

re3data.org  
REGISTRY OF RESEARCH DATA REPOSITORIES

NFDI4Ing Community Meeting CC-41

🔍 2022

## **NFDI4Ing Community Meeting CC-41**

Workshop: re3data – Indexing and discovering research data repositories for the engineering sciences

March 3rd, 2022

15:30 pm - 16:30 pm CET





# Introduction

**Rouven Schabinger**

Karlsruhe Institute of Technology

**Dorothea Strecker**

Humboldt-Universität zu Berlin

**Robert Ulrich**

Karlsruhe Institute of Technology

**Nina Weisweiler**

Helmholtz Open Science Office

## Agenda

Introduction to the service

Indexing workflow

Subject classification

Repositories of NFDI4Ing institutions

Please ask for our  
Brunch menu!





## Mission

global registry of research data repositories

covers all academic disciplines

presents repositories and portals for the permanent storage and access of research data sets to researchers, funding bodies, publishers and scholarly institutions.

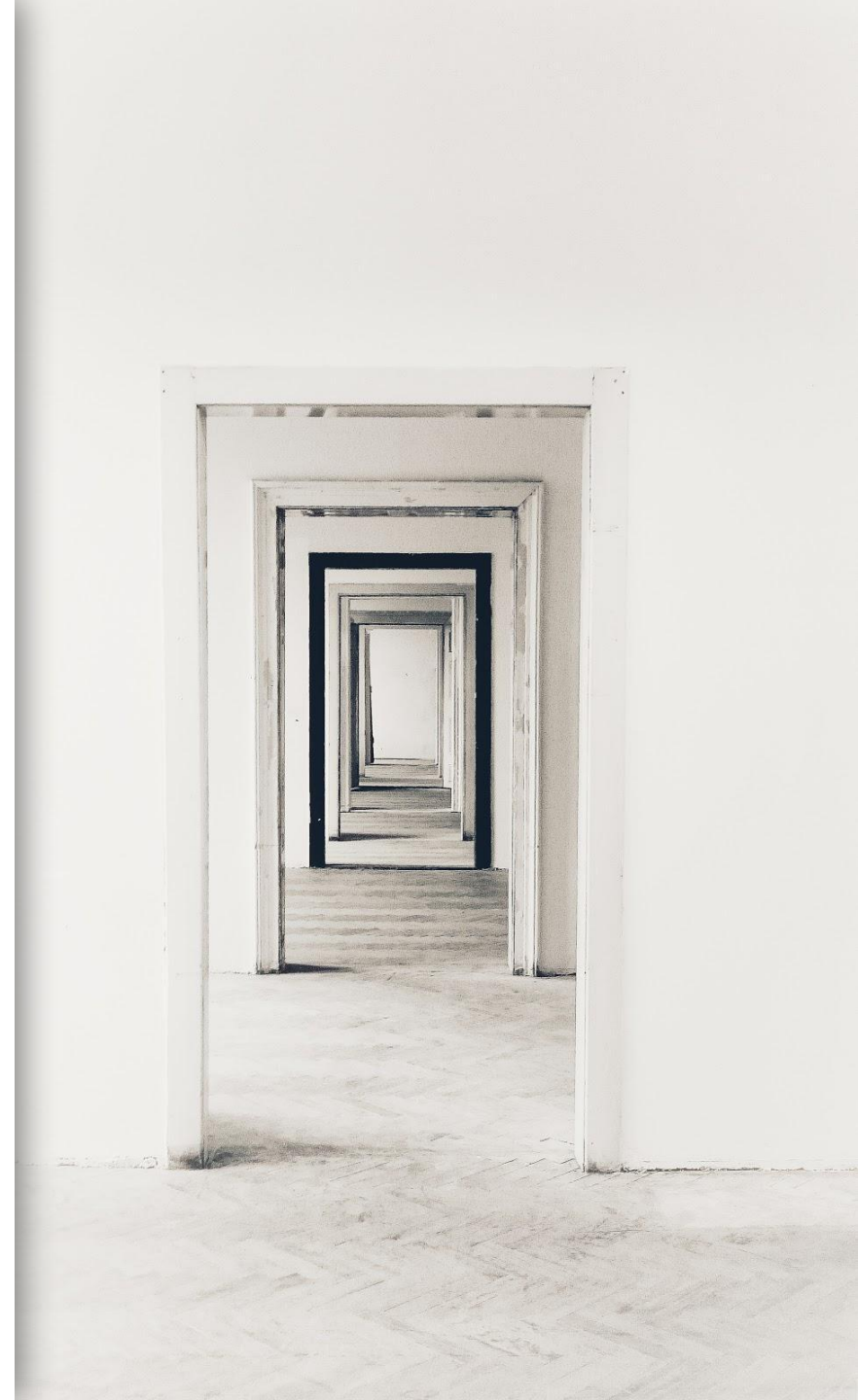
promotes a culture of sharing, increased access and better visibility of research data

