(AC)$^3$ Data Policy

Draft, Version 24.07.2020

Preamble

“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.

Handling of different types of data

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 and the processing is involved enough to warrant the specification as a distinct data set, instead of solely documenting the processing steps. If the processing cannot be repeated without expert knowledge, the data has to be published.

**Model Data Products** are results from model runs that may involve primary or secondary data as input but the computational costs or expert knowledge required are high enough to justify publication as a distinct dataset.

Guidance in

Preparation for Publication

Collaboration

Data publication

embargo end of project

closed - within (AC)3 preliminary - pangaea password Projektlaufzeit 1 Ack: soll DOI citation

satellite Model - paper citation

To formalize the requirements stated above and to specify them in the context of the (AC)$^3$ 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 Definitions

**Primary Data Products** are quality controlled “raw” data products without further post-processing applied to them. All primary data collected within the (AC)³ are subject to this data policy. It is assumed that primary data from third party sources outside the frame of (AC)³ is archived and made accessible by the third party already.

**Secondary Data Products** are derived from primary data products and the processing is involved enough to warrant the specification as a distinct data set, instead of solely documenting the processing steps.

**Model Data Products** are results from model runs that may involve primary or secondary data as input but the computational costs or expert knowledge required are high enough to justify publication as a distinct dataset.

**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.

**Participant** refers to a person that signed this document and/or was granted participation by approval of the responsible PI after signing this document in Appendix A.

§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 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 Provision of data, Licensing and Sharing

(1) Provision of data during campaigns

Each group will provide quick looks of their data to the participants shortly after each flight or as soon as possible after recording, unless technical reasons prohibit the creation of timely data sets. The data of the individual scientific instruments along with the corresponding documentation to understand the data shall be made available to all participants as soon as possible.

Six months after the observations were made the data and the documentation must be made available to all participants. 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.

During the duration of the (AC)³ project the data can be put under an embargo by the data provider. If a dataset is under embargo, it is then only available to the participants of the project and their collaborators. An individual scientist wishing to collaborate can still apply to get access to the embargoed data of interest by contacting the responsible PI and provision of a short written description of her/his aims and agreement to the terms in Appendix A.

When data are shared, 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 must decide if the data shall be released to individual persons or to the public by one of the following options:

- **Closed:** A potential data user can request the data of the instrument or model by contacting the respective PI. The data can be made available to that person under the conditions defined in this data policy. The agreement in Appendix A has to be signed by this person and returned to the respective PI. The final decision to share or not to share the data with that person remains by the respective PI.

- **Open:** The data is made publicly available under a suitable open access license such as one of the Creative Commons licenses.

In either case the users of the data are obliged to accept the provisions as stipulated in §4 et seq. All collected scientific data finally has to be made publicly available by the end of the project at the latest.

§4 Publications/conference contributions/public meetings

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.

(1) 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.

(2) 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 project must be acknowledged in a suitable way. (AC)³ needs to be acknowledged by the following phrase:

We gratefully acknowledge the funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Projektunmner 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 as well as commitments to other projects and networks 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 the (AC)³ data policy.

Name: .................................................................

Address: ........................................................................

E-mail: .................................................................

Date, place: .................................................................

Signature of participant: .................................................................
Appendix A

Agreement to receive data that was licensed according to §§ 2 and 3

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: ..........................................................

Date, place: ..........................................................

Signature of the receiver of the data set: .................................