

re3data.org

REGISTRY OF RESEARCH DATA REPOSITORIES

Making research data repositories
visible and discoverable

Robert Ulrich – Karlsruhe Institute of Technology

Outline

- Background
- Mission
- Project
- Impact
- Sustainability
- Lessons learned & best practices

Background

- Research data are valuable and ubiquitous
- New technologies facilitate data-intensive science
- Broad discussion about the permanent access to research data
- Increasing requirements from funders to make research data openly available
- Growing demand for trustable and sustainable research data repositories
- Trend: data journals

Background

Funding organizations: Data Policies

Example: european commission

29.3 Open access to research data

[OPTION for actions participating in the open Research Data Pilot: Regarding the digital research data generated in the action ('data'), the beneficiaries must:

- (a) deposit in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate — free of charge for any user — the following:*
 - (i) the data, including associated metadata, needed to validate the results presented in scientific publications as soon as possible;*
 - (ii) other data, including associated metadata, as specified and within the deadlines laid down in the 'data management plan' (see Annex 1);*
- (b) provide information — via the repository — about tools and instruments at the disposal of the beneficiaries and necessary for validating the results (and — where possible — provide the tools and instruments themselves).*

European Commission. (2014). Horizon 2020 Annotated Model Grant Agreements. Version 1.6.2 .Retrieved from http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf

Background

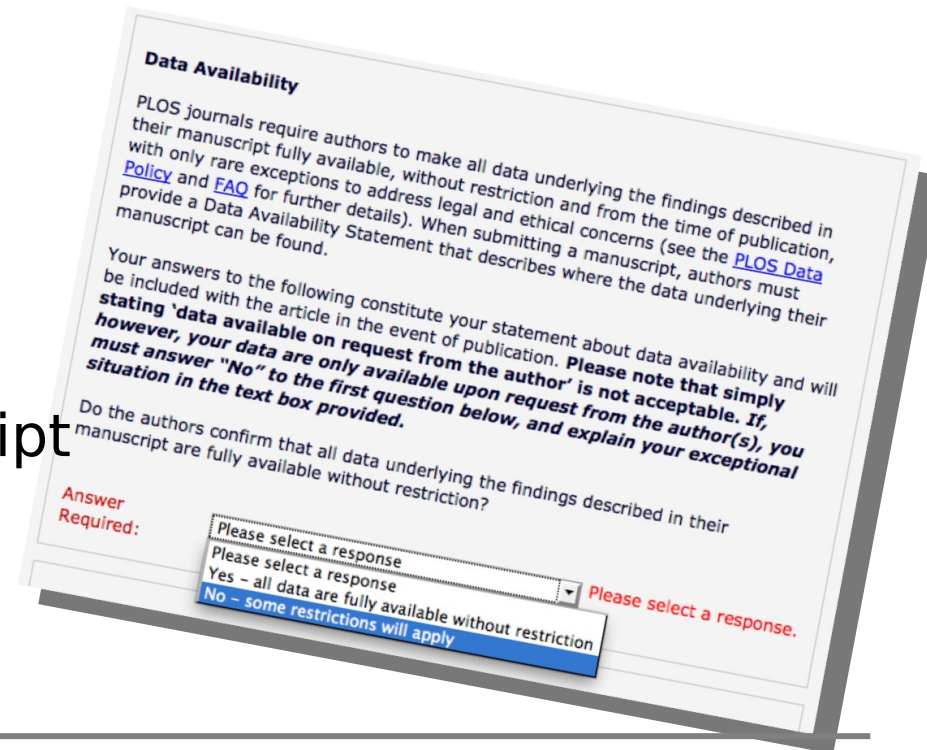
Journals: Data Policies

Example: Nature Publishing Group

“[...] authors are required to make materials, data and associated protocols promptly available to readers without undue qualifications. “

Example: PLOS

“PLOS journals require authors to make all data underlying the findings described in their manuscript fully available without restriction, with rare exception.”



Data Availability

PLOS journals require authors to make all data underlying the findings described in their manuscript fully available, without restriction and from the time of publication, with only rare exceptions to address legal and ethical concerns (see the [PLOS Data Policy](#) and [FAQ](#) for further details). When submitting a manuscript, authors must provide a Data Availability Statement that describes where the data underlying their manuscript can be found.

Your answers to the following constitute your statement about data availability and will be included with the article in the event of publication. **Please note that simply stating 'data available on request from the author' is not acceptable. If, however, your data are only available upon request from the author(s), you must answer "No" to the first question below, and explain your exceptional situation in the text box provided.**

Do the authors confirm that all data underlying the findings described in their manuscript are fully available without restriction?

