Guidance on DOIs for NERC data

Digital Object Identifiers (DOIs) are the Natural Environmental Research Council's (NERC) chosen permanent identifier for data deposited with its Data Centres. A digital object identifier (DOI) is a string of characters which uniquely identifies an object. DOIs are often used for journal papers but can also be used for other publication and data resources (e.g., datasets, model code, software).

The NERC Environmental Data Service (EDS) works with DataCite in collaboration with the British Library to assign DOIs to data resources generated by NERC funding. We can only assign DOIs to data held by one of the five NERC EDS Data Centres.

How to get a DOI for your data.

1 Deposit data with the relevant NERC Data Centre with required metadata and relevant documentation in good time

2 Your data and any documentation will be checked by Data Centre staff and deposited when suitable for long-term archival

3 You may need to request a DOI. This should be done as early as possible (minimum 2 weeks) ahead of when it is required

4 The NERC Data Centre confirm metadata details and will assign a DataCite NERC DOI to your data

5 The DOI and citation details are added to a public landing page or catalogue record

6 Cite your data in any subsequent journal publications

i Your data will be made publicly available straight away or after an agreed embargo period

Why get a DOI for your data?

DOIs enable your data to be fully citable for the long-term, providing a level of permanence that other citation options are unable to offer. Consequently, DOIs for data are increasingly being sought by journals and these aid tracking data usage in the literature.

Giving your data a DOI means:

- that usage of your data can be followed as others use and cite it.
- the data are uniquely identifiable.
- you will always be identified as the creator of the cited data
- your data/publication can always be located with a simple web search

NERC fully support the publication and citation of data, as explicitly stated in their data policy:

> "NERC has a policy on data in order to: …
>
> "Help in the formal publication of data sets, as well as enabling the tracking of their usage through citation and data licences"

They believe data should be treated as a first-class research output and researchers who create it should get an appropriate level of academic credit for their work. NERC have a DOI policy which details how it manages DOIs for data.

Who can get a NERC data DOI?

Researchers can request a DOI for data archived within a NERC EDS data centre provided that their data and associated metadata meet a minimum set of requirements (see below). The NERC EDS does not assign DOIs to data stored elsewhere, or for outputs that aren't data resources e.g., theses. Both new and legacy data held in a NERC data centre are eligible for DOI assignment.

Note: Although the NERC EDS is principally for NERC-funded research data, we may also accept data resources that are funded by other means, as per the NERC EDS data centres' acquisition policies. Such data are similarly eligible for NERC DOIs.

Minimum requirements to obtain a DOI:

1. **All DOI requests require agreement from all authors of the data resource.** This agreement must be obtained from all parties by the data provider or researcher responsible for the data (e.g., PI) as outlined in the associated Data Management Plan.

2. **The dataset must be a stably defined data resource**. That is, no major additions or changes are expected to be made to it. In most cases this corresponds to the dataset being complete. Where errata are required, this will result in a new version of the dataset being released with a fresh DOI. For long-term datasets, e.g., instruments which have been in place for years and are anticipated to be producing well defined, stable data products for the foreseeable future, your NERC EDS Data Centre will be able to advise the most appropriate approach to curation and DOIs. Note, subsequent updates to completed datasets can be made as new versions of the datasets with dedicated DOIs.

3. **The dataset should be in an appropriate format for long-term preservation and usage**. Appropriate formats are defined by the EDS Data Centre.

4. **The dataset must be held by a NERC EDS Data Centre**. By assigning a DOI the NERC EDS is committing to make the dataset available for the foreseeable future. The NERC EDS will ensure that the DOI resolves to a data description page (landing page) associated with the data in perpetuity.

5. **The dataset should be of good technical quality**. The data should be appropriately structured, reusable, and well described to the standard required by the NERC EDS Data Centre. The metadata describing the data should include the mandatory elements as defined by [DataCite](#) as well as any discipline specific elements mandated by the NERC EDS.

When to get a DOI?

Increasingly, journals require any data referred to in a journal publication to be deposited in an approved repository and have an associated DOI. In order for researchers to meet this requirement in good time, proper data management planning is required. The NERC EDS will always do their best to provide DOIs for datasets in a timely fashion. However, researchers must allow sufficient time to prepare the data and supporting material required for dataset publication (including any revisions required by the Data Centre) if they wish to obtain a DOI by a specific deadline.

