

(AC)³ Data Policy

Version 08.2020



Introduction

“ArctiC Amplification: Climate Relevant Atmospheric and SurfaCe Processes, and Feedback Mechanisms - (AC)³” (<http://www.ac3-tr.de>) is a transregional collaborative research center (TR172) and is funded by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft). It is constituted by a German research consortium (Universities of Leipzig, Bremen and Cologne, Alfred-Wegener-Institute for Polar and Marine Research, Leibniz Institute for Tropospheric Research) and investigates why the Arctic region has been warming more rapidly than the rest of the world over the past three decades.

The (AC)³ partners are committed to handle generated scientific data according to the [Guidelines for Safeguarding Good Research Practice](#) and the FAIR (Findable, Accessible, Interoperable, Reusable) principles [[paper](#), [webpage](#)].

All scientific data collected within the framework of (AC)³ will be made publicly available. This document specifies the process of data publication and sets guidelines to the handling of research data before the final publication.

This data policy aims to:

- Ensure the fair use of the data
- Encourage the rapid dissemination of the scientific results
- Uphold the rights of the individual scientists
- Ensure the visibility and the integrity of the project
- Have all involved researchers treated equitably
- Encourage an orderly and timely analysis and publication of the data
- Define a central, open access repository of the data to be published

§1 Subjects of this data policy

Data collected in (AC)³ originates in a variety of different sources and therefore requires appropriate approaches in data archiving to ensure traceability and reproducibility of scientific results. Broadly the data can be divided into 3 different categories.

Primary Data Products are quality controlled “raw” data products without further post-processing applied to them. Generally, all (raw) sensor data falls under this category. All primary data collected within (AC)³ have to be published.

Secondary Data Products are derived from primary data products. If the processing is involved enough to warrant the specification as a distinct data set, instead of solely documenting the processing steps or if the processing cannot be repeated without expert knowledge, the data has to be published as well.

Model Data Products are results from model runs that may involve primary or secondary data as input. If the computational costs or expert knowledge required are high enough to justify publication as a distinct dataset, these data sets have to be published.

Further definitions

Data Provider is the person in charge of the instrument used to measure the data, or the person responsible for the retrieval or model run, respectively.

Early access refers to the access to a dataset before publication.

Quick Look refers to a visualization of preliminary data that is produced shortly after the respective data set was created.

Data Publication refers to either a published data set, or a scientific journal article with the data set appended to it.

§2 Data Management

Support

During all steps of the data publication process, support can be requested from the (AC)³ INF project (data@ac3-tr.de).

Data access

The data shall be made publicly available without access restrictions. The central repository for the publication of the data is [PANGAEA](#), operated by the Alfred-Wegener-Institute for Polar and Marine Research (AWI). The data can be published as a stand-alone data set or as a supplement to a publication. In case of conflicting commitments to prior or secondary data publication requirements, the data can also be published through third party services, given that the FAIR principles are met.

Citation

A persistent unique identifier such as a Digital Object Identifier (DOI) will be assigned to the data published in Pangaea to ensure correct attribution and reference to the data by potential users.

Data formats

The data must be provided in a form that ensures accessibility without proprietary tools/software and follows standards of the scientific community (e.g. ASCII, NASA-Ames, or NetCDF format, etc.). Metadata, sufficient to independently understand the data must accompany the submission. The usage of international standards for the metadata delivery (e.g. CF conventions) is encouraged.

Data quality

The analysis of the scientific data has to be performed according to the [Guidelines for Safeguarding Good Research Practice](#) by the DFG. Quality control of the data sets submitted for publication is the responsibility of the data provider.

§3 Provisioning of data, Licensing and Sharing

(1) Provisioning of data

The data of the individual scientific instruments along with the corresponding documentation to understand the data shall be made available to all (AC)³ partners as soon as possible, no later than 6 months after collection (Early access). In case of campaigns, data sets shall be made available to project partners shortly after their collection (Quick Look), unless technical reasons prohibit the creation of timely data sets.

