

MOSAiC Data Publication

- [In detail: MOSAiC data publication guideline](#)
- [MOSAiC Raw Data Publication](#)
- [Reuse and republication of MOSAiC data - guideline](#)

In short (*dos*, *don'ts* and *keep in minds*)

Why do I have to publish “my” MOSAiC data?

Each PI / lead author assures both raw and primary (processed) data are published, which will guarantee their Findability, Accessibility, Interoperability and Reusability (FAIR data principles).

See MOSAiC [Data Policy](#) for details.

What is data publication?

Data publication is a published data set or data collection equipped with a complete set of metadata. It is fully citable by having a title, authors, abstract and a persistent identifier (usually DOI). It can have but need not have a reference to a scientific paper publication.

Data publication isn't sharing data at MOSAiC Central Storage (MCS).

Data publication isn't adding a data table as a supplement to a published scientific paper.

Where do I publish “my” MOSAiC data?

The default repository for MOSAiC is PANGAEA.

My national funding agency requires depositing data in a special national repository. What should I do? These cases are handled as exceptions (see [Data Policy](#)) and a legitimate reason for not publishing the data in PANGAEA. At the moment, written agreements have been signed with several repositories: [Arctic Data Center \(ADC\)](#), [Atmospheric Radiation Measurement \(ARM\) data center](#), [British Oceanographic Data Centre \(BODC\)](#), [UK Polar Data Centre](#) and [Centre for Environmental Data Analysis \(CEDA\)](#). When archiving data in these repositories, it is always important to acknowledge the MOSAiC project (see [Data Policy](#)). The agreements assure FAIR data publication and future findability of MOSAiC data from a single access point (portal to project and data).

Other exceptions are possible for special data types (e.g., genomics, source code, high volume model data), for which PANGAEA is not a suitable repository and a dedicated community repository exists which fulfills the FAIR criteria. When archiving data in these repositories, it is always important to acknowledge the MOSAiC project (see [Data Policy](#)). Otherwise, findability of MOSAiC data from a single access point (portal) cannot be assured in the future. If you are unsure if this applies, [contact the PANGAEA team](#).

PANGAEA does not assign or link own DOIs to data sets published somewhere else.

When do I submit and publish “my” MOSAiC data?

Submit your quality controlled data sets as early as possible and before they are used for a paper. In PANGAEA, the data can be password protected, which means only metadata are accessible, but data itself cannot be viewed or downloaded. However, PANGAEA editors can already provide a temporary access key for reviewers. While data set status is “in review”, the content can still be changed, and the DOI is provided already for use in your manuscript.

Cite your data sets in any paper which is using them. Remember, data sets have full citations which can be used just like any other references.

Like paper publication, data publication involve editorial work, which requires time (sometime up to several weeks). Do not wait with the submission of data for publication for the last minute. No data citation will be possible before the data are actually archived in the repository.

How do I approach data publication?

[Raw data publishing](#) will be a semi-automatic process and the responsible PIs should provide necessary metadata and initiate the data archivation from MCS to PANGAEA.

Downloads and Related Links

- PANGAEA webinar "MOSAiC research data publishing" on 2020-11-10 [Slides \[4.6 MB, pdf\]](#)
- PANGAEA webinar "MOSAiC research data publishing" on 2020-07-08 [Slides \[4.43 MB, pdf\]](#)
- PANGAEA webinar "MOSAiC research data publishing" on 2020-11-10 [recording](#) (1:30 min)

- MOSAiC data policy <https://spac.es.awi.de/x/4pcpEg>
- PANGAEA wiki (official PANGAEA documentation) <https://wiki.pangaea.de>
- Videotutorial: [How to publish FAIR Data with PANGAEA](#) (5 min)

Primary data publication (calibrated data, data ready made for a paper publication) needs to be always initiated by the authors by opening a [data submission ticket in PANGAEA](#) or other designated repository, if exceptions apply.

If the raw data wasn't published with PANGAEA at the time of primary data publication yet, and is needed, [contact the PANGAEA team](#) for initiating the raw data publication.

During data publication instruct the editors in your data repository to create links to other versions of data (e.g., raw data), especially when they were or are being published in another repository.

All published data must include a funding acknowledgment of MOSAiC in the following form: "Multidisciplinary drifting Observatory for the Study of the Arctic Climate (MOSAiC)" with the tag "MOSAiC20192020". Additionally, the Project ID given for specific expedition must be mentioned. For the Polarstern expedition this is "AWI_PS122_00". Additional attributions like specific award/grant numbers might be added.
