

TIB

TECHNISCHE
INFORMATIONSBIBLIOTHEK

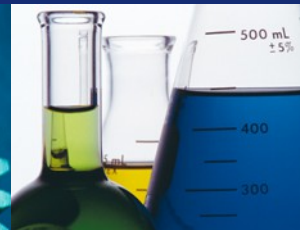
DataCite und DOI für Forschungsdaten

Dr. Janna Neumann

Persistent Identifier für Kultur und Wissenschaft

104. Bibliothekartag 2015

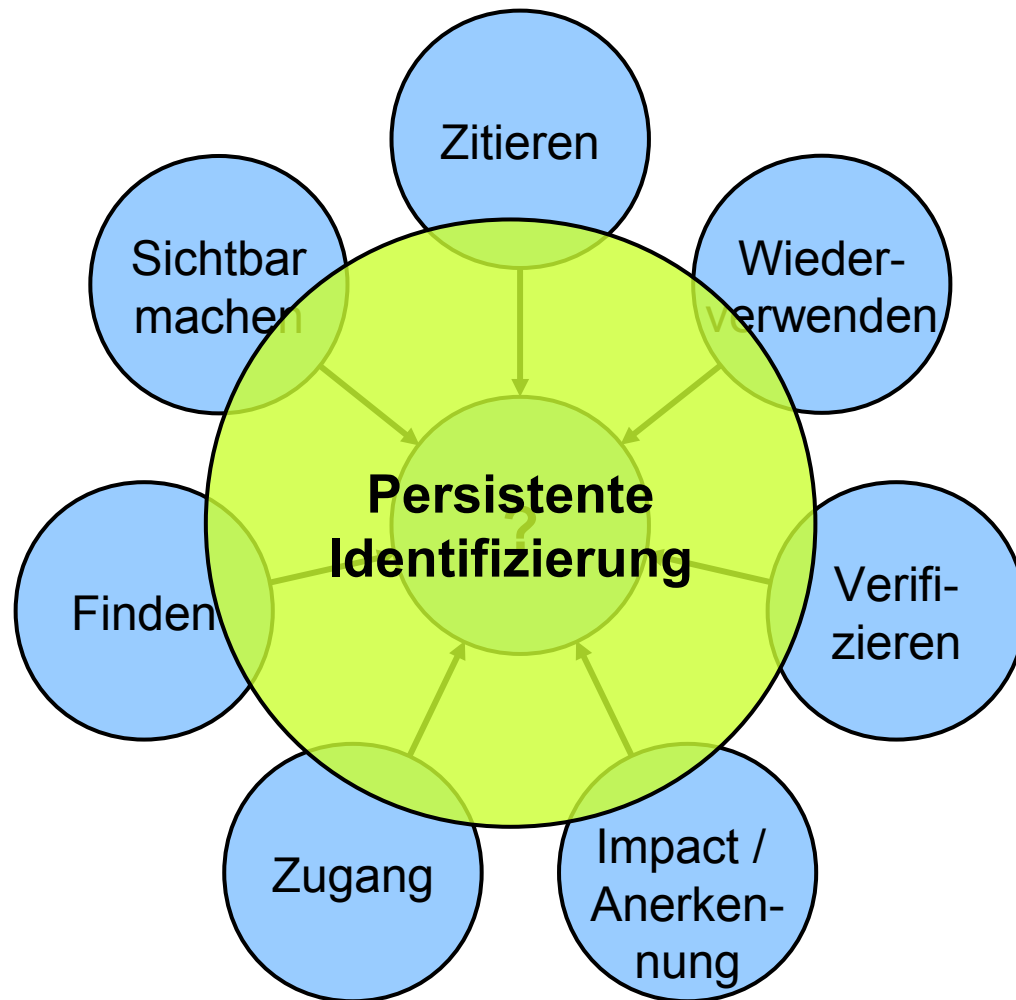
27.05.2015 Nürnberg



Agenda

- Persistente Identifizierung
- DOI System
- DataCite
- DOI für Forschungsdaten
- DOI-Service der TIB

Persistente Identifizierung



Persistente Identifizierung

Eigenschaften

- **Identifizieren** einer Ressource.
- Die Ressource wird **eindeutig referenzierbar** und **zitierfähig**.
- **Dauerhaft**, d.h. ggf. auch über die Lebensdauer des identifizierten Objektes hinaus.
- Strikte Trennung von Identifikation der Ressource und der Standortreferenz.