The data should contain qualifiers of the status of the analysis (e.g. preliminary, final). Corrections and amendments to preliminary data must be made available to all participants as soon as possible.

(2) Data licensing and sharing

Only the data provider can share their data. Sharing of data without the consent of the data provider is not permitted.

Individual scientists wishing to collaborate can apply to get access to the data of interest before publication by contacting the responsible PI and provision of a short written description of her/his aims and agreement to the terms in Appendix A. Users of the data are obliged to accept the provisions as stipulated in § 4 et seq.

When data are shared before publication, it should be ensured that research should not duplicate already ongoing work by the PI and/or their collaborators. This holds especially when the topic is central to an ongoing PhD thesis. Further §7 applies.

The data provider can choose to password-protect the data uploaded to PANGAEA before final publication, e.g. during the peer-review process of a paper in which the data is used.

All collected scientific data have to be made publicly available no later than one year after collection.

§4 Publications/conference contributions/public meetings

(1) Co-authorship

If data of the project is used for publications (scientific articles), conference contributions and/or in public meetings the lead author has to invite the data provider to contribute to the publication. Co-authorship has to follow the [Guidelines for Safeguarding Good Research Practice](#).

(2) Right to refuse

The data provider has the right to refuse the usage of the data in publications, conference contributions and/or in public meetings prior to her/his publication of that work.

(3) Co-authorship

- The manuscript for any journal publication must be sent to the co-authors at least three weeks before submission with a copy to the PIs involved.
- Abstracts to contributions to conferences and/or public meetings must be sent to the co-authors at least one week before submission with a copy to the PIs involved.
- The lead author has to take into account substantial comments from the co-authors before the submission of the manuscript or abstract. The co-authors and the instrument or model PIs involved shall be kept informed of the progress of the review and publication.

(4) Usage of data

The lead author must have the explicit agreement of the data providers to use the data in publications, conference contributions and/or public meetings.

(5) Recent version of data

The PI must ensure that the data used in publications, conference contributions and/or public meetings is the best available at that time.

(6) Acknowledgement

The (AC)³ project must be acknowledged in scientific publications. E.g. by the following phrase:

We gratefully acknowledge the funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Projektnummer 268020496 - TRR 172, within the Transregional Collaborative Research Center "Arctic Amplification: Climate Relevant Atmospheric and Surface Processes, and Feedback Mechanisms (AC)³"

§5 Press releases

- Publications and/or press releases where the project itself is the main subject (so called overview articles) must be approved by all PIs before submission.
- Press releases related to data, created within (AC)³, must be approved by the data provider(s).

§6 Collaborations

Effective and productive collaborations between the participants of the (AC)³ and other projects are encouraged.

§7 Avoidance of disputes

Any disputes about the use of the data in particular with respect to publications will be resolved by a committee of scientists suggested by the PIs involved and approved by the Scientific Steering Team (SST).

With the aim to avoid redundant work and parallel efforts which are both sources of potential disputes, scientific meetings on a regular basis are encouraged, where all participants of the mission can present their ongoing work and their plans for future work with the data of the mission. This holds especially during the funding period of the project, when the data is still not public.

§8 Constraints

National and international laws or regulations, as well as commitments to other projects or networks, might constrain the agreements of this data policy. In that case the principal investigator has to inform other involved PIs about those constraints. The involved PIs shall find a mutual solution on a case-by-case basis.

The signatory agrees to the conditions of this data policy.

Name:

Address:

E-mail:

Date, place:

Signature:

Appendix A

Agreement to receive data licensed under the (AC)³ Data Policy

I will receive a copy of a data set of the project

I agree to the provisions stipulated in the (AC)³ Data Policy and the following common principle:

"If data of the project is used for a publication, conference contribution and/or public meeting, co-authorship has to be offered to the data provider(s). I will not distribute this data set or parts of it."

Name:

Affiliation and Address:

E-mail Address:

Date, Place:

Signature of the receiver of the data set: