re3data.org – Registry of Research Data Repositories

Heinz Pampel | GFZ German Research Centre for Geosciences, LIS
Maxi Kindling | Humboldt-Universität zu Berlin, Berlin School of Library and Information Science
Frank Scholze | Karlsruhe Institute of Technology, KIT Library

RDA-Deutschland-Treffen 2015| Potsdam, November 26, 2015
Background

- Funders’ 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).}

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: Public Library of Science - PLOS
  • “PLOS journals require authors to make all data underlying the findings described in their manuscript fully available without restriction, with rare exception.“

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: Public Library of Science - PLOS
    - "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


re3data.org – Mission

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

The research data repository provides additional information on its service.

The research data repository is either certified or supports a repository standard.

The terms of use and licenses of the data are provided by the research data repository.

The research data repository provides open/restricted/closed access to its data.

The research data repository uses a persistent identifier system to make its provided data persistent, unique, and citable.
Search for Repositories (1350 reviewed repositories)

Subject: Add subjects
Content Type: Add content types
Country: Add countries

binary: Certificates
Open Access
Persistent Identifier

Results: 3 results (filtered) (1 – 3)
Sort by: Weight

PANGAEA
Data Publisher for Earth and Environmental Science

Subjects: Atmospheric Science and Oceanography, Biology, Geochemistry, Mineralogy and Crystallography, Geology and Paleontology, Geophysics
Content types: Archived data, Audiovisual data, Images, Plain text, Standard office documents
Countries: Germany

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.
### PANGAEA

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name of repository</td>
<td>PANGAEA</td>
</tr>
<tr>
<td>Additional name(s)</td>
<td>Publishing Network for Geoscientific and Environmental Data</td>
</tr>
<tr>
<td>Repository URL</td>
<td><a href="http://www.pangaea.de">http://www.pangaea.de</a></td>
</tr>
<tr>
<td>Subject(s)</td>
<td>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</td>
</tr>
<tr>
<td>Description</td>
<td>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.</td>
</tr>
<tr>
<td>Content type(s)</td>
<td>Standard office documents, Images, Plain text, Archived data, Audiovisual data</td>
</tr>
<tr>
<td>Keyword(s)</td>
<td>Earth Science, Environmental Science</td>
</tr>
<tr>
<td>Repository type(s)</td>
<td>disciplinary</td>
</tr>
<tr>
<td>Research data repository</td>
<td>eng</td>
</tr>
<tr>
<td>language(s)</td>
<td></td>
</tr>
<tr>
<td>Data and/or service provider</td>
<td>dataProvider</td>
</tr>
<tr>
<td>Data availability</td>
<td></td>
</tr>
</tbody>
</table>
## PANGAEA

<table>
<thead>
<tr>
<th>Institution name</th>
<th>Alfred Wegener Institute for Polar and Marine Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additional name(s)</td>
<td>AWI</td>
</tr>
<tr>
<td></td>
<td>Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://www.awi.de/en/home/">http://www.awi.de/en/home/</a></td>
</tr>
<tr>
<td>Contact(s)</td>
<td><a href="mailto:hgrobe@pangaea.de">hgrobe@pangaea.de</a></td>
</tr>
<tr>
<td>Country</td>
<td>Germany</td>
</tr>
<tr>
<td>Type(s) of responsibility</td>
<td>general</td>
</tr>
<tr>
<td></td>
<td>technical</td>
</tr>
<tr>
<td>Type of institution</td>
<td>non-profit</td>
</tr>
</tbody>
</table>

| 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                                           |
# PANGAEA

## Repository details

<table>
<thead>
<tr>
<th>General</th>
<th>Institutions</th>
<th>Terms</th>
<th>Standards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policies (1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Policy Name</td>
<td>Data policy of the information system PANGAEA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Database access</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of access to research data repository</td>
<td>open</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data access (1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of access to data</td>
<td>open</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data licenses (1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data License</td>
<td>CC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://wiki.pangaea.de/wiki-License">http://wiki.pangaea.de/wiki-License</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data upload (1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of data upload</td>
<td>restricted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data upload restriction type(s)</td>
<td>registration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data upload license (1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data upload license name</td>
<td>Data Submission</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data upload license URL</td>
<td><a href="http://wiki.pangaea.de/wiki/Data_submission">http://wiki.pangaea.de/wiki/Data_submission</a></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Repository details