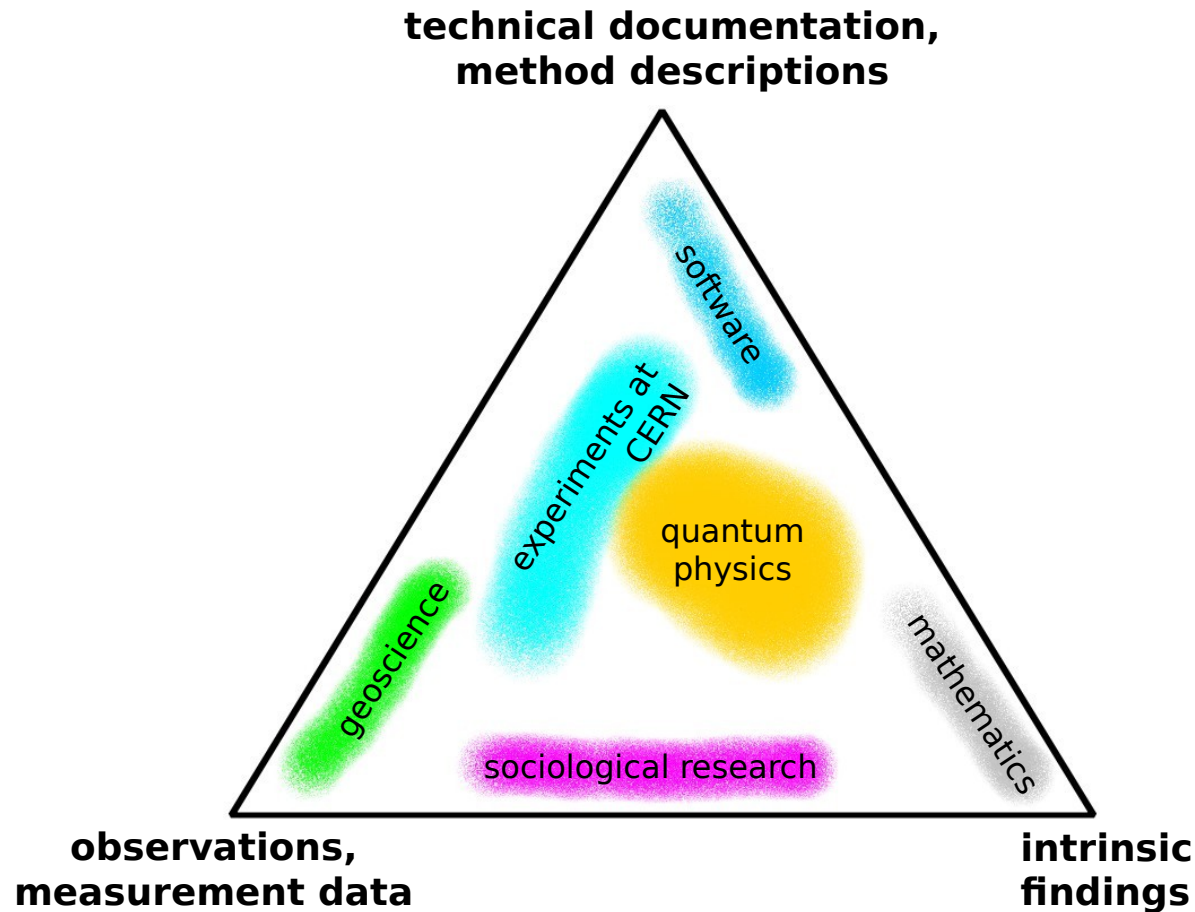
Answer Required:

Please select a response
Please select a response
Yes - all data are fully available without restriction
No - some restrictions will apply

Please select a response.

Background

Generic range of research data (with examples)



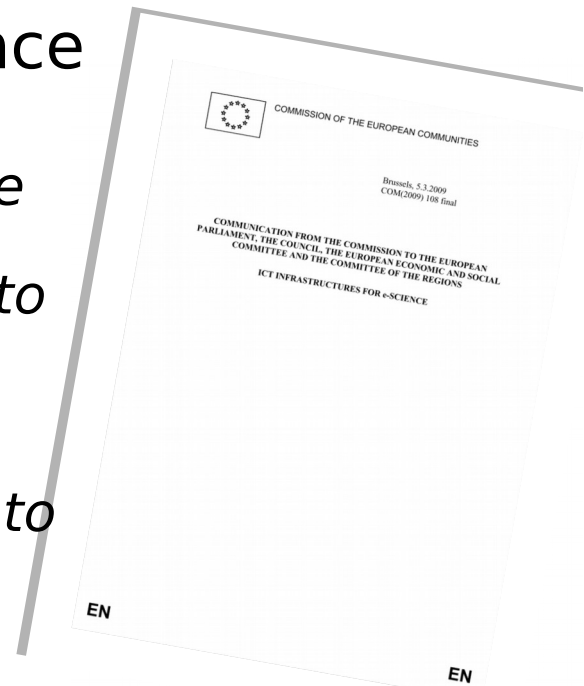
Background

- Research data are of most varied nature.
- Research data can only imperfectly be treated by an information management like conventional information/library media.
- Research data repositories (RDRs) often represent an essential stage of compression, abstraction and summary of research data, authorized and authenticated by the producers.
- RDRs can be operated centrally (institutional RDRs) or locally (disciplinary RDRs).
- In particular local or disciplinary RDRs are very popular in science because they represent a kind of a bottom up approach in research data management by the research groups themselves.

Research Data Repositories

- Highly heterogeneous landscape of research data repositories
- Different communities and different approaches
- EC (2009): ICT infrastructures for e-science

„The landscape of data repositories across Europe is fairly heterogeneous, but there is a solid basis to develop a coherent strategy to overcome the fragmentation and enable research communities to better manage, use, share and preserve data.“



The RDR Landscape



funders



research data
repositories



journals



scientists



universities and
research labs

[RRZE Icon Set \(CC: BY-SA\)](#)

The RDR Landscape

Investigators are expected to share their data!



funders



research data
repositories

Underlying data must be accessible!



journals

Where can I find data and store mine?

How can we set up repositories?



scientists

Should we offer repositories for all disciplines?



universities and
research labs

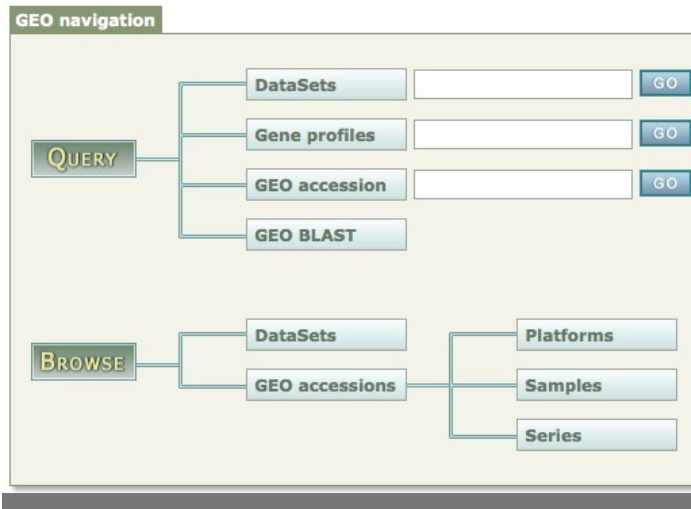
[RRZE Icon Set \(CC: BY-SA\)](#)

Research Data Repository

- Example: Disciplinary repositories



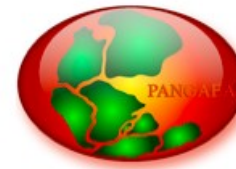
Gene Expression Omnibus: a public functional genomics data repository supporting submissions. Array- and sequence-based data are accepted. Tools are provided to help with experiments and curated gene expression profiles. [More information »](#)



