



Implementation of the Data Seal of Approval

The Data Seal of Approval board hereby confirms that the Trusted Digital repository DKRZ-LTA complies with the guidelines version 2014-2017 set by the Data Seal of Approval Board.

The afore-mentioned repository has therefore acquired the Data Seal of Approval of 2013 on June 29, 2016.

The Trusted Digital repository is allowed to place an image of the Data Seal of Approval logo corresponding to the guidelines version date on their website. This image must link to this file which is hosted on the Data Seal of Approval website.

Yours sincerely,

The Data Seal of Approval Board

Assessment Information

Guidelines Version:	2014-2017 July 19, 2013
Guidelines Information Booklet:	DSA-booklet_2014-2017.pdf
All Guidelines Documentation:	Documentation
Repository:	DKRZ-LTA
Seal Acquiry Date:	Jun. 29, 2016
For the latest version of the awarded DSA for this repository please visit our website:	http://assessment.datasealofapproval.org/seals/
Previously Acquired Seals:	None
This repository is owned by:	Deutsches Klimarechenzentrum GmbH

T +49 40 460094-0
F +49 40 460094-270
E thiemann@dkrz.de
W <http://www.dkrz.de/>

Assessment

0. Repository Context

Applicant Entry

Self-assessment statement:

General Description:

The mission of the “Long Term Archive” (DKRZ LTA) at Deutsches Klimarechenzentrum GmbH (DKRZ) is to provide central data preservation support for the climate modelling research community. For this service the DKRZ LTA has received a certification from the ICSU’s “World Data System” (WDS) in 2003 (renewed in 2011), and acts as well as “DKRZ - WDC Climate” (WDCC).

The services of the DKRZ LTA comprise among others:

- Data ingest, archival storage, management, administration and access operated by the CERA (Climate and Environmental Retrieval and Archiving) data and information system. The typical archiving duration is 10 years.
- A DOI data publication service for the direct integration of research data in scientific publications has been developed and implemented together with the Technical Information Library (TIB) in Hannover. The editorial process at DKRZ-LTA ensures the quality of metadata and research data in collaboration with the data producers. A DOI (Digital Object Identifier) and a citation code are provided and afterwards registered with the DOI registration agency in the TIB's library catalogue under DataCite's (datacite.org) regulations.
- Hosting of various data centers / collections , e.g. the IPCC Data Distribution Center (IPCC-DDC) and collaboration within scientific or technical projects.

Data processing, storing, copying and analysis facilities are available for the full data life cycle.

Data objects from climate research usually are homogeneous and very large (in the petabyte range). Members of DKRZ-LTA actively participate in the development of netCDF/CF, the most widely used data format in climate modelling. Special emphasis is put on the quality checking of data and access to subsets of data objects, both particularly in view of large datasets.

Submission Information Packages (SIP) as ingest are accepted in electronic form, specifics have to be discussed with DKRZ-LTA staff. Access to the databases is possible via Internet browser and API, or other media on request. An user account is required to access the data. This account is free of charge and available on request. CERA Metadata are mapped to ISO19115, DIF and DataCite metadata format.

Activities:

- Attending major conferences/meetings like AGU, EGU, GO-ESSP and RDA
- Producing articles or research papers
- Organise trainings and workshops
- Attending WDS groups
- Attending RDA groups

We do not outsource any of our functions.

Supporting Links and Documentation

- DKRZ <http://www.dkrz.de/>
- Data preservation services are listed on subpages from http://www.dkrz.de/daten/data-services/langzeitarchivierung?set_language=de
- Description of CERA https://www.dkrz.de/daten-en/cera?set_language=en
- Access to CERA <http://cera.wdc-climate.de>
- DKRZ-LTA preservation and storage policy <http://cera-www.dkrz.de/docs/DKRZ-LTA-PreservationAndStoragePolicy.pdf>
- ICSU World Data System <https://www.icsu-wds.org/>
- List of projects http://www.dkrz.de/daten-en/wdcc/projects_cooperations
- Publications <https://www.dkrz.de/about-en/media/pub/wiss-pub>
- Talks <https://www.dkrz.de/about-en/media/downloads/vortraege>

Data Seal of Approval Board

W www.datasealofapproval.org

E info@datasealofapproval.org

- Conference posters:
<https://www.dkrz.de/about-en/media/downloads/posters/wiss-poster>
- Workshops <https://www.dkrz.de/Nutzerportal-en/workshops>
- RDA <https://rd-alliance.org/>

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

1. The data producer deposits the data in a data repository with sufficient information for others to assess the quality of the data, and compliance with disciplinary and ethical norms.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Evidence:

DKRZ-LTA helps the data producers to comply with rules of good scientific practice as called for by the WDS Data Policy, DFG (German Research Foundation), MPG (Max Planck Society) or others.

