Outline

• Background
• Research Data Repositories
• Mission
• Status quo
Background

1. Research data are valuable and ubiquitous. New technologies facilitate data-intensive science.

2. Broad discussion about the permanent access to research data. Increasing requirements from funders to make data openly available.

3. Growing demand for trustable and sustainable research data repositories.
Figure 1-1: Types of scholarly activities and framework conditions

„Scientists should communicate the data they collect and the models they create, to allow free and open access, and in ways that are intelligible, assessable and usable for other specialists in the same or linked fields wherever they are in the world. Where data justify it, scientists should make them available in an appropriate data repository. Where possible, communication with a wider public audience should be made a priority, and particularly so in areas where openness is in the public interest.“
EC: Recommendations for the Memberstates

Further develop e-infrastructures underpinning the system for disseminating scientific information by:

- Supporting scientific data infrastructures for dissemination of knowledge, research institutions and funding entities to address all stages of the data life cycle. These stages should include acquisition, curation, metadata, provenance, persistent identifiers, authorisation, authentication and data integrity. Approaches need to be developed to provide a common look and feel to data discovery across disciplines, thus reducing the learning curve required to achieve productivity;

Research-Data-Repositories

• EC: 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.“

Vision 2030

• “Researchers and practitioners from any discipline are able to find, access and process the data they need. They can be confident in the data, and they can evaluate the degree to which that data can be trusted.”

• “Producers of data benefit from opening it to broad access, and prefer to deposit with confident and reliable repositories. A framework of repositories is guided by international standards, to ensure they are trustworthy.”
Key Drivers for the Vision 2030

- Incentives for scholars
- Training the researcher and users
- Building the infrastructure

„The three main challenges in developing an ecosystem of data repositories are (1) gaps in the present data infrastructure and (2) connectivity issues (between the workflow of researchers and the institutional data infrastructure and between institutional and national data infrastructures) and (3) long-term financial basis.”

- Funding the infrastructure

The Research Data Repositories Landscape

Investigators are expected to share their data!

Where can I store my data?
- funders
- scientists

How is it to organize?
- research data repositories

Underlying data must be accessible!

Where can I find data?
- journals
- universities and research labs

Should we offer repositories for all disciplines?
Research Data Repositories

PANGAEA, http://www.pangaea.de
Research Data Repositories

Scientific Drilling Database
Data from Deep Earth Sampling and Monitoring

Dataset Description

Citation: Domory, Francois, Nowaczyk, Norbert, Witt, Andre. Intercomparison of down-core variations of the high-resolution magnetic susceptibility curve for CON01-803 and CON01-805. Scientific Drilling and Scientific Data Base (SDDB). DOI: 10.25918/SDDB.D01003

Title: Intercomparison of down-core variations of the high-resolution magnetic susceptibility curve for CON01-803 and CON01-805

Abstract: Down-core variations of the high-resolution magnetic susceptibility curve for CON01-803 and CON01-805 have been obtained for kasten and pilot cores from sites ES01-1004 (Vityaz Shoulder) measured in 2001 and ES01-804 (Vityaz Shoulder) measured in 2001 and 2002. Comparison of the susceptibility curves from these two sites highlights major discrepancies in the magnetic susceptibility curve.

SDDB, http://www.scientificdrilling.org

The Bern Digital Pantheon Project

Repository

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

BDPP, http://www.digitalpantheon.ch

SDDB, http://www.scientificdrilling.org

BDPP, http://www.digitalpantheon.ch
Research Data Repositories

Phosphorus and Potassium Influence on Alfalfa Nutrition

By Jeffrey J Velasco

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

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

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

re3data.org

• will be a global registry of research data repositories
• will cover research data repositories from all academic disciplines
• will help 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
General Workflow
The research data repository was reviewed by the re3data.org Team.

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

The research data repository uses a persistent identifier system to make its provided data persistent, unique and citable.

The research data repository provides additional information on its service.

The research data repository provides a policy.

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

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

- Indexing and reviewing
  - 31 metadata elements
  - 25 child elements
  - 22 controlled vocabularies
- Icon set to easily grasp basic information
  - Useful for institutions and researchers
Suggest

Do you miss a repository in re3data.org? We are open to suggestions. Please use the following form:

**Suggest a repository to be indexed:**

* Required

Repository name: *

URL of the repository: *

Comment:

Submit
Reviewed Results

Search for repositories (alpha version)

- mineralogy

- Subject: Add Subjects
- Content Type: Add Types
- Country: Add Countries

- Certificates
- Open Access
- Persistent Identifier
- Repository reviewed by re3data.org

Search Results (4 results)

Search terms: mineralogy

Crystallography Open Database

Physics

Open-access collection of crystal structures of organic, inorganic, metal-organic compounds and minerals, excluding biopolymers