GEO,
<http://www.ncbi.nlm.nih.gov/geo/>

PANGAEA®

Data Publisher for Earth & Environmental Science



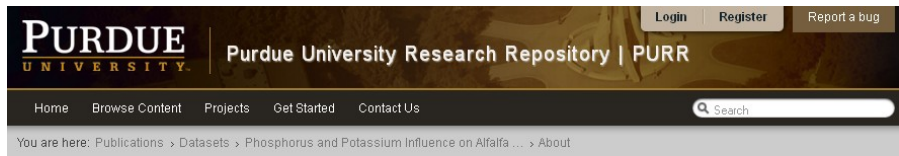
[About](#) – [Submit Data](#) – [Projects](#) – [Software](#) – [WDC-MARE](#) – [Contact](#)

This work is licensed under a [Creative Commons License](#)

PANGAEA, <http://www.pangaea.de>

Research Data Repository

- Example: Institutional repositories



Phosphorus and Potassium Influence on Alfalfa Nutrition

By Jeffrey J Volenec

Data from several studies on the influence of phosphorus and potassium nutrition on alfalfa. It includes numeric data such as yield, plant mass, plant counts, and tissue concentration of various nutrients

Listed in Datasets

Download (XLS)

Additional materials available (1)




Version 1.0 - published on Feb 22, 2012
doi:10.5072/FK2NG4S08 - cite this

0.0 RANKING

0 review(s) (Rev)

0 questions (Ask)

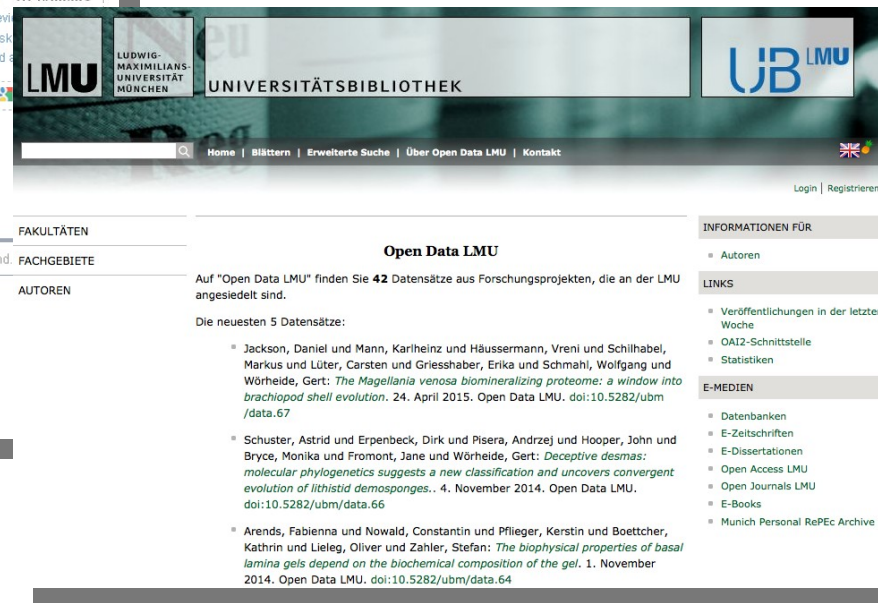
0 wish(es) (Add)

Share:   

About Supporting Docs Versions Reviews Questions Wish List

Abstract This is a master file of merged excel files with data from several studies on the influence of phosphorus and potassium nutrition on alfalfa. It includes numeric data such as yield, plant mass, plant counts, and tissue concentration of various nutrients. Conventional wisdom states that plants that are poorly fertilized and do not have adequate phosphorus and especially low potassium will become winterkill. This study showed that plants did not necessarily die in winter with low potassium stress.

PURR, <http://research.hub.purdue.edu>



Home | Blättern | Erweiterte Suche | Über Open Data LMU | Kontakt

SEE ALSO

FAKULTÄTEN

FACHGEBIETE

AUTOREN

Open Data LMU

Auf "Open Data LMU" finden Sie 42 Datensätze aus Forschungsprojekten, die an der LMU angesiedelt sind.

Die neuesten 5 Datensätze:

- Jackson, Daniel und Mann, Karlheinz und Häussermann, Vreni und Schilhabel, Markus und Lüter, Carsten und Griesshaber, Erika und Schmah, Wolfgang und Wörheide, Gert: *The Magellania venosa biomineralizing proteome: a window into brachiopod shell evolution*. 24. April 2015. Open Data LMU. doi:10.5282/ubm/data.67
- Schuster, Astrid und Erpenbeck, Dirk und Pisera, Andrzej und Hooper, John und Bryce, Monika und Fromont, Jane und Wörheide, Gert: *Deceptive desmas: molecular phylogenetics suggests a new classification and uncovers convergent evolution of lithistid demosponges.*. 4. November 2014. Open Data LMU. doi:10.5282/ubm/data.66
- Arends, Fabienna und Nowald, Constantin und Pflieger, Kerstin und Boettcher, Kathrin und Lieleg, Oliver und Zahler, Stefan: *The biophysical properties of basal lamina gels depend on the biochemical composition of the gel*. 1. November 2014. Open Data LMU. doi:10.5282/ubm/data.64

INFORMATIONEN FÜR

- Autoren

LINKS

- Veröffentlichungen in der letzten Woche
- OAIZ-Schnittstelle
- Statistiken

E-MEDIEN

- Datenbanken
- E-Zeitschriften
- E-Dissertationen
- Open Access LMU
- Open Journals LMU
- E-Books
- Munich Personal RePEc Archive

Open Data LMU, <http://data.ub.uni-muenchen.de/>

Research Data Repository

- Example: Project focused repositories

Scientific Drilling Database

Data from Deep Earth Sampling and Monitoring

- + Home
- + About SDDB
- + News
- + Data Publications
- + Catalogue
- + Authors
- + Dataset
- + Research Programs
- + Sampling Gear
- + Analytical Parameters
- + Publication Process
- + Admin

GFZ POTSDAM icdp

Dataset Description

Citation: Demory, Francois; Nowaczyk, Norbert; Witt, Annette; Oberhänsli, Hedi; (2006): Intercorrelation of down-core variations of the high-resolution magnetic susceptibility for CON01-603 and CON01-605. *Scientific Drilling Database*. <http://dx.doi.org/10.1016/j.gloplacha.2006.05.003>

[Download Citation \(EndNote\)](#)

Title: Intercorrelation of down-core variations of the high-resolution magnetic susceptibility for CON01-603 and CON01-605

Abstract: Down-core variations of the high-resolution magnetic susceptibility (0.1 cm) were obtained for kasten and pilot cores from sites CON01-603 and CON01-605. AMS 14C dating performed on kasten cores to the pilot cores to the pilot cores subjected to paleomagnetic investigation allowed the transfer of AMS 14C dating performed on kasten cores to the pilot cores subjected to paleomagnetic investigation. Discrepancies in the magnetic susceptibility curves from the pilot core for 01-605 (Vydrino Shoulder) measured in 2001 and 2003, respectively (right large peaks visible in the first measurement from 2001 (dashed lines) disappear after a 2-year-long storage. This is a first hint for the presence of the ferromagnetically unstable greigite.

[Show in Google Earth](#)

Related Publications:

- N. Fagel, L.Y. Alleman, L. Granina, F. Hatert, E. Bozso, R. Cloots, André, Vivianite formation and distribution in Lake Baikal sediments. *Planet. Change* 46 (2005), pp. 315-336. doi:10.1016/j.gloplacha.2005.05.003

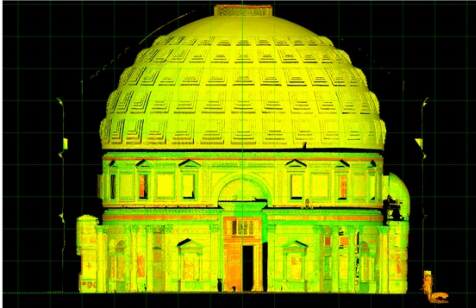
The Bern Digital Pantheon Project

Home The Building Research Publications The Team Repository Contact Search

You are here: Home > Repository > BDPP0019

Repository

first previous Record 198 of 776 next last



The Bern Digital Pantheon Project: Transverse cross-section, facing north

The Bern Digital Pantheon Project. Transverse cross-section, facing north. In: Gerd Grabhoff, Markus Wäfler, Jon Albers and Christian Berndt (ed.) Digital Repository of the Bern Digital Pantheon Project. Bern 2009, BDPP0019 (Published online at: <http://www.digitalpantheon.ch/BDPP0019>).

BibTeX:

```
@other(BDPP0019,
  author = (The Bern Digital Pantheon Project),
  title = (Transverse cross-section, facing north),
  year = (2009),
  keywords = (Medium Visualisation Building Interior),
  owner = (berndt),
  timestamp = (2009.05.28),
  editor = (Grabhoff, Gerd and Wäfler, Markus and Albers, Jon and Berndt, Christian),
  address = (Bern),
  repository = (Digital Repository of the Bern Digital Pantheon Project)
)
```

Usage of the repository

SDDB, <http://www.scientificdrilling.org>

BDPP, <http://www.digitalpantheon.ch>

Research Data Repository

- Example: Generic repositories

zenodo Research. Share.

Search Communities Browse Upload Get started - Sign In

08 May 2015 Dataset Embargoed access

Amber fossils demonstrate deep-time stability of Caribbean lizard communities

Sherratt, Emma; Castañeda, María del Rosario; Garwood, Russell; Mahler, D. Luke; Sanger, Thomas J.; Herrel, Anthony; de Queiroz, Kevin; Losos, Jonathan B.

(show affiliations)

This is the dataset for the paper Sherratt et al. 2015 (PNAS, doi:). Included are the X-ray microtomography (micro-CT) TIFF stacks for the amber encased anole lizard fossils studied. Also included are the morphometric data for the fossil and 100 modern specimens, and the phylogenetic character matrix for the fossils and 181 species (91 morphological characters, 4873 DNA bases).

Abstract: Whether the structure of ecological communities can exhibit stability over macroevolutionary timescales has long been debated. The similarity of independently evolved *Anolis* lizard communities on ecologically similar Greater Antillean islands supports the notion that community evolution is deterministic. However, a dearth of Caribbean *Anolis* fossils - only 3 have been described to date - has precluded direct investigation of the stability of anole communities through time. Here we report on an additional 17 fossil anoles in Dominican amber dating to 15-20my before the present. Using data collected primarily by X-ray microtomography, we demonstrate that the main elements of Hispaniolan anole communities were in place in the Miocene. Phylogenetic analysis is consistent with the hypothesis that the

Publication date: 08 May 2015
Embargoed
Files available as Open Access after 12 July 2015
DOI: 10.5281/zenodo.17442
Keyword(s): Anolis, micro-Computed Tomography, amber, Adaptive radiation, Ecomorph, Hispaniola
Collections: Communities, Datasets, Open Access
License (for files): Creative Commons Attribution-NonCommercial ShareAlike 4.0 International
Uploaded by: esherratt (on 08 May 2015)

Zenodo, <http://zenodo.org>

figshare My data search figshare (titles, tags, authors, etc.) Browse Upload H. Pampel

Main Dataset for "Evolution of Popular Music: USA 1960-2010"

	A	B	
1	recording_id	artist_name	artist_name
2	1	Suzanne Vega	SUZANNEY
3	2	Janet Jackson ft Carly Simon	JANETJAC
4	3	Neneh Cherry	NEHECHI
5	4	Neneh Cherry	NEHECHI
6	5	Junior M.A.F.I.A. Featuring The Notorious B.I.G.	JUNIORMPA
7	6	Junior M.A.F.I.A.	JUNIORMPA
8	7	Missy "Misdemeanor" Elliott Featuring Da Brat	MISSYMISI
9	8	Nas	NAS

EvolutionPopUSA_MainData

510 Views 0 shares

Published on 06 Apr 2015 - 20:15 (GMT)
Filesize is 18.81 MB

Categories

- Anthropology
- Applied Computer Science
- Media Studies
- Performing Arts
- Evolutionary Biology

Authors

Matthias Mauch

Tags

- music
- Billboard Hot 100
- charts
- cultural evolution
- evolution

License (what's this?)

