

# Metadaten und ihre Bedeutung für Data Discovery am Beispiel von GFZ Data Services

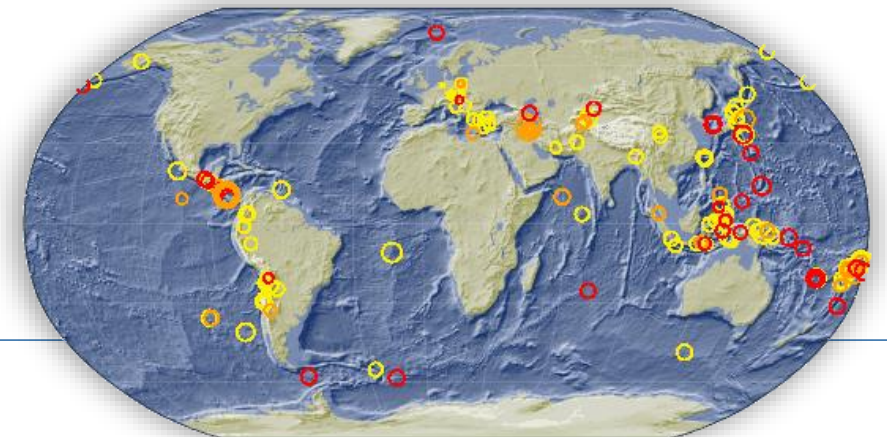
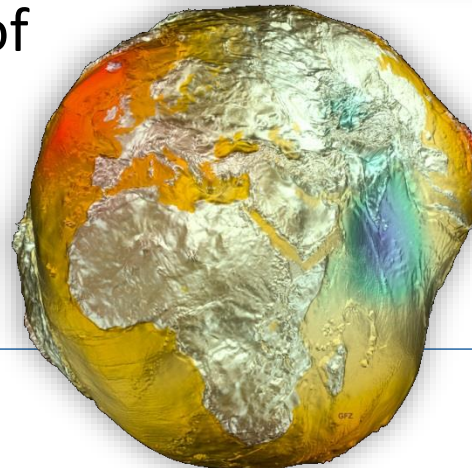
Kirsten Elger; Damian Ulbricht  
Deutsches GeoForschungsZentrum GFZ  
Potsdam, Germany

# Outline

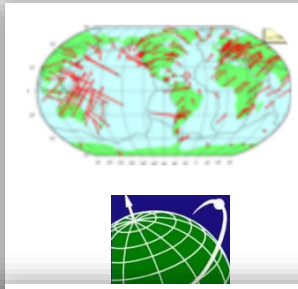
- GFZ und seine Daten
- GFZ Data Services: Domain Repository für die Geowissenschaften
- Metadaten
- Metadatenstandards und Vokabularien
- Harvesting

# GFZ German Research Centre for Geosciences

- National lab for solid Earth geosciences in Potsdam and part of the Helmholtz Association, Germany's largest scientific organisation
- ~1200 employees
- "from space to the Earth's core"
- The development and maintenance of data systems is an essential pillar of GFZ activities and service for the scientific community

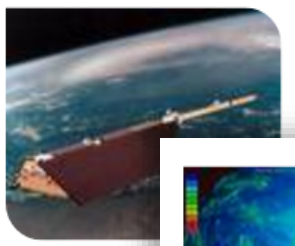


# Data at GFZ: satellites, global networks, observatories ...



## GNSS data

GFZ is operating an "International GNSS Service (IGS)". Data from more than 200 stations serve for the monitoring of tectonic movements, monitoring of Earth's magnetic field, and monitoring of Earth's gravity field.



## Satellite data

GFZ analyses data of various satellite missions, which describe the geometric and physical shape of the Earth or its magnetic field. Data from satellite missions include altimetry, gravity, and magnetic field data.



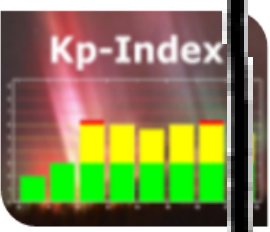
## Global gravity field data

The International Centre for Global Earth Models (ICGEM) makes all global gravity field models of the Earth, which are provided as sets of spherical harmonic coefficients, available to the public. This includes the most recent models back to historical data. The spherical harmonic coefficients are available in a standardised self-describing format. The models can be downloaded from the ICGEM website.



