WDCC User Guide for Data Publications
(Version 1.0)

Long Term Archive (LTA) group

May 10, 2022

Contents

1 Introduction

The World Data Center for Climate (WDCC) is hosted by the German Climate Computing Center (DKRZ) in Hamburg and is maintained by its Data Management (DM) department.

The WDCC collects, archives, and disseminates Earth System data and climate-related products relevant for climate and Earth System research with a focus on climate simulation data.

It is a FAIR repository (https://www.wdc-climate.de/WDCC/ui/cerasearch/info?site=fairness), certified according to the criteria of the Core Trust Seal (CTS) and is, as World Data Center for Climate (WDCC), accredited as regular member of the World Data System.

WDCC offers a DataCite DOI publication of the archived items to make them citable. DataCite is a leading global DOI provider for research data. The mission of the WDCC is to ensure long term data accessibility and re-usability of climate data.
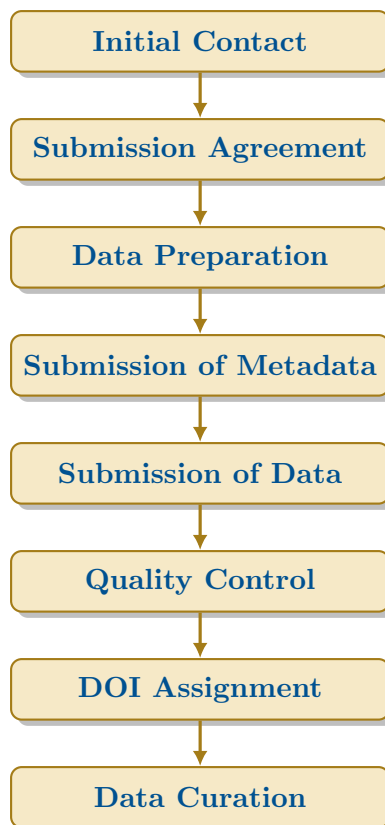
This document supports the WDCC Step-by-Step Guide (https://www.wdc-climate.de/WDCC/ui/cerasearch/info?site=guide_wdcc) and aims at guiding users through the process of publishing data in the WDCC.

1.1 Submission Steps

DKRZ's Data Management (DM) will support you, starting from your initial intent to archive data until they are published in the WDCC.

Before you start the submission process in WDCC, please make sure that you have the authorization to publish the data.

In WDCC the submission process includes the following steps, which are described in detail in the remainder of this document:

Initial Contact

↓

Submission Agreement

↓

Data Preparation

↓

Submission of Metadata

↓

Submission of Data

↓

Quality Control

↓

DOI Assignment

↓

Data Curation

2 Initial Contact

Whenever you wish to archive data in the WDCC — including or not a DataCite DOI — the first step is to get in contact with DKRZ Data Management (DM). Please write an email to data@dkrz.de including a short statement describing the data, your contact and project details, the approximate size of the data to be preserved and whether you wish to have a DataCite DOI assigned to the data or not. Based upon your description, DM will decide if your data fits into the scope of the WDCC.

More specifically, please provide the following information for data archival:

- Institute

- Scientific Project

- DKRZ-Project (with account number if available)

- WDCC-Project (with URL if available)

- Do you need a DOI? When is it needed?

- Where is the data (already on Mistral/Levante in /work or somewhere else)?
    - if on Mistral/Levante, please provide the path

- Approximate data size

- Approximate number of files

- Data organization: How is the data structured/organized (e.g. according to models, simulations, variables, . . .)?

- Format of files (e.g. netCDF)

- Does the data follow a specific community standard (e.g. CMOR, EASYDAB)? Is the data subject to a certain quality control?

- License: If you do not plan to choose a CC-License (https://creativecommons.org/), please make a note. (For more information on licenses see below)

- Will there be changes in the data later on? Do you plan to add data later in the process? Would you like something like an open time series, versioning or according data organization?

If you don't have all the information in the first email, you can add this in the process later on.