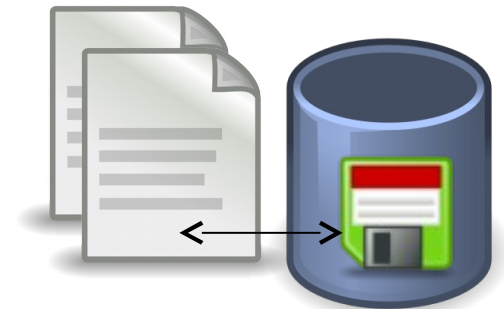
Figshare, <http://figshare.com>

Publishing strategies

- As independent information object
- As document within a reviewed “data paper”
- As addition to a reviewed article

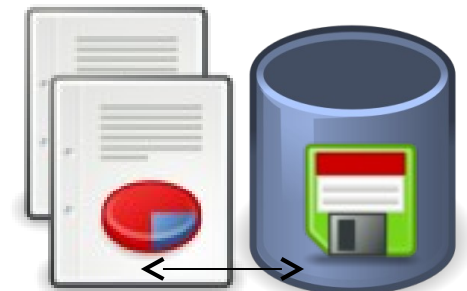


doi:XX.XXXX/XXX.XX



doi:XX.XXXX/XXX.XX

doi:XX.XXXX/XXX.XX



doi:XX.XXXX/XXX.XX

doi:XX.XXXX/XXX.XX

RRZE Icon Set (CC: BY-SA)

Mission

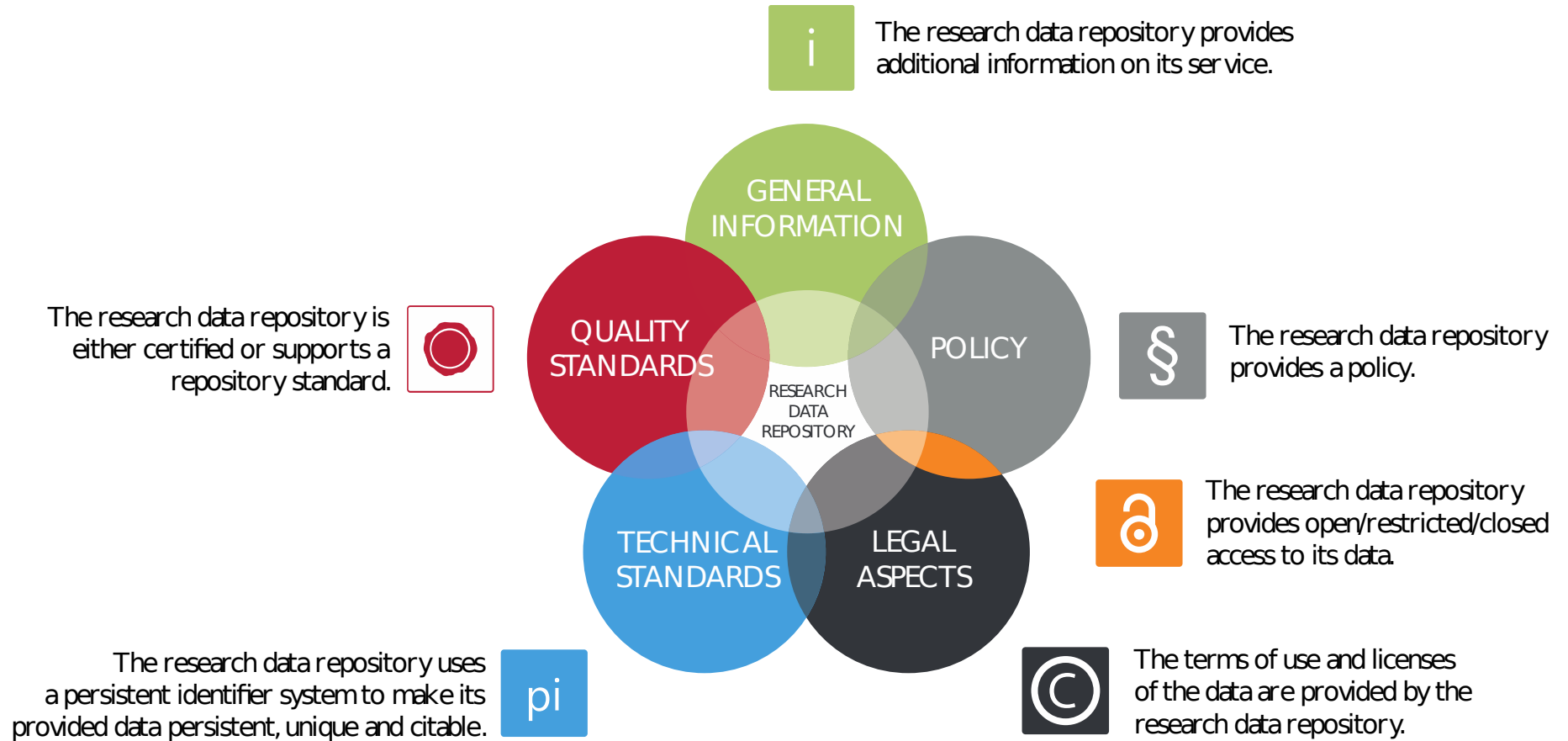
- is a global registry of research data repositories
- covers research data repositories from all academic disciplines
- helps researchers, funding bodies, publishers and scholarly institutions to find research data repositories
- aims to promote a culture of sharing, increased access and better visibility of research data

Schema

- general information (e.g. short description of the RDR, content types, keywords)
- responsibilities (e.g. institutions responsible for funding, content or technical issues)
- policies (e.g. policies of the RDR, incl. their URL)
- legal aspects (e.g. licenses of the database and datasets)
- technical standards (e.g. APIs, versioning of datasets, software of the RDR)
- quality standards (e.g. certificates, audit processes)