## Registration Policy

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**





# Metadata

42 properties on



General information

Responsibilities

Policies

Legal aspects

Technical standards

Quality standards

## Web

URL: <https://www.re3data.org/>

Search

Facettes & Filters

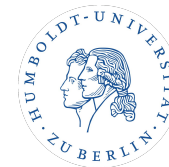
Widget & Badges

API





**Sustainability**



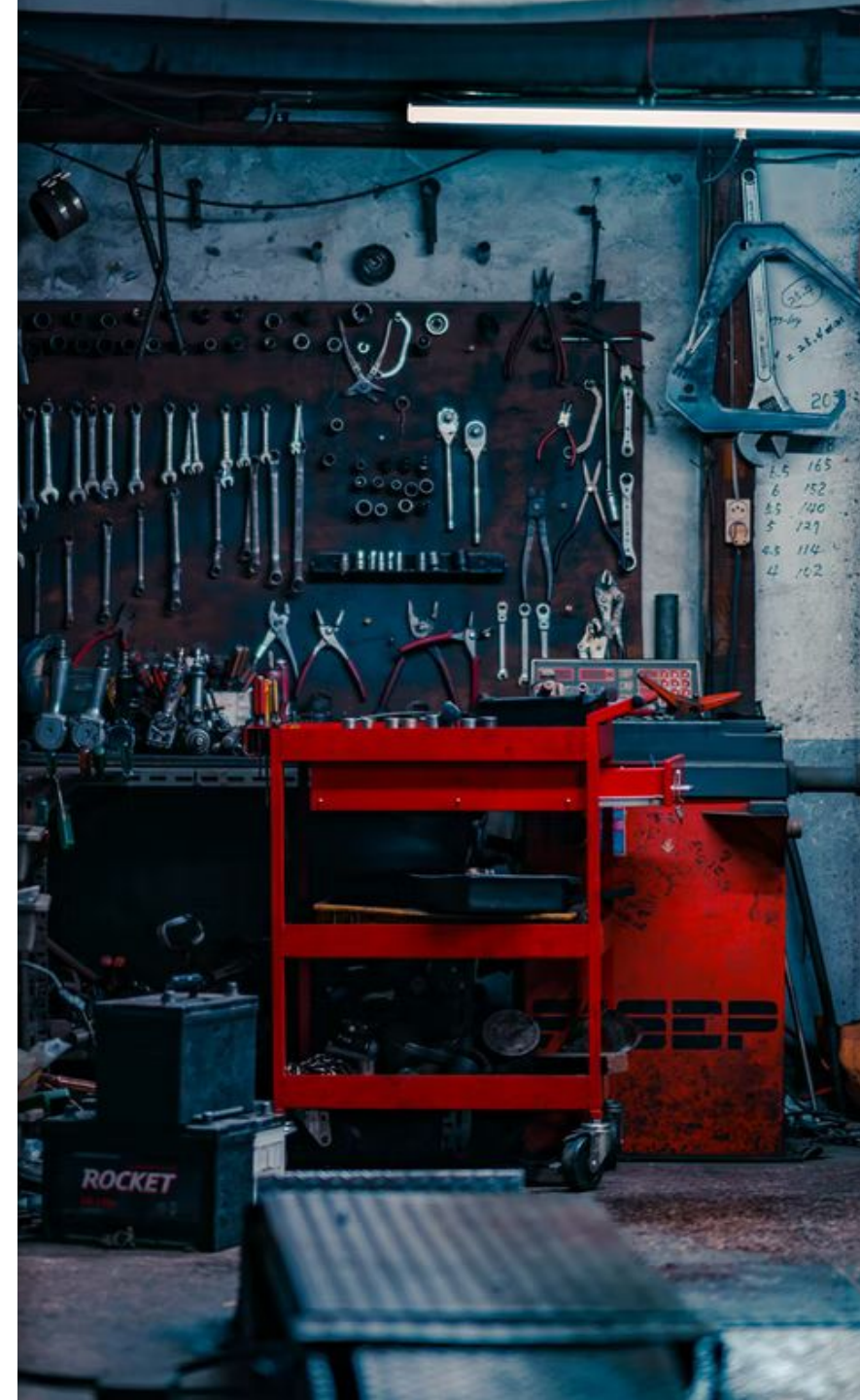


## Projects

FAIRsFAIR / FAIR-Impact

re3data COREF

(Community Driven Open Reference for  
Research Data Repositories)





