

# Making Research Data Repositories Visible – The re3data.org Registry

Paul Vierkant, Heinz Pampel | GFZ German Research Centre for Geosciences, LIS

GeoBerlin 2015 | Berlin, October 6, 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).*

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](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf)

# Background

- Journal Data Policies
  - Nature Publishing Group
    - “[...] authors are required to make materials, data and associated protocols promptly available to readers without undue qualifications.”
  - 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

NPG (2013). Availability of data and materials. Retrieved from <http://www.nature.com/authors/policies/availability.html>  
PLOS (2014). PLOS Editorial and Publishing Policies. Retrieved from <http://www.plosone.org/static/policies.action>

# Reproducibility and trust

From: [Infect Immun. 2011 October; 79\(10\): 3855–3859.](#)

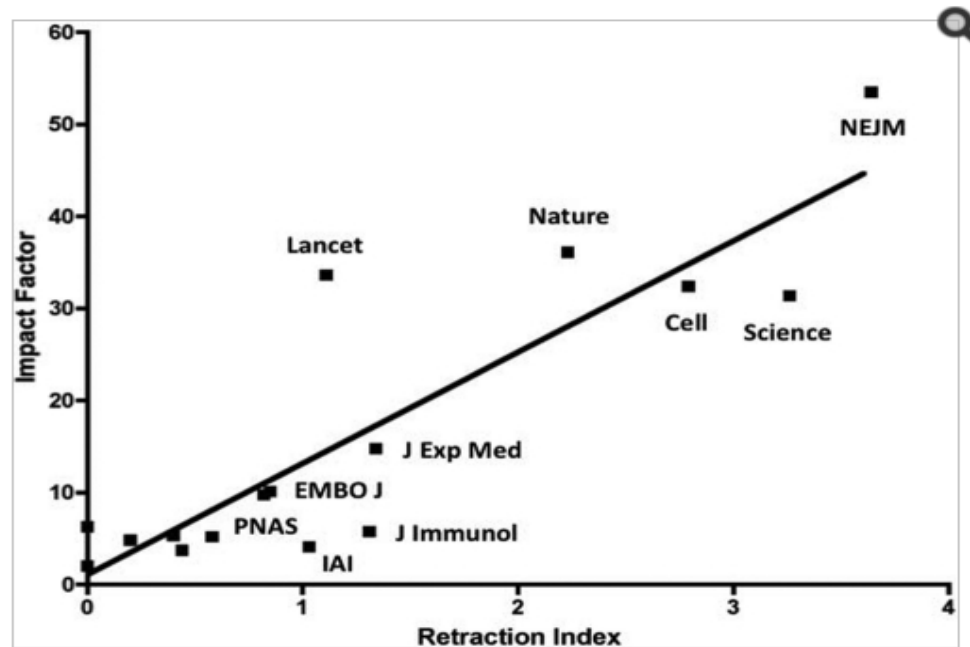
doi: [10.1128/IAI.05661-11](#)

<< Prev Next >>

[Copyright/License ▶](#)

[Request permission to reuse](#)

Fig. 1.



Correlation between impact factor and retraction index. The 2010 journal impact factor (37) is plotted against the retraction index as a measure of the frequency of retracted articles from 2001 to 2010 (see text for details). Journals analyzed were *Cell*, *EMBO Journal*, *FEMS Microbiology Letters*, *Infection and Immunity*, *Journal of Bacteriology*, *Journal of Biological Chemistry*, *Journal of Experimental Medicine*, *Journal of Immunology*, *Journal of Infectious Diseases*, *Journal of Virology*, *Lancet*, *Microbial Pathogenesis*, *Molecular Microbiology*, *Nature*, *New England Journal of Medicine*, *PNAS*, and *Science*.

# 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

# Schema



Properties

39

Based on Analyses,  
Feedback and Experience

Version

2.2

# Icons



The research data repository provides additional information on its service.

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



The research data repository provides a policy.



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.