- Vergabe wird von Registrierungsagenturen übernommen:
 - **Standards** für Struktur und Syntax der PI
 - **Resolving-Mechanismus**

Das DOI-System

- International DOI Foundation (IDF) 1998 gegründet
- DOI System
 - bietet langfristige Zugänglichkeit zu Objekten
 - basiert technisch auf dem Handle System.
 - im May 2012 als ISO Standard 26324 publiziert
 - garantiert vertrauenswürdige Verantwortlichkeiten, einheitliche Standards
- Teil der Qualitätskontrolle: obligatorische Metadaten für jedes mit einem DOI registrierte Objekt
- IDF besteht derzeit aus 9 Registrierungsagenturen (RA)
 - CrossRef, DataCite
- RA zuständig für PI-Vergabe und Pflege



DataCite

Historie

- **2003:**
DFG Projekt der TIB mit World DataCenters bzgl. Publikation von Forschungsdaten
- **2005:**
TIB wird erste DOI-Registrierungsagentur für Forschungsdaten. Von Anfang an wird auch graue Literatur registriert.
- **03.2009:**
Paris Memorandum bzgl. der Zusammenarbeit von 6 europäischen Informationseinrichtungen
- **12.2009:**
Gründung von DataCite e. V. mit 7 Mitgliedern in London



- Globales Konsortium getragen von lokalen Institutionen
- Ziel:
 - Publikationsinfrastruktur für Daten und andere nicht-textuelle Inhalte verbessern
- Dienstleistungsanbieter für Datenzentren/Inhaltsanbieter
- Entwicklung von Standards, Workflows und Best practice
- Basiert auf dem DOI System
- Infos auf <http://www.datacite.org/>
- In 2015:
 - 24 Mitglieder, 8 assoziierte Mitglieder
 - über 5 Millionen DOI

DataCite

Mitglieder

CISTI – Canada Institute for Scientific and Technical Information

California Digital Library, USA

Purdue University, USA

OSTI – Office of Scientific and Technical Information, USA

The British Library

TIB, Germany

ZB MED, Germany

ZBW, Germany

GESIS, Germany

SUB Göttingen, Germany

University of Tartu, Estonia

JaLC – Japan Link Center

DTIC – Technical Information Center of Denmark

Library of TU Delft, The Netherlands

Library of ETH Zürich, Switzerland

INIST – L'Institut de l'Information Scientifique et Technique, France

SND – Swedish National Data Service

ANDS – Australian National Data Service

NRCT – National Research Council of Thailand

The Hungarian Academy of Sciences

CRUI – Conferenza dei Rettori delle Università Italiane

SAEON – South African Environmental Observation Network

CERN – European Organization for Nuclear Research

BIBSYS – Library System, Norway



Affiliated members:

