GEOBONN 2018 Living Earth

Workshop

Publication of Research Data

associated with peer-reviewed research articles.

Andreas Hübner
Deutsches GeoForschungsZentrum GFZ

Inke Achterberg
Georg-August-Universität Göttingen

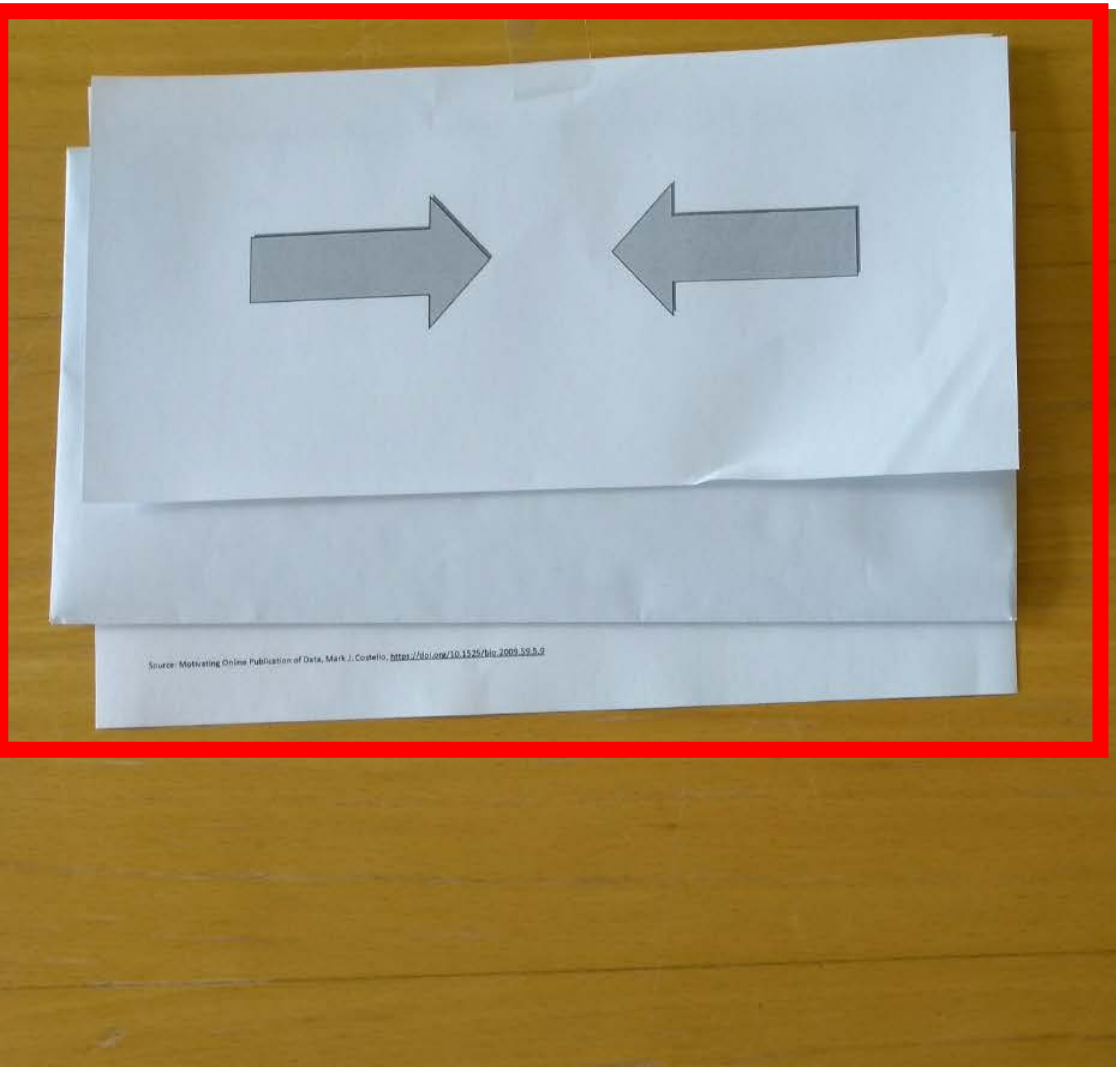
FID GEO
Fachinformationsdienst
Geowissenschaften der festen Erde

This Presentation is licensed under a Creative Commons Namensnennung 4.0 International Licence.

Agenda

- Introduction
- Why publish research data?
- How to publish?
- 5 min Break / 3 min video clip
- Licences
- GFZ Data Services: Metadata Editor
- Wrap up

Funded by DFG


Source: Motivating Online Publication of Data, Mark J. Costello, https://doi.org/10.1525/bio.2009.59.5.9

Forschungs daten
Publizieren
FIDGEO
Fachinformationsdienst Geowissenschaften der festen Erde
gefördert durch
DFG
E-PUBLIZIEREN
FORSCHUNGSDATEN
DIGITALISIERUNG
FID GEO
What
Why

Forschungs daten
Publizieren
FIDGEO
Fachinformationsdienst
Geowissenschaften der festen Erde
gefördert durch
DFG Deutsche Forschungsgemeinschaft
E-PUBLIZIEREN
FORSCHUNGSDATEN
DIGITALISIERUNG
FIDGEO
FIDGEO

Feedback to the workshop "Publication of research data related to scientific articles in the geosciences", GeoBonn 2018

What do you think about the **Workshop in general**?	very good	good	So, so	less good	not good
Selection of contents					
Complexity of the content					
Useful for your practical work					
Interest in the topic					
Excercises					

What do you think about the Workshoptopic "**Why publish research data**"?	very good	good	So, so	less good	not good
Selection of contents					
Complexity of the content					
Useful for your practical work					
Interest in the topic					

What do you think about the Workshoptopic "**How to publish**"?	very good	good	So, so	less good	not good
Selection of contents					

What do you think about the Workshoptopic "**Licences**"?	very good	good	So, so	less good	not good
Selection of contents					
Complexity of the content					
Useful for your practical work					
Interest in the topic					
What do you think about the Workshoptopic "**Metadata Editor**"?	very good	good	So, so	less good	not good
Selection of contents					
Complexity of the content					
Useful for your practical work					
Interest in the topic					
Your comments					