Inorganic Crystal Structure Database

Chemistry Geochemistry Materials Science and Engineering Mineralogy and Crystallography Physics

The most comprehensive database on fully determined inorganic crystal structures • Full structural data: cell parameters, atom positions for all entries, displacement parameters • Full bibliographic data: publication title, journal reference(s), author names • Full structure description: Structural formula, compositions, ANX formulae, structure types • High-quality data: extensive data evaluation and correction by senior experts • Web and PC-based software solutions, data updated twice a year • 25+ years of serving the scientific community

PANGAEA

Biology Geochemistry Geodesy Geoinformatics Geology and Palaeontology Geophysics Mineralogy and Crystallography Oceanography Remote Sensing

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.
The icons shall help users to identify important characteristics of a research data repository at first sight. The following table explains the meaning of the icons:

1. The research data repository provides additional information on its service.
2. The research data repository provides open access to its data.
3. The research data repository provides restricted access to its data.
4. The research data repository provides closed access to its data.
5. The terms of use and licenses of the data are provided by the research data repository.
6. The research data repository provides a policy.
7. The research data repository uses DOI to make its provided data persistent, unique and citable.
8. The research data repository uses URN to make its provided data persistent, unique and citable.
9. The research data repository uses ARK to make its provided data persistent, unique and citable.
10. The research data repository uses handle to make its provided data persistent, unique and citable.
11. The research data repository uses Purl to make its provided data persistent, unique and citable.
12. The research data repository uses a persistent identifier system to make its provided data persistent, unique and citable.
13. The research data repository is either certified or supports a repository standard.
14. The research data repository was reviewed by the re3data.org Team.
## General information

<table>
<thead>
<tr>
<th>Name of repository</th>
<th>PANGAEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additional name</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>Subjects</td>
<td>Biology, Geology, Geotechnology, Geology and Palaeontology, Geophysics, Mineralogy and Crystallography, Oceanography, Remote Sensing</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</td>
<td>disciplinary</td>
</tr>
<tr>
<td>Research data repository language(s)</td>
<td>eng</td>
</tr>
<tr>
<td>Data and/or service provider</td>
<td>dataProvider</td>
</tr>
</tbody>
</table>

## Responsible institutions (2)

<table>
<thead>
<tr>
<th>Institution name</th>
<th>Alfred Wegner Institute for Polar and Marine Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additional name</td>
<td>AWI</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://www.awi.de/">http://www.awi.de/</a></td>
</tr>
<tr>
<td>Contact</td>
<td><a href="mailto:lgrohe@pangaea.de">lgrohe@pangaea.de</a></td>
</tr>
<tr>
<td>Country</td>
<td>Germany</td>
</tr>
<tr>
<td>Type(s) of responsibility</td>
<td>general</td>
</tr>
</tbody>
</table>
Search for repositories (611 repositories)

Filter results

Archaeology Data Service
ads

Subjects: Ancient Cultures, Classical Archaeology, History
Content types: Archived data, Audiovisual data, Databases, Images, Plain text, Raw data, Scientific and statistical data formats, Standard office documents, Structured graphics, other
Countries: United Kingdom

The Archaeology Data Service supports research, learning and teaching with freely available, high quality and dependable digital resources. It does this by preserving digital data in the long term, and by promoting and disseminating a broad range of data in archaeology. The ADS promotes good practice in the use of digital data in archaeology, it provides technical advi...
Partners and Sustainability

- Berlin School of Library and Information Science
  Peter Schirmbacher (Principal Investigator), Maxi Kindling, Paul Vierkant, Jessika Rücknagel, Shaked Spier

- GFZ German Research Centre for Geosciences
  Roland Bertelmann (Principal Investigator), Jens Klump, Heinz Pampel

- Karlsruhe Institute of Technology (KIT), KIT Library
  Frank Scholze (Principal Investigator) Hans-Jürgen Goebelbecker, Jens Gundlach,

- All partners are actively involved in the German Initiative for Network Information (DINI).
- re3data.org is funded by the German Research Foundation DFG in the period 2012 to 2014
Welcome to re3data.org

Posted on March 16, 2012 by re3data.org team

Research data are valuable and ubiquitous. The permanent access to research data is a challenge for all stakeholders in the scientific community. The long-term preservation and the principle of open access to research data offer broad opportunities for the scientific community.

More and more universities and research centres are starting to build research data repositories allowing permanent access to data sets in a trustworthy environment. Due to disciplinary requirements, the landscape of data repositories is very heterogeneous. Thus it is difficult for researchers, funding bodies, publishers and scholarly institutions to select appropriate repositories for storage and search of research data.

Thanks for your attention!

With the exception of all photos and graphics, this slides are licensed under the “Creative Commons Attribution 3.0 Germany (CC BY 3.0)“ Licence.
Quotation of References