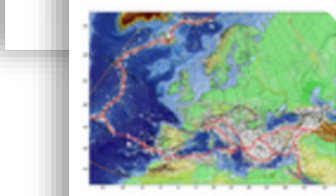
## Seismological data

The global seismic network GEOFON offers seismic data and earthquake data. It runs a data center and a data archive. The data center records and distributes real-time data from GEOFON stations as well as of several boundary observatories. It also arranges temporary experiments and serves as a platform for data exchange.



## Kp-Index: Indices of global geomagnetic activity

The Kp index is a three-hour-range Kp index was introduced in 1959 and is derived from the standardized geomagnetic indices at observatories. It is designed to measure geomagnetic effects. The geomagnetic activity is measured in terms of international Q-days and Kp-index.



## Earthquake catalogue Mediterranean

The European-Mediterranean Seismological Centre (EMSC) maintains a catalogue of some 45,000 entries of earthquakes (from 1900 to 2006). Data within the catalogue are available through the EMEC Earthquake Catalogue web service. The web service also enables the creation of seismicity maps.



## World Stress Map WSM

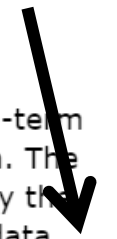
The World Stress Map (WSM) is the global compilation of information on the present-day stress field of the Earth. It contains 21,750 stress data records in its current WSM database (as of 2008). The database is open access via the WSM website.



## Physical experiment data

The International Geophysics Experiment and Data Archive is the platform for long-term storage and dissemination of physical experiment data and its dissemination. The archive contains metadata of all field experiments supported by the International Geophysics Experiment (IGE) and the "Geophysical Pool Potsdam" (GIPP), and contains data from controlled-source tomography and magnetic field measurements.

...to long tail data



# Open Research Data @ GFZ

**GFZ**  
Helmholtz Centre  
POTSDAM

HELMHOLTZ CENTRE POTSDAM  
**GFZ GERMAN RESEARCH CENTRE  
FOR GEOSCIENCES**

## Guidelines on Research Data at the GFZ German Research Centre for Geosciences

- We acknowledge the principles of open access to knowledge, results and technology.
- We uphold the guidelines and rules of good scientific practice.
- We provide our infrastructure to the geoscientific community and contribute to national and international services.
- We initiate and coordinate national and international geoscientific networks.

*(Extract, Mission Statements – Strategy Paper GFZ 2014)*

Quality-assured research data form a basic pillar of scientific knowledge and - regardless of the actual original purpose of the research - the data obtained can often provide the basis for the initiation of further research. The sustainable protection of and access to research data, thus, not only serves the assessment of previous research results but, to a large extent, also the achievement of future results, with the objective of enhancing the quality, productivity and competitiveness and, in this way, forms an elementary basis for knowledge transfer.

(March 2016)

Scientific  
Recognition

Archiving and  
Access

**Citability and  
Publication  
Licences**

Research Data  
Infrastructure

Qualification


Implementation


Responsibility


**Recommendation:  
Data/Software  
publications with DOI**


Recommendation:  
Open Licences  
(CC, Open Source)

**FAIR** Prinzipien

Findable 

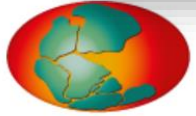
Accessible 

Interoperable 

Reusable 



Deposit  
Once



PANGAEA.

Data



# Best Practice: Data Publication

**Publication of datasets as individual publications (with assigned persistent Identifier; DOI) through data repositories**

## Data Repositories:

- permanent archives and access point to research data
- Open Access
- disciplinary, institutional, general
- persistent identifier (ideally DOI)

- **Findable:** integration of standardised metadata in external data portals (e.g. DataCite, EUDAT B2Find)
- **Accessible:** persistent data storage and access guaranteed by the publisher (= data repository)
- **Documented:** with metadata for discovery and reuse
- **Citable:** DOI-referenced datasets are citable just as journal articles (→ credit for the researcher)

# Coalition on Publishing Data in the Earth and Space Sciences

Data Publications are citable in research articles (COPDESS Statement of Commitment)



## STATEMENT OF COMMITMENT

(January 2015)

- data should be stored in appropriate domain repositories.
- citations of data sets should be included within reference lists.
- include in research papers concise data availability statements.
- links to data sets in publications and corresponding links to journals in data facilities

<http://www.copdess.org/statement-of-commitment/>


Search the Research Data Repository of GFZ Data Services below and read [here](#) how to publish data.