Participants Workshop „Publication of Research Data“, 05 Sept. 2018, GeoBonn

Name	Institution	Email
1		
5		

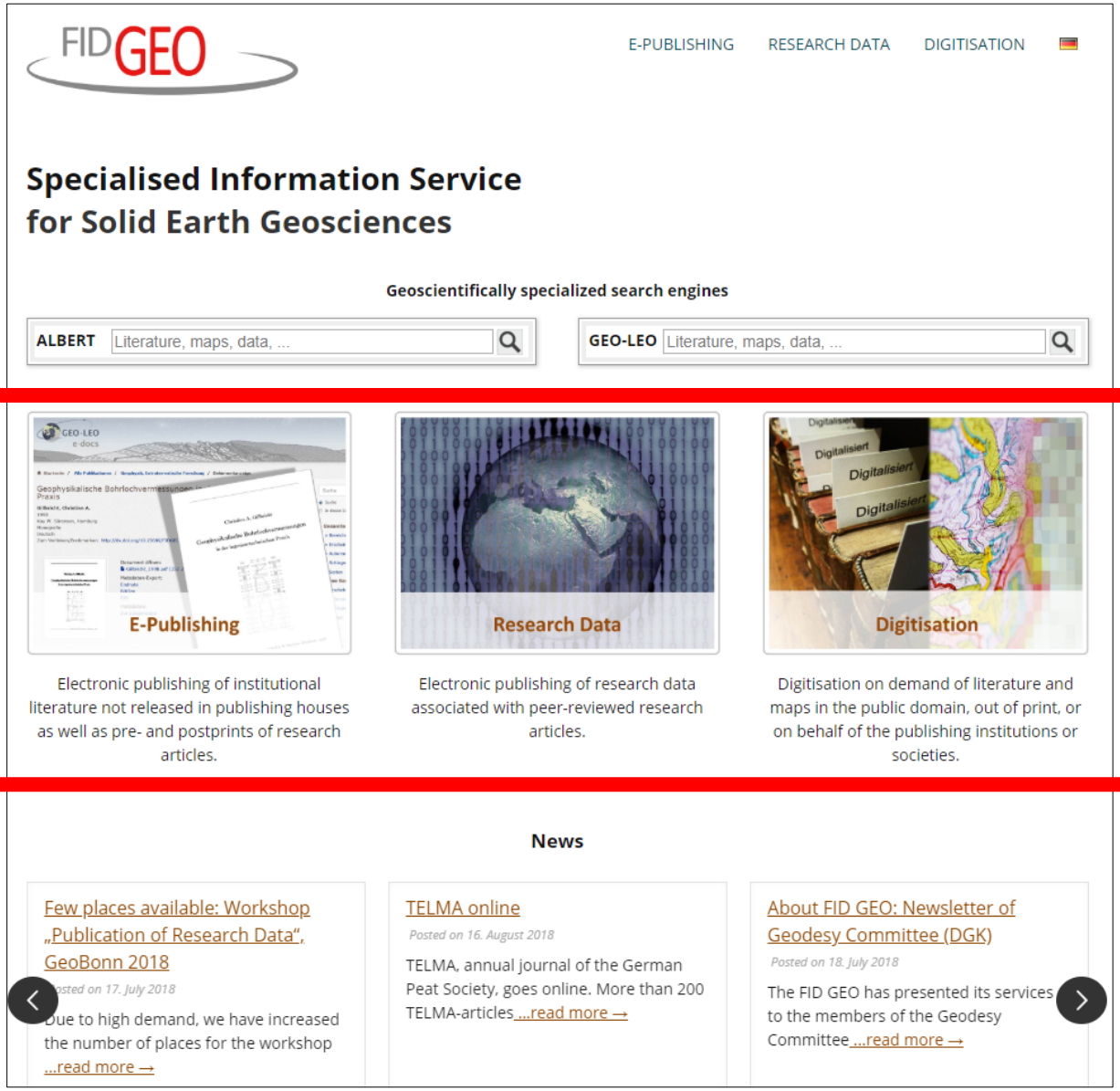
DFG Fachinformationsdienste (FID)

- Nationwide initiative to improve information infrastructures in research institutions
- FID GEO Partner:

SUB | NIEDERSACHSISCHE STAATS- UND UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

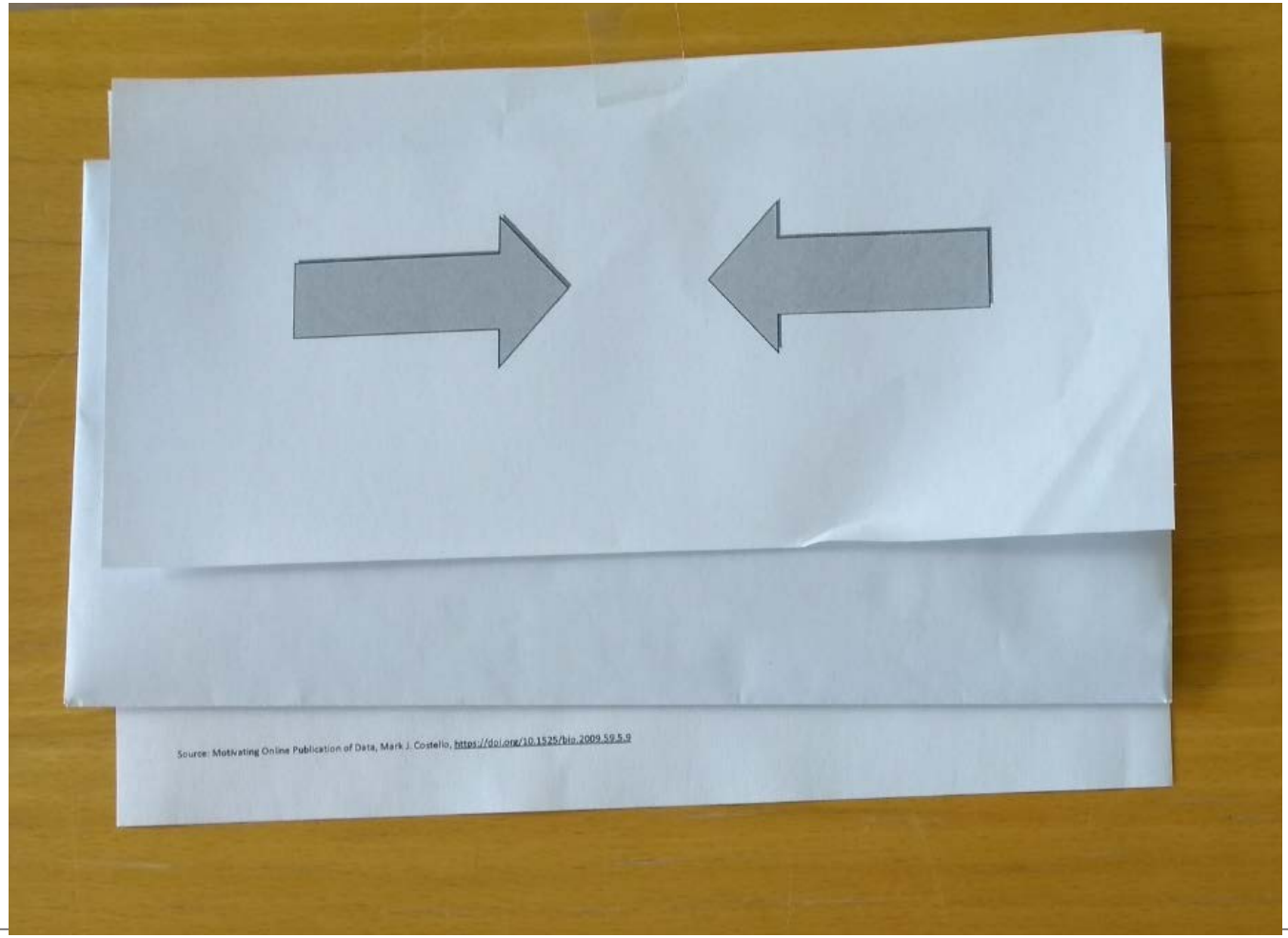
GFZ Helmholtz-Zentrum Potsdam

- Start June 2016, Website: **fidgeo.de**


FIDGEO
E-PUBLISHING
RESEARCH DATA
DIGITISATION
Specialised Information Service
for Solid Earth Geosciences
Geoscientifically specialized search engines
ALBERT
Literature, maps, data, ...
GEO-LEO
Literature, maps, data, ...
E-Publishing
Research Data
Digitisation
Electronic publishing of institutional literature not released in publishing houses as well as pre- and postprints of research articles.
Electronic publishing of research data associated with peer-reviewed research articles.
Digitisation on demand of literature and maps in the public domain, out of print, or on behalf of the publishing institutions or societies.
News
Few places available: Workshop „Publication of Research Data", GeoBonn 2018
Posted on 17. July 2018
Due to high demand, we have increased the number of places for the workshop ...read more →
TELMA online
Posted on 16. August 2018
TELMA, annual journal of the German Peat Society, goes online. More than 200 TELMA-articles ...read more →
About FID GEO: Newsletter of Geodesy Committee (DGK)
Posted on 18. July 2018
The FID GEO has presented its services to the members of the Geodesy Committee ...read more →

fidgeo.de

Introduce yourself



Funded by DFG

Introduce yourself

- name
- working place
- research topic
- career level
- …and: the „personal thing(s)“

Funded by DFG

Agenda

Introduction

Why publish research data?

How to publish?

5 min Break / 3 min video clip

Licences

GFZ Data Services: Metadata Editor

Wrap up

Funded by DFG

Good Research Practice

Figure: The University of Sheffield, CC-BY 3.0

Good Research Practice

Figure: The University of Sheffield, CC-BY 3.0

Political perspective

"…all researchers should be able to deposit, access and analyse scientific data across disciplines and at the global scale,…

G7 Science Ministers Communiqué, Turin

http://www.g8.utoronto.ca/science/2017-science-communique.html

Funded by DFG

Funder perspective

"…open access is the default setting for research data generated in Horizon 2020."

