be accepted, the deposit will need to meet the following acceptance criteria:

- The dataset has been created in the course of research conducted at least in part at the University of Reading.
- The dataset meets the definition of restricted data given above.
- The purpose of archiving is to enable the data to be shared outside the University, subject to relevant safeguards.
- The information provided to participants and their consent to participate in the original research project provide a viable basis for the long-term preservation and managed sharing of the research data.
- There is a permanent member of staff who is willing to act as the UoR PI for the dataset and to be a member of the Data Access Committee formed to consider any access requests.
- If the dataset is not owned by the University of Reading, or is jointly owned by the University and one or more third parties, any non-University owner of the dataset has agreed to its being archived at the University of Reading, and either has agreed that the UoR PI is empowered to authorise access to the dataset on its behalf, or has nominated one or more employees of its organisation who are able to authorise access to the dataset.
- There is not a more suitable place to archive the dataset.

If the dataset appears to meet the acceptance criteria, a formal deposit request can be submitted.

2. Submit a restricted dataset deposit request

A restricted dataset deposit request should be submitted by email, containing the information outlined below.

We advise you do not invest too much work in preparing a deposit until firm acceptance has been confirmed.

To submit a deposit request, please provide:

- A provisional title and a brief description of the dataset, including the type(s) of data contained, types and numbers of files, and the volume of the dataset if greater than 1 GB;
- The names, affiliations and status (member of staff or student) of the creators of the dataset. If a creator's current status was different at the time of their creative contribution to the dataset (for example, if a current member of staff was a student when they contributed to the dataset), please indicate this. Please also identify the UoR PI for dataset and if possible a back-up in the absence of the PI. The PI and any back up must be empowered to authorise access to the dataset as part of a Data Access Committee. Please confirm that all creators are aware of and agree to the proposed archiving of the dataset as a restricted dataset at the University;

- Why you believe the dataset is suitable for archiving as a restricted dataset, and the Restricted or Highly Restricted information as defined by the University Classification Policy the dataset contains;
- Contacts for the dataset at any other organisation involved in the creation of the dataset (i.e. primary contacts for notification and permission purposes), identified by job title and role in relation to the dataset. These persons must have confirmed in principle in writing that they are willing for the dataset to be archived with the University, and must be empowered to authorise access to the dataset on the part of their organisations. If possible, please indicate whether the lead contact (i) gives permission for the University to authorise access to the dataset on its behalf, (ii) wishes to exercise a veto over any data access request as part of a Data Access Committee formed to consider the request, (iii) requires a data storage agreement with the University authorising the University to store and manage the dataset on its behalf;
- If a student involved in creating the dataset was sponsored by another organisation or has assigned their IP to the University, a copy of the sponsorship or IP assignment agreement;
- Any relevant information sheets and consent forms (as blank copies), and any other documentation necessary to establish an ethical basis for archiving the data.

3. Review and acceptance

Once we have reviewed the deposit request we will notify you if we agree to accept the deposit. This will initiate the deposit process.

We will ask Legal Services to prepare the data access agreement. If the data are owned in whole or part by another organisation, we may also need to draw up a data storage agreement authorising the University to manage the data on the organisation's behalf.

Drafting of agreements can take a few weeks. Preparation of the dataset can be continued during this time, but please note that if a data storage agreement with the University is required to empower the University to store the data on behalf of another organisation, we will not be able to accept the data this agreement has been executed.

4. Prepare the deposit

We recommend you refer to University [guidance on preparing data for archiving](#) when preparing your dataset and supporting documentation.

A readme or documentation file for the dataset must be included, which will be made publicly available from the Archive. This is to provide basic information about the dataset and to help people assess whether the data would be of use to them, so that they can decide if they want to apply for access. We have a readme template depositors can download from the [Archive web page](#).

The readme/documentation file must include a Terms of Use section (there is one in the readme template). This must include the following information:

Copyright [year] [name(s) of rights-holder(s)]

This dataset is available to authorised users subject to a Data Access Agreement between the University of Reading and a recipient organisation. A copy of the University of Reading Data Access Agreement is included with this item.

To request access to the dataset, please complete a data access request at <https://redcap.link/data-request>.

Do not include in the public documentation any information that is not suitable for making publicly available. You can if need be prepare a more detailed version of the documentation file for inclusion with the restricted dataset.

The dataset documentation may include other documents, such as a copy of an interview schedule or a survey questionnaire. If these documents are suitable for making publicly available, they can also be uploaded to the public Archive record.

5. Create the metadata record and public files

When the data are ready to be deposited, create a metadata record in the Research Data Archive and upload any files that can be made public (i.e. the readme/documentation file and any other supporting documents), following the standard deposit process described in the Data Deposit Instructions, available from the [Archive web page](#).

