

THE EPSRC EXPECTATIONS FOR RESEARCH DATA

Essential information for funded researchers

What are the EPSRC expectations?

The *EPSRC Policy Framework on Research Data* (2011) specifies nine expectations for research data¹ which apply to both researchers and organisations in receipt of EPSRC funding.

Researchers and **research students** funded by EPSRC are required to **preserve, record, and enable access** by third parties to data collected or generated by them that directly support published research findings.

The EPSRC monitors organisations and researchers for compliance with these expectations. If it appears that proper sharing of research data is being obstructed, the EPSRC reserves the right to impose appropriate sanctions.

How do the EPSRC expectations apply to me?

As a researcher in receipt of EPSRC funding you are subject to the following key requirements:

- research data collected or generated in the funded research activity, and which support published research findings, must be securely preserved for **a minimum of 10 years**;
- appropriately structured metadata describing the research data must be published (normally within 12 months of the data being generated) and made freely accessible on the internet;
- published papers must include a statement describing how and on what terms any supporting research data (digital and non-digital) may be accessed.

How can I comply with the EPSRC expectations?

You or your project PI will need to take three key steps:

1. **Preserve the data** (digital and non-digital) in an accessible location for a minimum of 10 years;
2. **Publish an online metadata record** describing the data, and register the data with the University;
3. **Include a statement in research publications** indicating where the data are and how they can be accessed.

1. Preserve the data in an appropriate location where they can be accessed

The data should be preserved in a location and form suitable for access by others for **a minimum of 10 years** from the project completion date.

¹ <https://www.epsrc.ac.uk/about/standards/researchdata/expectations/>.

EPSRC expectations for research data

For digital data, use either a suitable external data repository service, or the University's own **Research Data Archive**.² The Research Data Archive is a service available for use by research staff and students to register, preserve and provide access to data arising from funded research activities.

You should deposit your data in a suitable repository before publication of your findings, so that they can be referenced from your publications. Many repositories, including the University Archive, allow deposit of data under embargo pending publication of findings.

Your data must be preserved in a suitable institutional service, even if they are sensitive or confidential and cannot be openly shared. The University Archive can restrict access to authorised users only under a data use agreement.

Suitable external services may include relevant disciplinary data repositories, which will preserve and publish data, and free-to-use general data sharing services, such as figshare and Zenodo.³ Guidance on external services is provided on our **Where to archive your data** web page.⁴

You must also ensure relevant non-digital data are stored securely for the long-term under appropriate stewardship, or digitised for electronic storage, such that the data can be provided in the event of a valid request for access being made. If you are not aware of any secure archival storage facility for non-digital data in your Department or School, contact the Research Data Manager for advice.

2. Publish online a structured metadata record for the data/register data with the University

Wherever your data are preserved, you are required to **register the data with the University** by creating a metadata record in the **Research Data Archive**, normally within twelve months of the data being generated. This requirement applies even if you have deposited your data and created a metadata record in an external data service.

You can create a metadata record in the process of depositing data in the Research Data Archive. You can also create a metadata record in the Archive that references data held externally. You will need to do this for digital data stored in another online service and for non-digital data held offline.

The primary metadata record for your data (whether this is in the University Research Data Archive or an external online service) must include a persistent unique identifier (such as a DOI) and sufficient information to allow others to understand what research data exist, why, when and how they were generated, and how to access them.

Where access to the data is restricted, the metadata record should also give the reason and summarise the conditions which must be satisfied for access to be granted. For example 'commercially confidential' data, in which a business organisation has a legitimate interest, might be made available to others subject to a suitable legally enforceable non-disclosure agreement.

3. Include in any published paper that relies on the data a statement describing how and on what terms the data may be accessed.

If you are the lead or corresponding author of a publication that relies on the data, this is your responsibility.

The statement should be in the form of a reference to an online metadata record using a DOI. When you deposit data in the University's **Research Data Archive** your dataset will be assigned a DOI and you will be provided with a citation for your dataset that you can copy and paste into a publication.

If legal or ethical reasons exist to protect access to the data these should be noted in the statement.

² <http://www.reading.ac.uk/reas-RDArchive.aspx>.

³ <https://figshare.com/> and <https://zenodo.org/>.