Digital Curation Center, UK

Microsoft Research, USA

ICPSR – Interuniversity Consortium for Political and Social Research, USA

KISTI – Korea Institute of Science and Technology Information

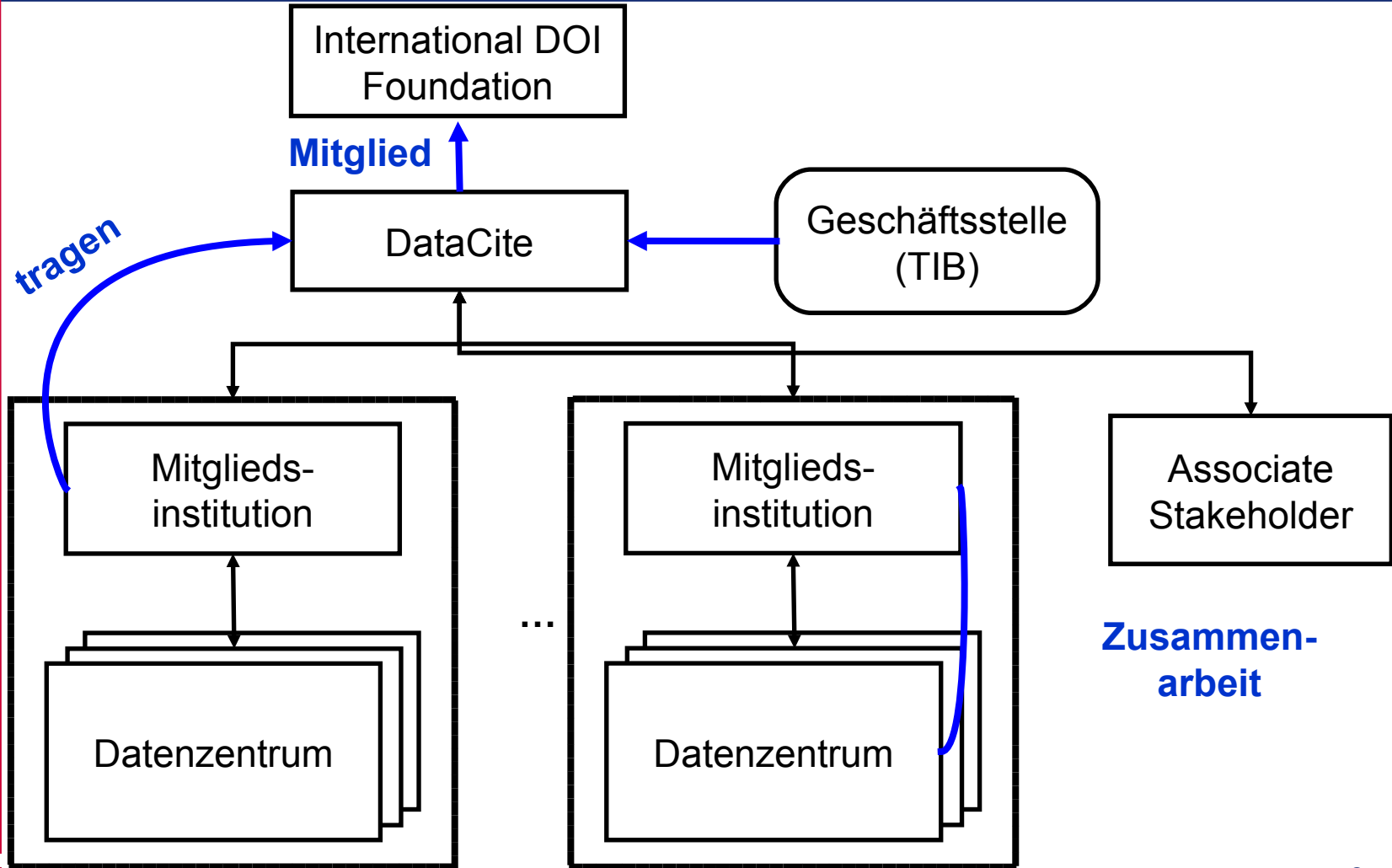
BGI – Beijing Genomic Institute, China

IEEE, USA

Harvard University Library, USA

GWDG, Germany

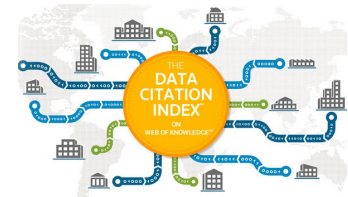
DataCite Struktur



- Mitglieder und assoziierte Mitglieder
 - Bibliotheken, Informations- und Datenzentren
 - DataCite „in Deutschland“
- Working Groups
 - Metadaten
 - Best Practices
- Weitere Services
 - Metadata Store
 - Search, Stats, OAI Provider
 - <http://www.datacite.org/services>

- Thomson Reuters - Data Citation Index
 - Harvesten von Metadaten über DataCite
 - Vorteile für Kunden:
Zugriff auf die Statistiken des DCI

- ORCID – ODIN und THOR Projekt
 - Aufnahme von Datensätzen in Publikationslisten
 - Nachverfolgung von Nutzung von Datensätzen
 - Verlinkung von Datensätzen mit zugehörigen Artikeln, Lizenzen und allen beteiligten Personen



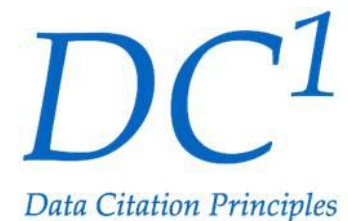
- re3data.org / Databib Fusion
 - Anbieter von Repositorylisten
 - fusionieren unter Schirmherrschaft von DataCite



- MoU mit RDA:
 - DataCite wird ein “organizational member”



- Endorsement of the Force11
„Joint Declaration of Data Citation Principles“



- In Zusammenarbeit mit CrossRef:



- <http://crosscite.org/citeproc/>

Der Citation Formater stellt Zitierungen in über 100 Formaten bereit.

- <http://crosscite.org/cn/>

Über die Content Negotiation kann maschinell direkt auf verschiedene (vorher hinterlegte) Medienformate eines Objektes zugegriffen werden.