Your first email to data@dkrz.de will open a ticket in a request tracker. DM will contact you shortly with further instructions. Please do not write a new email, but only answer to this thread. Otherwise a new ticket would be opened.

3 Submission Agreement

The submission agreement regulates how your data should be archived. You will discuss with DM the structure of your data and organization into hierarchical elements as well as licensing issues and authorship.

Data Management at DKRZ will advise you on how to best structure your data for publication in WDCC.

3.1 License

You need to decide under which license to publish your data. We suggest CC-BY 4.0 for data submissions. A short overview is given here.

3.1.1 Creative Commons Attribution 4.0 International

→ https://creativecommons.org/licenses/by/4.0/

→ https://creativecommons.org/licenses/by-nd/4.0/

→ https://creativecommons.org/licenses/by-nc/4.0/

→ https://creativecommons.org/licenses/by-nc-nd/4.0/

→ https://creativecommons.org/licenses/by-nc-sa/4.0/

→ https://creativecommons.org/licenses/by-sa/4.0/

Summary (human readable): To see which terms apply to which Creative Common license in detail, please refer to table 1.

You are free to (green in table 1):

- *Share* – copy and redistribute the material in any medium or format

- *Adapt* – remix, transform, and build upon the material.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms (blue in table 1):

- *Attribution* – You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

- *NoDerivatives* – If you remix, transform, or build upon the material, you may not distribute the modified material.

- *NonCommercial* – You may not use the material for commercial purposes.

- *ShareAlike* – If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

No additional restrictions: You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices: You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

License	Share	Adapt	Attribution	NoDerivatives	NonComercial	ShareAlike
CC BY 4.0	■	■	■			
CC BY-ND 4.0	■		■	■		
CC BY-NC 4.0	■	■	■		■	
CC BY-NC-ND 4.0	■		■	■	■	
CC BY-NC-SA 4.0	■	■	■		■	■
CC BY-SA 4.0	■	■	■			■

Table 1: Summary of Creative Commons Attribution 4.0 International

3.2 Hierarchy Structure Description

Data in WDCC is organized in multiple layers with a hierarchical structure. These layers are named: *Projects*, *Experiments*, *Dataset groups* and *Datasets*. The relationships between them can be described with a tree structure as in figure 1.

For the top-level elements only single-parent (mono-hierarchical) relationships are supported. That means that only one parent is allowed for each Experiment and Dataset group, but a parent can have multiple children-elements. Datasets and *Additional Infos* (which are a bit of an exception and will be described later in this section) need to have at least one parent, but it is possible that they have more than one.

The Project, Experiment and Dataset groups contain the overall description of the data, i.e. the metadata. Metadata is the description of the data itself. Detailed metadata is essential for later reuse of the data and therefore for long-term archiving. These elements (Project, Experiment and Dataset groups) can therefore be used to structure the data. The data itself is located in the Datasets (which also contain description of the data).

Each Dataset is part of a hierarchical tree with an Experiment and a Project, while Dataset groups are optional elements that are only used for more complex structures. The most minimalistic data to be archived needs to contain at least one Project, one Experiment and one Dataset in the hierarchical structure.

The only hierarchical elements that can get a DataCite DOI are Experiments and Dataset groups. It is not possible to get a DOI for Projects or Datasets.

Project: A Project is the top hierarchical element. It is a scientific research activity and can consist of one or many experiments. It can be used to organize collections of data from a number of potentially different data sources (e.g. different models or different institutes) that still belong to the same scientific project.

