

Guidelines for a good Data Access Statement

The UKRI Open Access policy states that all publications **acknowledging funding from UKRI or any of its councils** must have a data access statement that informs readers where the underlying research materials associated with a paper are available, and how the research materials can be accessed. Research materials have a very broad definition including data, code, software, images, and sounds. These guidelines are for data (including model outputs).

The statement should include:

- where the data can be accessed (i.e. which data repository);
- a unique persistent identifier, such as a Digital Object Identifier (DOI) or accession number, or a link to a permanent record for the data;
- details of any restrictions on accessing the data and a justifiable explanation (e.g. for legal, ethical or commercial reasons).

The steps below describe how to create a clear and complete Data Access Statement for all environmental data which was fully, or partially, funded by UKRI or any of its councils (e.g. NERC) that will also comply with the NERC data policy.

Steps to follow for Data

Step 1: Ensure data are submitted to a NERC data centre (or an alternative repository on the advice of a NERC data centre). The Polar Data Centre will be the most relevant NERC data centre for much of the environmental data produced at by BAS scientists. The data centre will mint a DOI for each data set which provides a persistent identifier for the data, provide a citation for the dataset and an appropriate licence (typically an Open Government Licence) which stipulates how the data can be reused. Images and sounds stored as digital files can also be submitted to a NERC data centre. Ideally this step should be undertaken prior to submission of the paper – the PDC can embargo the data until a paper is published and provide password controlled access to the data for reviewers.

Step 2: Be specific in the statement so that you identify the data used in different elements of the manuscript – e.g. The data supporting Fig.2 and Table 1 are publicly available in the UK Polar Data Centre at: <https://doi.org/10.5285/6184B04A-CB7B-4D94-AF9B-9B20406A4E96> (Dunn et al. 2021). The data supporting Tables 2 & 3 are publicly available in the UK Polar Data Centre at: <https://doi.org/10.5285/A5174472-204D-45D4-862C-3BEFC32B52CA> (Trathan et al. 2021) and so on.

Step 3: Include the citation of each dataset in the reference list for your manuscript. This enables tools that automatically pick up dataset citations to find it.

Step 4: If there are any data you cannot make available, then clearly state why in the statement. For example, this might be to protect personal data, the location of protected species or because you have used data provided with a licence that does not enable you to share it with others. For example: the data supporting Table 3 have not been made available because ...

Step 5: If there are no data underlying your manuscript then state this in the statement e.g. “No datasets were generated or analysed as part of this study”.