EU Horizon2020
OPEN RESEARCH DATA PILOT

EC's Guide on Open Access to Scientific Publications and Research Data in Horizon 2020 (updated August 25, 2016)
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

Funder perspective

"…research data should be made available as soon as possible."

German Science Foundation
Guidelines for handling research data
Appeal to use open licences

http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf

http://www.dfg.de/foerderung/info_wissenschaft/2014/info_wissenschaft_14_68/

Publisher perspective

Supporting data must be made available to editors and peer reviewers at the time of submission for the purposes of evaluating the manuscript. All manuscripts reporting original research published in Nature journals must include a data availability statement …

nature
International journal of science

After publication, all data and materials necessary to understand, assess, and extend the conclusions of the manuscript must be available to any reader of a *Science* Journal.

Science

https://www.nature.com/authors/policies/availability.html
http://www.sciencemag.org/authors/science-journals-editorial-policies

Funded by DFG

Publisher perspective

Instructions for Authors

Authors should therefore archive the primary data to their papers at PANGAEA (http://www.pangaea.de) or a similar information system.

Instructions for Authors

We encourage research data to be archived in data repositories wherever possible.

https://www.springer.com/earth+sciences+and+geography/geology/journal/531
https://www.springer.com/earth+sciences+and+geography/paleontology/journal/12542

COPDESS Coalition for Publishing Data in the Earth and Space Sciences

http://www.copdess.org/

GOAL

OPEN DATA in the EARTH and SPACE SCIENCES

43 SIGNATURES (Dec. 2017)

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.

Enabling FAIR Data

15 September 2017

AGU is convening a partnership in the Earth and space science community to develop the standards to connect researchers, publishers, and data repositories.

The partnership currently includes **AGU**, Earth Science Information Partners (**ESIP**) and Research Data Alliance (**RDA**), and has support from the Proceedings of the National Academy of Sciences, Nature, Science, National Computational Infrastructure, AuScope, the Australian National Data Service, and the Center for Open Science.

https://eos.org/editors-vox/enabling-findable-accessible-interoperable-and-reusable-data

Obstacles

Rights of other scientist
with Co-authorship, all authors can only jointly decide on the reuse or publication.

Secrecy agreements
In third-party funded projects or by instruction of employer.

Patents
When the research data describe a patentable invention and this invention is to be filed for a patent.

Personalised data
Must be anonymised before publication.

Concerns

Lack of time to curate data
Data are only useful if they are understandable, adequate documentation is time-consuming.

Personal investment
Data collection takes time and trouble, other researcher are conceived as freeloaders.

Scooping
Fear of competition and resulting reduced publication opportunities.

Errors
Fear that errors being found in the data.

Advantages

Individual scientist

- Additional publications
- Greater citation rate
- Wider recognition among peers
- Invitations to meetings
- Invitations to collaborate
- Invitations to provide consultancy

- Creators of data are known from citation and so are contactable for more information
- Citation of data sources adds authority that indicates their quality

Mark J. Costello, Motivating Online Publication of Data https://doi.org/10.1525/bio.2009.59.5.9

Funded by DFG

Advantages

Individual scientist

- Additional publications
- Greater citation rate
- Wider recognition among peers
- Invitations to meetings
- Invitations to collaborate
- Invitations to provide consultancy

- Creators of data are known from citation and so are contactable for more information
- Citation of data sources adds authority that indicates their quality

Mark J. Costello, Motivating Online Publication of Data https://doi.org/10.1525/bio.2009.59.5.9

Editors, peer reviewers

- Independent verification and qualify-cation of research findings is possible

Scientific community

- Data can be reused for similar and new purposes
- Data can be integrated with other data to create new data resources

Funding agencies

- Better financial return from research investment as a data can be used again

Society

- Better science

Funded by DFG

Choose and rank the five most important advantages

Individual scientist

- ☐ Additional publications
- ☐ Greater citation rate
- ☐ Wider recognition among peers
- ☐ Invitations to meetings
- ☐ Invitations to collaborate
- ☐ Invitations to provide consultancy

- ☐ Creators of data are known from citation and so are contactable for more information
- ☐ Citation of data sources adds authority that indicates their quality

Editors, peer reviewers

- ☐ Independent verification and qualification of research findings

Scientific community

- ☐ Data can be reused for similar and new purposes
- ☐ Data can be integrated with other data to create new data resources

Funding agencies

- ☐ Better financial return from research investment

Society

- ☐ Better science

Funded by DFG

Choose and rank the five most important advantages

Individual scientist

- ☐ 2 Additional publications
- ☐ Greater citation rate
- ☐ 1 Wider recognition among peers
- ☐ Invitations to meetings
- ☐ 4 Invitations to collaborate
- ☐ Invitations to provide consultancy

- ☐ Creators of data are known from citation and so are contactable for more information
- ☐ Citation of data sources adds authority that indicates their quality

Editors, peer reviewers

- ☐ Independent verification and qualification of research findings

Scientific community

- ☐ 3 Data can be reused for similar and new purposes
- ☐ Data can be integrated with other data to create new data resources

Funding agencies

- ☐ Better financial return from research investment

Society

- ☐ 5 Better science

Funded by DFG

Agenda

Introduction

Why publish research data?

How to publish?

5 min Break / 3 min video clip

Licences

GFZ Data Services: Metadata Editor

Wrap up

Funded by DFG

Contents lists available at ScienceDirect

Tectonophysics

journal homepage: www.elsevier.com/locate/tecto

Properties of granular analogue model materials: A community wide survey

M. Klinkmüller [a], G. Schreurs [a,1], M. Rosenau [b], H. Kemnitz [b]

[a] Institute of Geological Sciences, University of Bern, Baltzerstrasse 1 +3, CH-3012 Bern, Switzerland
[b] Helmholtz-Zentrum Potsdam, GFZ Deutsches GeoForschungsZentrum, Telegrafenberg, D-14473 Potsdam, German

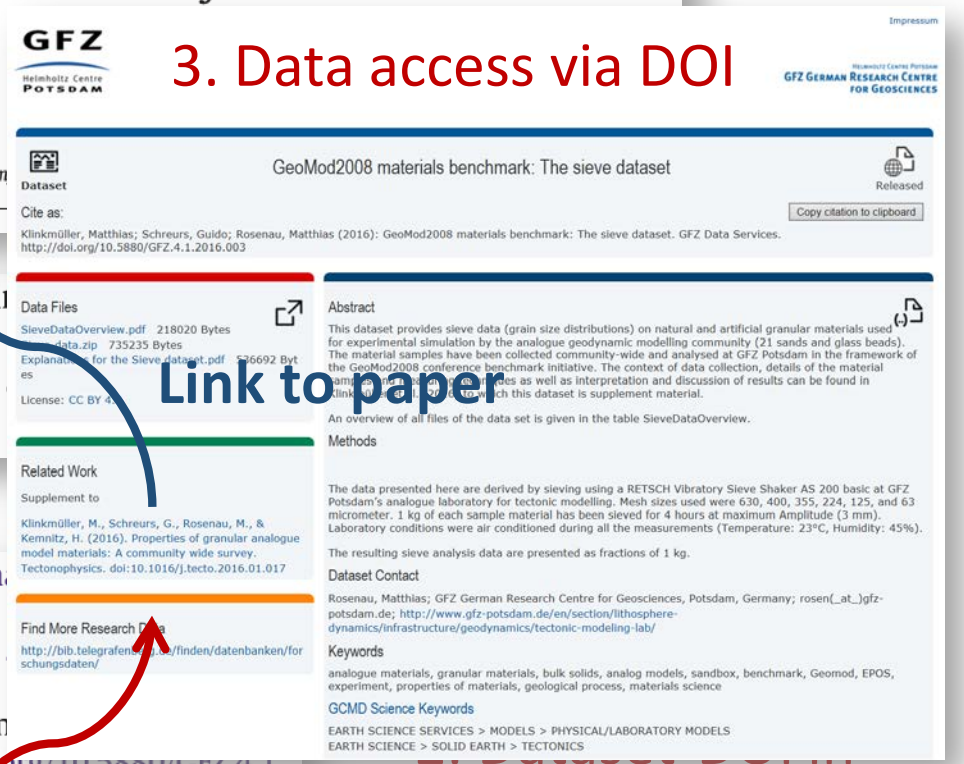
sented as grain size distribution curves, in which particle grai
plotted against cumulative weight percentage (Fig. 2).