The “Data Submission Preparation Guide” contains a checklist for depositing, which helps the data providers to verify and organise their materials, to carefully consider copyrights, download permissions, licenses, etc. This guide also contains links to further documents, an important one being a link to the “Depositor agreement” which lays out the legal basis for the deposit.

Different levels of quality assessment ensure the quality of data and metadata and the consistency between those two. Strong emphasis is put on unique identification of data and possible (re)use of data within the designated community. Further details on the concept of the data publication process and quality control can be found at the DOI publication process webpage.

The different quality levels of quality assessment correspond to the different project requirements. The DKRZ-LTA developed a common quality assessment with the Quality Maturity Matrix method.

Independent from specific project requirements, the DKRZ-LTA asks for a bundle of information that should be deposited. Required are the core metadata of CERA2. Information about methodology, quality checks and validation against independent data should be deposited at the latest in the DOI publication process step. A standard set of checks is applied to all submissions but further project specific checks can be defined and implemented.

Supporting Links and Documentation

- DKRZ-LTA services https://www.dkrz.de/daten-en/data-services/long_term_archiving
- WDS Data Policy <http://www.icsu-wds.org/organization/data-policy>
- DFG Rules of good scientific practice
http://www.dfg.de/en/research_funding/principles_dfg_funding/good_scientific_practice/index.html
- Rules of good scientific practice at MPG <http://www.mpg.de/197494/rulesScientificPractice.pdf>
- Data Submission Preparation Guide <http://cera-www.dkrz.de/docs/DataSubmissionPreparationGuide.pdf>
- Depositor Agreement <http://cera-www.dkrz.de/docs/DKRZ-LTA-DepositorAgreement.pdf>
- DOI publication process:
<https://www.dkrz.de/daten-en/data-services/Datapublication/quality-assurance-of-data>
- Quality Maturity Matrix: https://www.dkrz.de/pdfs/poster/Hoeck_et_al_EGU2015_maturitymatrices_15apr.pdf

Data Seal of Approval Board

W www.datasealofapproval.org

E info@datasealofapproval.org

- Core metadata CERA2: https://www.dkrz.de/daten-en/cera/data_model/tables-attributes-1

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

2. The data producer provides the data in formats recommended by the data repository.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Evidence:

As a thematic data center the number of formats is limited. DKRZ-LTA supports the data producers in the conversion of data into an accepted format e.g. NetCDF/CF and WMO GRIB.

Both formats are open standard.

For other formats the DKRZ-LTA asks the data producers to provide a description of the model including a section "how to read the data". If the data is binary we ask to convert the data into an ascii version (txt).

Formats are checked by the repository before data is archived. NetCDF/CF checkers are available for the data producers on the internet.

Supporting Links and Documentation

- DKRZ-LTA preservation and storage policy:
<http://cera-www.dkrz.de/docs/DKRZ-LTA-PreservationAndStoragePolicy.pdf>
- DKRZ-LTA format policy: <http://cera-www.dkrz.de/docs/DKRZ-LTA-Formats.pdf>
- Data Preparation Guide: <http://cera-www.dkrz.de/docs/DataSubmissionPreparationGuide.pdf>
- NetCDF <http://www.unidata.ucar.edu/netcdf>
- CF convention <http://www.cfconventions.org>
- WMO GRIB <http://www.wmo.int/pages/prog/www/WDM/Guides/Guide-binary-2.html>
- CF checker <http://cfconventions.org/compliance-checker.html>

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

Data Seal of Approval Board

W www.datasealofapproval.org

E info@datasealofapproval.org

3. The data producer provides the data together with the metadata requested by the data repository.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Data will only be archived by DKRZ-LTA if sufficient metadata is provided by the data producers. Data and metadata will be checked by DKRZ-LTA. For the creation of metadata DKRZ-LTA provides a graphical user interface for the data producers. Depending on the type of project and data this metadata will be supplemented by metadata automatically generated from the data (which can be on the level of files) and external repositories, e.g. CIM model documents or citation details for scientific publications via CrossRef. Review of metadata is part of our data publication process. The metadata are stored in the CERA database.

