

Terrestrial Environmental Observatories (TERENO)

Data Policy

This Data Policy was developed by the TERENO Coordination Team Data Management and has been approved by the TERENO Scientific Steering Committee. It outlines the principles adopted by the TERENO initiative regarding its data policy. The detailed implementation of these principles is further described in the TERENO Data Management Plan. The data rights statement in this document provides general directives and rules on how to share, re-use and cite all research data in order to

- Ease collaboration among the participants to the TERENO initiative
- Ensure timely submission of data for the use within the consortium
- Protect the researchers' intellectual property rights and rights to publish their results
- Provide rules for use of the data within the consortium and by third parties
- Build an homogeneous body of valuable scientific information on results observed within TERENO network
- Provide the broader scientific community with an easy access to integrated knowledge created by TERENO
- Enhance access of TERENO to data derived from the use of TERENO data

The detailed implementation of these principles is further described in the TERENO Data Management Plan.

Document version

V 2.0

Document Date:

2014-06-25

Authors:

- Ralf Kunkel, Research Centre Jülich, 52425 Jülich, Germany, E-Mail: r.kunkel@fz-juelich.de
- Martin Abbrent, Helmholtz-Centre for Environmental Research – UFZ, 04318 Leipzig, Germany, E-Mail: martin.abbrent@ufz.de
- Heye Bogena, Research Centre Jülich, 52425 Jülich, Germany, E-Mail: h.bogena@fz-juelich.de
- Anusuriya Devaraju, Research Centre Jülich, 52425 Jülich, Germany, E-Mail: a.devaraju@fz-juelich.de
- Jan Friesen, Helmholtz-Centre for Environmental Research – UFZ, 04318 Leipzig, Germany, E-Mail: janfriesen@ufz.de
- Rainer Gasche, Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research – Atmospheric Environmental Research (IMK-IFU), 82467 Garmisch-Partenkirchen, Germany, E-Mail: rainer.gasche@kit.edu
- Jens Klump, former GFZ German Research Centre for Geosciences, 14473 Potsdam, Germany, now CSIRO ARRC, 26 Dick Perry Avenue, Kensington, WA 6151, Australia, E-Mail: Jens.Klump@csiro.au
- Frank Neidl, Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research – Atmospheric Environmental Research (IMK-IFU), 82467 Garmisch-Partenkirchen, Germany, E-Mail: frank.neidl@kit.edu
- Karsten Rink, Helmholtz-Centre for Environmental Research – UFZ, 04318 Leipzig, Germany, E-Mail: karsten.rink@ufz.de
- Thomas Schnicke, Helmholtz-Centre for Environmental Research – UFZ, 04318 Leipzig, Germany, E-Mail: thomas.schnicke@ufz.de
- Matthias Schroeder, GFZ German Research Centre for Geosciences, 14473 Potsdam, Germany, E-Mail: matthias.schroeder@gfz-potsdam.de
- Jürgen Sorg, FZJ, Research Centre Jülich, 52425 Jülich, Germany, E-Mail: j.sorg@fz-juelich.de
- Vivien Stender, GFZ German Research Centre for Geosciences, 14473 Potsdam, Germany, E-Mail: vivien.stender@gfz-potsdam.de
- Ute Wollschläger, Helmholtz-Centre for Environmental Research – UFZ, 04318 Leipzig, Germany, E-Mail: ute.wollschlaeger@ufz.de
- Steffen Zacharias, Helmholtz-Centre for Environmental Research – UFZ, 04318 Leipzig, Germany, E-Mail: steffen.zacharias@ufz.de

Document history:

- | | |
|------------|--|
| 2012-10-31 | First version 1.0, approved by the TERENO Scientific Steering Committee |
| 2013-12-20 | First draft of refined version according to the negotiations at the TERENO Coordination Team Data Management meeting on 2013-11-26 |
| 2014-05-28 | Agreement of document by the TERNEO scientific steering committee meeting |
| 2014-06-25 | Merging comments from law departments of FZJ, UFZ and GFZ into the document |

Contents

1 Definition of terms 4

2 Introduction..... 6

3 Scope and Objectives 6

4 Provision of Data 7

5 Data Sharing..... 8

5.1 Access to Data 8

5.2 Reuse of Data outside of TERENO 9

5.3 Terms of Use 9

5.4 Citation and Acknowledgment 9

6 Data Privacy 10

7 Adherence and Review 11

8 Statement of Agreement..... 11

1 Definition of terms

For the purposes of this document the following definitions have been adopted:

