GFZ Data Services – A domain repository for geosciences data



Florian Ott , <u>Kirsten Elger</u>, Simone Frenzel GFZ German Research Centre for Geosciences, Potsdam, Germany









Data Publications – best practice for FAIR sharing data

Publication of datasets as individual publications (with assigned persistent Identifier, e.g., DOI) through domain repositories

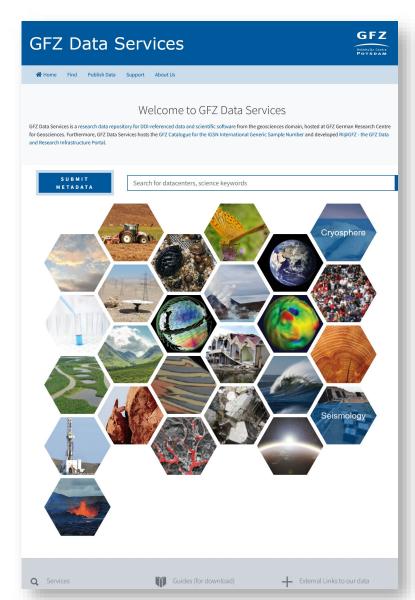
- **Findable:** integration of standardised machine readable metadata in external data portals (e.g. DataCite, B2Find, Google Dataset Search)
- **Accessible:** via DOI, persistent data storage and access guaranteed by the publisher (= data repository)
- **Documented:** with metadata for discovery and reuse \rightarrow curation
- Citable: DOI-referenced datasets are citable just as journal articles
 (> credit for researcher and institution)

GFZ Data Services: Research Data Repository

Profile

- Domain repository for the Geosciences since 2006
- DOIs for Data and software
- Data: real-time data streams, tables, maps, model data, ...
- Data curation by domain scientists
- Partner of:





https://dataservices.gfz-potsdam.de

FAIR data

- International metadata standards (human & machine readable)
- Controlled vocabularies for "rich" metadata
- PIDS ORCID

 Connecting Research and Researchers

 Crossref Funder Registry

 in prep
- Open Licences for data and software
- OAI-PMH interface
- schema.org → Google
 Dataset Search

GFZ Data Services: Profile

Focus:

1. curation of long-tail data

variable, difficult to standardize and curate. But represent large portion of total research data





GFZ Data Services: Profile

Focus:

POTSDAM

- 1. curation of long-tail data
- 2. DOI minting services for global monitoring networks/observatories in geodesy and geophysics and collaborative projects.





Different layouts for DOI Landing Pages

What do I need for a data publication?

- Data
- Metadata



Contextual Metadata

Definition of data labels



highly variable
between the
disciplines but key
information for
data reuse

GF.

TYPE: Température moyenne du sol (Degré celsius (°C)) / Average ground temperature (Degree Celsius (°C))

Helmholtz-Zentrum

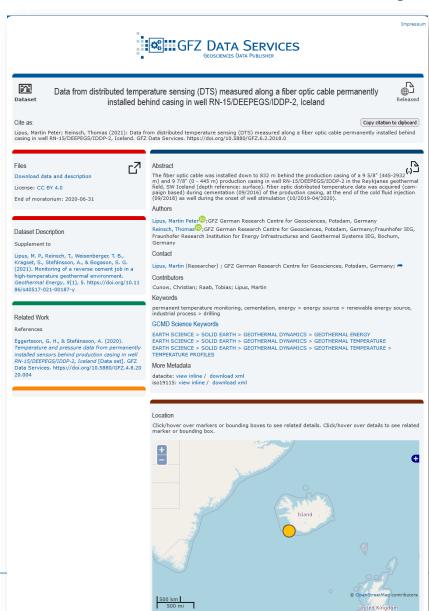
Metadata for data discovery: example DOI Landing Page

title citation

download data

key paper

related work



description/abstract

authors/ ORCID (iii)

keywords/controlled vocabularies

spatial coverage

Essential for data discovery, DOI registration, etc: international standards across all disciplines