The metadata consists of descriptive, structural and administrative blocks. The quality control of metadata is covered by the answer to guideline 8. The metadata elements are derived from NASA DIF, ISO 19115, DataCite and CF netcdf standards.

Relevant metadata items for supporting resource discovery are project, experiment, entry and topic names, keywords, codelists and topic hierarchy.

Supporting Links and Documentation

- GUI for metadata creation http://cera-www.dkrz.de/LTA_metadata
- Data publication process <http://www.dkrz.de/daten-en/Datapublication>
- Technical description of CERA http://www.dkrz.de/daten-en/cera/data_model?set_language=en
- NASA DIF <http://gcmd.gsfc.nasa.gov/add/difguide/index.html>
- ISO 19115 http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020
- DataCite <http://schema.datacite.org/meta/kernel-3/>
- CF standard <http://cfconventions.org/>
- CERA GUI <http://cera-www.dkrz.de/WDCC/ui/Index.jsp>
- CIM <https://verc.enes.org/data/enes-model-data/cim-metadata-standard>

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

Data Seal of Approval Board

W www.datasealofapproval.org

E info@datasealofapproval.org

4. The data repository has an explicit mission in the area of digital archiving and promulgates it.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

The archive (DKRZ - WDC Climate) as run by the DKRZ is a regular member of the ICSU World Data System and has been a member of the former ICSU World Data Centres since 2003. The mission of the DKRZ - WDC Climate is to provide central support for the climate modelling research community. Data for and from climate research are collected, stored and disseminated. The WDCC is restricted to model data products. No raw data storage of data, for example, from satellites or climate models is planned in the WDCC itself.

The DKRZ-LTA is hosted by the Data Management department of the DKRZ which is funded by the German Government and by the four shareholders Max-Planck-Society, University of Hamburg, Helmholtz-Zentrum Geesthacht and the Alfred-Wegener-Institut. The DKRZ as host of DKRZ-LTA has robust funding commitments. The DKRZ-LTA is a thematically oriented archive. A continuous adjustment of this integration is established. Fundamental components of this includes the execution of consortial runs (e.g. CMIP) and long-term archiving (e.g. IPCC-DDC).

The DKRZ actively participates in national, European and global projects and initiatives like EUDAT and RDA, which contribute to integrated data infrastructures in climate research and beyond.

The DKRZ-LTA as WDCC is a recommended data repository included on the Nature repositories list which meets the Nature requirements for data access, preservation and stability. It is also listed on the registry of research data repositories re3data.org.

Supporting Links and Documentation

- ICSU <http://www.icsu.org>
- DKRZ - WDC Climate <https://www.dkrz.de/daten-en/wdcc>
- DKRZ - Partner for Climate Research <http://www.dkrz.de/about-en/aufgaben>
- RDA <https://rd-alliance.org/>
- EUDAT <http://www.eudat.eu/>
- Recommended Data Repositories <http://www.nature.com/sdata/data-policies/repositories>
- Registry of research data repositories: <http://www.re3data.org>

Reviewer Entry

Accept or send back to applicant for modification:

Data Seal of Approval Board

W www.datasealofapproval.org

E info@datasealofapproval.org

Accept

Comments:

5. The data repository uses due diligence to ensure compliance with legal regulations and contracts including, when applicable, regulations governing the protection of human subjects.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

The archive is run by the DKRZ which is a Limited Liability Company. Shareholders are the Max-Planck-Society, University of Hamburg, Helmholtz-Zentrum Geesthacht and the Alfred-Wegener-Institut. The organizational structure of DKRZ is described at <https://www.dkrz.de/about-en/Organisation>. Contracts with data producers are not required if they are a member of one of our shareholders. For other data producers regulations as layed out in the depositor agreement or individual contracts based on WDS regulations apply. The terms of use of the archive are published and apply to all clients of DKRZ-LTA.

Data with disclosure risks are not handled. Few submissions are under embargo to allow data producers first use of data.