Experiment: An Experiment is the hierarchical child element of a Project and a compilation of Dataset groups or Datasets. It is possible to get a DataCite DOI for an entire Experiment.
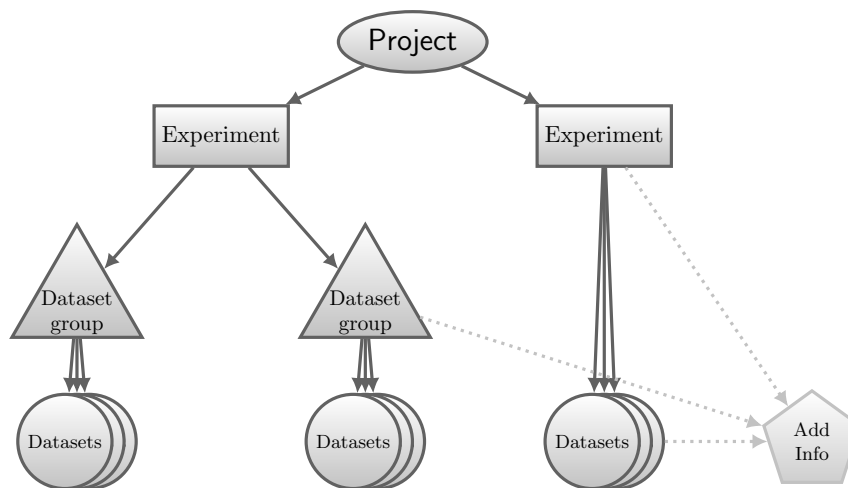
Figure 1: WDCC hierarchical structure elements

It has to be noted that "Experiment" is just a name for a hierarchical element and should not be confused with a model run for example.

Dataset group: A Dataset group is a compilation of Datasets or Dataset groups. This is an optional element that does not need to be used for smaller projects, but can be helpful for bigger or more complex projects to sort them into finer structures. It is possible to get a DataCite DOI for an entire Dataset group. A Dataset group is a child of an Experiment or another Dataset group.

Dataset: A Dataset consists of the data itself and its description (the metadata) needed to make the data understandable. A Dataset is a child of an Experiment or of a Dataset group.

Additional Info: There is another element a bit outside of the hierarchical tree in figure 1. An Additional Info is an optional element that is located in the same layer as a Dataset, i.e. is the child of an Experiment or a Dataset group (but can also be the child of a Dataset) and can have more than one parent. It is defined as a compilation of documents or plots enhancing further understanding of the data. An Additional Info can also receive a DataCite DOI.

3.3 Nomenclature

DM will help you to organize your data into the hierarchical structure described above. In this context the nomenclature of the elements will also be discussed. You have to give each element a name which is the title for this entry and will also be used in the WDCC search interface. Hints for good entry names will be given in the section "Submission of Metadata".

You can decide together with DM which structure and nomenclature makes sense for your data.

4 Data Preparation

Please prepare your data files for submission. In the following, you can check what to take into consideration.

Most likely DM will ask you for some example files (or the location of your files if they are accessible to DM) to do some first checks of their own.

4.1 Data Format

Only open source data formats are accepted in WDCC. The preferred file format for archiving is NetCDF (Network Common Data Format). More information on NetCDF can be found at https://www.unidata.ucar.edu/software/netcdf/.

Other supported data formats are for example GRIB (GRIdded Binary), CSV, and ASCII. If you need to archive your data with another (open source) format, please contact DM to discuss this.

There can be only one unique file format for each Dataset. If you have multiple file formats in your data to be archived you need to divide them into different Datasets.

4.1.1 Additional Infos

If you want to upload extra material, please convert them to PDF, which is the preferred format for Additional Info files. In some cases it might be sensible to upload other formats, depending on the content. Just contact DM to discuss this.

4.2 CF Conventions

We strongly recommend the use of the CF Metadata Conventions (http://cfconventions.org/) for NetCDF data. The CF conventions define metadata that provide a definitive description of what the data in each variable represents and the spatial and temporal properties of the data.

An important part of the CF conventions are the CF standard names for variables. Please refer to http://cfconventions.org/Data/cf-standard-names/current/build/cf-standard-name-table.html for a current list of CF standard names and check if there are appropriate ones for your variable(s). These should be added to the NetCDF file headers (as `standard_name`). Additionally (or in case you cannot find an appropriate CF name in the list) you can add a `long_name` attribute to the file headers to describe your variable(s).