Icons

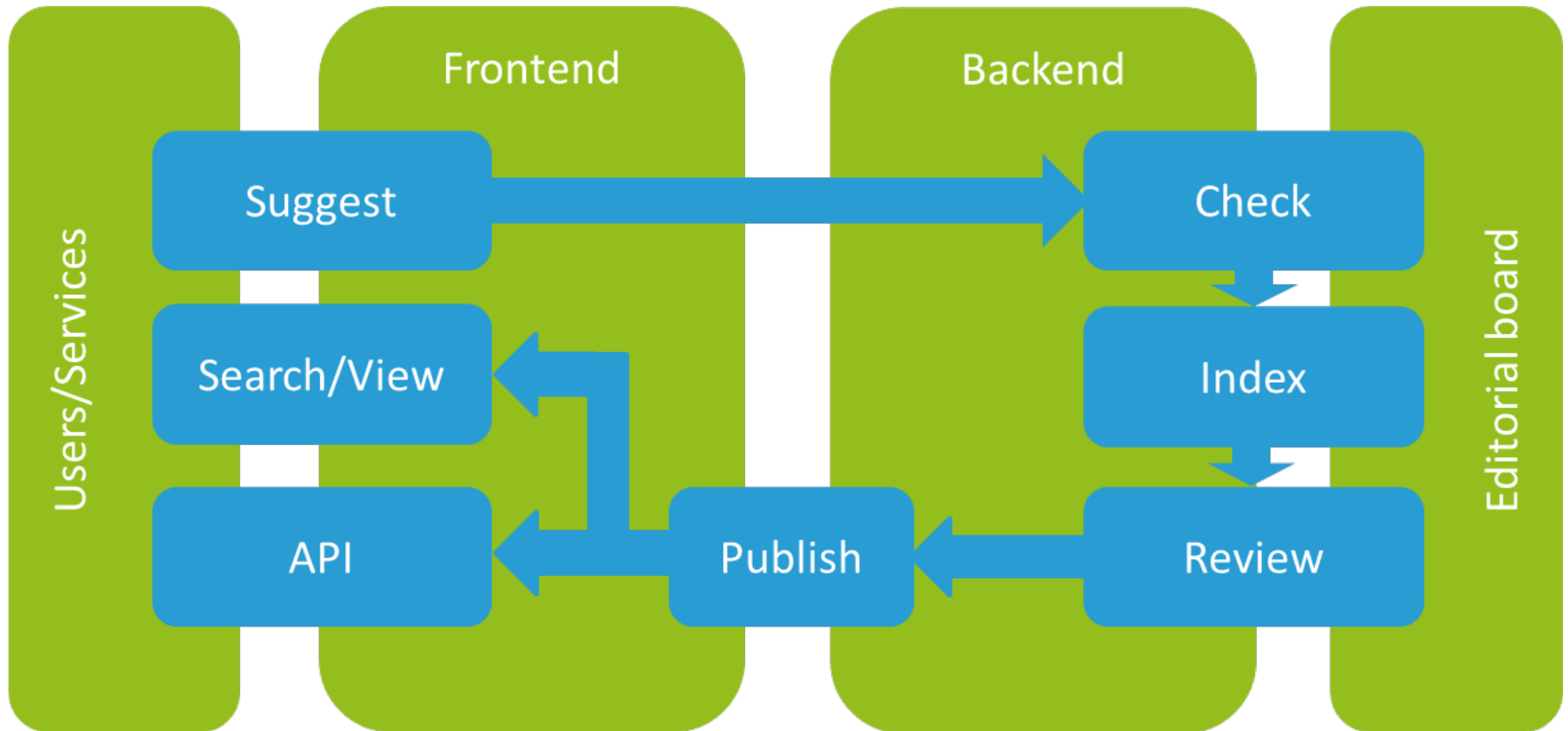


Quality

Requirements

- Run by a legal entity, such as a sustainable institution (e.g. library, university)
- Clarify access conditions to the data and repository as well as the terms of use
- Have focus on research data
- (Have an english graphical user interface)

Workflow



Interface

search

Search for Repositories (1234 Reviewed Repositories)

Search

filters

Subject

Add subjects

Content Type

Add content types

Country (of the responsible institutions)

Add countries

Certificates

Open Access

Persistent Identifier

Reset filter

results

1234 results (1 - 25)

Sort by weight

«	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	»
---	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	---

icons

3TU.Datacentrum

3TU.DC



Subjects: Agriculture, Forestry, Horticulture and Veterinary Medicine; Analytical Chemistry, Method Development (Chemistry); Basic Biological and Medical Research; Bioinformatics and Theoretical Biology; Biology; Biophysics; Chemistry; Computer Science; Computer Science, Electrical and System Engineering; Construction Engineering and Architecture; Engineering Sciences; Geochemistry, Mineralogy and Crystallography; Geography; Geophysics; Geophysics and Geodesy; Geosciences (including Geography); Life Sciences; Materials Science; Materials Science and Engineering; Mathematics; Natural Sciences; Physics; Soil Sciences; Systems Engineering; Traffic and Transport Systems, Logistics; Urbanism, Spatial Planning, Transportation and Infrastructure Planning, Landscape Planning; Water Research

Content types: Archived data; Audiovisual data; Images; Plain text; Raw data; Scientific and statistical data formats; Standard office documents; Structured text

Interface

Repository details

PANGAEA



General

Institutions

Terms

Standards

Name of repository

PANGAEA

Additional name(s)

Publishing Network for Geoscientific and Environmental Data

Repository URL

<http://www.pangaea.de>

Subject(s)

Oceanography **Geology and Palaeontology** **Geophysics** **Geochemistry, Mineralogy and Crystallography**
Biology **Atmospheric Science and Oceanography** **Geosciences (including Geography)** **Natural Sciences**
Geology and Palaeontology **Geophysics and Geodesy** **Geochemistry, Mineralogy and Crystallography**
Life Sciences

Description

The information system PANGAEA is operated as an Open Access library aimed at archiving, publishing and distributing georeferenced data from earth system research. The system guarantees long-term availability of its content through a commitment of the operating institutions.

Content type(s)

Standard office documents **Images** **Plain text** **Archived data** **Audiovisual data**

Keyword(s)

Earth Science **Environmental Science**

Repository type(s)

disciplinary

Research data repository
language(s)

eng

Data and/or service provider

dataProvider

Interface

Repository details

PANGAEA



General

Institutions

Terms

Standards

Institution name	Alfred Wegener Institute for Polar and Marine Research
Additional name(s)	AWI Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung
URL	http://www.awi.de/en/home/
Contact(s)	hgrobe@pangaea.de
Country	Germany
Type(s) of responsibility	general technical
Type of institution	non-profit

Institution name	Center for Marine Environmental Sciences (MARUM)
URL	http://www.marum.de/
Contact(s)	mdiepenbroek@pangaea.de
Country	Germany
Type(s) of responsibility	general
Type of institution	non-profit

Interface

Repository details



PANGAEA

General

Institutions

Terms

Standards

Policies (1)