Supporting Links and Documentation

- DKRZ structure <https://www.dkrz.de/about-en/Organisation>
- Terms of Use <http://cera-www.dkrz.de/WDCC/ui/docs/TermsOfUse.html>
- Depositor Agreement <http://cera-www.dkrz.de/docs/DKRZ-LTA-DepositorAgreement.pdf>
- WDC regulations <https://www.icsu-wds.org/files/wds-certification-summary-11-june-2012.pdf>
- WDS Data Policy https://www.icsu-wds.org/files/WDS_Data_Policy.pdf

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

Data Seal of Approval Board

W www.datasealofapproval.org

E info@datasealofapproval.org

6. The data repository applies documented processes and procedures for managing data storage.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

The DKRZ-LTA long term archive preservation and storage policy is published and is relevant for all DKRZ-LTA staff and users.

Supporting Links and Documentation

- DKRZ-LTA preservation and storage policy
<http://cera-www.dkrz.de/docs/DKRZ-LTA-PreservationAndStoragePolicy.pdf>

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

7. The data repository has a plan for long-term preservation of its digital assets.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

3. In progress: We are in the implementation phase.

Self-assessment statement:

Evidence:

DKRZ-LTA only accepts long-lasting open source formats for preservation, e.g. NetCDF/CF and GRIB, both being accepted standards in the climate community. Due to their property of being “Open Source” and a huge number of available open source analysis tools these formats will not become obsolete within the promised storage duration of 10 years.

NetCDF/CF brings in the CF convention which is ‘self-contained. So for example the CF Standard Names attempt to be as general and well-defined as possible, so the reader does not have to access outside sources to understand the terms.’

However, should such a very unlikely case of format obsolescence for the supported formats in the promised storage duration occur, DKRZ-LTA will provide either an on-the-fly format conversion into an actual and accepted format or will do the format conversion offline. This is possible although likely time consuming. The second would be typically done within the framework of media conversion to avoid unnecessary media mounts.

We have implemented a DOI data publication process and are assigning DOIs to quality assured objects in our repository with citation information. After DOI assignment the data are persistent and long-term preserved.

Bit stream preservation (including data redundancy) is already covered by our answer to guideline 6. Due to close collaboration with the relevant scientific communities we stay informed about developments of formats and implement these if necessary by active involvement of our group.

With regards to long-term usability we are aiming at grade 3 and 4 within the “Quality Maturity Matrix”.

Data Seal of Approval Board

W www.datasealofapproval.org

E info@datasealofapproval.org

Through a variety of active partnerships with both data providers and data users, the DKRZ is in a position to identify relevant changes of both technical and scientific nature and to translate them if necessary into requirements for migrations and implement them if required. In the history of DKRZ-LTA this can be proven e.g. by the migration of the HSM system from Unitree to HPSS or by the addition of another data format (netCDF).

Supporting Links and Documentation

- DKRZ-LTA supported formats <http://cera-www.dkrz.de/docs/DKRZ-LTA-Formats.pdf>
- Poster “Quality Maturity Matrix”
https://www.dkrz.de/pdfs/poster/Hoeck_et_al_EGU2015_maturitymatrices_15apr.pdf
- CF conventions: http://cfconventions.org/faq.html#when_started
- DKRZ-LTA preservation and storage policy
<http://cera-www.dkrz.de/docs/DKRZ-LTA-PreservationAndStoragePolicy.pdf>
- Data Publication Process <http://www.dkrz.de/daten-en/Datapublication>
- Search for DOI in CERA <http://cera-www.dkrz.de/WDCC/ui/FindDoiPublications.jsp>

Data Seal of Approval Board

W www.datasealofapproval.org

E info@datasealofapproval.org

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

8. Archiving takes place according to explicit work flows across the data life cycle.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

The DKRZ LTA workflows across the data life cycle are published.

In the following we address topics as listed in “Guidelines v2”

Selection process and type of data: The DKRZ-LTA is a thematically oriented archive. The focus is on output from numerical climate models. The archival process is integrated into and adapted to the workflow of the scientists. Archiving could take place at varying phases of the workflow. Either at a very early stage in order to make data reusable and properly citable or at a very late stage where the main motivation is the application of the Rules of Good Scientific Practice.