- International Association of STM Publishers :



- Verbesserung des Zugriffs auf und der Findbarkeit von Forschungsdaten.
- Förderung von bidirektionalen Verlinkungen zwischen Datensätzen und Publikationen in Datenarchiven.
- Erhöhung der Sichtbarkeit von Links zwischen Publikationen und Datensätzen.

DataCite

Inhaltstypen

- Derzeit ca. 5.000.000 DOI-Namen von DataCite vergeben für:
 - Forschungsdaten (~45%)
 - graue Literaturobjekte (~40%)
 - Bilder (~10%)
 - Medizinische Fallstudien
 - Videos
 - Landkarten
 - Lernobjekte

DOI für Forschungsdaten

Warum?

- Eindeutige Referenzierung und Zitierfähigkeit
- Verlinkung der Daten mit anderen Publikationen
- Erhöhte Sichtbarkeit
- Transparente Forschung
- Verbesserter Zugang
- Vermeidung von Doppelarbeit
- Förderung der wiss. Zusammenarbeit
- Motivation für neue Forschung

DOI Registrierung

Anforderungen

- Datenpersistenz gewährleisten
- Metadaten liefern
- Landing page bereitstellen
- Granularität (zitierwürdig?)
- DOI syntax:
 - Eindeutige Zeichenfolge
 - Zu nutzende Zeichen: A-Z a-z 0-9 . : - _ /
- <http://www.datacite.org/resources> -> DataCite Business Modells Principles

DataCite Metadaten Schema

Pflichtfelder

- Identifier (*mit type Attribut*)
- Creator (*mit type und nameIdentifier Attributen*)
- Title (*mit optionalem type Attribut*)
- Publisher
- PublicationYear

- Zitierung:
Creator (PublicationYear): Title. Publisher. Identifier

Zitierung mit DOI-Namen (I)

Die DOI-Registrierung ermöglicht eine elegante Verlinkung zwischen einem wissenschaftlichen Artikel und den im Artikel analysierten Forschungsdaten.

So wird beispielsweise der Datensatz:

Kuhlmann, H et al. (2009):

Age models, iron intensity, magnetic susceptibility records and dry bulk density of sediment cores from around the Canary Islands.

[doi:10.1594/PANGAEA.727522](https://doi.org/10.1594/PANGAEA.727522),

in folgendem Artikel analysiert:

Kuhlmann, Holger; Freudenthal, Tim; Helmke, Peer; Meggers, Helge (2004):

Reconstruction of paleoceanography off NW Africa during the last 40,000 years: influence of local and regional factors on sediment accumulation.

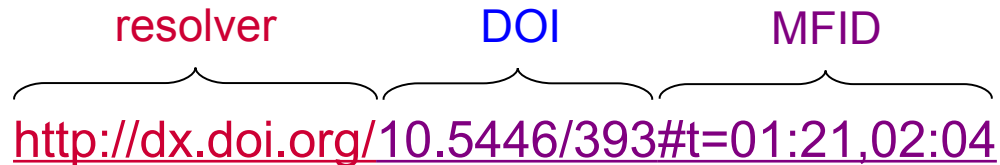
Marine Geology, 207(1-4), 209-224,

[doi:10.1016/j.margeo.2004.03.017](https://doi.org/10.1016/j.margeo.2004.03.017)

Zitierung mit DOI-Namen (II)

Media Fragment Identifier

- Punktgenaue Zitierung von Videos


<http://dx.doi.org/10.5446/393#t=01:21,02:04>

- Auch für andere Medien einsetzbar, wenn die Fragmentierung unterstützt wird
 - z.B. PDF: <http://doi.org/10.5438/0010#page=9>

Dynamische Datensätze

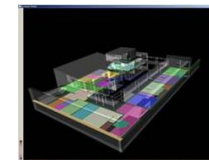
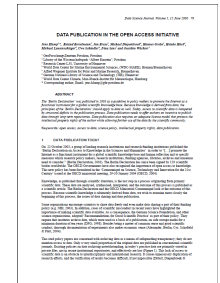
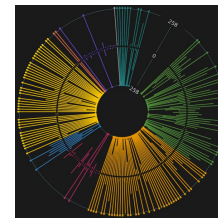
- Zitierung eines bestimmten Abschnitts eines Datensatzes mit eigenem DOI.
- Zitierung eines Snapshots (ganzer Datensatz zu einem bestimmten Zeitpunkt) mit eigenem DOI.
- Zitierung eines kontinuierlich aktualisierten Datensatzes mit Zugriffszeitpunkt mit dem DOI des Datensatzes.
 - Nur möglich, wenn die alten Daten nicht verändert werden!