Before you transfer your data to DKRZ, please use a CF checker to make sure that your data does not produce any errors. For single file check you can use (for example): https://pumatest.nerc.ac.uk/cgi-bin/cf-checker.pl. To check an entire directory, you can, for example, use the CF checker that is part of the atmodat data checker (https://github.com/AtMoDat/atmodat_data_checker).

4.3 File Size

The size of each individual file should not be larger than 4GB. If your files are larger than 4GB, please split them into smaller files (If in doubt, you can discuss with DM, which would be the best way to divide your files).

4.4 Compression

Normally, it is not necessary to compress your files. However, if you have many small files it can be useful to create a tarball or a zip-archive. A single Dataset should not contain more than 1000 individual files. Please discuss with DM, if this is applicable to your data.

4.5 Labeling of files and directories

Files should be labeled in a consistent manner. Special characters and spaces in directory and file names should be avoided. The file names have to be unique (even if spread over multiple directories).

Please sort your files so that you have one directory for each Dataset.

4.6 Temporal Aggregation

The temporal aggregation of your data should be included in the file header.

In each Dataset there can only be one type of temporal aggregation, so if you have different aggregations, you need multiple Datasets for that.

5 Submission of Metadata – MetaXA Guide

The metadata belonging to your data can be entered via a graphical user interface (GUI), MetaXA (https://cera-www.dkrz.de/LTA_metadata). Before you can use this application, you need an WDCC account. If you do not have one already, please apply for an account here: https://www.wdc-climate.de/WDCC/ui/cerasearch/register.

In MetaXA you can enter metadata for the different hierarchical elements: Projects, Experiments, Dataset groups, and Datasets. Furthermore, you can add (or edit) a Person, add an Additional Info or a Citation/Reference. However, please wait for instructions from DM which elements to fill out (and when).

Only marked fields are required, but try to fill out as many details as possible. Help texts are available directly in MetaXA (via the little ?) and more details for each field are given in the Appendix in the respective tables.

5.1 Person

Please start with the Person entry. Here you should enter all persons relevant to your data (e.g. all authors who should be listed in the citation). Via the search form you can check if the person in question is already listed in WDCC. Please do so to avoid duplicate entries. If you find the person in WDCC, check if all information is still correct. Make corrections, if necessary.

If you cannot find the person, you need to click the button "Add new person". Please enter the details belonging to the person here.

External Identifier: We would like to get the ORCID from all users (even though it is not a required field). If you do not have an ORCID, you can get it here: https://orcid.org. Alternatively (or additionally) you can also add the ID for ResearchGate or ResearcherID.

Please enter it in the "External Identifier" field. For that you need to choose the correct identifier type in the first row and enter the corresponding ID in the second row.

Institute: When entering information for a person, please choose an institute from the list. If you do not find the correct institute, click on "Other (new) Institute" and enter the details for the institute here.

5.2 Project

As a next step, please enter your project in MetaXA. If the project already exists, you can search for it here and make changes if necessary. If it does not exist yet, click on "Create new Project" and enter the project acronym, name, and description. For details, please see table A1.

5.3 Experiment

Entering an Experiment in MetaXA you can either start from scratch or you can prefill the form from an existing experiment (e.g. if you have entered a similar experiment before and just need to edit a few fields). To choose an experiment to prefill the form from, just click on the button "Prefill from existing or edit entry" and enter the name of the respective experiment.

Once you have entered all the information (at least all required fields), you can create and save your entry with "Create Entry". In case you want to make some more changes you can come back to the entry later.

Once you have finished editing your metadata and want to inform DM that you have completed all entries for the experiment (or any other entry), click on "Notify Completion". If, later on, changes need to be made, this is still possible.

Metadata Entry: Please refer to table A2 for a detailed description.

Citation Information: Please refer to table A3 for a detailed description.

Temporal Coverage: This block is only available if the box "Data has no temporal information" has NOT been checked. Please refer to table A4 for a detailed description.

Spatial Coverage: This block is only available if the box "Data has no spatial information" has NOT been checked. Please refer to table A5 for a detailed description.

Quality: Please refer to table A6 for a detailed description.