A continuous adjustment of this integration is established.

The DKRZ actively participates in national, European and global projects. Fundamental components of this includes the execution of consortial runs (e.g. CMIP) and long-term archiving (e.g. IPCC-DDC).

Approach towards data that do not fall within the mission: The WDCC is member of KomFor. KomFor supplies information where data can be submitted. Data repositories with different scope are members of KomFor. Its services are connected with existing ones provided by:

- Re3data
- World Data System

In cooperation with partners from these networks we support data producers to find the most appropriate repository.

The repository team maintains a dialogue with the researchers who want to deposit data. The information about the “Terms of Use” and the archiving workflow is part of this negotiation.

Guarding privacy of subjects: see answer to guideline 5.

Decision-making process for archival data transformation:

Data transformation within the archive only takes place in close cooperation with data producers (as long as they are available) or in cooperation with the relevant scientific community.

Skills of employees: The team has expertise in the areas of data management, data archiving & preservation, data formats, controlled vocabularies, quality assessment, data usage, data publication and registration. LTA staff strives to improve its knowledge in these areas by close cooperation in scientific and technical projects, training and participation in a technology watch team. Members of DKRZ-LTA actively participate in working and interest groups of RDA and WDS.

Supporting Links and Documentation

- DKRZ LTA Workflows across data life cycle
 - o https://www.dkrz.de/pdfs/poster/SC11_P1_DataLifeCycle_klein.pdf
 - o https://www.dkrz.de/pdfs/poster/DataServices_ENG.pdf
- DKRZ-LTA preservation and storage policy
<http://cera-www.dkrz.de/docs/DKRZ-LTA-PreservationAndStoragePolicy.pdf>

Data Seal of Approval Board

W www.datasealofapproval.org

E info@datasealofapproval.org

- Data Preparation Guide: <http://cera-www.dkrz.de/docs/DataSubmissionPreparationGuide.pdf>
- Rules of Good Scientific Practice <http://www.mpg.de/197494/rulesScientificPractice.pdf>
- List of projects https://www.dkrz.de/daten-en/wdcc/projects_cooperations
- KomFor <http://www.komfor.net/>
- KomFor submission <http://www.komfor.net/submit-data.html>
- Re3data <http://service.re3data.org/search/>
- WDS members <http://www.icsu-wds.org/community/membership>
- Terms of Use (<http://terms.wdc-climate.de>)
- Climate Model Comparison Projects (CMIP) <http://cmip-pcmdi.llnl.gov/>
- Data Distribution Centre of the Intergovernmental Panel on Climate Change (IPCC-DDC)
<http://www.ipcc-data.org/>
- RDA <https://rd-alliance.org/>

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

9. The data repository assumes responsibility from the data producers for access and availability of the digital objects.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Evidence:

All archived data objects are available online. The data producers specify the access and usage permissions, the DKRZ-LTA enforce them accordingly. Contracts with data producers are not required if they are a member of one of our shareholders, for other data producers there will be individual depositor agreements based on regulations as laid out in the depositor agreement

As a regular member of WDS and thereby accepting its regulations, WDCC at DKRZ takes over responsibilities according to these rules.

The DKRZ-LTA enforces the Creative Commons License as far as those conditions are not in any way modified.

Precautions against data loss are described in the DKRZ-LTA preservation and storage policy.

Established arrangements and precautions, such as redundant off-site data storage at Garching / Munich, make it possible in the event of a crisis, to re-enable access to the data within reasonable time.

The disclaimer of DKRZ-LTA, as written in the “Depositor Agreement”, states that DKRZ-LTA is not liable for the loss or damage of data, although every care is taken to preserve the objects.

Supporting Links and Documentation

- Terms Of Use <http://cera-www.dkrz.de/WDCC/ui/docs/TermsOfUse.html>
- Creative Commons License <http://creativecommons.org/licenses/by-nc-sa/2.0/de/deed.en>
- DKRZ-LTA preservation and storage policy
<http://cera-www.dkrz.de/docs/DKRZ-LTA-PreservationAndStoragePolicy.pdf>
- Depositor Agreement <http://cera-www.dkrz.de/docs/DKRZ-LTA-DepositorAgreement.pdf>
- WDS regulations <https://www.icsu-wds.org/files/wds-certification-summary-11-june-2012.pdf>

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

Data Seal of Approval Board

W www.datasealofapproval.org

E info@datasealofapproval.org

10. The data repository enables the users to discover and use the data and refer to them in a persistent way.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Data formats see guideline 7.