Typical metdata standards for data discovery:

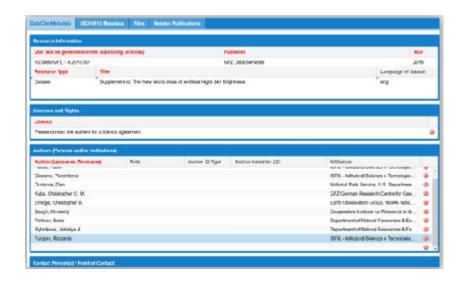
DataCite, ISO19115,

Dublin Core



Tools for data publications by GFZ Data Services

1. Discovery Metadata: via GFZ Metadata Editor

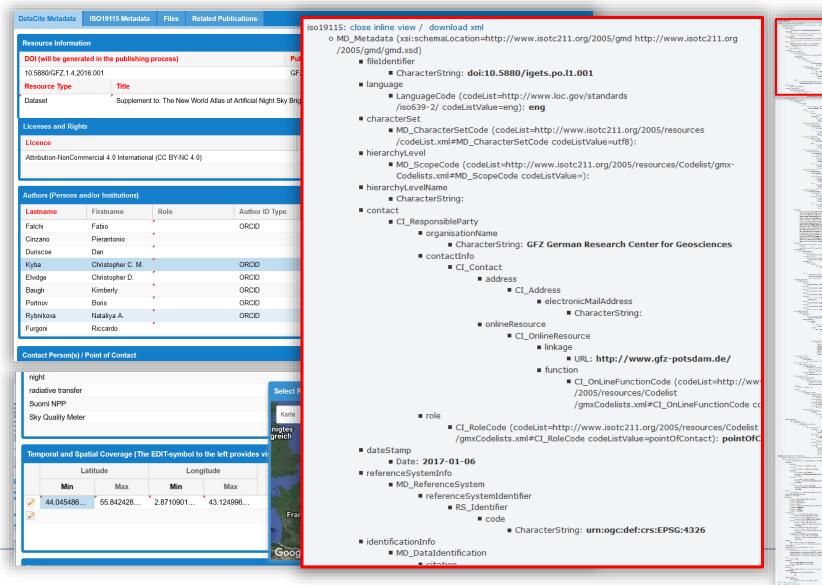




GFZ Metadata Editor (Java Script "translator")

Input: provided by researchers

Output: XML
(Extensible
Markup
Language):
Metadata
exchange format







GFZ Metadata Editor (Java Script "translator")

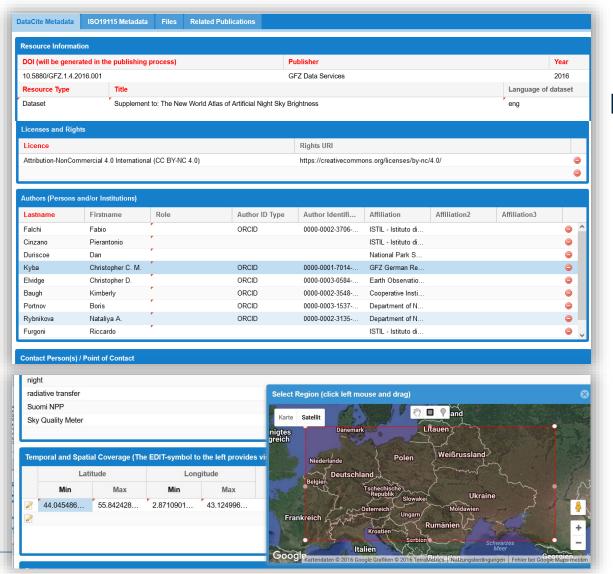
Input:

provided by researchers



"Special" Features:

- Interactive map
- Searchable vocabulary lists



Output:

- Standardised XML files (Datacite, ISO 19115, Dublin Core)
- DOI Landing Pages
- Data Catalogue