## Search

(press ESC to close suggestions)

## Spatial Filter

[Close Map](#)


Found 6259 datasets.

### **gms-vis: a web-based visual-analytics approach for input data assessment, job parameter definition and progress monitoring for the GeoMultiSens platform**

**Authors:** Eggert, Daniel; Sips, Mike; Dransch, Doris

**Abstract:** gms-vis is a web-based implementation of our visual-analytics approach for assessing remote-sensing data. It is implemented based on the GWT framework. Once deployed through a webserver it acts as the user interface for the GeoMultiSens (GMS) platform. Within the interface users can [more](#)

### **The Iquique Local Network and PicArray**

**Authors:** Cesca, Simone; Sobiesiak, Monika; Tassara, Arturo et al.

**Abstract:** The Iquique Local Network (ILN), a temporal network of broadband and short period seismic stations has been operating in Northern Chile since 2009. The aim of this installation was to locally densify the permanent seismic installation of the Integrated Plate Boundary Observatory in Chile (IPOC), [more](#)

## Datacenters

- CRC 1211 Database
- EnMAP
- GEOFON Seismic Events
- GEOFON Seismic Networks
- GFZ German Research Centre for Geosciences
- GIPP Geophysical Instrument Pool Potsdam
- ICGEM International Centre for Global Earth Mo...
- IGETS International Geodynamics and Earth Tid...
- ISDC Information System and Data Center
- PIK Potsdam Institute for Climate Impact Resea...
- SDDB Scientific Drilling Database
- SFB806 and CRC806-Database
- TERENO
- TR32DB Transregio 32 Database
- WDS World Stress Man

## Categories

- earth science
- earth science services

## Top Subjects

# GFZ Data Services

- Research Data Repository for the **Geosciences domain**
- **DOIs** for data and software
- Optional: Data description via **Data Reports** (DOI, internally reviewed)
- **OAI-PMH Interface**
- **Data:** real-time data streams (e.g. seismic waveforms, climate stations, observatories), tables, maps, model data (input/ output data of climate models, analogue and numerical modelling data), data products, ....



# Data access via DOI Landing Pages

- Citation information
- PIDs: ORCID, Fundref
- Contributors listed
- Controlled Vocabularies: NASA GCMD Science Keywords, GeoSciML....
- XML metadata for download: ISO19115, DataCite, .dif, Dublin Core
- Related References to papers, reports, data, software, IGSN
- Map

**GFZ**  
Helmholtz-Zentrum  
Potsdam

Supplementary material for analogue experiments on the interactions of two indenters, and their implications for curved fold-and-thrust belts

Cite as:  
Reiter, Karsten; Kukowski, Nina; Ratschbacher, Lothar; Rosenau, Matthias (2016): Supplementary material for analogue experiments on the interactions of two indenters, and their implications for curved fold-and-thrust belts. GFZ Data Services. <http://doi.org/10.5880/GFZ.4.1.2016.007>

**Data Files**

Explanations_Reiter-et-al-2016.pdf	527520 Bytes
list-of-files-Reiter-et-al-2016.pdf	238166 Bytes
Experimenting.avi	82477450 Bytes
gb70-pictures.pdf	509078 Bytes
gb40-3Dview-30-34.avi	6096554 Bytes
gb50-3Dview-30-33.avi	5947146 Bytes
gb55-3Dview-30-32.avi	6397110 Bytes
gb60-3Dview-30-29.avi	6697512 Bytes
gb70-3Dview-30-30.avi	6247434 Bytes
gb80-3Dview-30-31.avi	6187996 Bytes
si60-3Dview-30-38.avi	5941054 Bytes

**Dataset**

Supplementary material for analogue experiments on the interactions of two indenters, and their implications for curved fold-and-thrust belts

Released

Cite as:  
Reiter, Karsten; Kukowski, Nina; Ratschbacher, Lothar; Rosenau, Matthias (2016): Supplementary material for analogue experiments on the interactions of two indenters, and their implications for curved fold-and-thrust belts. GFZ Data Services. <http://doi.org/10.5880/GFZ.4.1.2016.007>