All data can be searched and downloaded free of charge through one of our web interfaces

<http://cera-www.dkrz.de/>. Browse and search are available. Metadata is also offered via OAI-PMH. DKRZ-LTA offers a DataCite DOI data publication for long-term archived data. Published data is granted permanent access by the assigned unique identifier DOI, data and citation metadata remain unchanged.

As a regular member of WDS and thereby accepting the regulations DKRZ - WDC Climate takes over responsibilities according to these rules.

Supporting Links and Documentation

- OAI-PMH <http://c3grid1.dkrz.de:8080/oai/oaisearch.do>
- DOI publications <http://cera-www.dkrz.de/WDCC/ui/FindDoiPublications.jsp>
- WDS regulations (e.g. rule 1.5 and 3.4)
<https://www.icsu-wds.org/files/wds-certification-summary-11-june-2012.pdf>

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

11. The data repository ensures the integrity of the digital objects and the metadata.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Guidelines regarding integrity of data objects and metadata are covered in the DKRZ-LTA Preservation and Storage policy.

Checksums md5 are used for the data objects.

The quality assurance of the integrity of the data and metadata is part of the publication process and documented in the CERA database. Changes to data objects are not allowed, new versions of data objects are possible but they are treated as independent objects. Old versions are kept in the archive. This also applies to the corresponding metadata entries.

Supporting Links and Documentation

- DKRZ-LTA Preservation and Storage Policy:

<http://cera-www.dkrz.de/docs/DKRZ-LTA-PreservationAndStoragePolicy.pdf>

- Publication Process:

<http://www.dkrz.de/daten-en/data-services/Datapublication/konzept-datacite-tib-metadatakernel>

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

12. The data repository ensures the authenticity of the digital objects and the metadata.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

3. In progress: We are in the implementation phase.

Self-assessment statement:

Data delivery procedures can be divided into two cases:

1. Data is delivered to us from clients who are known to us personally or are designated by someone whom we know personally.
2. Data from projects with their own data infrastructure (like ESGF/CMIP5) in which DKRZ-LTA participates are long-term archived according to specific workflows. Those are agreed on by the project and DKRZ-LTA. The authenticity underlies the project regulations.

Once the data has been delivered to us, changes to it are not allowed anymore. If changes are required, a new version will be put into the system while keeping the old one. The usage of checksums is already described in our answer to guideline 11. Archiving of data is only allowed if metadata is provided. The qualification of metadata follows the standard as described in <https://www.dkrz.de/daten-en/Datapublication/quality-assurance-on-data>, Quality Maturity Matrix level 3 and level 4 are considered to be sufficient. If needed, links to external resources (data and/or metadata) are kept as well.

The review of the metadata is part of the DOI data publication process. The assistant system Atarrabi provides a key part of the review process by supporting the authors' and reviewers' final quality assurance of metadata. Data authors and reviewers are required to register and login for the Atarrabi review process. After DOI assignment the complete data is accessible by the registered DOI.

Supporting Links and Documentation

- Publication Process

<http://www.dkrz.de/daten-en/data-services/Datapublication/konzept-datacite-tib-metadatakernel>

- ESGF <http://esgf.llnl.gov/>
- CMIP5 <http://cmip-pcmdi.llnl.gov/cmip5/index.html?submenuheader=0>
- Metadata qualification <https://www.dkrz.de/daten-en/Datapublication/quality-assurance-on-data>
- Atarrabi <http://cera-www.dkrz.de/atarrabi2>
- "Quality Maturity Matrix"

https://www.dkrz.de/pdfs/poster/Hoeck_et_al_EGU2015_maturitymatrices_15apr.pdf

Reviewer Entry

Accept or send back to applicant for modification:

Data Seal of Approval Board

W www.datasealofapproval.org

E info@datasealofapproval.org

Accept

Comments:

13. The technical infrastructure explicitly supports the tasks and functions described in internationally accepted archival standards like OAIS.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

Applicant Entry

Statement of Compliance:

3. In progress: We are in the implementation phase.

Self-assessment statement:

The WDCC at DKRZ follows the guidelines given in the OAIS reference model across the whole workflow. There is an exception with regards to preservation planning where plans are adjusted to administrative, infrastructure and project requirements.