Standardised API













Metadata harvesting to other portals



Tools for data publications by GFZ Data Services

- 1. Discovery Metadata: via GFZ Metadata Editor
- Contextual Metadata: via Data Description Templates (or data reports)

Paleosol-derived data used for the reconstruction of environmental conditions during the Holocene in the upper part of the Kali Gandaki valley, Central Nepal (http://doi.org/10.5880/GFZ.4.6.2019.001)

Johanna Menges¹, Niels Hovius¹, Christoff Andermann¹, Michael Dietze¹, Charlie Swoboda¹, Kristen Cook¹, Basanta Adhikari², Andrea Vieth-Hillebrand¹, Stephane Bonnet³, Tony Reimann⁴, Andreas Koutsodendris³, Dirk Sachse¹

- 1. GFZ German Research Centre For Geosciences, Telegrafenberg, 14473 Potsdam, Germany
- 2. Department of Civil Engineering, Pulchowk Campus, Institute of Engineering, Tribhuvan University, Nepal
- 3. GET CNRS Univ Toulouse, UMR 5563, Toulouse, France
- Soil Geography and Landscape group & Netherlands Centre for Luminescence dating, Wageningen University, The Netherlands
- Heidelberg University Institute of Earth Sciences, Heidelberg, Germany



Data Description Templates

- Many users are unaware of what a data publication represents and what to include in description
- Increase the quality of metadata
- Reduces curation workload
- Uniform format aids comprehension
- Template via gfzpublic (https://gfzpublic.gfzpotsdam.de/pubman/item/item_5007103)

Paleosol-derived data used for the reconstruction of environmental conditions during the Holocene in the upper part of the Kali Gandaki valley, Central Nepal (http://doi.org/10.5880/GFZ.4.6.2019.001)

Johanna Menges¹, Niels Hovius¹, Christoff Andermann¹, Michael Dietze¹, Charlie Swoboda¹, Kristen Cook¹, Basanta Adhikari², Andrea Vieth-Hillebrand¹, Stephane Bonnet³, Tony Reimann⁴, Andreas Koutsodendris³, Dirk Sachse¹

- 1. GFZ German Research Centre For Geosciences, Telegrafenberg, 14473 Potsdam, Germany
- 2. Department of Civil Engineering, Pulchowk Campus, Institute of Engineering, Tribhuvan University, Nepal
- 3. GET CNRS Univ Toulouse, UMR 5563, Toulouse, France
- Soil Geography and Landscape group & Netherlands Centre for Luminescence dating, Wageningen University, The Netherlands
- 5. Heidelberg University Institute of Earth Sciences, Heidelberg, Germany

1. Licence

Creative Commons Attribution 4.0 International License (CC BY 4.0)



2. Citation

These data are freely available under the Creative Commons Attribution 4.0 International Licen 4.0).

When using the data please cite:

Menges, J.; Hovius, N.; Andermann, C.; Dietze, M.; Swoboda, C.; Cook, K.; Adhikari, B.; Vieth-Hilli A.; Bonnet, S.; Reimann, T.; K., Andreas; Sachse, D. (2019): Paleosol-derived data used for the reconstruction of Holocene environmental conditions during in the upper Kali Gandaki valley, Ce Nepal. GFZ Data Services. http://doi.org/10.5880/GFZ.4.6.2019.001

The data are supplementary to:

Menges, J., Hovius, N., Andermann, C., Dietze, M., Swoboda, C., Cook, K. L., ... Sachse, D. (2019). Late



- 2. Citation
- 3. Data Description
 - Sampling method
 - Analytical procedure
 - Data processing
- 4. File description
 - File inventory
 - File naming convention
 - Description of data tables