The original sieve data have been published open access
available in Klinkmüller et al. (2016b).

References

Heilbronner, R., Keulen, N., 2006. Grain size and grain shape an
Tectonophysics 427, 199–216.
Hubbert, M.K., 1951. Mechanical basis for certain familiar geologic
Am. Bull. 62, 1259–1273.
Klinkmüller, M., Schreurs, G., Rosenau, M., 2016a. GeoMod2008 m
The ring shear test data set. GFZ Data Services. http://dx.doi.org/10.5880/GFZ.4.1.
2016.002.
Klinkmüller, M., Schreurs, G., Rosenau, M., 2016b. GeoMod2008 materials benchmark:
The sieve data set. GFZ Data Services. http://dx.doi.org/10.5880/GFZ.4.1.2016.003.
Klinkmüller, M., Kemnitz, H., Schreurs, G., Rosenau, M., 2016c. GeoMod2008 materials
benchmark: The SEM image data set. GFZ Data Services. http://dx.doi.org/10.5880/
GFZ.4.1.2016.004.

3. Data access via DOI



Link to paper

2. Dataset DOI in the References

Folie von Kirsten Elger

Repositories

re3data.org

REGISTRY OF RESEARCH DATA REPOSITORIES

Repository = (online accessible) database for the recording and publication of research data, texts and other digital objects[1]

[1] Einstieg ins Forschungsdatenmanagement in den Geowissenschaften, https://doi.org/10.2312/lis.14.01

Funded by DFG

Repositories

Institutional Repository

- members of the institution
- many disciplines

Deposit Once

Repository for Research Data and Publications

mediaTUM – der Medien- und Publikations-server der Technischen Universität München

Publizieren, archivieren und recherchieren Sie Hochschulschriften und andere wissenschaftliche Publikationen, Bild- und Videokollektionen sowie Forschungsdaten. Derzeit sind auf mediaTUM **mehr als 230.000 Datensätze** öffentlich zugänglich und werden **von vielen Diensten indexiert**, darunter die Deutsche Nationalbibliothek und Google Scholar.

Domain-specific Repository

- Researchers worldwide
- discipline-specific

GFZ Data Services

PANGAEA.

Data Publisher for Earth & Environmental Science

EarthChem

Generic Repository

- Researchers worldwide
- all disciplines

zenodo

figshare

- Domain-specific metadata
- Connected to domain-specific Data portals
- Better quality-control
- extra services, e.g. integration of IGSN

Funded by DFG

Description of data

Ensuring that data is „independently understandable“ is crucial.

Funded by DFG

Data Supplements

Links to data files

Link to journal article

Abstract

Data Journals

Peer-reviewed articles with the description of datasets, data collections, data infrastructures, etc.

No Inter-pretation!

Folie von Kirsten Elger

Agenda

Introduction

Why publish research data?

How to publish?

5 min Break / 3 min video clip

Licences

GFZ Data Services: Metadata Editor

Wrap up

Funded by DFG

xtranormal

Data Sharing and Management Snafu in 3 Short Acts

https://www.youtube.com/watch?v=66oNv_DJuPc

Agenda

Introduction

Why publish research data?

How to publish?

5 min Break / 3 min video clip

Licences

GFZ Data Services: Metadata Editor

Wrap up

Funded by DFG

Copyright

- Machine-generated and unprocessed raw data are not copyrightable.
- For most other data you should assume that data is protected by intellectual property rights.

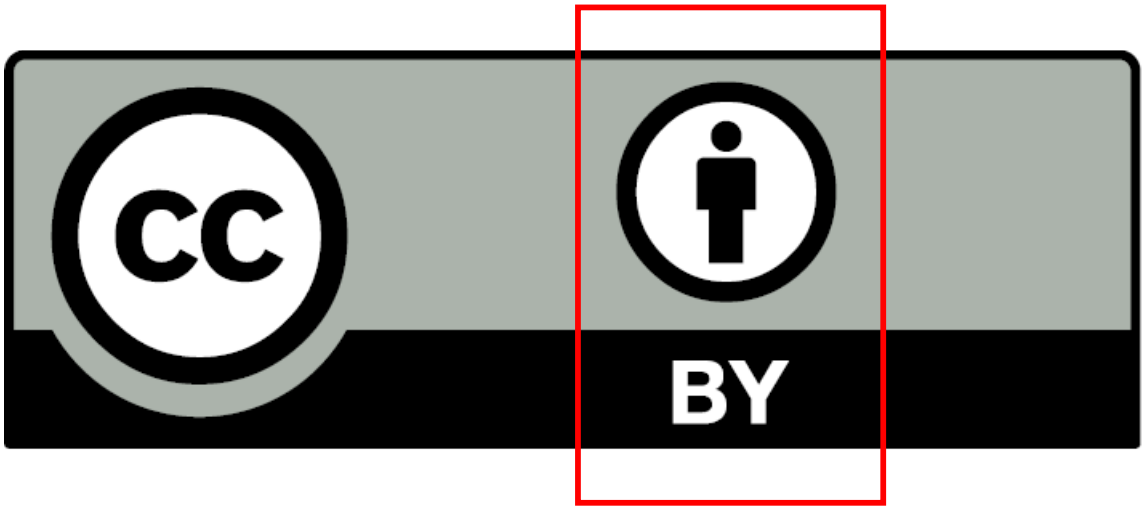
Übersetzt aus: Gutachten zu den rechtlichen Rahmenbedingungen des Forschungsdatenmanagements (2018)
https://tu-dresden.de/gsw/jura/igewem/jfbimd13/ressourcen/dateien/publikationen/DataJus_Zusammenfassung_Gutachten_12-07-18.pdf

Funded by DFG

Reuse of data

https://creativecommons.org/

Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung 4.0 International Lizenz.