**Data Files**

Explanations_Reiter-et-al-2016.pdf	527520 Bytes
list-of-files-Reiter-et-al-2016.pdf	238166 Bytes
Experimenting.avi	82477450 Bytes
gb70-pictures.pdf	509078 Bytes
gb40-3Dview-30-34.avi	6096554 Bytes
gb50-3Dview-30-33.avi	5947146 Bytes
gb55-3Dview-30-32.avi	6397110 Bytes
gb60-3Dview-30-29.avi	6697512 Bytes
gb70-3Dview-30-30.avi	6247434 Bytes
gb80-3Dview-30-31.avi	6187996 Bytes
si60-3Dview-30-38.avi	5941054 Bytes

**Abstract**

This data publication includes animations and figures of eight scaled analogue models that are used to investigate the evolution of a curved mountain belt akin to the Pamir and Hindu Kush orogenic system and adjacent Tadjik basin. Crustal deformation is simulated by means of indentation of two basement blocks into a sedimentary sequence and the formation of a curved fold-and-thrust belt. The experimental set-up has two adjacent rigid indenters representing the basement blocks moving in parallel with a velocity difference (Figure 1). The slow indenter moves with a relative velocity ranging from 40 to 80% of that of the fast one. A layer of quartz sand in front of the indenters, 1 by 1 meter in size and 1.5 cm thick, represents the sedimentary basin infill. A basal detachment layer is made up of low-friction glass beads or viscous silicone oil representing weak shale or evaporites layers, respectively. The surface evolution by means of topography and strain distribution is derived from 3-D particle image velocimetry (PIV). This allows visualizing and analysing the development of the model surface during the complete model run at high spatio-temporal resolution. All details about the model set-up, modelling results and interpretation can be found in Reiter et al. (2011).

**Data Description**

Reiter, K., Kukowski, N., & Ratschbacher, L. (2011). The interaction of two indenters in analogue experiments and implications for curved fold-and-thrust belts. *Earth and Planetary Science Letters*, 302(1-2), 132-146. doi:10.1016/j.epsl.2010.12.002

**Contributors**

GFZ German Research Centre for Geosciences

**Keywords**

two indenter tectonics, particle image velocimetry, fold-and-thrust belts, Tadjik basin, Pamir, 4D analogue experiments, EPOS, European Plate Observing Systems, mountain building, continental collision, sandbox model, digital elevation model, analogue model, tectonic setting, collisional setting, foreland setting, crust setting, upper continental crustal setting, tectonic process, detachment fault, oblique slip fault, reverse fault, wrench fault, thrust fault, tectonic and structural features, mountain, Microsphere, Glassy, Silicon, Sand, Quartz, Sandbox (cm scale), Time lapse camera, Sectioning, tectonics, fault, geological process, digital land model

**Related Work**

Adam, J., Urai, J. L., Wieneke, B., Oncken, O., Pfeifer, K., Kukowski, N., ... Schmatz, J. (2005). Shear localisation and strain distribution during tectonic faulting—new insights from granular-flow experiments and high-resolution optical image correlation techniques. *Journal of Structural Geology*, 27(2), 283-301. doi:10.1016/j.jsg.2004.08.008

**Find More Research Data**

<http://bib.telegrafenberg.de/finden/datenbanken/fo-rschungsdaten/>

**Location**

Click/hover over markers or bounding boxes to see related details. Click/hover over details to see related marker of bounding box.

**Data Description**

Reiter, K., Kukowski, N., & Ratschbacher, L. (2011). The interaction of two indenters in analogue experiments and implications for curved fold-and-thrust belts. *Earth and Planetary Science Letters*, 302(1-2), 132-146. doi:10.1016/j.epsl.2010.12.002

**Contributors**

GFZ German Research Centre for Geosciences

**Keywords**

two indenter tectonics, particle image velocimetry, fold-and-thrust belts, Tadjik basin, Pamir, 4D analogue experiments, EPOS, European Plate Observing Systems, mountain building, continental collision, sandbox model, digital elevation model, analogue model, tectonic setting, collisional setting, foreland setting, crust setting, upper continental crustal setting, tectonic process, detachment fault, oblique slip fault, reverse fault, wrench fault, thrust fault, tectonic and structural features, mountain, Microsphere, Glassy, Silicon, Sand, Quartz, Sandbox (cm scale), Time lapse camera, Sectioning, tectonics, fault, geological process, digital land model

