

To document, or not to document?

How to write a readme file and what is metadata.

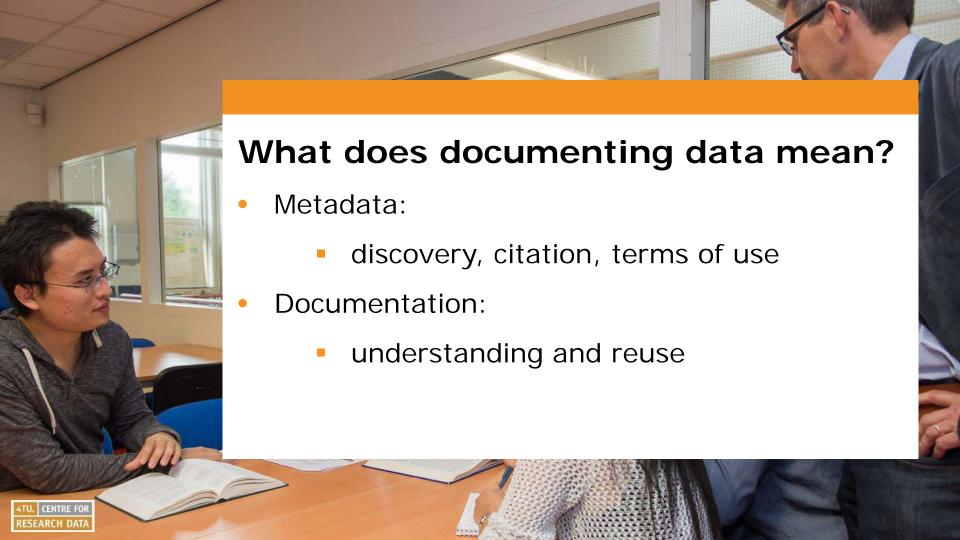




Why documenting data?

- Content
 - Data files
- Context
 - Data collection
 - Data processing





Dataset Pan-European data sets of annual number of situations conducive to thunderstorm formation under present and future climate

▶ ▶ ▶ ▶ Link/cite as https://doi.org/10.4121/uuid:425f1f99-dee7-4a32-b5c6-8ad820af5f69 | full citation

▼go to DATA section▼		
title	?	Pan-European data sets of annual number of situations conducive to thunderstorm formation under present and future climate
creator	?	Pucik, T. (Tomas)
creator	?	Groenemeijer, P. (Pieter)
contributor	?	European Severe Storms Laboratory
contributor	?	info:eu-repo/grantAgreement/EC/FP7/608166
date accepted	?	2016-05-02
date created	?	2014 through 2016
date published	?	2016
description	?	NetCDF files containing gridded annual number of situations conducive for thunderstorm formation for present day and the future climate. The fields are multi model means of 15 regional climate model simulations (CORDEX).
language	?	en en
publisher	?	European Severe Storms Laboratory
subject	?	Climate ◊ Climate change ◊ Climate hazards ◊ CORDEX (Coordinated Regional Climate Downscaling, Experiment) ◊ Future climate ◊ Meteorology ◊ Risk analysis ◊ Thunderstorm ◊ Thunderstorms probability
▲ in collection	?	RAIN: Pan-European gridded data sets of extreme weather probability of occurrence under present and future climate
spatial coverage	?	<u>Europe</u>
time coverage	?	<u>years 1971-2100</u>
licence	?	☑ General terms of use

DATA

readme_thunderstorms.pdf (application/pdf)

? Direct access through OPeNDAP

Total 643 KiB

[Move over links for options and info]