BY

Attribution (BY): You must give credit to author(s) or licencesor

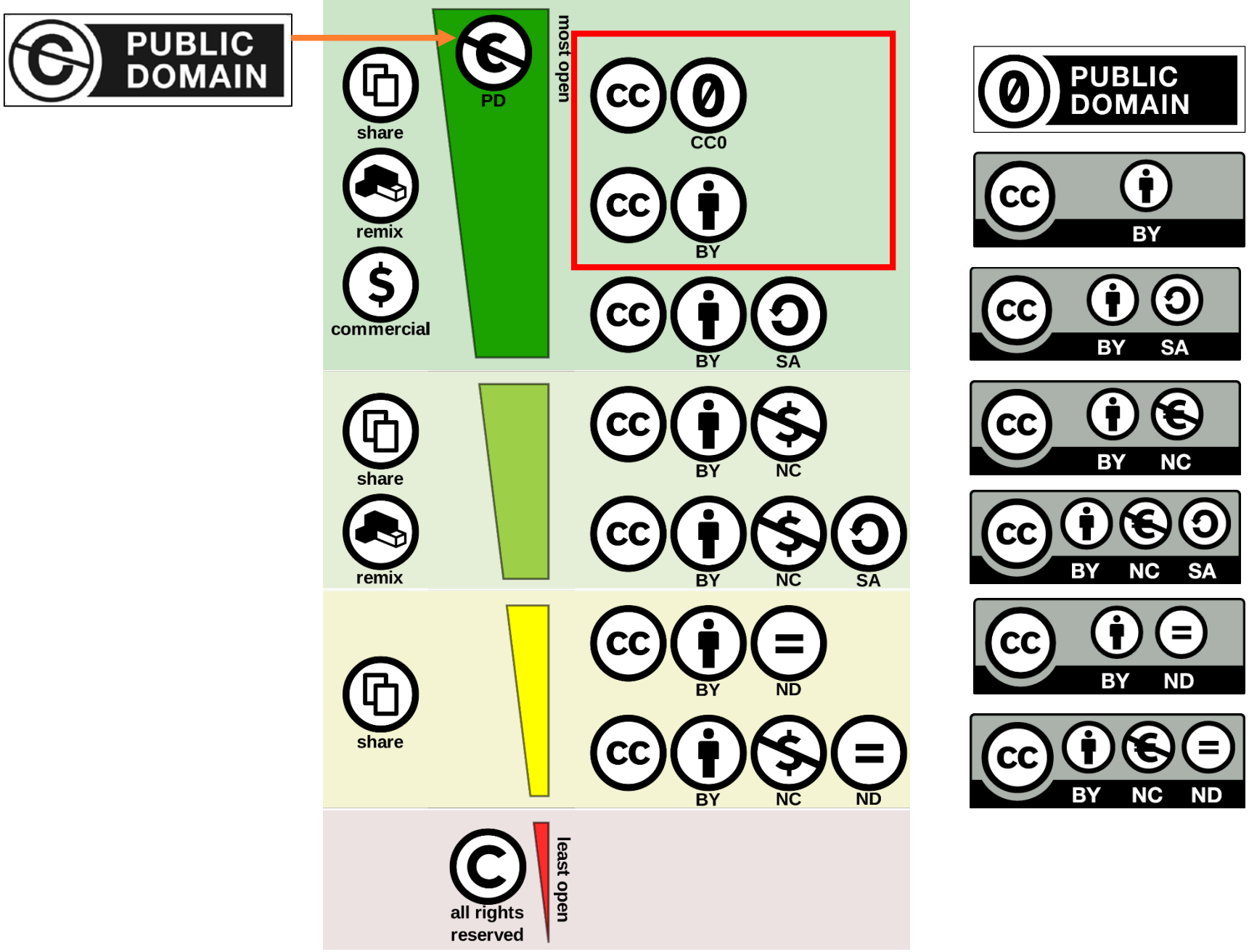
Non-Commercial (NC): reuse not for commercial purposes

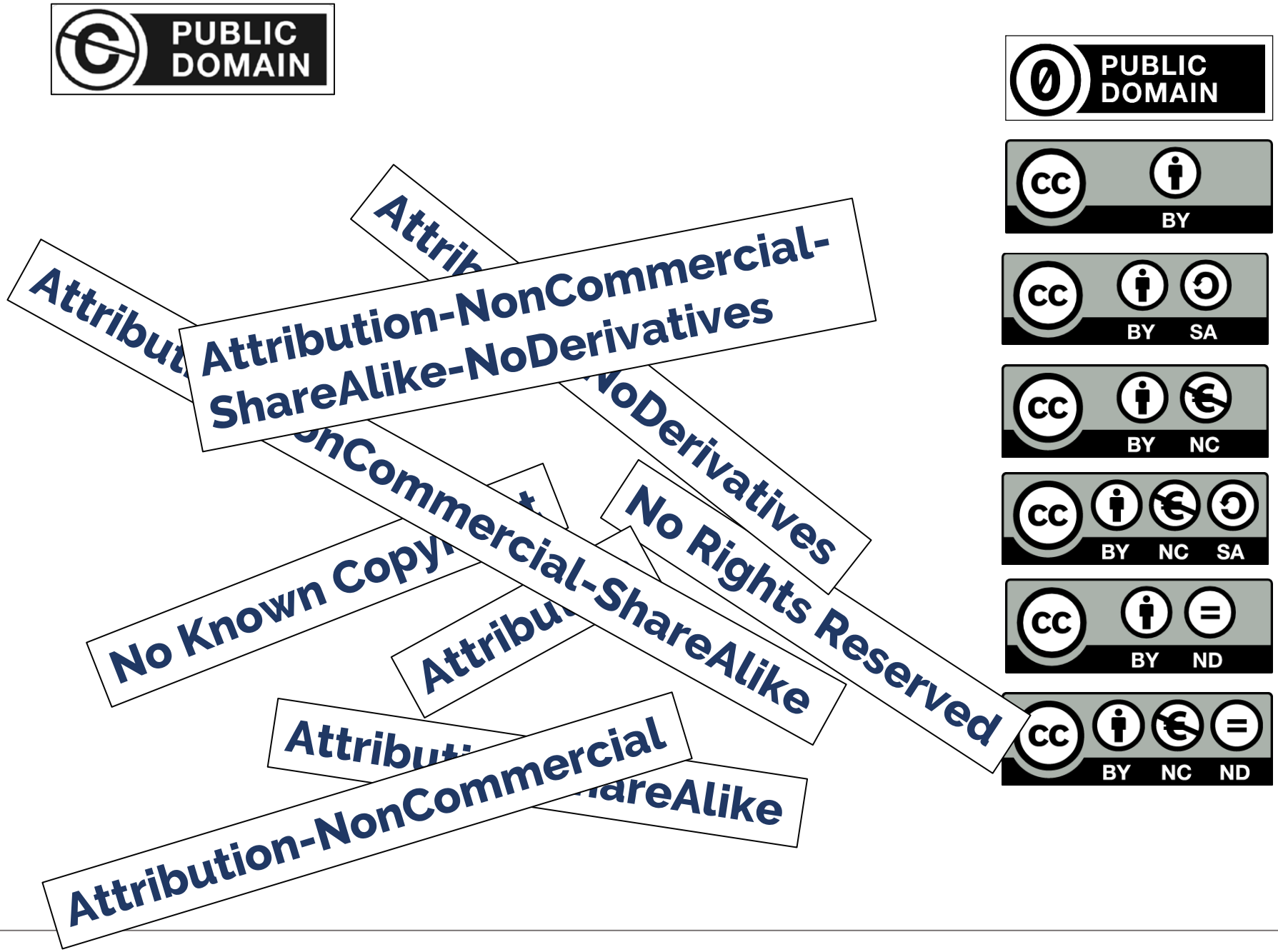
Share Alike (SA): share only with an identical licence