Policy Name Data policy of the information system PANGAEA

URL <http://www.pangaea.de/curator/files/pangaea-data-policy.pdf>

Database access

Type of access to research data repository open

Data access (1)

Type of access to data open

Data licenses (1)

DataLicense CC

URL <http://wiki.pangaea.de/wiki/License>

Data upload (1)

Type of data upload restricted

Data upload restriction type(s) registration

Interface

Repository details

PANGAEA



General

Institutions

Terms

Standards

Name of the repository software other

Versioning yes

Persistent identifier system(s) DOI

Quality management yes

Certificates and Standards WDS

Application programming interfaces (1)

API type OAI-PMH

URL <http://ws.pangaea.de/oai/>

Alerting services (1)

Type of alerting service RSS

Alerting service <http://www.pangaea.de/tools/latest-datasets.rss>

Remarks

Remarks Data of World Data Center for Marine Environmental Sciences (WDC-MARE) are available via the data library PANGAEA which will be operated as a member of the new WDS (World Data System)

Guidelines

“Datasets are more likely to be seen, reused, and have impact if they can be found where potential reusers are likely to look. If you are unsure where that might be, the Registry of Research Data Repositories (re3data.org) provides a list of repositories organised by subject, content type and country.” Alex Ball (DCC) and Monica Duke (DCC) - How to Measure the Impact of Research Data / A Digital Curation Centre ‘working level’ guide

Guidelines

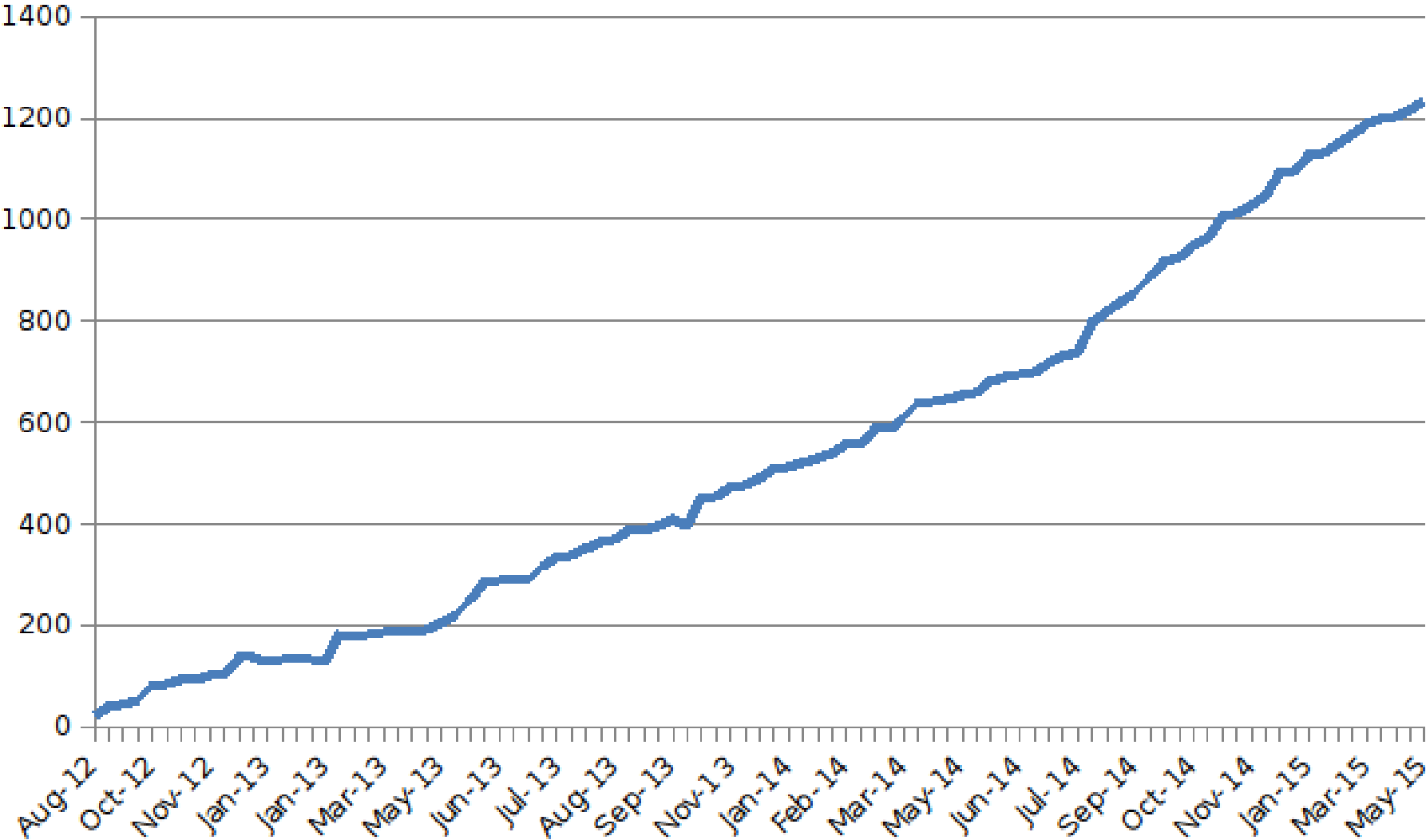
European commission

1) **Step 1:** participating projects are required to deposit the research data described above, preferably into a research data repository. 'Research data repositories' are online archives for research data. They can be subject-based/thematic, institutional or centralised. Useful listings of research data repositories include the Registry of Research Data Repositories (www.re3data.org) and Databib (<http://databib.org>). In addition, it is expected that the Open Access Infrastructure for Research in Europe (OpenAIRE) will become an entry point for linking publications to underlying research data.

Scientific Data (NPG)

„Physics, astrophysics, astronomy and geoscience databases should be registered with re3data.org.“

Growth



Four dimensions of sustainability



Sustainability - Technology

- Open interfaces

RESTful API

OpenSearch

Documentation: <http://www.re3data.org/api/doc>

Various usage scenarios, e.g. by OpenAIRE

- Open metadata

Documentation: <http://www.re3data.org/schema/>

Sustainability - Legal

- Open licenses
- CC BY for web page content
- CC 0 for metadata

LICENSES



Except where otherwise noted, content on this site is licensed under a [Creative Commons Attribution 4.0 International License](#).