**Related Work**

Adam, J., Urai, J. L., Wieneke, B., Oncken, O., Pfeifer, K., Kukowski, N., ... Schmatz, J. (2005). Shear localisation and strain distribution during tectonic faulting—new insights from granular-flow experiments and high-resolution optical image correlation techniques. *Journal of Structural Geology*, 27(2), 283-301. doi:10.1016/j.jsg.2004.08.008

**Find More Research Data**

<http://bib.telegrafenberg.de/finden/datenbanken/fo-rschungsdaten/>

# DOI Services/ Datacentres



## Datacenters

- EnMAP
- GEOFON Seismic Events
- GEOFON Seismic Networks
- GFZ German Research Centre for Geosciences
- GIIP Geophysical Instrument Pool Potsdam
- ICGEM International Centre for Global Earth M...
- IGETS International Geodynamics and Earth Ti...
- ISDC Information System and Data Center
- PIK Potsdam Institute for Climate Impact Rese...
- SDDB Scientific Drilling Database
- SFB806 and CRC806-Database
- TERENO
- TR32DB Transregio 32 Database
- WDS World Stress Map

- DOI Services for ICGEM, IGETS (IAG), partner institutes, large collaborative projects → Organisation in Data Centres
- Project-specific layout of DOI Landing Pages

# IGSN - International Geo Sample Number

- Globally unique identifier for physical samples and materials
- Central registration based on the Handle system
- QR Code on the sample

- Online sample description online via IGSN Landing Pages/ catalogue
- IGSN citation in papers possible

**General Identifiers**

Program:	ICDP
Expedition:	ICDP 5054
Type:	Core
Name:	5054_1_A_3_Z
IGSN:	ICDP5054EC4Q001 (Open)
Parent IGSN:	ICDP5054EEW1001
Release Date:	2017-3-1

**Sampling Location**

Latitude:	63.4063
Longitude:	13.203057
Coordinate System:	WGS84
Elevation:	415.74
Final Depth:	412.61
Location Type:	N/A
Location Name:	Åre, Jämtlands län, Sweden
Location Description:	COSC-1 is located in the vicinity of the abandoned Fröå mine
Country:	Sweden
Province:	Jämtlands län
County:	N/A
City:	Åre

**Geology**

Material:	Rock
Rock Classification:	N/A
From Corrected Depth:	106.26
To Corrected Depth:	109.39
Depth Reference:	meter below ground level
Geological Age:	mid-paleozoic
Geological Unit:	N/A

**Methods**

MSCL	yes
XRF	yes
Lithological Description	yes
Core Overview	yes
Core Section Scan	yes
Core Catcher Scan	no

**Drilling**

Drilling Method:	Coring>RockCorer wireline diamond coring, HQ and NQ bit size
Operator:	Lund University, Engineering Geology Larsson Drilling Consulting AB
Funding Agency:	Swedish Research Council (Vetenskapsrådet)
Total Length:	2400.1m
Comments:	N/A
Platform Type:	drill rig

**Sample Family**

- 5054\_1\_A\_1\_Z
- 5054\_1\_A\_2\_Z
- 5054\_1\_A\_3\_Z
  - 5054\_1\_A\_3\_Z\_1
  - 5054\_1\_A\_3\_Z\_2
  - 5054\_1\_A\_3\_Z\_3
  - 5054\_1\_A\_3\_Z\_4

Legend: ⊕=Hole, □=Core, ▭=Core-Section, ▩=Core-Sample

The Sample Family shows a sub-sampling graph. Select entries to navigate samples. Core-Samples are issued to scientists on request. The naming convention for a Core-Sample is: Expedition\_Site\_Hole\_Core\_Section, from-to(cm). Hole, Core, and Core-Section are following the same schema respectively.