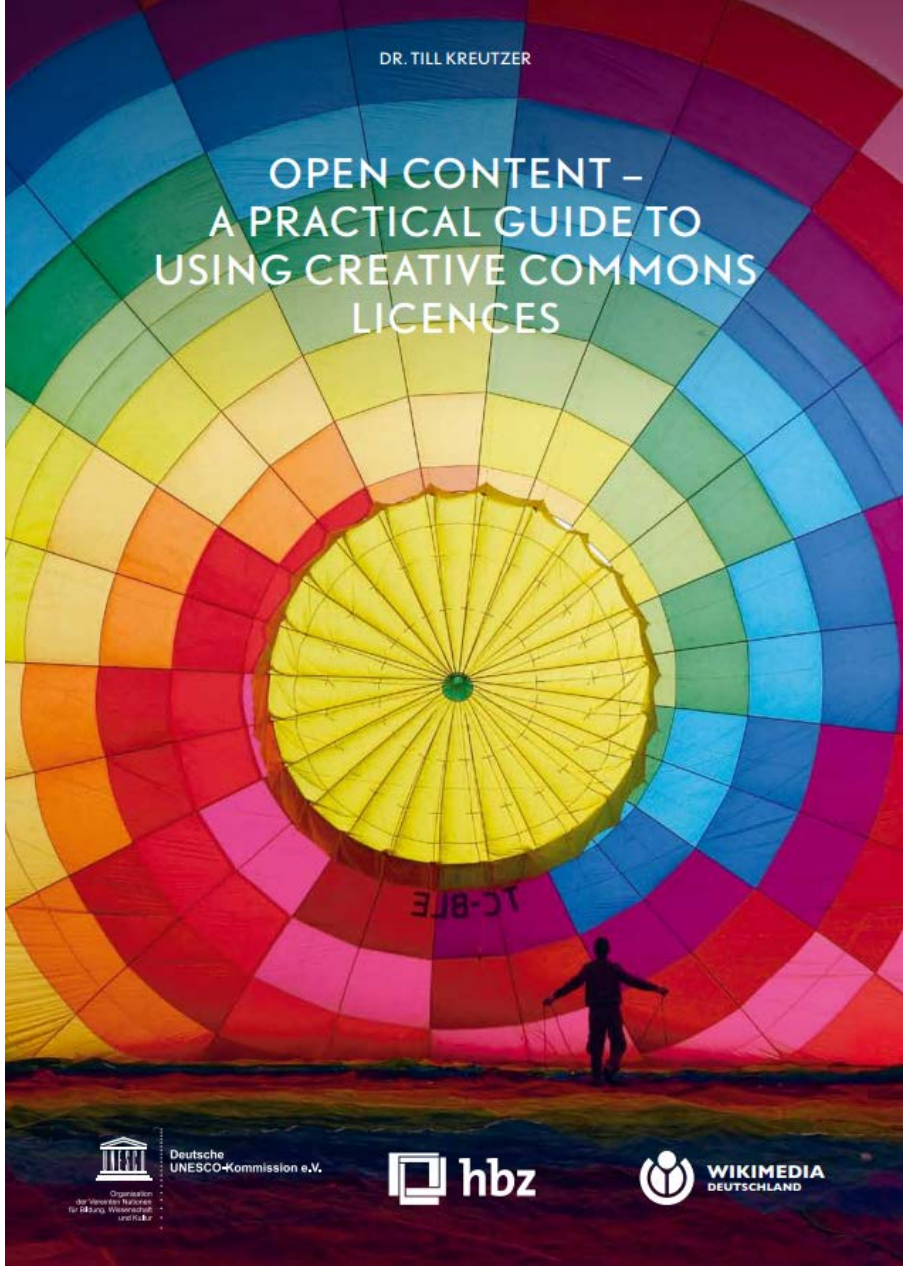
No Derivatives (ND): work must not be changed.

No rights reserved (0): No Rights Reserved

Public Domain (PD): no known copyright restrictions


PUBLIC DOMAIN
most open
PD
share
remix
commercial
CC0
BY
BY SA
BY NC
BY NC SA
BY ND
BY NC ND
all rights reserved
least open


PUBLIC DOMAIN
Attribution-NonCommercial-ShareAlike-NoDerivatives
No Known Copyright
No Rights Reserved
Attribution-NonCommercial-ShareAlike
Attribution-NonCommercial
PUBLIC DOMAIN
BY
BY SA
BY NC
BY NC SA
BY ND
BY NC ND



https://irights.info/wp-content/uploads/2014/11/Open_Content_A_Practical_Guide_to_Using_Open_Content_Licences_web.pdf

Agenda

Introduction

Why publish research data?

How to publish?

5 min Break / 3 min video clip

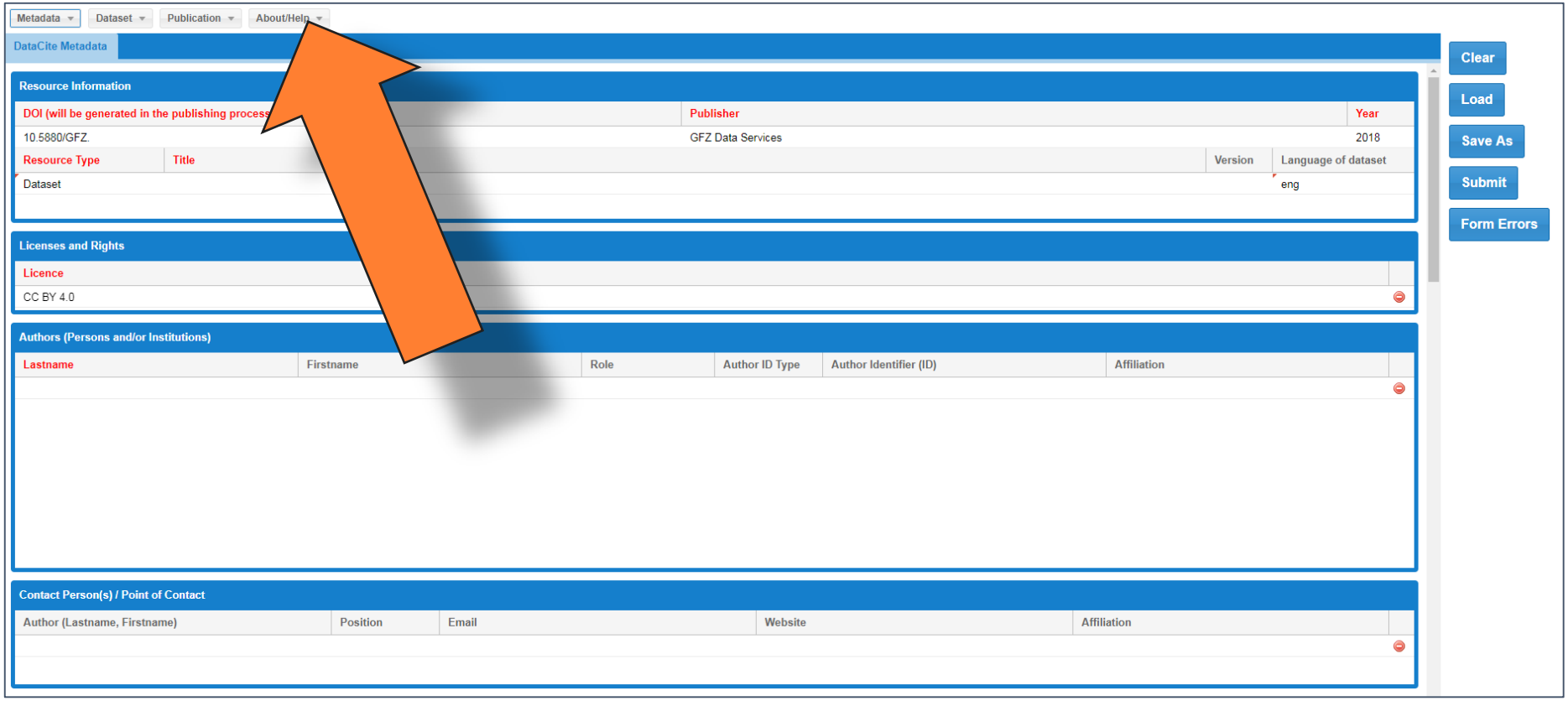
Licences

GFZ Data Services: Metadata Editor

Wrap up

Funded by DFG

Metadata-Editor



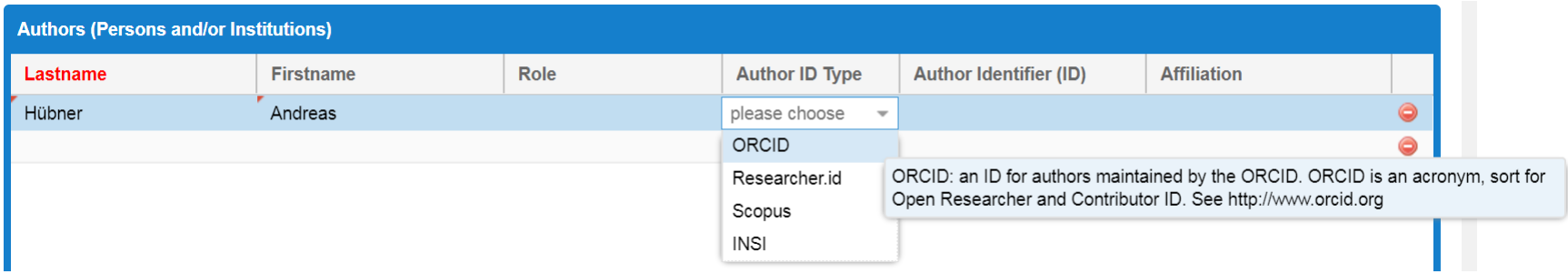
http://pmd.gfz-potsdam.de/panmetaworks/metaedit/