To the extent possible under law, [re3data.org](#) has waived all copyright and related or neighboring rights to the database entries of re3data.org.

Sustainability - Partners

- Berlin School of Library and Information Science
- GFZ German Research Centre for Geosciences
- Karlsruhe Institute of Technology (KIT), KIT Library
- Purdue University, Purdue Libraries



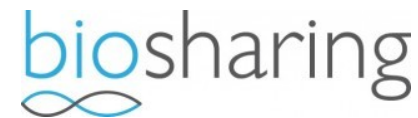
Funded by

- German Research Foundation
- Institute of Museum and Library Services (IMLS)



Sustainability – Cooperations

- German Initiative for Network Information (DINI)
- DataCite (MoU, April 2012)
- OpenAIRE (MoU, October 2013)
- BioSharing (MoU, November 2013)
- Databib (MoU, March 2014)
- RDA



Sustainability - Join forces

Databib and re3data.org have agreed to the following five principles for successful cooperation:

- Openness
- Optimal quality assurance
- Development of innovative functionalities
- Shared leadership
- Sustainability

Sustainability - Organization

- Databib and re3data.org merged in spring 2015
- Will become a service of DataCite

The screenshot shows the Databib website interface. At the top, the Databib logo is on the left, and navigation links for DataCite (ABOUT DATACITE, DATACITE FOR YOU, PARTICIPATE, EVENTS, NEWS, JOIN DATACITE) are on the right. Below the navigation, there's a breadcrumb trail: Home • DataNews • DataCite to Manage and Develop re3data.org. The main article is titled "DataCite to Manage and Develop re3data.org" and is dated May 04, 2015, by Frauke Ziedorn. The article text states: "The DataCite General Assembly met at the British Library last week and approved the inclusion of the re3data.org registry of research data repositories into the portfolio of services that are managed by the DataCite organization. Over 1,200 data repositories have been indexed by re3data.org and can be searched and accessed at its website or by using its API. Bringing this service together with DataCite, who mints and manages Digital Object Identifiers for datasets, will yield new opportunities to explore in combining a registry of data repositories with information about persisted datasets to create new value for the research community. A new re3data.org Working Group within DataCite will be co-chaired by Michael Witt and Frank Scholze, who will also lead the editorial board that evaluates new additions and maintains the records in the registry. Members of the working group will be appointed by the DataCite Executive Board and include equal participation from the original re3data.org project partners and DataCite member institutions. The working group will convene in September to migrate the operation of". To the right of the article is a yellow diamond-shaped road sign with a black arrow pointing up and a black figure of a person walking, set against a blue sky background. Below the article, there are sections for "Events" (Keep an eye out for the upcoming future events) and "News" (Persistent Identifiers in Paris – Mark Your Diaries and Calendars, Job offer: Technical Architect for THOR Project (DataCite)).

Sustainability - Finance

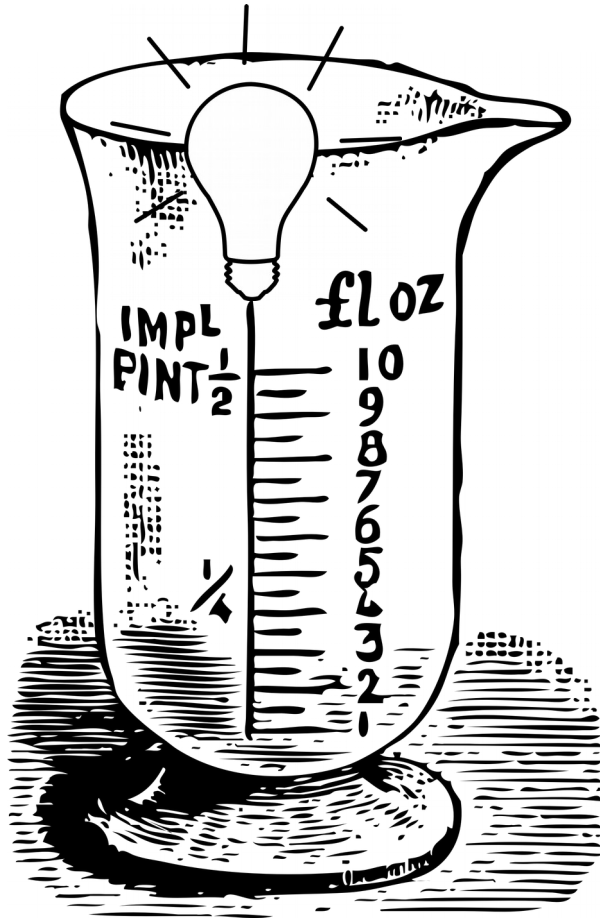
- Finance for hosting and maintenance by DataCite
- Future management by DataCite
- May be further third party funding



Lessons learned

- Openness as paradigm works well
(“open science” & schema development)
- Cooperation is worth the effort
(Databib & DataCite are reliable partners)
- Quality assurance by international editorial board

Raise awareness

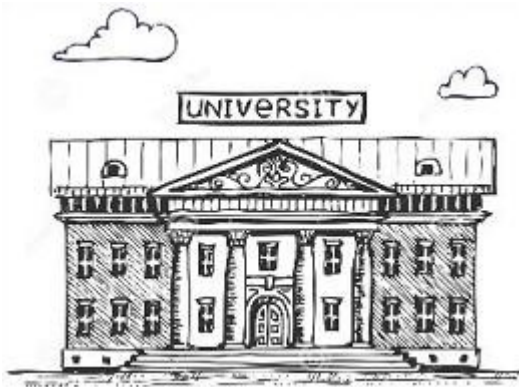


- Among service units & researchers
- Teach students & young scientists
- Not during lectures
- But during projects & thesis's

Clarify responsibilities



Responsibility: Correctness of Data
Responsible: Scientist



Responsibility: Data management
Responsible: Service units

Thank you for your attention!

WHERE DO YOU STORE
YOUR RESEARCH DATA?

- USB DRIVE
- DROPBOX
- RESEARCH DATA REPOSITORY

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

re3data.org

REGISTRY OF RESEARCH DATA REPOSITORIES

info@re3data.org

<http://re3data.org>



With the exception of all photos and graphics, this slides are licensed under the “Attribution 4.0 International (CC BY 4.0)”
Licence: <http://creativecommons.org/licenses/by/4.0/>