**Publications & Datasets**

Lorenz, H., Rosberg, J.-E., Juhlin, C., Bjelm, L., Almqvist, B. S. G., Berthet, T., ... Tsang, C.-F. (2015). COSC-1 – drilling of a subduction-related allochthon in the Palaeozoic Caledonide orogen of Scandinavia. *Sci. Dril.*, 19, 1–11. doi:10.5194/sd-19-1-2015

Lorenz, Henning; Rosberg, Jan-Erik; Juhlin, Christopher; Bjelm, Leif; Almqvist, Bjarne; Berthet, Théo; Conze, Ronald; Gee, David G.; Klonowska, Iwona; Pascal, Christophe; Pedersen, Karsten; Roberts, Nick; Tsang, Chinfu; (2015): COSC-1 operational report - Operational data sets; GFZ Data Services. <http://dx.doi.org/10.1594/GFZ.SDDB.ICDP.5054.2015>

# What is Metadata?

Metadata is „data about data“ or „information about information“

**Metadata is key to ensuring that resources will survive and continue to be accessible into the future.**

Descriptive metadata	For finding or understanding a resource
Administrative metadata <ul style="list-style-type: none"><li>- Technical metadata</li><li>- Preservation metadata</li><li>- Rights metadata</li></ul>	<ul style="list-style-type: none"><li>- For decoding and rendering files</li><li>- Long-term management of files</li><li>- Intellectual property rights attached to content</li></ul>
Structural metadata	Relationships of parts of resources to one another
Markup languages	Integrates metadata and flags for other structural or semantic features within content

# Descriptive Metadata

## Descriptive Metadata types:

- **Contextual Metadata:** e.g. transfer function, instruments used, processing steps, resolution (header, readme, data report, data paper, etc.)  
→ highly variable between the disciplines but key information for data reuse
- **Metadata for Data Discovery**  
Who? What? When? Where? Why?  
→ essential for data discovery, DOI registration, etc.: international standards/  
schemes across all disciplines → Data Portals, Repositories

## Dublin Core Metadata (oai\_dc)

<b>Title</b>	PyRQA - Tool for the fast Recurrence Quantification Analysis (RQA) of long time series based on the OpenCL framework
<b>Author or Creator</b>	Rawald, Tobias
<b>Author or Creator</b>	Sips, Mike
<b>Author or Creator</b>	Dransch, Doris
<b>Publisher</b>	GFZ German Research Centre for Geosciences
<b>Date</b>	2018
<b>Resource Identifier</b>	<a href="http://dx.doi.org/10.5880/GFZ.1.5.2018.002">http://dx.doi.org/10.5880/GFZ.1.5.2018.002</a>
<b>Relation</b>	doi:10.1007/978-3-319-09531-8_2
<b>Relation</b>	urn:urn:nbn:de:0074-1330-0
<b>Description</b>	Abstract
<b>Description</b>	PyRQA is a tool to conduct recurrence quantification analysis (RQA) and to create recurrence plots in a massively parallel manner using the OpenCL framework. It is designed to process very long time series consisting of hundreds of thousands of data points efficiently.
<b>Language</b>	eng
<b>Resource Type</b>	Software
<b>Resource Type</b>	Dataset
<b>Format</b>	43864 Bytes
<b>Format</b>	application/octet-stream
<b>Rights Management</b>	Apache License Version 2.0
<b>Rights Management</b>	<a href="https://www.apache.org/licenses/LICENSE-2.0">https://www.apache.org/licenses/LICENSE-2.0</a>

# Metadaten Schemata

← Dublin Core – kleinster gemeinsamer Nenner  
DataCite – „erweitertes“ Dublin Core Schema

```
<creator>
  <creatorName>Sips, Mike</creatorName>
  <givenName>Mike</givenName>
  <familyName>Sips</familyName>
  <nameIdentifier nameIdentifierScheme="ORCID">0000-0003-3941-7092</nameIdentifier>
  <affiliation>GFZ German Research Centre for Geosciences, Potsdam, Germany</affiliation>
</creator>
```

```
<relatedIdentifiers>
  <relatedIdentifier relationType="IsReferencedBy"
    relatedIdentifierType="DOI">10.1007/978-3-319-09531-8_2</relatedIdentifier>
  <relatedIdentifier relationType="IsReferencedBy"
    relatedIdentifierType="URN">urn:nbn:de:0074-1330-0</relatedIdentifier>
</relatedIdentifiers>
```

ISO 19115/19139 – super für Geospatial Data

Maschinen-lesbar!!