tstm 1971 2000.nc

<u>tstm_rcp45_2021_2050.nc</u>

tstm_rcp45_2071_2100.nc

tstm_rcp85_2021_2050.nc

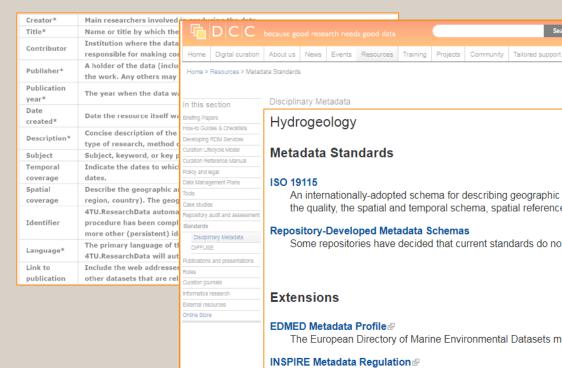
tstm_rcp85_2071_2100.nc



Access on our OPeNDAP server: xml, html

Creator*	Main researchers involved in producing the data.
Title*	Name or title by which the dataset is known.
Contributor	Institution where the data was created or collected. A person or organization responsible for making contributions to the dataset.
Publisher*	A holder of the data (including archives appropriate) or institution which submitted the work. Any others may be listed as contributors.
Publication year*	The year when the data was or will be made publicly available.
Date created*	Date the resource itself was put together; this could be a date range or a single date.
Description*	Concise description of the contents of the dataset. Describe the research objective, type of research, method of data collection and type of data.
Subject	Subject, keyword, or key phrase describing the resource.
Temporal coverage	Indicate the dates to which the data refer. Enter the year, or beginning and ending dates.
Spatial coverage	Describe the geographic area to which the data refer (e.g. municipality, town/city, region, country). The geographic coordinates of the area may be included, if desired.
Identifier	4TU.ResearchData automatically assigns a DOI to a dataset once the entire deposit procedure has been completed. In some cases, a dataset may be known by one or more other (persistent) identifiers.
Language*	The primary language of the resource. When no language is added, 4TU.ResearchData will automatically assign 'English'.
Link to publication	Include the web addresses or DOIs for any publication, important internal reports or other datasets that are related to your dataset.





An internationally-adopted schema for describing geographic information and services. It provides information about the identification, the extent, the quality, the spatial and temporal schema, spatial reference, and distribution of digital geographic data.

Some repositories have decided that current standards do not fit their metadata needs, and so have created their own requirements.

The European Directory of Marine Environmental Datasets metadata scheme, which is a profile of ISO 19115.

A profile of ISO 19115:2003, adopted in 2007 as the common metadata standard for the Infrastructure for Spatial Information in the European Community (INSPIRE). The other profiles of ISO 19115 in use in European Member States have been made compliant with INSPIRE.

Use Cases

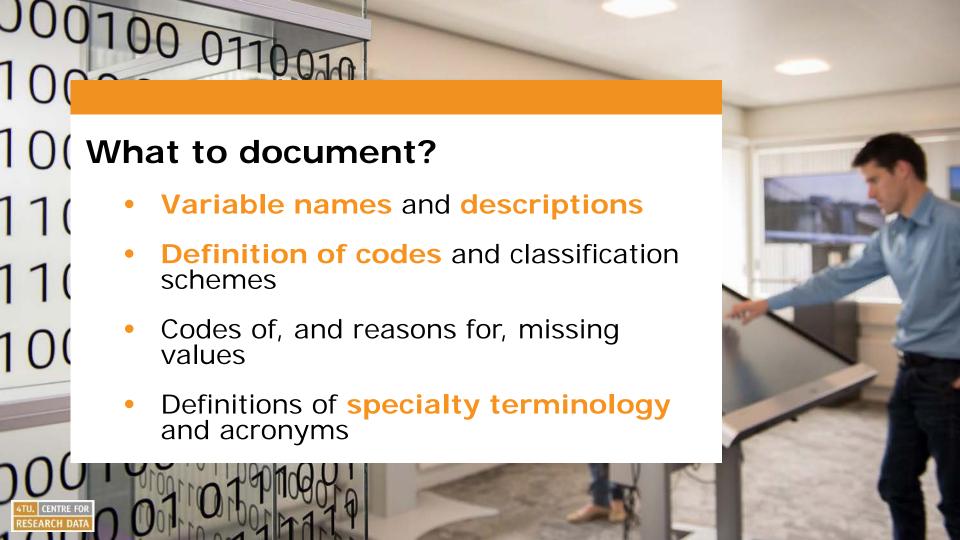
BAS PDC - British Antarctic Survey Polar Data Centre