Metadata-Editor

1. ORCID
2. Contributors
3. Related work
4. Embargo

Funded by DFG

ORCID


Authors (Persons and/or Institutions)
Lastname
Firstname
Role
Author ID Type
Author Identifier (ID)
Affiliation
Hübner
Andreas
please choose
ORCID
Researcher.id
Scopus
INSI
ORCID: an ID for authors maintained by the ORCID. ORCID is an acronym, sort for Open Researcher and Contributor ID. See http://www.orcid.org

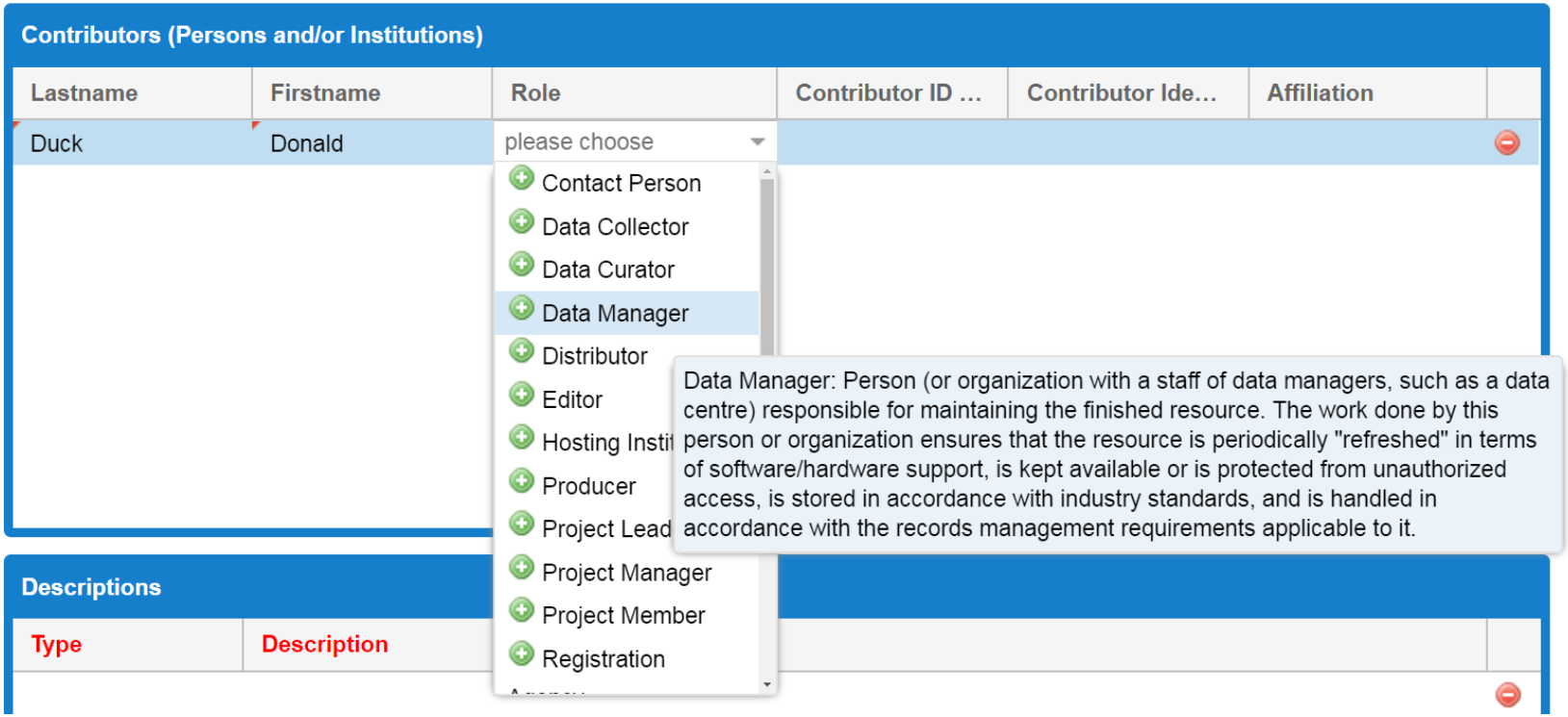
ORCID

https://orcid.org
http://www.orcid-de.org/

The **Open Researcher and Contributor ID** distinguishes you from every other researcher.

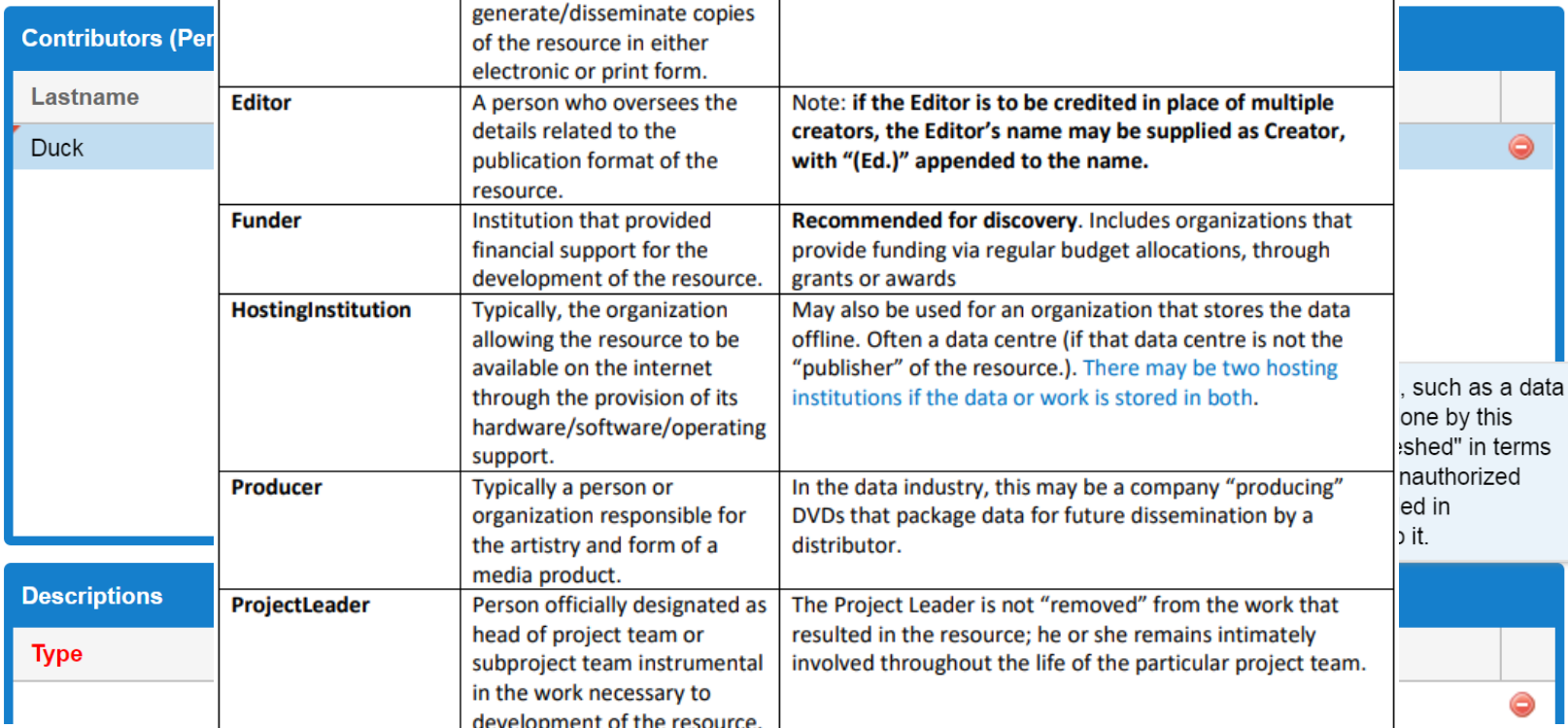
Contributors

With “contributor” you have the possibility to acknowledge additional persons or institutions related to the dataset but which you would normally not mention as authors. These are not named in the citation, but always related with the dataset and searchable as all the other metadata fields.