# FAIRsFAIR

## 1. Foster FAIR implementations

Create incentives for repositories to implement FAIR enabling functions and get certified to be visible and promoted as FAIR enabling in DataCite Commons and re3data.

## 2. Promote FAIR culture

Increasing visibility and promoting FAIR enabling repositories within the scientific community. re3data is the common tool to guide researchers in the repository landscape and providing metrics

## 3. Support FAIR ecosystem

DataCite Commons will enable users & services to explore the broader context of FAIR enabling repositories via connected PIDs

# re3data COREF

## Service Model

Embed the registry within the research community and the infrastructure landscape to meet the emerging needs for a trusted repository reference.

## Metadata schema

Revision and enhancement of the re3data metadata schema on research data repositories.

## Technical infrastructure

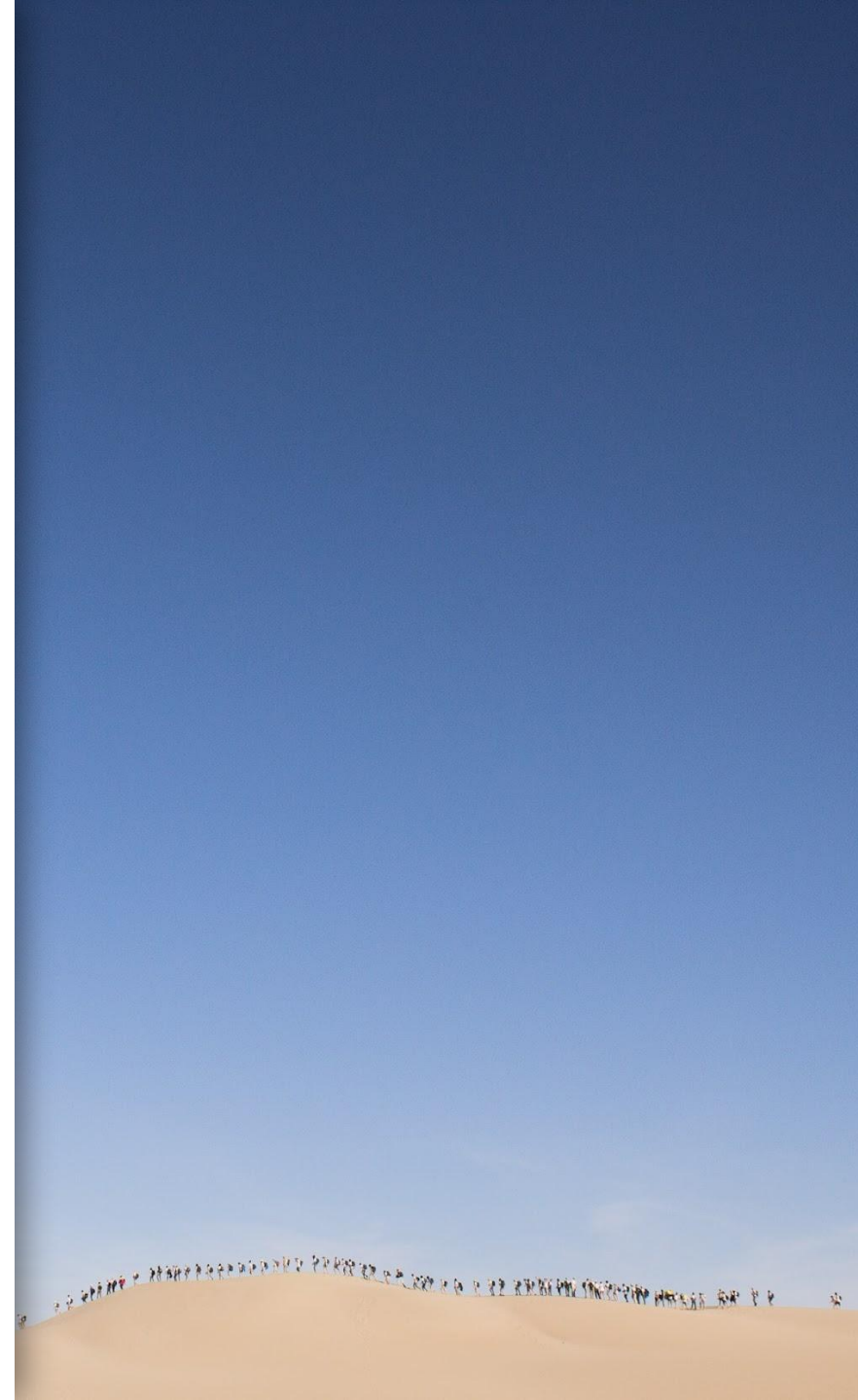
Implement the developed concepts, update the metadata schema implementation and provide the technical infrastructure and services needed.