# Ergänzung: kontrollierte Vokabularien

- Jedes Schema ist ein Kompromiss mit individuellen Stärken und Schwächen
- Schemas gibt es vor allem für Data Discovery (Titel, Autor, Jahr, Identifier, spatial domain, related identifier ... → interdisziplinär)
- Ergänzung von disziplinärem Kontext durch Integration von kontrollierten Vokabularien

# GFZ Metadata Editor (Java Script „translator“)

XML (Extensible Markup Language):  
Metadata exchange format

The screenshot displays the GFZ Metadata Editor interface. On the left, a form is visible with sections for Resource Information, Licenses and Rights, Authors, and Temporal and Spatial Coverage. On the right, an XML preview window shows the metadata structure for 'iso19115'. The XML includes fields for fileIdentifier, language, characterSet, hierarchyLevel, hierarchyLevelName, contact, dateStamp, referenceSystemInfo, and identificationInfo. The contact information is specifically for the GFZ German Research Center for Geosciences, with a URL of http://www.gfz-potsdam.de/ and a role of pointOfContact. The dateStamp is 2017-01-06, and the referenceSystemInfo includes the EPSG:4326 coordinate system.

```
iso19115: close inline view / download xml
o MD_Metadata (xsi:schemaLocation=http://www.isotc211.org/2005/gmd http://www.isotc211.org/2005/gmd/gmd.xsd)
  ■ fileIdentifier
    ■ CharacterString: doi:10.5880/igets.po.11.001
  ■ language
    ■ LanguageCode (codeList=http://www.loc.gov/standards/iso639-2/ codeListValue=eng): eng
  ■ characterSet
    ■ MD_CharacterSetCode (codeList=http://www.isotc211.org/2005/resources/codeList.xml#MD_CharacterSetCode codeListValue=utf8):
  ■ hierarchyLevel
    ■ MD_ScopeCode (codeList=http://www.isotc211.org/2005/resources/Codelist/gmx-Codelists.xml#MD_ScopeCode codeListValue=):
  ■ hierarchyLevelName
    ■ CharacterString:
  ■ contact
    ■ CI_ResponsibleParty
      ■ organisationName
        ■ CharacterString: GFZ German Research Center for Geosciences
      ■ contactInfo
        ■ CI_Contact
          ■ address
            ■ CI_Address
              ■ electronicMailAddress
                ■ CharacterString:
            ■ onlineResource
              ■ CI_OnlineResource
                ■ linkage
                  ■ URL: http://www.gfz-potsdam.de/
                ■ function
                  ■ CI_OnlineFunctionCode (codeList=http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_OnlineFunctionCode codeListValue=):
          ■ role
            ■ CI_RoleCode (codeList=http://www.isotc211.org/2005/resources/Codelist/gmxCodelists.xml#CI_RoleCode codeListValue=pointOfContact): pointOfContact
    ■ dateStamp
      ■ Date: 2017-01-06
    ■ referenceSystemInfo
      ■ MD_ReferenceSystem
        ■ referenceSystemIdentifier
          ■ RS_Identifier
            ■ code
              ■ CharacterString: urn:ogc:def:crs:EPSG:4326
    ■ identificationInfo
      ■ MD_DataIdentification
        ■ citation
```

Access via: <http://dataservices.gfz-potsdam.de/portal/about.html> „Publishing step by step“



# GFZ Metadata Editor (Java Script „translator“)

## „Special“ Features:

- Interactive map
- ORCID and Fundref
- Controlled vocabularies
- Multiple affiliations for authors

The screenshot displays the GFZ Metadata Editor interface with several sections:

- DataCite Metadata** (selected tab)
- Resource Information**: DOI (will be generated in the publishing process), Publisher (GFZ Data Services), Year (2016), Resource Type (Dataset), Title (Supplement to: The New World Atlas of Artificial Night Sky Brightness), Language of dataset (eng).
- Licenses and Rights**: Licence (Please contact the authors for a licence agreement).
- Authors (Persons and/or Institutions)**: Table with columns: Author (Lastname, Firstname), Role, Author ID Type, Author Identifier (ID), Affiliation. Authors listed include Cinzano, Pierantonio; Duriscoe, Dan; Kyba, Christopher C. M.; Elvidge, Christopher D.; Baugh, Kimberly; Portnov, Boris; Rybnikova, Nataliya A.; Furgoni, Riccardo.
- Contact Person(s) / Point of Contact**: night, radiative transfer, Suomi NPP, Sky Quality Meter.
- Temporal and Spatial Coverage**: Table with columns: Latitude (Min, Max), Longitude (Min, Max). Values: 44.045486..., 55.842428..., 2.8710901..., 43.124996....
- Interactive Map**: A map of Europe with a red bounding box around Germany and Poland. Text: "Select Region (click left mouse and drag)".