5.4 Dataset Group

Entering a Dataset group is very similar to an Experiment. Please do not forget to go to "Notify Completion" once you are finished.

Metadata Entry: Please refer to table A7 for a detailed description.

Citation Information: This block is only available if the box "Citation information have been provided on parent level" has NOT been checked. Please refer to table A3 for a detailed description.

Temporal Coverage: This block is only available if the box "Data has no temporal information" has NOT been checked. Please refer to table A4 for a detailed description.

Spatial Coverage: This block is only available if the box "Data has no spatial information" has NOT been checked. Please refer to table A5 for a detailed description.

Quality: Please refer to table A6 for a detailed description.

5.5 Dataset

There are different methods to enter metadata for Datasets. Please discuss with DM whether MetaXA or CSV-lists are better for your case, before entering any metadata.

5.5.1 MetaXA

The following tables will help you to enter metadata into MetaXA. Please do not forget to go to "Notify Completion" once you are finished.

Metadata Entry: Please refer to table A8 for a detailed description.

Temporal Coverage: This block is only available if the box "Data has no temporal information" has NOT been checked. Please refer to table A4 for a detailed description.

Spatial Coverage: This block is only available if the box "Data has no spatial information" has NOT been checked. Please refer to table A5 for a detailed description.

Quality: Please refer to table A6 for a detailed description.

5.5.2 CSV-Lists

In case you have many Datasets (or Dataset groups) that do not differ strongly from each other, it might be more convenient to create a CSV-list containing all relevant metadata. DM will assist you with this and will give you further instructions in case this applies to your data.

5.6 Additional Info

Fewer fields are needed for the Additional Info. Please do not forget to go to "Notify Completion" once you are finished.

Metadata Entry: Please refer to table A9 for a detailed description.

Citation Information: Please refer to table A10 for a detailed description.

File upload: The Additional Info can consist of one or more files. You can upload the file(s) directly in MetaXA. The preferred file format for Additional Infos is PDF. If you cannot upload your file, this is most probably caused by an unsupported file format. Please convert it to PDF or contact DM to discuss if it possible to upload your preferred file format.

5.7 Citation / Reference

You can add references to your entries. Use the upper left field to search for existing citations and update them if necessary.

If you want to add new citations, use the button in the upper right corner "Create Reference Template". This opens the form to enter metadata for a new Citation/Reference. If a DOI is available, please enter it in the respective field in the format `doi:10.1045/january2015-brase` and click on "Insert from CrossRef". Otherwise, fill in the fields below (required fields are marked with a red asterisk).

If you edit or add a reference via MetaXA, you can also connect it with your entry. If you have already entered the Experiment, Dataset group, or Dataset in MetaXA you will find it in a list and can choose one or more of the entries. Otherwise contact DM and they will assist you in connecting the reference to your entries.

5.8 Notify Completion

Once you have completed entering all relevant metadata for an entry and want to notify DM, please do not forget to go to "Notify Completion". This needs to be done for Experiments, Dataset groups, Datasets, and Additional Infos. Persons, Projects, and Citations can be added without "Notify Completion".

Check all appropriate entries and click on "Finalize Selected Entries". After you have finished your work with "Notify Completion", DM will further process your entries.

Changes can still be made as long as a DOI has not been assigned. Please contact DM, if you find that amendments are needed.

6 Submission of Data

Once all metadata has been published in WDCC and the data files are prepared for submission, the data needs to be handed over to DM. If the data is already located on Mistral/Levante, please provide DM with the path.

Otherwise, it depends on the size and granularity of your data which method is best suited for the transfer. DM will advise you how to proceed.

7 Quality Control

Before metadata and data are published quality assurance and control will take place. It is therefore possible that DM will get back to you and some modifications need to be performed.

An even more thorough check will be made during the DOI assignment. Details will therefore be discussed in the next section.

8 DOI Assignment

If a DataCite DOI assignment was agreed upon (which happens mostly in the first emails), DM will contact you after your data has been published in WDCC and guide you through the process.