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

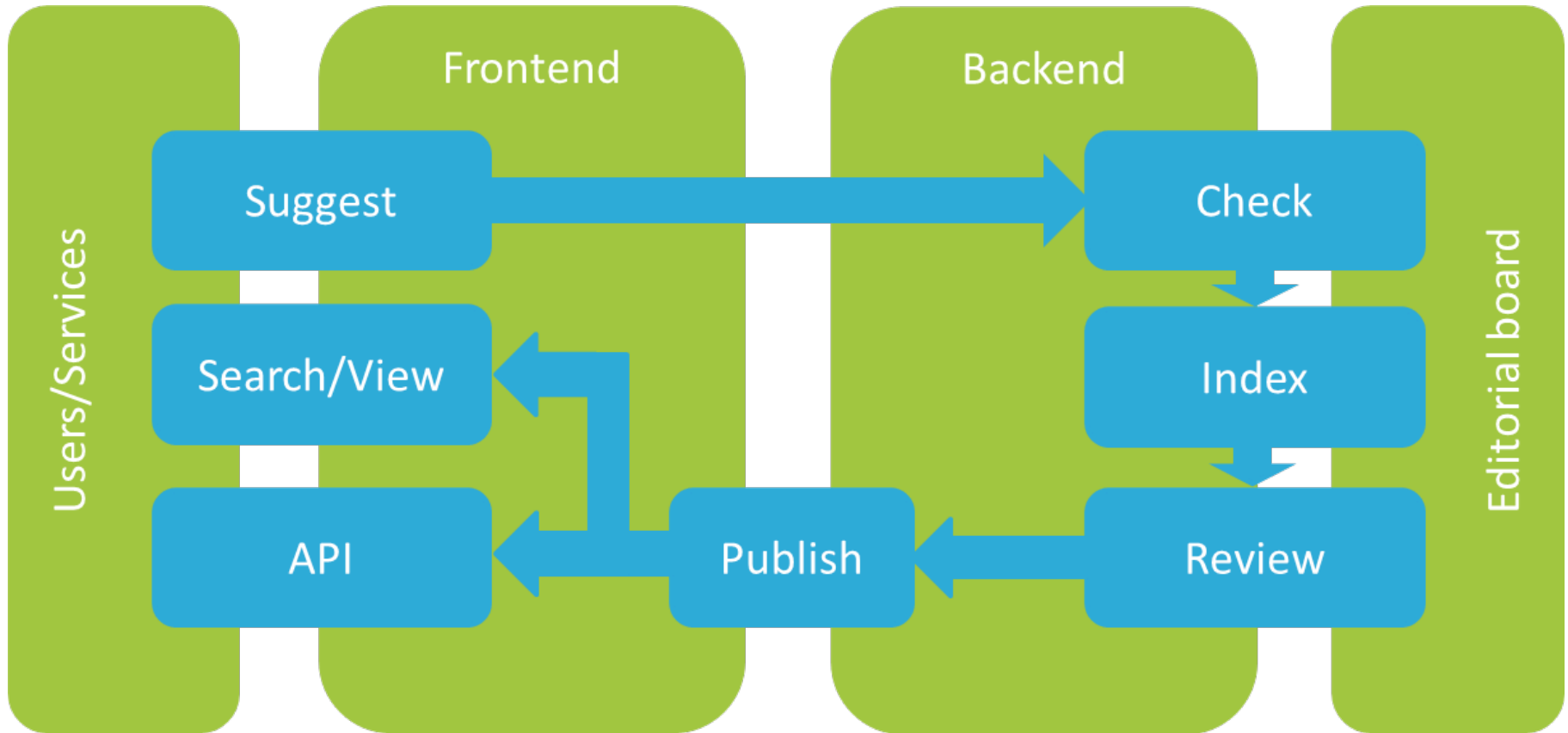
# Quality

## Requirements:

- 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 **focus on research data**



# Workflow



## Search for Repositories (1275 Reviewed Repositories)

geosciences

Search

Subject

Content Type

Country (of the responsible institutions)


Add subjects

Add content types

Add countries

Germany

Certificates

 Open Access

 Persistent Identifier

Reset filter

3 results (filtered) ( 1 - 3 )

Sort by **weight**

### PANGAEA

Publishing Network for Geoscientific and Environmental Data



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

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.

simple search box

filters

results

icons

Repository details



## PANGAEA

General Institutions Terms Standards

Name of repository	PANGAEA
Additional name(s)	Publishing Network for Geoscientific and Environmental Data
Repository URL	<a href="http://www.pangaea.de">http://www.pangaea.de</a>
Subject(s)	<a href="#">Oceanography</a> <a href="#">Geology and Palaeontology</a> <a href="#">Geophysics</a> <a href="#">Geochemistry, Mineralogy and Crystallography</a> <a href="#">Biology</a> <a href="#">Atmospheric Science and Oceanography</a> <a href="#">Geosciences (including Geography)</a> <a href="#">Natural Sciences</a> <a href="#">Geology and Palaeontology</a> <a href="#">Geophysics and Geodesy</a> <a href="#">Geochemistry, Mineralogy and Crystallography</a> <a href="#">Life Sciences</a>
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)	<a href="#">Standard office documents</a> <a href="#">Images</a> <a href="#">Plain text</a> <a href="#">Archived data</a> <a href="#">Audiovisual data</a>
Keyword(s)	<a href="#">Earth Science</a> <a href="#">Environmental Science</a>
Repository type(s)	disciplinary
Research data repository language(s)	eng
Data and/or service provider	dataProvider

# Integration of re3data in Guidelines

- 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](http://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.