Tools for data publications by GFZ Data Services

- 1. Discovery Metadata: via GFZ Metadata Editor
- 2. Contextual Metadata: via Data Description
 Templates (or data reports)
- 3. Data Discovery and access via the Data Portal

https://dataservices.gfz-potsdam.de





Special Features: "Data in Review" Links

- Link: https://dataservices.gfzpotsdam.de/panmetaworks/review/9c5de649b6b3 0c588f9fecad56a1c71dd56d1fb4f68ada89b934000 2ff84abb7/
- Allows access to still unregistered data (for review purposes)
- DOI is reserved and citable
- Data can still be changed
- DOI registration when paper is accepted



of the paper once known.

est the utility of tourmaline as an indicator lled NW-SE zone. Mineralization comprises ed porphyry host rocks and disseminated overview of the geology, geochronology tourmaline breccias are given by Frikken ere comprises in-situ chemical analyses of

nicroprobe (EPMA) as well as in-situ boron-isotope analyses of tourmaline in the same samples by SIMS. Tourmaline was analysed in 12 samples including 8 from mineralized breccia bodies (Sur-Sur: 4, La Americana: 4), and 2 samples each from barren breccia and nearby granite-hosted tourmaline nodules in the Diamante area

(MS Excel)

Table 1 - List of all che SIMS B- isotope anal

analyses of tourma

Freiberg and Oo

Frederic Couffign

Make Data

AGU100 ADVANCING

the Chilean Commission for Scientific and Technological exchange Service (DAAD), We thank Thomas Seifert and ing expenses for field visits, sample preparation and h the help of Bernhard Schulz and Joachim Krause in analyses in Potsdam were done with the expert help of

jurmaline electron microprobe (MS Excel) Table 2 - Summary of

aline (MS Excel) Table 3 - List of all SIMS boron isotope

JOURNAL REQUIREMENTS

Cites

Catanzaro, E.J., Champion, C.E., Garner É. L.,

erence materials: US NBS Special Publication 216-17, 70pp. https://nvlpubs.nist.gov/nistpubs/L

egacy/SP/nbsspecialpublication260-17.pdf

Dyar, M. D., Wiedenbeck, M., Robertson, D.,

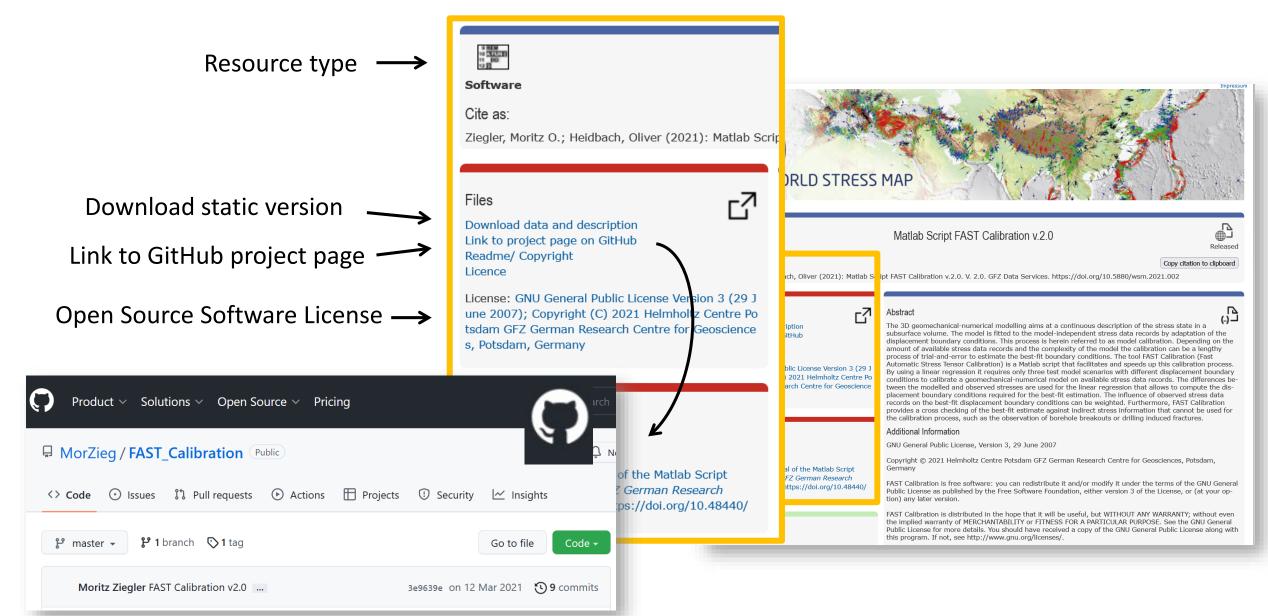
Marienko, O., Sappenfield, K.M., and Shields, W.R.,