The DOI assignment at WDCC is a key process to make (meta)data FAIR (https://www.wdc-climate.de/WDCC/ui/cerasearch/info?site=fairness).

The only hierarchical elements that can get a DataCite DOI are Experiments, Dataset groups, and Additional Infos (more details see section on "Hierarchy Structure Description"). It is not possible to get a DOI for Projects or Datasets.

The DOI resolves to a landing page hosted by WDCC, which summarizes basic data-related information. The page functions as an entry point for data access and detailed information on the data. A complete list of data published by WDCC is accessible at: https://www.wdc-climate.de/WDCC/ui/cerasearch/q?ref_type_name_s=Citation-DOI.

For the DOI publication at WDCC, the data has to meet certain quality requirements to ensure the following benefits:

- ensuring a high data quality and long-term usability for an interdisciplinary user community (data entities are fixed and remain unchanged)

- ensuring a permanent and persistent access and long term availability of data and metadata via the persistent identifier (DOI)

- enabling easy and clear formal data citations for use in scientific publications and scientists' lists of publications in order to give credit to the data creators

8.1 Quality Assurance

Quality tasks and responsibilities are divided between:

- Data Publishing Agency (WDCC) performing automated checks on metadata consistency (*Technical Quality Assurance – TQA*)

- Data Provider controlling the scientific data quality (*Scientific Quality Assurance – SQA*).

Contents of quality checks (especially in SQA) as well as definitions of quality levels and the overall quality procedures vary significantly between data types and scientific projects.

8.1.1 Technical Quality Assurance at WDCC

During the cross- and double-checks of WDCC's publication agent at least the following criteria are checked:

- Number of data files is correct and > 0

- Size of every data file is > 0

- The data files and corresponding metadata are accessible

- The data file sizes are controlled and correct

- The spatial-temporal coverage description (metadata) is consistent to the data, time steps are correct and the time coordinate is continuous

- The format is correct

- Variable description and data are consistent

To avoid problems in the TQA, it is advisable to do also these checks yourself before handing over your data.

9 Data Curation

After your data has been published and (possibly) a DOI has been assigned, data curation takes place in close cooperation between you (the data provider) and WDCC. Main tasks include:

Add references and cross-references to journal publications: If your data is used for a journal publication and/or is cited in a journal, please inform DM so that this reference can be added to your WDCC data publication.

Add data to "open time series": If this was agreed upon with DM in the beginning of your request, it is possible to store your data as an open time series. This is for data that comes in a time series where new data is continuously created. The requirement is that the method to generate this data must not change. In this case it is possible to even assign a DOI to this data even if data is appended to this time series later on.

Creation of new versions: If an open time series is not the right format for you (e.g. because the method to generate your data has changed or because you did not foresee that new data would come or simply because it is not a time series) but you want to update your data, a new version can be created.

Errata: In case you notice errata in your data that have already been published in WDCC, please inform DM immediately. In most cases it will not be possible to simply exchange the data files anymore once it is completely published in WDCC. However, it is possible to withdraw the data and create a new version. In that case DM will also ask you for an explanatory text to add to the old and new versions' summary.

Appendix A: Description of MetaXA fields

List of Tables

Table A1: **Projects**

Field	Required	Comment
Project Acronym	✓	The project acronym can be 31 characters or less and only alphanumeric characters as well as underscores and hyphens are allowed (no white spaces). The project acronym will be displayed together with the project name.
Project Name	✓	The project name can have a maximum length of 250 characters and should meet the following criteria: • The first letter should be in upper-case • Do not use underscores if not necessary • Whitespaces are the preferred word delimiters • It should provide proper information about the Entry • Give non-scientists an idea about the content • Avoid constant change of letters and numbers as well as constant change of upper and lower case letters. This will worsen the searchability. • If a version number needs to be provided, use the following format: Title of the entry (Version 2.1)
Project Description	✓	In the description please summarize the project, describe its scientific ideas and name its funders. Please also provide a URL to the homepage of the project, if there is one. Maximum length: 2000 characters

