

RDM Best Practice

September 17, 2026

1 Introduction

Research data underlying any publication should be made publicly available in a persistent repository. The preferred repository is Zenodo (<https://zenodo.org/>), which provides persistent identifiers (DOIs), versioning, metadata, and long-term access & control to the research data.

In addition to research data, software may also be published through Zenodo. This provides a convenient way of making a particular software release associated with a persistent DOI, and thus easily citable.

The purpose of these guidelines is to provide a common minimum for research data associated with publications.

2 Recommended Procedure

For each paper:

1. Prepare the research data underlying the figures and tables of the manuscript.
2. Organise the data according to the structure described below.
3. Upload to Zenodo.
4. Where appropriate, link the Zenodo record to the relevant Zenodo community.
5. Check the metadata, files, and DOI before making the record public.
6. Add a data availability statement to the manuscript.

A data availability statement should briefly describe what has been deposited and provide the DOI of the corresponding Zenodo record, for example:

Data availability — The data underlying the figures of this article are available in Zenodo [1].

3 Preparing the Data

3.1 What to include?

The repository should contain at minimum the data underlying all quantitative figures and tables in the article.

The deposited data should be sufficient for a reader to recreate the numerical content of the published plots without having to manually extract data from the figures.

For example, if a figure contains several curves, the numerical data for each curve should be provided. If a figure contains a heatmap, the underlying two-dimensional data should be provided.

3.2 File Formats for Data

Whenever practical, use open, well-documented, and machine-readable file formats.

For small numerical datasets, the preferred format is generally CSV (comma-separated values).

Other suitable formats include:

- csv for simple tabulated numerical data;
- textttHDF5 or other well-documented formats for larger or more complex datasets.

Do *not* use a file format that is native to only a specific language!

3.3 Metadata

For any numerical dataset, metadata should be included at the beginning of the CSV file. At a minimum, metadata should specify:

- The figure or table to which the dataset belongs;
- the meaning of each column;
- physical units, if applicable;
- relevant normalization conventions;
- parameter values required to interpret the data;

Example: CSV file

```
# Fig. 1: Dynamics of magnetization and correlations
# First column: time Jt
# Second column: magnetization M(t)
# Third column: Autocorrelation C(t)
# Units: time is measured in 1/J
# Initial state: ...
time, Mt, Ct
0.0, 1.000, 0.000
0.1, 0.982, 0.013
...
```

Example: Heatmap

```
# Fig. 4: Spatiotemporal buildup of correlations
# x-axis: distance l
# x-axis range: -10 to 10
# x-axis spacing: 1
# y-axis: time Jt
# y-axis range: 0 to 10
# y-axis spacing: 0.1
# Matrix entry (i,j): correlation at x_i and t_j
```

Important - Using # for metadata lines ensures that the file remains easily machine-readable.

3.4 Directory Organization

The data should be organised so that the relationship between files and manuscript figures is immediately clear. For example:

```
figures/  
  Fig1/  
    Fig1a.csv  
    Fig1b.csv  
  Fig2/  
    Fig2a.csv  
    Fig2b.csv  
  Sup_FigS1/  
    Sup_FigS1.csv
```

The names of the files should make their relation to the manuscript clear.

3.5 Upload

Data should be uploaded as a `.zip` file. This allows Zenodo to read it and provide a clear file structure to others without them needing to open the data.

4 Auxiliary Information

The repository should contain auxiliary information about the project. It should include at least:

- title: written as "Dataset for 'ARTICLE NAME'"
- All authors, including:
 - Affiliations
 - ORCID; where available
- Short description of the repository.
- Funding agencies.
- Related works - connection to relevant article DOIs.

5 Software

Zenodo is also very useful for software packages, giving them a clear DOI, while also allowing clear author lists and versioning.

By connecting a GitHub repository to Zenodo, software releases can be automatically archived and assigned a DOI, making them easier for others to use and cite. For example: [2].

6 General Points

6.1 Earmarking a DOI pre-publication

If you plan on releasing your manuscript before releasing the associated data archive, Zenodo offers an option for earmarking DOIs. To do this, you should open a repository, and click on **Get a DOI Now!** under **Digital Object Identifier**.

A full BibTeX is generated and available under the **Preview** window.

6.2 Embargo

If preferred, an embargo can be placed on the repository to keep it private for a limited amount of time.

References

- [1] Tom Ben-Ami, Markus Heyl, and Roderich Moessner. Dataset for 'breakdown of multipole expansion in emergent electromagnetism', September 2026.
- [2] Griffin Hosseinzadeh, K. Azalee Bostroem, Tom Ben-Ami, and Sebastian Gomez. Light curve fitting v0.10.0, May 2024.