CB	The TERENO Coordination Board, which is a group of four coordinators from the research centres FZJ, UFZ, KIT and GFZ, manages the TERENO initiative. They will develop plans for implementing the TERENO activities and ensure that the plans are carried out. The Coordination Board will plan and facilitate integration and synthesis exercises, including workshops, meetings and conferences. Furthermore they supervise the capacity building, knowledge transfer and external communication in conjunction with the Scientific Steering Committee.
CTMC	Central TERENO Metadata Catalogue, which stores and publishes all metadata managed by TEODOOR
Data	Distinct units of information such as facts, numbers, letters, symbols, usually formatted in a specific way, stored on a computer device and suitable for processing by a computer
Data management	Development and execution of architectures, policies, practices and procedures to properly manage the full data lifecycle needs of an institutional unit or research project
Data provider	Individual or institution, which distributes data to other parties
Data rights management	Regulations of access to data by consideration of data ownership, intellectual property rights and other legal restrictions
Data user	An individual authorized to access and use data
DMP	Data Management Plan: formal document that outlines how project data are managed during research time and after project completion. The goal of a data management plan is to consider the many aspects of data management, metadata generation, data preservation and analysis before the project begins. This ensures that data are well-managed at present, and prepared for preservation in the future.
DOI	Digital Object Identifier: character string used to uniquely identify an object such as an electronic document. Metadata about the object is stored in association with the DOI name and this metadata may include a location where the object can be found. The DOI for a document is permanent, whereas its location and other metadata may change.

Institutional data manager	Responsible Person for the operational management and processing of data. Data managers are nominated for each institution contributing data to TEODOOR (see TERENO DMP).
Metadata	Data providing descriptive information on data to enable users to query valuable information to decide about the usability of the data for its own use and purposes
SSC	The TERENO Scientific Steering Committee consists of representatives from the involved research centres. The committee decides on the basic configuration of the observatories, including the selection of the instruments, the networking with external partners, and the integration of the infrastructure in national and international programs as well as the strategic planning future activities in the framework of TERENO.
TEODOOR	TERENO Data Online Repository: Spatial data infrastructure for the management and exchange of data from the TERENO project. It consists of individual data infrastructures operated from the institutions contributing to TERENO, a common data portal to access the data, metadata catalogues providing descriptive information about the data and standardized OGC services for the communication between the infrastructures and with the common data portal.
TEODOOR administrator	Group of persons managing and administering the TERENO Portal
TERENO	TERrestrial ENvironmental Observatories
TERENO Community	Group of institutions contributing to TERENO
TERENO portal	Web-portal of TERENO (http://www.tereno.net). This web front-end allows querying, visualizing and accessing data and metadata from the different observatories in a standardized way.
SDI	Spatial Data Infrastructure: data infrastructure implementing a framework of geographic data, metadata and tools that are interactively connected in order to use spatial data in an efficient and flexible way.
UUID	Universally Unique Identifier: identifier standard from the Open Software Foundation (OSF) in order to enable distributed systems to uniquely identify information without significant central coordination.

2 Introduction

The establishment of TERrestrial ENvironmental Observatories (TERENO) is an initiative funded by the large research infrastructure programme of the Helmholtz Association. The main goal of TERENO is to create observation platforms on the basis of an interdisciplinary and long-term aimed research program to investigate the consequences of global change for terrestrial ecosystems and the socioeconomic implications. Special attention is paid in bridging the different temporal and spatial scales by covering the investigated systems – land surface, biosphere, lower atmosphere and anthroposphere – hierarchical from the local to the regional scale and from direct observable time periods (years) to long time periods (centennials) derived from geo-archives. TERENO provides series of system variables (e.g. precipitation, runoff, groundwater level, soil moisture, water vapor and trace gases fluxes) for the analysis and prognosis of global change consequences using integrated model systems, which will be used to derive efficient prevention, mitigation and adaptation strategies.

3 Scope and Objectives

The TERENO Data Policy relates to the following kinds of data:

1. **Measurement Data** is provided by the institutions owning the instruments and comprises the following:
 - a. raw data collected with the use of TERENO instruments and yet to be processed
 - b. processed data is raw data that has undergone certain transformation procedures
 - c. quality controlled data is processed data that has passed quality assurance procedures such as routine estimation of timing and sensor calibration or visual inspection and removal of obvious errors.
2. **Derived Data** is data products generated from the Measurement Data by means of scientific and technical interpretation. Processed Data is provided by the institution employing the individual generating the Processed Data.
3. **External Data** is data obtained from third parties for use in TERENO related research.

The TERENO Data Policy addresses issues related to the provision, exchange, availability, maintenance and use of data produced during TERENO related research activities or acquired from third parties as a support to this research. It contains rules regarding the provision of data, use and access-rights within the TERENO data infrastructure (TEODOOR), published via the TERENO Data Portal.

In order to guarantee continuity of the research activities within TERENO, it is essential that data produced by one research group is made available as soon as possible to the other research groups or individuals, and respectively to external partners or members of the worldwide scientific community. Institutional Data Managers refer to the TERENO DMP when producing datasets to be exchanged.