Table A2: **Experiments – Metadata Entry**

Field	Required	Comment
Name	✓	The entry name uniquely identifies the metadata entry (Experiment). Maximum length: 160 characters. The name should meet the following criteria: • The first letter should be in upper-case • Do not use underscores if not necessary • Whitespaces are the preferred word delimiters • It should provide proper information about the Entry • Give non-scientists an idea about the content • Avoid constant change of letters and numbers as well as constant change of upper and lower case letters. This will worsen the searchability. • If a version number needs to be provided, use the following format: Title of the entry (Version 2.1)
Summary	✓	It should be described here which part is covered by the Experiment. Give information on the entry that has not been entered into any other fields (e.g. motivation for conducting the experiments, models used with model version, sensors, sources, technical aspects, basic state variables, aggregation, the grid used, the part of the project, ...). The summary may contain links to further information. Maximum length: 4000 characters.
Project	✓	Select a project from the list (all projects are listed here, which were entered before through the Project tab).
Use Constraint (License)		"Use Constraint" reflects constraints regarding the use of the data. Refer to the section about licenses for further details.
Data has no temporal/spatial information		Check these boxes (or one of these) if the data has no temporal and/or no spatial information.
Keywords		Please choose a few keywords to improve the findability of your data. You can use the upper field to search for keywords and the lower field to suggest new keywords.

Table A3: **Experiments and Dataset groups – Citation Information**

Field	Required	Comment
Citation Title		The title of the citation. Please follow the same rules as for the entry names. If left empty, the entry name will be used.
Authors List	✓	The authors of this data for the citation.
Contact Author	✓	The contact author is the contact of the primary data for the DOI process. Usually they are the first author. The contact author is responsible for the data and they are the contact person on the DOI internet page. A maximum number of three contact authors can be chosen.

Table A4: **Experiments, Dataset groups, and Datasets – Temporal Coverage**

Field	Required	Comment
Start Year	✓	The first year of the considered time period.
Month	✓	The first month of the considered time period.
Day	✓	The first day of the considered time period.
Stop Year	✓	The last day of the considered time period.
Month	✓	The last day of the considered time period.
Day	✓	The last day of the considered time period.
Currentness Reference		Currentness Reference reflects the construction of the time axis. Choose the correct calendar from the list. Use "not filled" if no entry matches.

Table A5: **Experiments, Dataset groups, and Datasets – Spatial Coverage**

Field	Required	Comment
Min Lon	✓	The minimum longitude of the considered domain [0.0, 360.0].
Max Lon	✓	The maximum longitude of the considered domain [0.0, 360.0].
Min Lat	✓	The minimum latitude of the considered domain [-90.0, 90.0].
Max Lat	✓	The maximum latitude of the considered domain [-90.0, 90.0].
Min Altitude		Minimum altitude. Example for ocean/atmosphere data: -6200m (ocean) to 10hPa (atmosphere)
Min Alt Unit		Unit for minimum altitude.
Max Altitude		Maximum altitude. Example for ocean/atmosphere data: -6200m (ocean) to 10hPa (atmosphere)
Max Alt Unit		Unit for maximum altitude.

Table A6: **Experiments, Dataset groups, and Datasets – Quality**

Field	Required	Comment
Accuracy Report		The sole responsibility for quality control lies with the authors. The accuracy report should contain information on the procedure of data quality checking and its findings. For model data the accuracy report should contain a link to the homepage of the model documentation, short descriptions of the resolution, parameterization, boundary, conditions, input data, and constants of the experiment. For observational data it should contain the data level, description of campaigns, supersites, resolution in time and space, instruments, and platforms. The accuracy report should be completed with the confirmation: "SQA - Scientific Quality Assurance 'approved by author'". Information that does not fit into this space can be uploaded as an Additional Info. Maximum length: 2000 characters.
Completeness Report		Please insert a summary of the temporal, spatial and parameter completeness with description about how missing values are indicated (if applicable). Maximum length: 2000 characters.