⁴ <http://www.reading.ac.uk/reas-WheretoArchive.aspx>.

Guidance on referencing data from publications is provided on our [Data Access Statements](#) web page.⁵

Further information and support

Robert Darby, Research Data Manager

researchdata@reading.ac.uk / r.m.darby@reading.ac.uk or x6161

<http://www.reading.ac.uk/reas-ResearchDataManagement.aspx>

FAQs

Do these expectations apply to me if I am not the project PI?

Yes: you are funded by the EPSRC, so these expectations apply to you, and you share a joint responsibility for meeting them with all members of the project team.

But the project PI bears primary responsibility for meeting these obligations, and should ensure that appropriate measures are in place to meet them. If you are not aware of any such measures, you should discuss this with your PI.

What if the University is not the lead organisation?

If the data are generated or collected by you while you are employed by the University of Reading, they are intellectual assets of the University, and you and the University have a responsibility to ensure they are managed in accordance with the terms of the EPSRC grant.

I was funded by the EPSRC as a research student. Do the expectations apply to me?

Yes: the expectations apply to all researchers funded by the EPSRC, including research students. As a research student, by default you own the data you collect, but rights in the data may be shared or assigned if they are collected as part of a collaboration or under the terms of a funding or employment contract. If you do not have sole rights in the data this may affect whether you can make them available to others. If you are in doubt, seek advice from your supervisor or the Research Data Manager.

What is meant by 'research data'?

The EPSRC defines research data as 'recorded factual material commonly retained by and accepted in the scientific community as necessary to validate research findings; although the majority of such data is created in digital format, all research data is included irrespective of the format in which it is created'.⁶

What data are in scope?

The EPSRC expectations apply to research data collected or generated in the funded research activity that directly support published research findings.

Do the expectations apply to secondary or derived data?

If you use data from an existing source you would not be expected to store and publish the data yourself, although you should reference the source of the data properly in your published findings.

⁵ <http://www.reading.ac.uk/reas-DataAccessStatements.aspx>.

⁶ <https://www.epsrc.ac.uk/about/standards/researchdata/scope/>.

EPSRC expectations for research data

If your research modifies data from an existing source, such that a new dataset is produced, for example by adding new data points, or by deriving new data through additional processing, then there may be a need to share the new dataset. In such a case you should ensure you have permission from the data source to share their data or any outputs derived from them.

What is meant by data that 'directly support' published research findings?

How to interpret this will be a matter of judgement to be applied in each individual case. You are not necessarily expected to make all your data accessible.

Ask yourself this question: what would a researcher who read the paper reporting the project's findings need in order to be able to validate or replicate them? The answer should tell you what you will need to preserve and make accessible. It may, for example, be sufficient to share the end result of your data processing, if the source data are readily available or easily generated, and you record the steps by which the data were arrived at in detail.

This is typically the case in computer simulation research, where the software code and input parameters may be more important than the outputs of any given execution of the model. If your research involves writing software code or scripts, you should also ensure that these are preserved and made accessible. For more information you can consult the **Guide to publishing research software** (download from the right hand side of the page).⁷

In other cases, for example where data have been obtained by observation or experiment, it may be important to preserve and share the raw data, especially if they could be of use for further research purposes, whether or not these are related to the original purpose for which they were created.

Can I store the data on my computer/hard drive and ask people to contact me for access?

No: you will not meet the EPSRC requirements if you store data on a local device such as a PC or external hard drive in your office, or your University personal network drive, and state in your paper that anyone requiring access to the data can contact the author. A local, personally-owned storage solution will not preserve the data or guarantee access in the long term. Who would provide access to the data if you ceased to act as their custodian or could no longer be contacted?

You must use an organisational storage service, such as a data centre or repository for digital data, or an organisational archive for non-digital data, which can manage the data over the long term and enable access to them on your behalf. Such a service will also allow you to create an online metadata record for your data with a unique identifier (such as a DOI), ensuring their continued discoverability. The University's Research Data Archive is an example of a service that meets the EPSRC requirements.

When should I make data available?

The EPSRC expects data supporting research findings to be made available so that they can be scrutinised and if necessary validated by others. At the least a metadata record describing the research data relied on in published findings should be accessible **no later than the date of publication of the findings**. Unless there is good reason to restrict access to the data, they should also be stored and made accessible by this date.

Until the date of publication, the principle of 'privileged access' applies, under which those researchers involved in the generation of the data are entitled to enjoy the first fruits of publication. But once findings have been placed on the public record, supporting data must be made accessible, freely and openly wherever possible, and on a restricted basis only if there is a legal, ethical or commercial reason why data cannot be shared openly.

⁷ This can be downloaded from <http://www.reading.ac.uk/reas-ResearchDataManagement.aspx> (on the right of the page).

Am I obliged to share all relevant data?

Not necessarily. The EPSRC identifies a number of legitimate exceptions:

You may refuse access to data supporting published research findings if their disclosure would compromise any intellectual property exploitation that is planned or in progress, such as licensing or patent registration. But this exploitation should take place during your 'privileged access' period; you cannot indefinitely refuse access to data on the basis that they may one day have commercial value.

You can also restrict access to data if there is a legitimate commercial interest on the part of third parties that have collaborated in the research. But research organisations and researchers have a responsibility to ensure that publicly-funded research involving third parties is planned and executed in such a way that published findings can be scrutinised and if necessary validated by others. Third parties that collaborate in publicly-funded research should be made aware of the importance of ensuring that the outputs of research can be made as openly available as possible.

Lastly, some data cannot be shared for legal or ethical reasons. Personal data should not be put into the public domain without the explicit consent of the person to which it relates. Sensitive information should not be placed in the public domain. Sensitive information is information the release of which would compromise unprotected intellectual property or which, in the judgement of the security services, would result in unacceptable risk to the citizens of the UK or its allies.

Can I share data selectively instead of making everything available to everyone?

If you have good reason to restrict access to data, because they either contain sensitive information or are subject to third-party restrictions, there are means to apply restrictions on both access and use.

If you use the University's Research Data Archive to store the data, you can apply a **Restricted** access setting to any files you upload and license the files using the **University of Reading Licence for Restricted Data**,⁸ which allows data to be used, subject to authorisation, in confidence, for non-commercial research and learning purposes only. Restricted access files cannot be accessed directly via the Archive: anyone wishing to consult such files would have to submit an access request, which may be granted or refused by an authorising agent (typically the research PI).

Alternatively, you can make data openly available under a legally-enforceable licence defining the terms on which they can be used. The University's Research Data Archive provides a number of standard licence options (including Creative Commons licences and Open Source software licences) that can be used to define different sets of permissions for use of data stored in the service. You could, for example, license your data for non-commercial use only.

For more information about these options, consult the Archive's **Data Policy**.⁹ Other data services may have similar options for placing restrictions on access to data and their use.

Will EPSRC know if I don't share my data?

The EPSRC has stated that it will compliance with its research data policy, and will investigate non-compliance. It also states that if it appears proper sharing of research data is being obstructed, it reserves the right to impose appropriate sanctions.¹⁰

It is known that the EPSRC has monitored publications acknowledging its funding to check whether a data access statement has been included and the supporting data/metadata have been made available.

⁸ http://researchdata.reading.ac.uk/licence_restricted_data.html.

⁹ https://researchdata.reading.ac.uk/data_policy.html.

¹⁰ <https://www.epsrc.ac.uk/about/standards/researchdata/impact/>.

Does the University monitor compliance?

Yes: we monitor all Research Council-funded publications as accepted manuscripts are deposited in CentAUR¹¹ and final versions are published to check that:

- they include appropriate data access statements;
- any data referred to are suitably preserved and made accessible where possible.

Where the EPSRC requirements have not been met we may contact authors to see how we can plan to meet the EPSRC expectations as soon as is practicable.

Do I need to create a data management plan for my project?

In short: Yes. This is not compulsory, and unlike other funders, the EPSRC does not ask for a data management plan as part of a grant application. But the EPSRC does state that 'project specific data management policies and plans... should exist for all data'.¹² It is also a matter for good practice to write a data management plan for your research, which can help you and other members of the project team manage data effectively and prepare data for preservation and sharing towards the end of the project.

¹¹ <http://centaur.reading.ac.uk/>.

¹² <https://www.epsrc.ac.uk/about/standards/researchdata/principles/>.