It is also a concern to ensure protection of the rights of the data providers, their intellectual property and the personal privacy of data providers and data users on the level of all types of digital data. Because sharing and re-use of data is subject to legal regulations, data rights management is of crucial importance in the data management framework, and of high priority in each data management infrastructure.

In order to ensure that the objectives are reached, the TERENO partners and TERENO undertake the following:

- a. Each TERENO partner institution provides all available Measurement Data and Processed Data to TEODOOR and grants use licenses if necessary (see Section 4).
- b. TERENO, represented by the SSC, is authorized to make the data accessible by means of TEODOOR under the terms contained in the TERENO Data Policy, and to grant use licenses as needed to stakeholders for scientific, policy-advice and decision making purposes respectively, by putting the data in the public domain (see Section 5.1).
- c. Access to data provided by a partner and under the control of TERENO due to its storage in TEODOOR may be limited to individual users or group of users, or temporarily be suspended if necessary (see Section 4e and 5.1).

4 Provision of Data

In consideration of the requirements and procedures mentioned in section 3, the following rules for data provision shall apply:

- a) The data providers contribute to the provision, sharing and exchange of data with good quality. They will settle conflicts in a friendly, mutually understanding manner to the benefit of all stakeholders. TERENO Scientific Steering Committee has the option to discuss these data provision issues with the individual partners.
- b) A description of all data, data products, documents, models, software code, which is of interest for in- and outside TERENO, will be created by the data owners or responsible individuals who have been assigned by the data providers and published by the CTMC.
- c) TERENO partners have to provide their data to TEODOOR as soon as it is quality controlled and prepared for exchange, and by following the standards defined in the DMP, but not later than three months after its acquisition.
- d) Certain forms of data may be protected by copyright. According to the “German Copyright Act” (“Urhebergesetz”) the results of scientific work, such as written works, speeches and computer programs, illustrations of a scientific or technical nature such as drawings, plans, maps, sketches, tables and three-dimensional representations may be subject to copyright provided that they meet the two requirements of originality (an intellectual creation) and individuality (i.e. having a personal feature). In case data is subject to copyright, or other forms of intellectual property rights, a license is required for use of the data. The provider of TERENO-related data that is subject to copyright or other intellectual property rights grants a non-exclusive, irrevocable, cost-free license to the TERENO community to use the data for non-commercial and scientific purposes upon uploading the data to TEODOOR.
- e) Data providers may temporarily restrict access to data of the TERENO community in general or specific groups in and outside of TERENO (“closed data sharing network”) for legitimate reasons. These restrictions and its durations have to be communicated to the CB and the data managers in charge.

- f) Individuals, who are working within TERENO and are funded by one of the institutions contributing to TERENO have priority rights to publish his/her data first in case the study topic is directly related to this particular instrumentation and it is approved by the SSC. However, the data will be uploaded to TEODOOR to make it available at least for the TERENO community. If this priority right to publish is not used, then after an adequate time span of (24 months) the funding TERENO institution has the exclusive right to use the data. The SSC may decide on an extension of this time span upon request.
- g) If the study is a Ph.D. study conducted at a partner university, the rules of the University will be applied.
- h) Except if otherwise prohibited by mandatory subsidy regulations, data of separate projects connected to TERENO are to be made available within three months after the project ends, or according to other allocated submission deadlines. It will be the Institutional Data Manager's responsibility to ensure that submission deadlines are observed and to chase possibly delayed or missing data.
- i) For data provided by external parties, individual agreements will be negotiated between TERENO and responsible persons from the data providing and using institutions. The external data providers are responsible for ensuring that the use of the data they supplied to the TERENO data infrastructure does not infringe any rights of third parties.

5 Data Sharing

5.1 Access to Data

The entry point to TERENO data is provided by the TERENO Online Data Portal, which provides a metadata catalogue application and other data query and data access tools. The responsible institutions for individual observatories may set up additional data portals, as long as they make also accessible the data provided by these portals through the TERENO Data Online Portal (see DMP for further details).

A fine-grained user and user group structure regulate the dissemination of data and data products by reflecting the right management demands as delineated in section 4e). Any user account and group membership application must be approved by the CB.

In addition, the following rules for TERENO data apply:

- a. Data subject to copyright or other intellectual property rights according to 4c) and 4d) and metadata hosted by TEODOOR will be made accessible via the TERENO Data Portal for scientific purposes as soon as quality control of the data has been performed (see DMP for further details). A user registration for data download is generally not required, however, the user must provide the data usage purpose, and his name, email address and institution. The TERENO Data Policy will be attached to the data as a link, and needs to be accepted by the user when downloading the data.
- b. Access to restricted data according to 4e) requires a registration. The users identified as a member of a closed data sharing network needs to be a registered user to access the data. The TERENO Data policy will be attached to the data as a link, and needs to be accepted by the user when downloading the data. When the re-

striction time period ends, the data will be made available to the public according to section 5.1. a.)