Immer wieder Thema, aber:
Keine allgemeingültigen Richtlinien.

Jedes Objekt, was zitiert werden soll, kann einen DOI
bekommen!

DOI-Service der TIB (I)

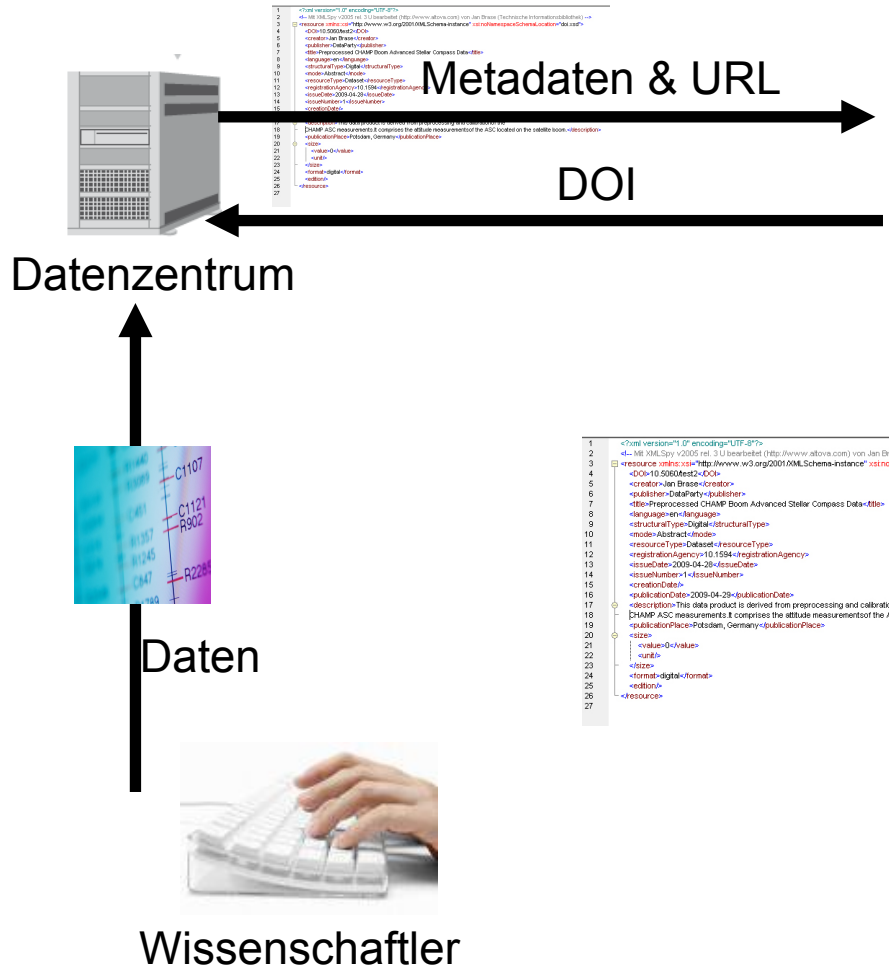
- TIB registriert Objekte aus dem Bereich Technik und Naturwissenschaften
 - Forschungsdaten
 - Graue Literatur
 - Open Access Artikel aus akademischen Einrichtungen
 - Videos
 - Bilder
 - Digitalisierte Kulturobjekte
 - Software
 - und weitere nicht-textuelle Objekte
- Zugang: über GetInfo



DOI-Service an der TIB (II)

- DOI Registrierungen (13.05.2015)
 - 1 018 624 DOI Namen (ca. 83 300 in 2015)
 - ~ 76% Forschungsdaten, 22% graue Lit., 0.25% AV-Medien
- 80 Datenzentren
 - 66 deutsche Datenzentren
 - 14 aus dem Ausland
 - Kürzlich neu dazugekommen sind:
 - Universitätsbibliothek Duisburg-Essen
 - Universitätsbibliothek der FU Berlin
 - Universitätsbibliothek Kassel