**PANGAEA**

<table>
<thead>
<tr>
<th>General</th>
<th>Institutions</th>
<th>Terms</th>
<th>Standards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name of the repository software</td>
<td>other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Versioning</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persistent identifier system(s)</td>
<td>DOI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality management</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Certificates and Standards</td>
<td>WDS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Application programming interfaces (1)**

<table>
<thead>
<tr>
<th>API type</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>OAI-PMH</td>
<td><a href="http://ws.pangaea.de/oai/">http://ws.pangaea.de/oai/</a></td>
</tr>
</tbody>
</table>

**Alerting services (1)**

<table>
<thead>
<tr>
<th>Type of alerting service</th>
<th>Alerting service</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSS</td>
<td><a href="http://www.pangaea.de/tools/latest-datasets.rss">http://www.pangaea.de/tools/latest-datasets.rss</a></td>
</tr>
</tbody>
</table>

**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)

**Entry date** 2012-07-16

**Last update** 2014-04-14
re3data.org – Quality

• Registration Policy

• To be registered in re3data.org a research data repository must
  • be 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 an English graphical user interface (GUI);
  • have focus on research data.
re3data.org – Indexed Repositories

Growth of re3data.org

Indexed Research Data Repositories
re3data.org – Cooperation

- Deutsche Initiative für Netzwerkinformation (DINI)
- DataCite (MoU, April 2012)
- OpenAIRE (MoU, October 2013)
- BioSharing (MoU, November 2013)
- Databib (MoU, March 2014)
- DataCite (Formal cooperation, March 2015)
re3data.org – Policies

• **Funder Example: 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.


• **Institutional Example: Bielefeld University**

  • „Verzeichnisse, wie das DFG-geförderte "Registry of Research Data Repositories", bilden die Grundlage für die Suche nach geeigneten Publikationsorten für die Forschungsdaten. “  
  Universität Bielefeld (2011): Resolution zum Forschungsdatenmanagement. https://data.uni-bielefeld.de/de/resolution

• **Publisher Example: Nature Publishing Group**

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

- Technical
- Financial
- Legal
- Organisational
re3data.org – Organizational sustainability

- Merger with DataBib under the auspices of DataCite
- re3data.org working group within DataCite
- International Editorial Board
- Cooperations within RDA and the research data repository community
- Community building and feedback loops during RFC phases (e.g. re3data.org schema)
re3data.org – Technical sustainability

• Open interfaces
  • RESTful API
  • OpenSearch
  • Documentation: http://www.re3data.org/api/doc
  • Used e.g. by OpenAIRE

• Open metadata
  • Documentation: http://www.re3data.org/schema/

• Long-term hosting commitment by KIT
re3data.org – Legal sustainability

- Open licenses
  - CC BY for the website
  - CC 0 for metadata
re3data.org – Financial sustainability

- Technical maintenance financed by DataCite from 2016
- Further development managed by DataCite
- Further project funding
re3data.org – Next steps

- re3data.org persistent identifier

Cite this re3data.org record:
re3data.org – Next steps

- Browsing
re3data.org – Next steps

• Change requests
Thanks to the team!

- Roland Bertelmann, Caludio Fuchs, Heinz Pampel, Paul Vierkant
  - GFZ German Research Centre for Geosciences, Library and Information Services (LIS)
- Florian Fritze, Maxi Kindling, Jessika Rücknagel, Peter Schirmbacher, Stephanie van de Sandt
  - Humboldt-Universität zu Berlin, Berlin School of Library and Information Science (BSLIS)
- Hans-Jürgen Goebelbecker, Gabriele Kloska, Evelyn Reuter, Frank Scholze, Edeltraud Schnepf, Angelika Semrau, Michael Skarupianski, Robert Ulrich
  - Karlsruhe Institute of Technology (KIT), KIT Library
- Michael Witt
  - Purdue University, Distributed Data Curation Center (D2C2)
WHERE DO YOU STORE YOUR RESEARCH DATA?

- USB DRIVE
- DROPBOX
- RESEARCH DATA REPOSITORY

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/