DO NOT upload the restricted dataset files or any documentation not suitable for being made publicly available.

Notify the Research Data Service team when you have created the dataset metadata record.

For the time being, maintain the metadata record in your Archive work area and do not deposit it.

6. Create the transfer folder and add the dataset files

In order to transfer the restricted dataset files, set up a OneDrive folder containing the files and share it with the member of the Research Data Service team who is supporting the deposit.

The files will not be transferred into secure storage at this stage, as they will first be reviewed, and may need to be edited.

The restricted dataset files should include:

- The data files;
- The readme or documentation file. This may be in a more detailed version than the version made available from the metadata record;
- Any other relevant supporting documentation, such as an interview schedule or survey questionnaire. These should be included with the restricted dataset even if they have been made publicly available from the metadata record.

- Copies of all relevant information sheets and consent forms. A single unsigned copy of a consent form can be supplied if the dataset does not contain the personal data of the data subjects. If the dataset contains personal data, signed copies of consent forms or other documentation of consent must be supplied for ALL data subjects whose personal data are included in the dataset. Consent documentation must be supplied in digital form. Note that personal data will include any ID codes used in the study that are linked to individual participant data in any existing records, even if the linking file is not included with the dataset files.
- Documentation of permissions, as necessary. This may include email confirmations from co-creators and data storage agreements.
- Copies of any relevant legal agreements, such as a student's sponsorship agreement or IP assignment agreement.

7. Dataset review and restriction

Once the dataset files have been added to the transfer folder and the metadata record has been created, the dataset will be reviewed by the Research Data Service team. We may get back to you with requests for changes or more information. As part of this process we will also edit the record to reflect the restricted status of the dataset. A PDF of the Data Access Agreement will be uploaded once it has been prepared by Legal Services.

8. Transfer and storage of the restricted dataset files

When the review has concluded and the dataset is ready for preservation and publication, the restricted dataset files will be transferred into our secure storage on the University network. As part of this process, we will health-check the files and let you know if any files have been identified as corrupted.

9. Completion of review, acceptance and publication

Once the preceding stages have been completed, the dataset will be formally accessioned and the metadata record will be published and the DOI assigned.

We will notify you and any nominated dataset PIs and contacts when this has been done. Dataset PIs and contacts will be advised of their roles and responsibilities as members of the Data Access Committee for the dataset.

Citing the dataset

A data availability statement with any associated publication (including a doctoral thesis if relevant) should cite the data using the name of the University of Reading Research Data Archive and the dataset DOI. It should explain the confidential nature of the data and the terms on which they may be accessed. For example:

Data supporting the findings reported in this article are held in the University of Reading Research Data Archive at [DOI]. Because of the confidential nature of the information

contained in this dataset, it can only be made available for bona fide research purposes subject to a data access agreement. Please refer to the dataset record for further information about the dataset and access conditions.

The University provides guidance on [data availability statements](#).

Access requests

Requests to access a restricted dataset must be submitted using the form at <https://redcap.link/data-request>.

The request will be triaged by the Research Data Manager. If it appears to be a legitimate and valid request a Data Access Committee will be convened to review the request. The Committee will include:

- The Research Data Manager
- The Data Protection Officer
- The Head of Quality Assurance in Research
- The UoR dataset PI or representative. In the absence of the PI and any back-up, this will be the Head of School or any nominated representative.
- The named contact at any external organisation if relevant. In the absence of the nominated contact, this will be the organisation's legal officer or any nominated representative.

The Data Access Committee will consider a number of factors when deciding whether to authorise or decline any data access requests. This will include:

- Is the applicant a bona fide researcher, affiliated to a research organisation?
- Does the applicant have a research purpose for requesting the data?
- Are the data likely to be useful for the stated purpose?
- Does the research organisation have appropriate information security policies and protocols?

Decisions will be made by consensus and subject to veto by any contact for an external organisation holding IP rights.

If a decision to grant access is made, the Research Data Manager will arrange for the applicant's organisation to sign a Data Access Agreement with the University. Once this is in place the dataset will be transferred by secure means to the applicant's organisation.

The Research Data Manager will undertake post-delivery administration, e.g. reminding users when the agreement is due to terminate and copies of the data must be destroyed, extending agreements if requested, etc.

Ongoing management

It is the responsibility of the dataset PI to ensure that they or anyone empowered to act on their behalf is able to participate in a Data Access Committee in a timely manner. If

the PI will no longer be able to perform this function, for example if they leave the University, they should arrange handover to another PI.