- c. For data provided by external partners (e.g. Met- or Hydro-Services) access restrictions, terms of use and data use licenses will be applied as agreed upon within bilateral agreements.

5.2 Reuse of Data outside of TERENO

Data access may be granted by TERENO to other organizations and programs working on comparable scientific agendas, such as the ISMN project and the GEOSS program, national political bodies and NGOs, upon request and approval by the CB. Institutional Data managers will have to further develop data management strategies fitting up-coming requirements in research and decision making, and considering new conceptual and technological developments relevant for the management of data (as far as this falls within the responsibility of the data managers).

5.3 Terms of Use

In information systems, however, where usable items under intellectual property can be downloaded, a license is included in the metadata and/or in the order process, so that the user accepts the license by downloading it , for instance by clicking a special checkbox. In TERENO the following terms of use/licensing rules apply:

- a. Users may copy, distribute, display and perform the work for non-commercial and scientific purposes provided that they follow the rules regarding citations and acknowledgements defined below.
- b. Users may make derivative works based on TERENO data. They agree to make their results available to TEODOOR as far as they are able to do so. In case they cannot provide their results (e.g. due to intellectual property rights of third parties such as publishers) they shall inform the CB about publications relating to TERENO data.
- c. The terms of use will be cited as abstracts within the metadata of the data.
- d. In the metadata catalogue the TERENO Data Policy will be included as a link in the order process for data. The user accepts the terms of the TERENO data policy by downloading the data.

5.4 Citation and Acknowledgment

When data that was made available by TEODOOR and obtained from the TERENO Data Portal is used within a publication, the source of the data is to be cited by including a persistent URL (either a DOI assigned to the data set or an UUID assigned to metadata as provided by the TERENO Data Portal), which addresses the data's metadata in the metadata catalogue. The link can either be included as a footnote or in the references/bibliography. If data is obtained from other colleagues within TERENO, these other colleagues must be cited properly according to the possible cases:

- a) If the data has been produced in actual fieldwork by one person within TERENO, this person should be acknowledged, cited or even become a co-author.

- b) If the data is available from the TERENO Data Portal, a citation of the data set's DOI or the URL to the metadata set in the Metadata Catalogue is sufficient, because the metadata contains the details from the data creator and contributors.
- c) If an institution has provided the data set, this institution must be cited.
- d) If considerable efforts by two or more persons have created one single data set, joint publishing (i.e. the naming of both as authors in a publication) is recommended and appropriate.
- e) As a rule, in joint publications, the person that has written the text of the publication (or over 50% of it) has the right to be the first author. All others should be listed as co-authors in case they have made a substantial contribution to the work being at the level of study preparation, execution, processing or interpretation. If the contribution of two authors is equal or disputed, the Scientific Steering Committee decides.
- f) If gathering or pre-processing of data was substantial supported by others (e.g. student assistants), they should be acknowledged in an acknowledgement.
- g) When a person intends to publish data, research results and information from the TERENO initiative especially in an international context (journal, book, conference, etc.), a reference to the donors and the executing institutes must be made. For example, the acknowledgement in the publication may state one of:

This work was supported by funding from the Helmholtz Association and the Federal Ministry of Education and Research (BMBF) in the framework of TERENO (Terrestrial Environmental Observatories).

6 Data Privacy

Personal data, which is data characterizing an individual human being (e.g. data with private/personal material or family details which lead back to the individuals), shall be kept confidential in accordance with the Federal Data Protection Act ("Bundesdatenschutzgesetz") and will not be forwarded to third parties. It may only be used and processed in a research framework, if the data is depersonalized. If this is not possible or meaningful for the particular research question, the data must be stored by separating characteristics of personal or material circumstances, which could be attributed to an identified or identifiable person. Personal data can only be used in publications with the written consent of the persons concerned.

Data downloads will be reported by E-Mail to the responsible person of the particular institution and/or to the Institutional Data Manager and/or to the TEODOOR administrators. The E-Mails may contain all personal information provided by the user by accepting the license. Above that, all personal information will be kept as private and confidential. Personal data may be used for statistical purposes. Personal details will not be transferred to third parties.

7 Adherence and Review

The degree of adherence to the plan is observed in the daily work of the Institutional Data Managers and checked by the CT Data Management. If necessary, this plan will be revised and approved by the SSC.

8 Statement of Agreement

This Data Policy has been agreed upon by the TERENO Scientific Steering Committee on 2014-08-28 and will be signed by each TERENO partner.