1970, Boric acid: isotopic and assay standard ref-

Frikken, P., Cooke, D.R., Walshe, J.L., Archibald,



Special Features: DOIs for software



Special Features: DOIs for software

Resource type ->

Download static version ____

Link to GitHub project page -->

Open Source Software License ->

Manual published as data report



Software

Cite as:

Ziegler, Moritz O.; Heidbach, Oliver (2021): Matlab Scrip

Files

[

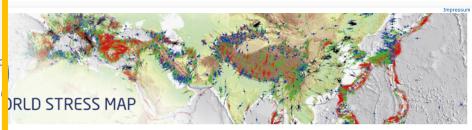
Download data and description Link to project page on GitHub Readme/ Copyright Licence

License: GNU General Public License Version 3 (29 J une 2007); Copyright (C) 2021 Helmholtz Centre Potsdam GFZ German Research Centre for Geoscience s, Potsdam, Germany

Software Description

Documented by

Ziegler, M. (2021). Manual of the Matlab Script FAST Calibration v2.0 . *GFZ German Research Centre for Geosciences*. https://doi.org/10.48440/ WSM.2021.002



Matlab Script FAST Calibration v.2.0



Copy citation to clipboard

nch, Oliver (2021): Matlab Seipt FAST Calibration v.2.0. V. 2.0. GFZ Data Services. https://doi.org/10.5880/wsm.2021.002

Ŋ

blic License Version 3 (29 J) 2021 Helmholtz Centre Po arch Centre for Geoscience

https://doi.org/10.48440/

al of the Matlab Script

FZ German Research

Abstract

The 3D geomechanical-numerical modelling aims at a continuous description of the stress state in a subsurface volume. The model is fitted to the model-independent stress data records by adaptation of the displacement boundary conditions. This process is herein referred to as model calibration. Depending on the amount of available stress data records and the complexity of the model the calibration can be a lengthy process of trial-and-error to estimate the best-fit boundary conditions. The tool FAST Calibration (Fast Automatic Stress Tensor Calibration) is a Matlab script that facilitates and speeds up this calibration process. By using a linear regression it requires only three test model scenarios with different displacement boundary conditions to calibrate a geomechanical-numerical model on available stress data records. The differences between the modelled and observed stresses are used for the linear regression that allows to compute the displacement boundary conditions required for the best-fit estimation. The influence of observed stress data records on the best-fit displacement boundary conditions can be weighted. Furthermore, FAST Calibration provides a cross checking of the best-fit estimate against indirect stress information that cannot be used for the calibration process, such as the observation of borehole breakouts or drilling induced fractures.

Additional Information

GNU General Public License, Version 3, 29 June 2007

Copyright © 2021 Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences, Potsdam, Germany

FAST Calibration is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License, or (at your option) any later version.

FAST Calibration is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details. You should have received a copy of the GNU General Public License along with this program. If not, see http://www.gnu.org/licenses/.

New Website launched in April 2022

- Responsive design
- Thematic search options
- Comprehensive **Publication instructions** (quick start guide, data file instructions, data description templates, ...)
- Portals:
 - GFZ Data Services catalogue:
 DOI-referenced data
 - RI@GFZ: discovery portal for data and research infrastructure at GFZ
 - GFZ Sample Catalogue (IGSN)



