

Depositing Data in the Energy Data Centre

Purpose: Guidance on how to deposit data

Audience: UKERC researchers who create research data

Summary

To make the deposit of data into the Energy Data Centre as smooth as possible we recommend that once your data is **intellectually complete enough to be shared** or at **the end** of your project you contact the EDC team to start the process.

You will need to have the **data documented** so that other people can understand it independently and the data will need a **license** and a **description**.

The EDC team will work closely with you to complete the submission.

Introduction

Data needs to have sufficient context for it to be usable by others, this guide establishes what is required when depositing datasets into the EDC.

Depositing data into a domain data centre ensures that it can be found and reused by other researchers and it is actively managed for the long-term. Publicly funded research data should be FAIR (Findable, Accessible, Interoperable and Reusable) and reproducible.

- **Findable:** Uniquely identified with good metadata in a searchable resource
- **Accessible:** Retrievable, with authorisation and authentication if required, or clear if withdrawn.
- **Interoperable:** Use of standards and vocabularies
- **Reusable:** Sufficient descriptive metadata, data license and provenance.

For UKERC funded researchers your Data Management Plan will have informed these requirements.

When to deposit?

Stable and persistent version

There needs to be a version which is intellectually complete enough to be shared. If you would like a Digital Object Identified assigned then there is an expectation that it will not change. This may be everything you have collected or a subset depending on the data creation/generation process.

If you are collecting data over a long time period, consider how best to organise and deposit this.

What do you need to prepare?

Documentation and file structures

Please provide as much supporting documentation as is appropriate. We recommend two README files: one describing the dataset and its collection. This includes topics such as what has been collected, units used, how missing data is indicated, timing conventions - basically everything that



an informed but un-involved person will need to be able to use the data. The other describes the folder structure.

File formats and names

Content in the EDC is preserved for the long-term. Some file formats are easier to preserve than others. The EDC has the following curation preferences:

- CSV rather than Excel. Excel can be provided for dissemination purposes.
- PDF-A rather than PDF
- Readmes – ASCII text files rather than Word

Ensure the file names are meaningful and unique. Please use standard alphanumeric characters and hyphen or underscores only in the file name. No spaces are allowed.

Data License

All data should have a clear license when it is released. Which license is chosen depends on many factors including data collection constraints and your institutional advice. The EDC can provide data access restrictions but does not hold sensitive data.

Metadata for discovery

Metadata is information about the data. It enables data discovery and gives an end-user the information they need to read and make sense of the data. All datasets in the EDC need both a title and description. There are options to add information about geographic and time coverage. This information is used to enable the dataset to be discovered, so the more information provided the better the retrieval will be.

Process

1. Contact us by emailing EDCManager@stfc.ac.uk
2. EDC team will review the deposit to ensure we understand it & any specific requirements
3. Create a metadata record using our [online form](#)
This gives the EDC permissions to curate the data for the long-term and is a required step. Ensure you have the rights to the data to be able to do this.
4. Working together we will ensure that the deposit format, metadata and documentation conforms to EDC requirements
5. Send a copy of the data & documentation
6. EDC team will load and publish the data and publicise it.

Further Reading

- More EDC advice from our service website [About Data Management pages](#)
- Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016).
<https://doi.org/10.1038/sdata.2016.18>