Contri

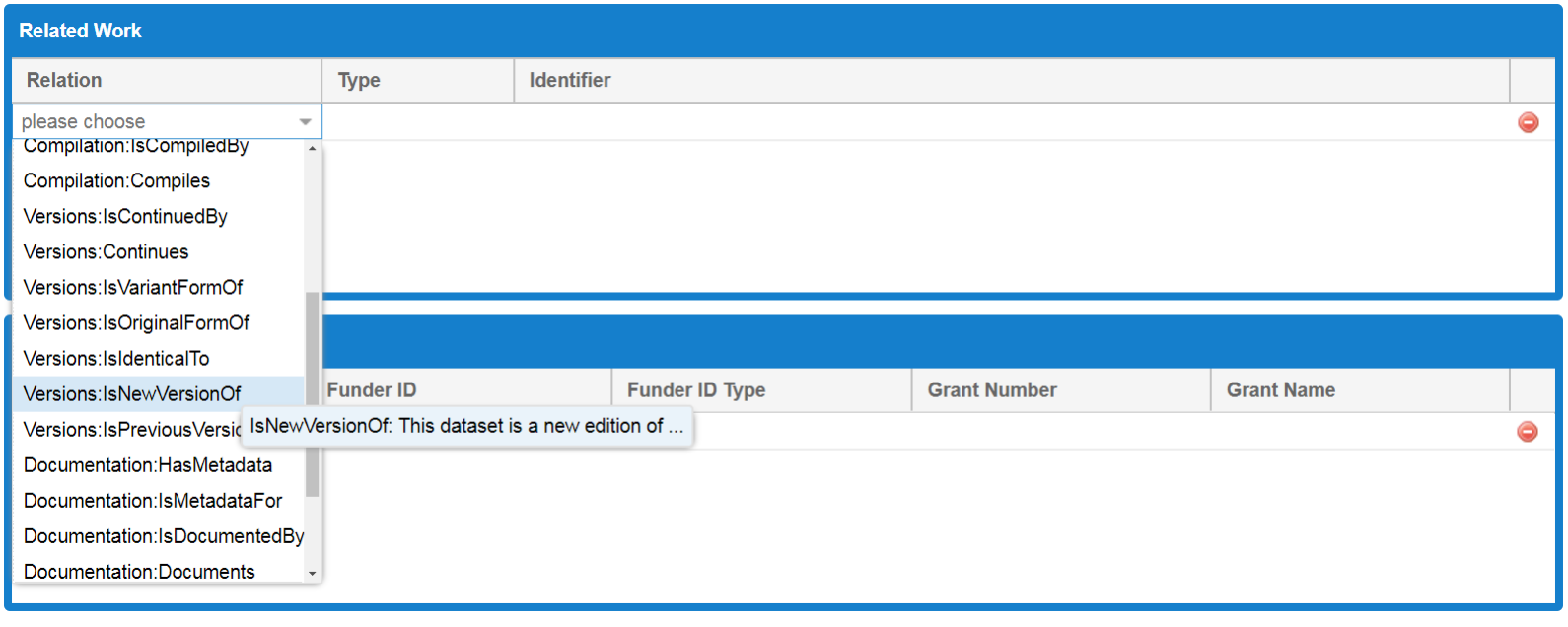
With "contr ...sons or
institutions ...tion as authors.
These are r ...nd searchable
as all the o



DataCurator	Person tasked with reviewing, enhancing, cleaning, or standardizing metadata and the associated data submitted for storage, use, and maintenance within a data center or repository	While the "DataManager" is concerned with digital maintenance, the DataCurators' role encompasses quality assurance focused on content and metadata. This includes checking whether the submitted dataset is complete, with all files and components as described by submitter, whether the metadata is standardized to appropriate systems and schema, whether specialized metadata is needed to add value and ensure access across disciplines, and determining how the metadata might map to search engines, database products, and automated feeds.
DataManager	Person (or organization with a staff of data managers, such as a data centre) responsible for maintaining the finished resource.	The work done by this person or organization ensures that the resource is periodically "refreshed" in terms of software/hardware support, is kept available or is protected from unauthorized access, is stored in accordance with industry standards, and is handled in accordance with the records management requirements applicable to it.
Distributor	Institution tasked with responsibility to generate/disseminate copies of the resource in either electronic or print form.	Works stored in more than one archive/repository may credit each as a distributor.
Editor	A person who oversees the details related to the publication format of the resource.	Note: **if the Editor is to be credited in place of multiple creators, the Editor's name may be supplied as Creator, with "(Ed.)" appended to the name.**
Funder	Institution that provided financial support for the development of the resource.	**Recommended for discovery**. Includes organizations that provide funding via regular budget allocations, through grants or awards
HostingInstitution	Typically, the organization allowing the resource to be available on the internet through the provision of its hardware/software/operating support.	May also be used for an organization that stores the data offline. Often a data centre (if that data centre is not the "publisher" of the resource.). There may be two hosting institutions if the data or work is stored in both.
Producer	Typically a person or organization responsible for the artistry and form of a media product.	In the data industry, this may be a company "producing" DVDs that package data for future dissemination by a distributor.
ProjectLeader	Person officially designated as head of project team or subproject team instrumental in the work necessary to development of the resource.	The Project Leader is not "removed" from the work that resulted in the resource; he or she remains intimately involved throughout the life of the particular project team.
ProjectManager	Person officially designated as manager of a project. Project may consist of one or many project teams and sub-teams.	The manager of a project normally has more administrative responsibility than actual work involvement.

Funded by DFG

Related Work



Related Work		
Relation	**IsCitedBy**	indicates that B includes A in a citation (recommended for discovery).
Relation	**Cites**	indicates that A includes B in a citation (recommended for discovery).
Relation	**IsSupplementTo**	indicates that A is a supplement to B (recommended for discovery).
Relation	**IsSupplementedBy**	indicates that B is a supplement to A (recommended for discovery).
Relation	**IsContinuedBy**	indicates A is continued by the work B
Relation	**Continues**	indicates A is a continuation of the work B
Relation	**HasMetadata**	indicates resource A has additional metadata B
Relation	**IsMetadataFor**	indicates additional metadata A for a resource B
Relation	**IsNewVersionOf**	indicates A is a new edition of B, where the new edition has been modified or updated
Relation	**IsPreviousVersionOf**	indicates A is a previous edition of B
Relation	**IsPartOf**	indicates A is a portion of B; may be used for elements of a series (recommended for discovery).
Relation	**HasPart**	indicates A includes the part B (recommended for discovery).
Relation	**IsReferencedBy**	indicates A is used as a source of information by B
Relation	**References**	indicates B is used as a source of information for A
Relation	**IsDocumentedBy**	indicates B is documentation about/ explaining A
Relation	**Documents**	indicates A is documentation about/B
Relation	**IsCompiledBy**	indicates B is used to compile or create A
Relation	**Compiles**	indicates B is the result of a compile or creation event using A
Relation	**IsVariantFormOf**	indicates A is a variant or different form of B, e.g. calculated or calibrated form or different packaging
Relation	**IsOriginalFormOf**	indicates A is the original form of B
Relation	**IsIdenticalTo**	indicates that A is identical to B, for use when there is a need to register two separate instances of the same resource
Relation	**IsReviewedBy**	indicates that A is reviewed by B
Relation	**Reviews**	indicates that A is a review of B
Relation	**IsDerivedFrom**	indicates B is a source upon which A is based
Relation	**IsSourceOf**	indicates A is a source upon which B is based

Embargo

Dates

	Date from	Date to
Created	YYYY-MM-DD	
Embargo until		YYYY-MM-DD
Valid	YYYY-MM-DD	YYYY-MM-DD

The date the resource is made publicly available. Use this field to indicate the end of an embargo period.

Funded by DFG

Agenda

Introduction

Why publish research data?

How to publish?

5 min Break / 3 min video clip

Licences

GFZ Data Services: Metadata Editor

Wrap up

Funded by DFG

Wrap up

- Please provide feedback (3 min)
- Meet us after the workshop

Funded by DFG