[http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-pilot-guide\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf)

- 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.“

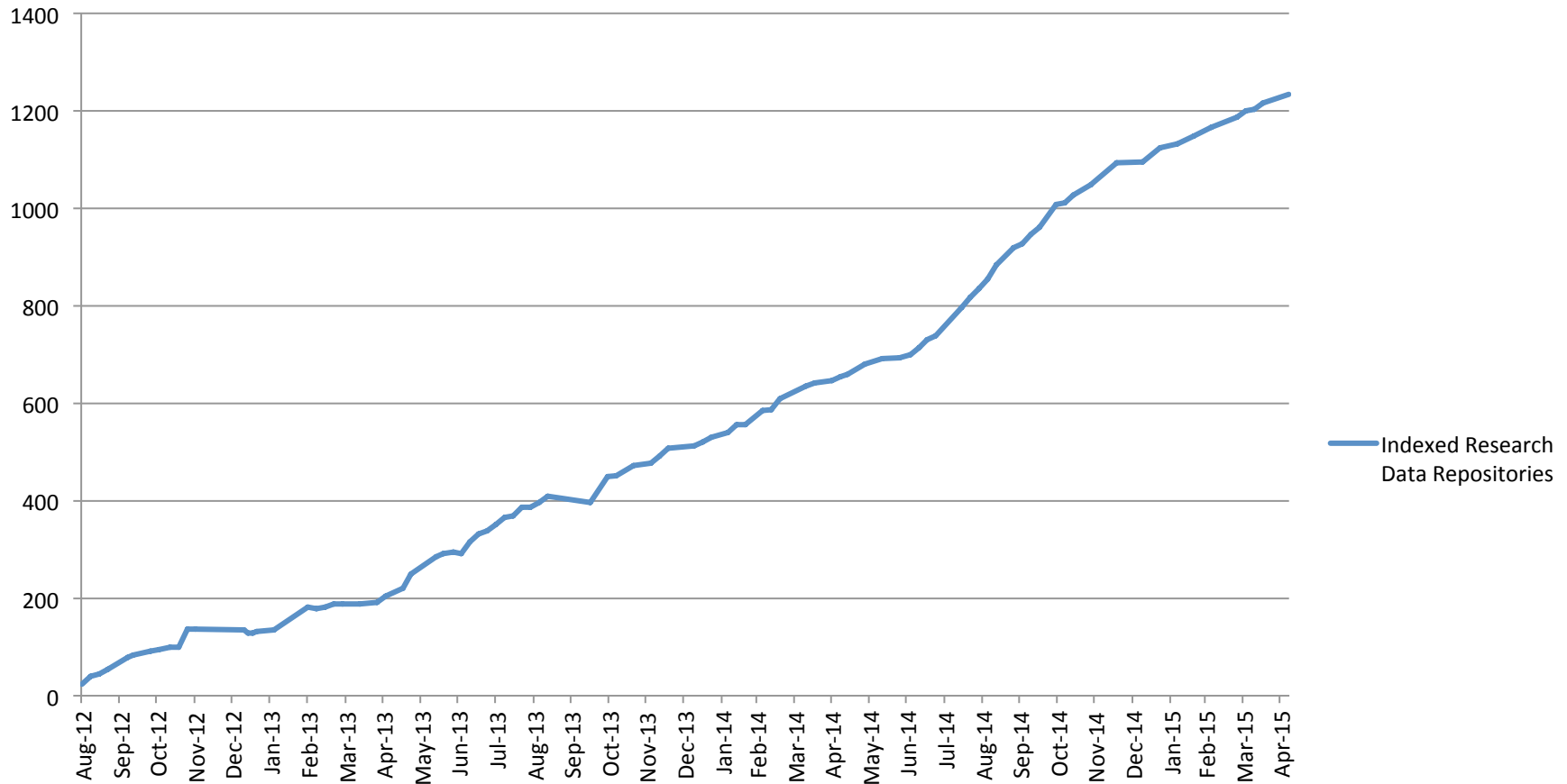
Universität Bielefeld (2011): Resolution zum Forschungsdatenmanagement. <https://data.uni-bielefeld.de/de/resolution>

- Publisher Example: Nature Publishing Group

- „Physics, astrophysics, astronomy and geoscience databases should be registered with [re3data.org](http://re3data.org).“

Scientific Data (2013): Data policies. <http://www.nature.com/sdata/data-policies>

# RDR indexed by re3data



# 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)



## 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/>