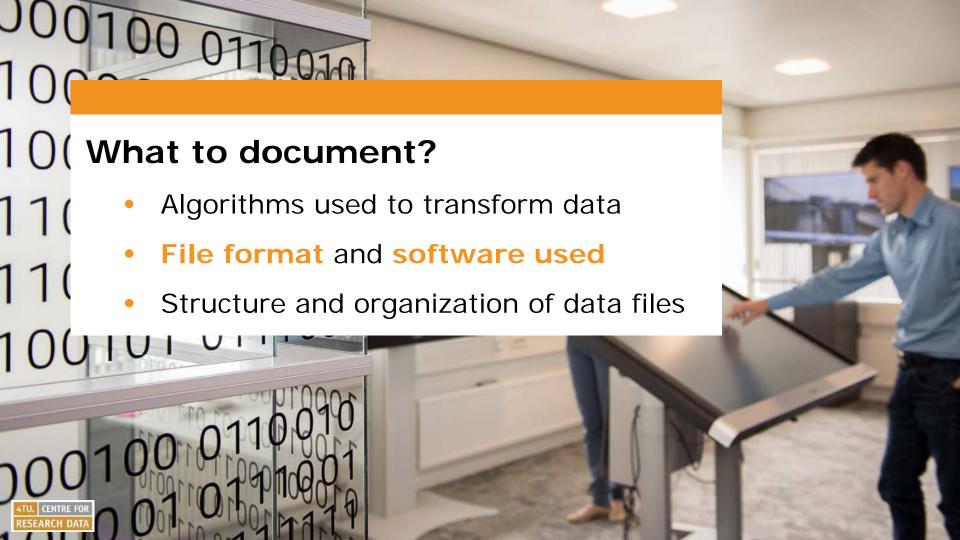
An organisation coordinating the management of data collected by UK-funded scientists in the polar regions, using an application profile that is harmonious with both ISO 19115 and DIF.

BGS NGDC - British Geological Survey National Geoscience Data Centre

The world's oldest national geological survey and the United Kingdom's premier centre for earth science information and expertise, with information access via an ISO 19115-compliant discovery metadata database.











Go to DANS (https://dans.knaw.nl/en/),

4TU.Centre for Research Data (https://data.4tu.nl/repository/)

or Registry of Research Data Repositories

(http://www.re3data.org/)

Search for a dataset which roughly fits in your field of research and contains a documentation or readme file.

Assess the metadata

Assess the documentation



Assess the metadata

Has been made use of a metadata scheme/standard? If yes, which one?

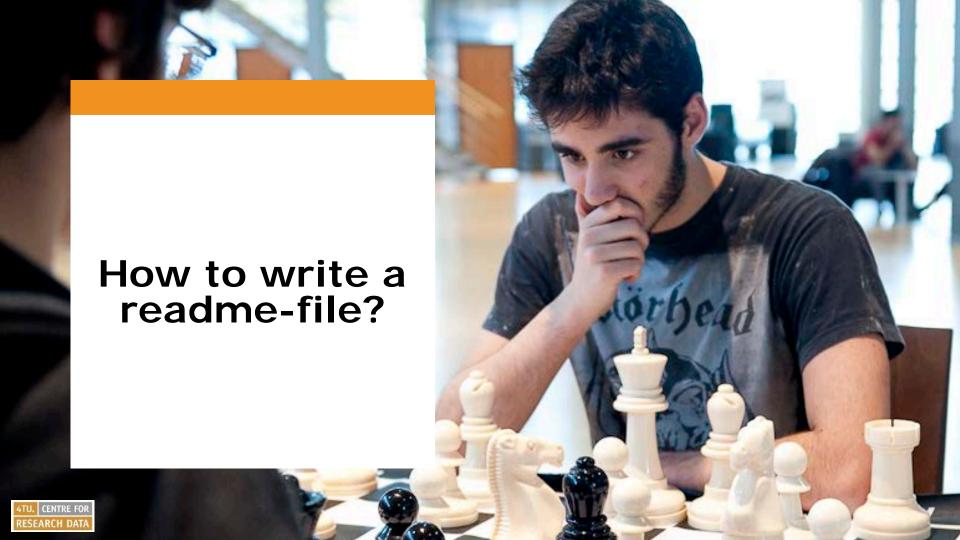
Are the metadata and values clear?