Table A7: **Dataset groups – Metadata Entry**

Field	Required	Comment
Name	✓	The entry name uniquely identifies the metadata entry (Dataset group). Maximum length: 160 characters. The name should meet the following criteria: • The first letter should be in upper-case • Do not use underscores if not necessary • Whitespaces are the preferred word delimiters • It should provide proper information about the Entry • Give non-scientists an idea about the content • Avoid constant change of letters and numbers as well as constant change of upper and lower case letters. This will worsen the searchability. • If a version number needs to be provided, use the following format: Title of the entry (Version 2.1)
Summary has been provided with parent metadata		Check this box, if the summary has already been provided with the parent element (Experiment or Dataset group)
Summary	✓	It should be described here which part is covered by the Dataset group. Give information on the entry that has not been entered into any other fields (e.g. motivation for conducting the experiments, models used with model version, sensors, sources, technical aspects, basic state variables, aggregation, the grid used, the part of the project, ...). The summary may contain links to further information. Maximum length: 4000 characters. *If the summary is left unfilled (if the box above is checked), the corresponding summary from the parent element will be shown.*
Parent	✓	Select an Experiment (or Dataset group) to be the parent of this entry. If you have already entered the element it should appear in this list.
Data has no temporal/spatial information		Check these boxes (or one of these) if the data has no temporal and/or no spatial information.
Citation information have been provided on parent level		Check this box, if you want to take the citation information from the parent element (Experiment or Dataset group).
Keywords		Please choose a few keywords to improve the findability of your data. You can use the upper field to search for keywords and the lower field to suggest new keywords.

Table A8: **Datasets – Metadata Entry**

Field	Required	Comment
Name	✓	The entry name uniquely identifies the metadata entry (Dataset). Maximum length: 160 characters. The name should meet the following criteria: • The first letter should be in upper-case • Do not use underscores if not necessary • Whitespaces are the preferred word delimiters • It should provide proper information about the Entry • Give non-scientists an idea about the content • Avoid constant change of letters and numbers as well as constant change of upper and lower case letters. This will worsen the searchability. • If a version number needs to be provided, use the following format: Title of the entry (Version 2.1)
Summary		It should be described here which part is covered by the Dataset. Give information on the entry that has not been entered into any other fields. The summary may contain links to further information. Maximum length: 4000 characters. *A Dataset does not necessarily need a summary of its own. If the summary is left unfilled, the corresponding summary from the parent element will be shown.*
Parent	✓	Select an Experiment (or Dataset group) to be the parent of this entry. If you have already entered the element it should appear in this list.
Format		Select format of data from the list. For details see section about Data Format.
Data has no temporal/spatial information		Check these boxes (or one of these) if the data has no temporal and/or no spatial information.

Table A9: **Additional Infos – Metadata Entry**

Field	Required	Comment
Name	✓	The entry name uniquely identifies the metadata entry (Additional Info). Maximum length: 160 characters. The name should meet the following criteria: - The first letter should be in upper-case - Do not use underscores if not necessary - Whitespaces are the preferred word delimiters - It should provide proper information about the Entry - Give non-scientists an idea about the content - Avoid constant change of letters and numbers as well as constant change of upper and lower case letters. This will worsen the searchability. - If a version number needs to be provided, use the following format: Title of the entry (Version 2.1)
Summary	✓	It should be described here which part is covered by the Additional Info. Give information on the entry that has not been entered into any other fields. The summary may contain links to further information. Maximum length: 4000 characters.
Parent	✓	Select an Experiment (or Dataset group) to be the parent of this entry. If you have already entered the element it should appear in this list.
Publication Type		If your publication is not really an "additional info", but something from this list fits better, you can specify that here.
Use Constraint (License)		"Use Constraint" reflects constraints regarding the use of the data. Refer to the section about licenses for further details.

Table A10: **Additional Infos – Citation Information**

Field	Required	Comment
Citation Title		The title of the citation. Please follow the same rules as for the entry names. If left empty, the entry name will be used.
Authors		The authors of this data for the citation. *This field is only required if a Data Cite DOI is requested.*