DOI Registrierung Workflow



```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <!-- XML Spy v2005 rel. 3 U bearbeitet (http://www.altova.com) von Jan Brase (Technische Informationsbibliothek) -->
3 <resource xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns="http://schema.datacite.org/meta" />
4 <DOI>10.5080/est2-DOI</DOI>
5 <creators>
6 <creator>Jan Brase</creator>
7 </creators>
8 <titles>
9 <title>Preprocessed CHAMP Boom Advanced Stellar Compass Data</title>
10 </titles>
11 <descriptions>
12 <description>This data product is derived from preprocessing and calibration of the CHAMP ASC measurements. It comprises the attitude measurement of the ASC located on the satellite boom.</description>
13 </descriptions>
14 <formats>
15 <format>Digital</format>
16 </formats>
17 <publicationInfo>
18 <publicationPlace>Potsdam, Germany</publicationPlace>
19 </publicationInfo>
20 <identifiers>
21 <identifier>
22 <value>O&value</value>
23 <url>
24 <value>http://www.w3.org/2001/XMLSchema-instance</value>
25 </url>
26 </identifiers>
27 </resource>
    
```

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <!-- XML Spy v2005 rel. 3 U bearbeitet (http://www.altova.com) von Jan Brase (Technische Informationsbibliothek) -->
3 <resource xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns="http://schema.datacite.org/meta" />
4 <DOI>10.5080/est2-DOI</DOI>
5 <creators>
6 <creator>Jan Brase</creator>
7 </creators>
8 <titles>
9 <title>Preprocessed CHAMP Boom Advanced Stellar Compass Data</title>
10 </titles>
11 <descriptions>
12 <description>This data product is derived from preprocessing and calibration of the CHAMP ASC measurements. It comprises the attitude measurement of the ASC located on the satellite boom.</description>
13 </descriptions>
14 <formats>
15 <format>Digital</format>
16 </formats>
17 <publicationInfo>
18 <publicationPlace>Potsdam, Germany</publicationPlace>
19 </publicationInfo>
20 <identifiers>
21 <identifier>
22 <value>O&value</value>
23 <url>
24 <value>http://www.w3.org/2001/XMLSchema-instance</value>
25 </url>
26 </identifiers>
27 </resource>
    
```

DataCite Metadata Store

Datensatz

Registriere neuen Datensatz
Zeige alle Datensätze
Suche nach DOI

Zeige

API Dokumentation

▼ **Registriere neuen Datensatz**

DOI-Latenzzeit: Bitte beachten Sie, dass es bis zu 24 Stunden dauern kann, bis ein DOI-Upload weltweit verteilt wurde. Neue DOIs sollten bereits nach etwa 5 Minuten auflosbar sein.

Für Testzwecke bitte ausschließlich unser spezielles Testpräfix 10.5072 nutzen.

DOI:

URL:

XML Upload: Keine Datei ausgewählt.

Bitte eine XML-Datei auswählen. Diese muss ein Schema benutzen, dessen URL folgendermaßen beginnt: <http://schema.datacite.org/meta/>

XML:

Katalogdaten/ Suchindex

Startseite | Sprache: |

GetInfo **TIB** GERMAN NATIONAL LIBRARY OF SCIENCE AND TECHNOLOGY

Home
Order without search
Collective search
Subject search
A-Z list
Computer Science
Chemistry
Mathematics
Physics
Engineering
TIB-Catalogue Search
Watch list
Prices
MyTIB.de
English

Detail view

TIK: SAF 00 Main-hole downtime logging data phase 1.2 1994-21239, year: 2004

Author(s): SAF 00

Published in: 2004

Publisher: GeoForschungsZentrum Potsdam(GFZ)(Potsdam, Germany)

Document type: Research Data

Language: English

DOI: 10.15144/GFZ/DCC081121

Abstract

GFZ 00 is motivated by the need to answer fundamental questions about the physics and chemical processes controlling faulting and earthquake generation within a major plate-bounding fault. SAF 00 will drill and instrument an intermediate-scale across the San Andreas Fault Zone to a depth of 12 km, targeting a repeating microearthquake source. The drill site is located west of the vertical San Andreas Fault on a segment of the fault that moves through a combination of aseismic creep and repeating microearthquakes. It lies at the advance northern end of the rupture zone of the 1982, magnitude 6.7 Parkfield earthquake, the most recent in a series of events that have ruptured the fault the times since 1957. This data set contains event rate geophysical waveform logging data from 1984-21239 measured (earthquake source location, which is 3.65m above ground level). This dataset is identifier 10.15144/GFZ/DCC081121.

Back | Add to watchlist | Go to data

Verteigert: 18 • 20167 Hannover • +49(0)51-76289-89

TIB

TECHNISCHE
INFORMATIONSBIBLIOTHEK

Vielen Dank für Ihre Aufmerksamkeit!

Fragen?