Supporting Links and Documentation

- Mapping of DKRZ-LTA components to OAIS:

http://cera-www.dkrz.de/docs/OAIS_DKRZLTA_mapping_20151012.pdf

- Workflow graphic https://www.dkrz.de/daten-en/data-services/long_term_archiving/LZAFlowchart_v2_en.jpg

- Published workflows are:

Quadt, F., Düsterhus, A., Höck, H., Lautenschlager, M., Hense, A.V., Hense, A.N. & Dames, M. (2012). Atarrabi - A Workflow System

for the Publication of Environmental Data. Data Science Journal, 11, 89-109

<http://dx.doi.org/10.2481/dsj.012-027>

Stockhause, M., Höck, H., Toussaint, F., Lautenschlager, M. (2012). Quality assessment concept of the World

Data Center for Climate and its application to CMIP5 data. Geosci. Model Dev., 5, 1023-1032, 2012,

[doi:10.5194/gmd-5-1023-2012](https://doi.org/10.5194/gmd-5-1023-2012)

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

14. The data consumer complies with access regulations set by the data repository.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Data consumers are required to register and login for data download. It is not necessary to register for metadata search. By using our system users accept the terms of use:

‘All data and pages available from DKRZ-LTA are licensed under a Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/2.0/de/deed.en>) as far as those conditions are not in any way modified by the following conditions or by any conditions specific to data or pages.’

For very few datasets special regulations for commercial use apply. In this case DKRZ-LTA acts as a mediator between the data producer and the data consumer.

DKRZ-LTA has no means to monitor how data is being used after it has been downloaded from our archive.

Information on use and/or access constraints are always provided within the metadata.

Supporting Links and Documentation

- Terms of use <http://terms.wdc-climate.de>
- Creative Commons <http://creativecommons.org/licenses/by-nc-sa/2.0/de/deed.en>

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

15. The data consumer conforms to and agrees with any codes of conduct that are generally accepted in the relevant sector for the exchange and proper use of knowledge and information.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

All data consumers of DKRZ-LTA must abide by the DKRZ-LTA ‘Terms of use’ which are in line with accepted data policies. Articles, papers, or written scientific works of any form, based in whole or in part on data supplied by DKRZ-LTA, shall contain an acknowledgment concerning the supplied data. An explicit citation specification is always provided by DKRZ-LTA. DKRZ-LTA promotes the usage of DOIs. DKRZ-LTA was one of the originators of the DOI initiative.

DKRZ-LTA is acting as an archive only, it does not carry out any research by itself. Due to the nature of research carried out by the data producers, there is no linkage to any sensitive personal information in the data and metadata. Data consumers accept the terms of use when using our system. In the rare cases where special usage or access constraints apply (e.g. with regards to data transfer to third parties) consumers are asked to explicitly accept these special conditions.

DKRZ-LTA has no means to monitor how data is being used after it has been downloaded from our archive.

However, DKRZ-LTA would take appropriate measures as soon as we gain knowledge of such fact (which has not been the case so far).

Supporting Links and Documentation

- Terms of use <http://terms.wdc-climate.de>
- J. Brase, M. Lautenschlager and I. Sens (2015): The Tenth Anniversary of Assigning DOI Names to Scientific Data and a Five Year History of DataCite. D-Lib Magazine Vol. 21, No. 1-2, doi:10.1045/january2015-brase

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

Data Seal of Approval Board

W www.datasealofapproval.org

E info@datasealofapproval.org

16. The data consumer respects the applicable licences of the data repository regarding the use of the data.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Self-assessment statement:

Applicable licenses regarding the use of research data are by default addressed in our terms of use and if necessary in the metadata of each data object.

All data and pages available from DKRZ-LTA are licensed under a Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/2.0/de/deed.en>) as far as those conditions are not in any way modified.

Whether data consumers comply with these licenses cannot be monitored by us effectively. DKRZ-LTA would take appropriate measures as soon as we gain knowledge of such a fact (which has not been the case so far).

Supporting Links and Documentation

- Terms of use <http://terms.wdc-climate.de>
- Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/2.0/de/deed.en>)

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments: