

DATA DEPOSIT CHECKLIST

University of Reading Research Data Archive

Before you start to deposit a new Dataset, use the following checklist to make sure you have everything prepared.

If you do not submit a valid Dataset or fail to provide all the information required, we will need to contact you to complete the deposit process before the Dataset can be published. To ensure your Dataset is eligible for deposit and that your submission meets the criteria for acceptance, check the Archive [Collection Policy](#) and [Submission Policy](#).

If you are creating a Metadata Record for a Dataset or Collection held in an external service or offline, but are not depositing data files, not all requirements will apply.

A more detailed [Data Preparation Guide](#) is available for download.

If you require advice or assistance, contact us at researchdata@reading.ac.uk or call Robert Darby, Research Data Manager, on extension 6161.

1. Check your permissions and rights

- Deposits into the Archive are made in accordance with our [Deposit Agreement](#), which grants the University a licence to manage the data on behalf of the rights-holders in the data.
- Before you can deposit data you must confirm the ownership of any intellectual property rights in the data and undertake that you and/or the University are lawfully entitled to distribute the data. In the absence of any contract stating otherwise, the University will own data created by its employees, and students will own data they have created. Rights in data may be jointly held.
- Make sure disclosure of the data will not compromise intellectual property interests, such as any confidential material or trade secrets, patent applications, licensing or other such activities.

2. Group your files into Datasets

- A Dataset is a set of files or a database constituting a discrete and clearly defined set of data and/or the means of generating data related to a research activity or project, or supporting one or more research publications.
- Large or multi-phase projects, or research that has given rise to multiple publications, can be decomposed into multiple Datasets, which can be grouped together as Collections.

3. Prepare your data files

- Check the total volume of the files you propose to deposit. If it is **greater than 20 GB**, [contact us](#) to discuss your deposit. Deposits exceeding 20 GB in volume are subject to a charge; unauthorised deposits above this limit will be rejected.
- Make sure your files are in the most suitable formats for preservation and sharing – check our [Recommended File Formats](#).

Data Deposit Checklist

- Give files meaningful names that reflect the file content, avoiding spaces and special characters.
- For large Datasets, group your data files in zip bundles according to their content or file format, to make upload and download easier. Keep a folder structure for the files in your zip bundle.
- The individual file/zip bundle size limit for web upload is 10 GB, but in practice may be less, depending on your browser and network connection. We recommend a maximum upload size of 4 GB. If you have trouble uploading files or have large amounts of data [contact us](#) to arrange alternative means of deposit.

4. Prepare your documentation

- Every Dataset must include at least one documentation file.
- The minimum requirement is a README.txt file with a basic description of the Dataset and its constituent files. This [README file template](#) can be used.
- Prepare essential documentation to upload with your data, e.g. collection and processing materials and methods, instrument user guides, lab notebook records, survey questionnaires, codebooks, information sheets, sample consent forms, etc.

5. Check consent and confidentiality

- Check that your data files contain no disclosive personal or confidential information.
- If you wish to deposit data obtained from research participants, make sure they have consented to their data being shared and clarify any restrictions that may apply to sharing of the data. You will be required to provide the Archive with evidence of consent obtained.
- Consider whether you may need to make data available on a Restricted Access basis, for example only to users who agree to an end-user agreement.

6. Decide your embargo and access control requirements

- You can place files in your Dataset under embargo for a specified period, e.g. until research findings have been published. Files should not be embargoed for more than 12 months from the date of deposit.
- File-level access can be set to: **Open**; **Registered users** (University staff and research students); or **Restricted**. **Registered users** can be used to restrict access to University members on a temporary basis, e.g. for internal review. **Restricted** files cannot be freely downloaded: the setting should only be used with a valid reason, e.g. based on the sensitive nature of the data.

7. Decide how to license your data

- Each file or zip bundle must have a licence statement applied.
- Use [Creative Commons](#) licences for open data. The University recommends you use [Creative Commons Attribution 4.0](#):
- For software code standard Open Source licence options are available, including Apache 2.0, MIT, GNU GPL 3.0 and LGPL 3.0. Guidance on licence options for software can be found in our [Guide to publishing research software](#), and at [choosealicense.com](#);
- For Restricted files use the [University of Reading Licence for Restricted Data](#) or upload your own licence.

8. Have your research project information to hand

- Be prepared to provide funder information and grant references where relevant.
- Have to hand references for related publications or other resources, e.g. websites.