Is there a description of the dataset and is it clear?
If not, what information is missing or unclear?

Assess the documentation

Is documentation enough for understanding AND/OR reusing the data?
If not, what information is missing or unclear?

Evaluate your answers with one or two workshop participants.





Guidelines for creating a README file

A readme file provides information about a dataset and is intended to help ensure that the data can be correctly interpreted, by yourself at a later date or by others when sharing or publishing data.

A readme file must be submitted along with the dataset file(s). The outline below should be completed with information relevant to the submitted dataset.

Best practices

- Create one readme file for each dataset
- . Name the file README; not readme, read_me, ABOUT, etc.
- Write your readme document as a plain text file; save as README.txt or README.md when writing in <u>Markdown</u>. Or use README.pdf when text formatting is important for your file.

Recommended content

Recommended minimum content is in bold.

- 1. Introductory information
 - Title of the dataset
 - For each file or group of similar files, a short description of what data it contains
 - o Explain the file naming convention, if applicable
 - o Format of the file if not obvious from the file name
 - If the data set includes multiple files that relate to each other, the relationship between the files or a description of the file structure that holds them
 - o Contact information; in case users have questions regarding the data files

2. Methodological information

- Method description for collecting or generating the data, as well as the methods for processing data, if data other than raw data are being contributed
- Any instrument-specific information needed to understand or interpret the data
- Software (including version number) used to produce, prepare, render, compress, analyze and/or needed to read the dataset, if applicable
- Standards and calibration information, if appropriate

1. Introductory Information

Dataset title and short description

2. Methodological Information

About collecting and processing

3. Data Specific Information

Formats, codes, units, definitions

4. Sharing and Access Information

Use license or restrictions

Home Profiles Research Units Research Output Activities Prizes Press / Media

Overlimiting current properties at ion exchange membranes

Joeri De Valenca

AWARDING INSTITUTION

Supervisors/Apvisors

PLACE OF PUBLICATION

DATE OF AWARD

PRINT ISBNs

Soft matter, Fluidics and Interfaces, Faculty of Science and Technology

ORIGINAL LANGUAGE English

University of Twente

Lammertink, Rob G.H., Supervisor Wagterveld, R.M., Advisor 10 Mar 2017

Enschede Universiteit Twente

978-90-365-4314-9

Published - 10 Mar 2017

Fingerprint

O Ion Exchange Membranes

Keywords

METIS-321651 | IR-103679

Cite this

and Technology, Soft Mapprocessing for Sustainable Water Te

ging to the publication: Infl Review Fluids 2, 033701

otentiometry ◊ Concentrat nvection

tion on electrokinetic insta exchange membranes (th software

netic instability in o

ctet-stream, application/xtext/x-matlab, video/avi, vi∥from Lambert instruments

/agterveld, R.M. (Martijn); Wood, J.A. (Jeffery); Lammertink, R.G.H. ection on electrokinetic instability in overlimiting current conditions. 4121/uuid:b55184a0-930d-4b5c-940b-fd3038e05d83

> MATLAB files (for version R2013b): *.mat MATLAB workspaces that contain data MATLAB script for calculations and data *.fig MATLAB figures that display data

Various image formats are used, which are all well known high resolution images *.jpg compressed images compressed images *.png vector image *.avi movie files compressed movie files compressed further *.mp4

compressed data using open source 7-zip

Data compression compressed data using open source 7-zip software

FLIM data

Data compression

*.fli FLIM data in commercial software LI-FLIM

De Valenca, J. (2017). Overlimiting current properties at ion exchange membranes Enschede: Universiteit Twente DOI: 10.3990/1.9789036543149





W researchdata.4tu.nl

+31 (0)15 27 88 600

researchdata@4tu.nl



twitter.com/4TUResearchData



https://openworking.wordpress.com/

j.k.boehmer@tudelft.nl

Maarten van Bentum m.vanbentum@utwente.nl