Whilst data submission to a NERC EDS Data Centre commonly occurs towards the end of a project, it is useful to start thinking about data citation when drafting the data management plan. For example, a data management plan should identify key datasets that will be created by the project and their suitability for citation should be assessed. Authoring, licensing and rights issues should also be considered when writing the data management plan, thus avoiding problems at a later stage when the dataset is ready to be deposited.

Data citation, DOIs and citation metrics

Data citation is analogous to the citation of other published work e.g., journal papers. Citing data gives credit to the creators of a data resource. Good data citation practices ensure scientific transparency and reproducibility by guiding other researchers to the original sources of information. Cite the data supporting your research (your data or data you have reused) using a standard citation in the reference section of your published work (as well as in any data access/availability statement). The NERC EDS Data Centre who assigned the DOI will provide you with a form of words for citing the data you deposited. These usually consist of the data authors, the date of publication, the title of the data, the publisher of the data and its DOI. Sometimes the data type will also be included e.g., 'Dataset'.

Examples:

Robinson, E. L., Blyth, E., Clark, D. B., Comyn-Platt, E., Finch, J., & Rudd, A. C. (2017). *Climate hydrology and ecology research support system meteorology dataset for Great Britain (1961-2015) [CHESS-met] v1.2* [Data set]. NERC EDS Environmental Information Data Centre. https://doi.org/10.5285/B745E7B1-626C-4CCC-AC27-56582E77B900

Schmidt, K., Graeve, M., Welteke, N., Hoppe, C., Fong, A., Hildebrandt, N., Castellani, G., Vortkamp, M., Belt, S., & Atkinson, A. (2023). *Fatty acid composition of particulate organic matter (POM) collected from surface waters and bottom sea-ice of the Central Arctic Ocean during the MOSAiC expedition in 2019/2020 (Version 1.0)* [Data set]. NERC EDS UK Polar Data Centre. https://doi.org/10.5285/d7708d08-4bd4-439a-99e2-307a175977ea

Watanabe, M. (2023): Chapter 7 of the Working Group I Contribution to the IPCC Sixth Assessment Report - data for Figure 7.17 (v20220721). NERC EDS Centre for Environmental Data Analysis, 10 July 2023. https://doi.org/10.5285/b9303c07edb24582b45088795f347ca9

Latas, M., Wade, B.S., Pearson, P.N. (2022). *Morphometric data for a new species of fossil planktonic foraminifera,* Globigerinoides rublobatus *n. sp, from IODP Site U1483A*. NERC EDS National Geoscience Data Centre. (Dataset). https://doi.org/10.5285/adb5b0be-a357-4416-846f-777300d78240

GEBCO Bathymetric Compilation Group 2023 (2023). *The GEBCO_2023 Grid - a continuous terrain model of the global oceans and land.* NERC EDS British Oceanographic Data Centre NOC. (Dataset). https://doi.org/10.5285/f98b053b-0cbc-6c23-e053-6c86abc0af7b

Different journals and publications may require different citation styles. This DOI Citation Formatter allows you to create more than 500 different citation styles.

Citing data you have used allows the NERC EDS to generate citation metrics for individual datasets. Citation metrics are based on the number of times a data resource has been cited. Citation metrics can demonstrate the full impact and influence of your research as data citation practices increase.

NERC EDS Data Centre contact details:

Data Centre	Area of interest	Contact
British Oceanographic Data Centre (BODC)	Marine science	enquiries@bodc.ac.uk
Centre for Environmental Data Analysis (CEDA) *	Atmospheric, earth observation, solar and space physics	data.management@ceda.ac.uk
Environmental Information Data Centre (EIDC)	Terrestrial & freshwater science	info@eidc.ac.uk
National Geoscience Data Centre (NGDC)	Earth sciences	ngdc@bgs.ac.uk
UK Polar Data Centre (PDC)	Polar and cryospheric science	PDCServiceDesk@bas.ac.uk

* The Centre for Environmental Data Analysis (CEDA) is run jointly with the Science & Technology Facilities Council