## Output:

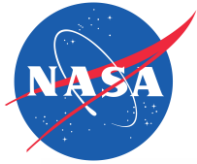
Standardised XML files (Datacite, ISO 19115, NASA GCMD DIF, Dublin Core)



GFZ Data Services Metadata Catalogue

EPOS, B2FIND, ENVRIplus, etc.

# Controlled Vocabularies in the Metadata Editor



**NASA Global Change Master Directory (GCMD) Keywords**

*Example: EARTH SCIENCE > SOLID EARTH > TECTONICS > VOLCANIC ACTIVITY > ERUPTION DYNAMICS*

The screenshot shows a 'Thesaurus' window with a search filter 'tecto'. The tree structure is as follows:

- NASA GCMD Science Keywords
  - EARTH SCIENCE
    - OCEANS
      - MARINE GEOPHYSICS
      - PLATE TECTONICS
    - SOLID EARTH
      - GEOMORPHIC LANDFORMS
      - TECTONICS
        - VOLCANIC ACTIVITY
          - ERUPTION DYNAMICS** (highlighted)
          - LAVA COMPOSITION/TEXTURE
          - MAGMA COMPOSITION/TEXTURE
          - PYROCLASTICS COMPOSITION/TEXTURE
          - ASH/DUST COMPOSITION
          - VOLCANIC GASES
          - LAVA SPEED/FLOW
          - MAGMA SPEED/FLOW
          - PYROCLASTIC PARTICAL SIZE DISTRIBUTION



**INSPIRE ISO 19115 Keywords**

**RDF-Version**



**GeoSciML: Geoscience Vocabularies for Linked Data**

# Harvesting: OAI-PMH Schnittstelle

## OAI 2.0 Request Results

[Identify](#) | [ListRecords](#) | [ListSets](#) | [ListMetadataFormats](#) | [ListIdentifiers](#)

You are viewing an HTML version of the XML OAI response. To see the underlying XML use your web browsers view source option. More i

**Datestamp of response** 2018-06-12T19:43:31Z

**Request URL** <http://doidb.wdc-terra.org/oaip/oaip>

Request was of type ListRecords.

**OAI Record: oai:doidb.wdc-terra.org:222**

### OAI Record Header

**OAI Identifier** oai:doidb.wdc-terra.org:222 [oai\\_dc](#) [oai\\_datacite](#) [datacite](#) [dif](#) [iso19139](#) [formats](#)

**Datestamp** 2012-03-21T13:37:26Z

**setSpec** DOIDB [Identifiers](#) [Records](#)

**setSpec** DOIDB.WSM [Identifiers](#) [Records](#)

### DataCite Metadata (datacite)

```
<resource xsi:schemaLocation="http://datacite.org/schema/kernel-2.2 http://schema.datacite.org/meta/kernel-2.2/metadata.xsd" >
  <identifier identifierType="DOI" >10.1594/GFZ.WSM.Rel2008</identifier>
  <creators>
    <creator>
```

- Austausch von Metadaten für Sichtbarkeit in externen Katalogen
- Im Jahr 2000 standardisiert und einfach gehalten
- Benötigt Identifier und XML-Metadaten
- Pro ID beliebige Metadaten, aber mindestens Dublin Core
- Inkrementelles Harvesting: "Änderungen seit gestern"

# conclusions

- Metadaten haben verschiedenen Einsatzzweck
- Konsens was auf welcher Granularität beschrieben wird, dafür ID vergeben (z.B. DOI für Daten)
- Metadaten dieser ID zuordnen
- Viele Kataloge basieren auf Dublin Core → Mapping zu einfachen Schemas
- Verschlagwortung über Vokabularien
  
- Schema.org → Von Google et al erkennbar

**Thank you for your attention!!**