## Quality measures

Conduct a comprehensive study on research data repository quality assurance amongst others to improve re3data metadata information; improve re3data metadata quality, trust and transparency.

## Community integration

Extensive activities for outreach, networking and policy development.



## Data provider

Open Science Monitor

SNF

DARIAH

...



# What makes a good registry record?

Search

Browse ▾

Suggest

Resources ▾

Contact

re3data.org  
REGISTRY OF RESEARCH DATA REPOSITORIES



Search...



Search

Rouven Schabinger

# Required information

Required information	
<b>Repository name</b>	<input type="text"/>
<b>Repository name language</b>	<input type="text" value="English"/> 
<b>Repository url</b>	<input type="text"/>
<b>Description</b>	<input type="text"/>
<b>Description language</b>	<input type="text" value="English"/> 
<b>Data licenses</b>	<a href="#">+ Add dataLicenses</a>
<b>Suggester's email</b>	<input type="text"/>

# General

- Additional names
- **Subjects**
- **Repository contacts**
- Content types
- **Certificates**
- Keywords
- **Repository identifiers**
- Size
- **Types**
- Mission statement URL
- Start date
- Repository languages
- Provider types

## PANGAEA

Data Publisher for Earth and Environmental Science

<https://www.pangaea.de/>

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

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.

<https://www.pangaea.de/contact/>

Source code   Standard office documents   Images   Plain text   Archived data   Audiovisual data

## CoreTrustSeal

lithosphere   paleontology   atmosphere   ecology   biosphere   land surface   cryosphere   fisheries  
agriculture   earth science   environmental science   biology

FAIRsharing\_doi:10.25504/FAIRsharing.6yw6cp

disciplinary

<https://www.pangaea.de/about/>

English

data provider

# Institutions

- Name
- Name language
- **Country**
- Type
- URL
- Responsibility start date
- Responsibility end date
- Additional names
- Contacts
- Identifiers
- **Responsibility types**

Alfred Wegener Institute - Helmholtz Centre for Polar and Marine Research

AWI

Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung

ROR:032e6b942

<https://www.awi.de/en/about-us/service/contact.html>

Germany

general

technical

non-profit



# Terms

- Policies
- Database access
- Data accesses
- Data uploads
- Data upload licenses

## Policies (2)

Policy Name [CoreTrustSeal assessment](#)

Policy Name [Data policy of the information system PANGAEA](#)

## Database access

Type of access to research data repository [open](#)

## Database licenses (1)

Database License [CC0](#)

## Data access (1)

Type of access to data [open](#)

## Data licenses (4)

DataLicense [CC](#)

DataLicense [other](#)

<i>Access (property)</i>	<i>Open Access</i>		<i>Restricted Access</i>		<i>Closed Access</i>
Access to Repository (20.1 databaseAccessType)	open		open or restricted		closed
Access to Data (22.1 dataAccessType)	open (embargoed, restricted, closed)		restricted (embargoed, closed)		closed
Data Upload (24.1 dataUploadType)	open or restricted	closed	open or restricted	closed	-

# Standards

- Software
- Versioning
- **PID systems**
- Citation guideline URL
- AID systems
- Enhanced publication
- Quality management
- **APIs**
- **Metadata standards**
- Syndications
- Remarks

DOI

other

yes

<https://wiki.pangaea.de/wiki/Citation>

ORCID

unknown

yes

## interfaces (1)

OAI-PMH

Darwin Core

DCC

ISO 19115

DCC

Dublin Core

DCC

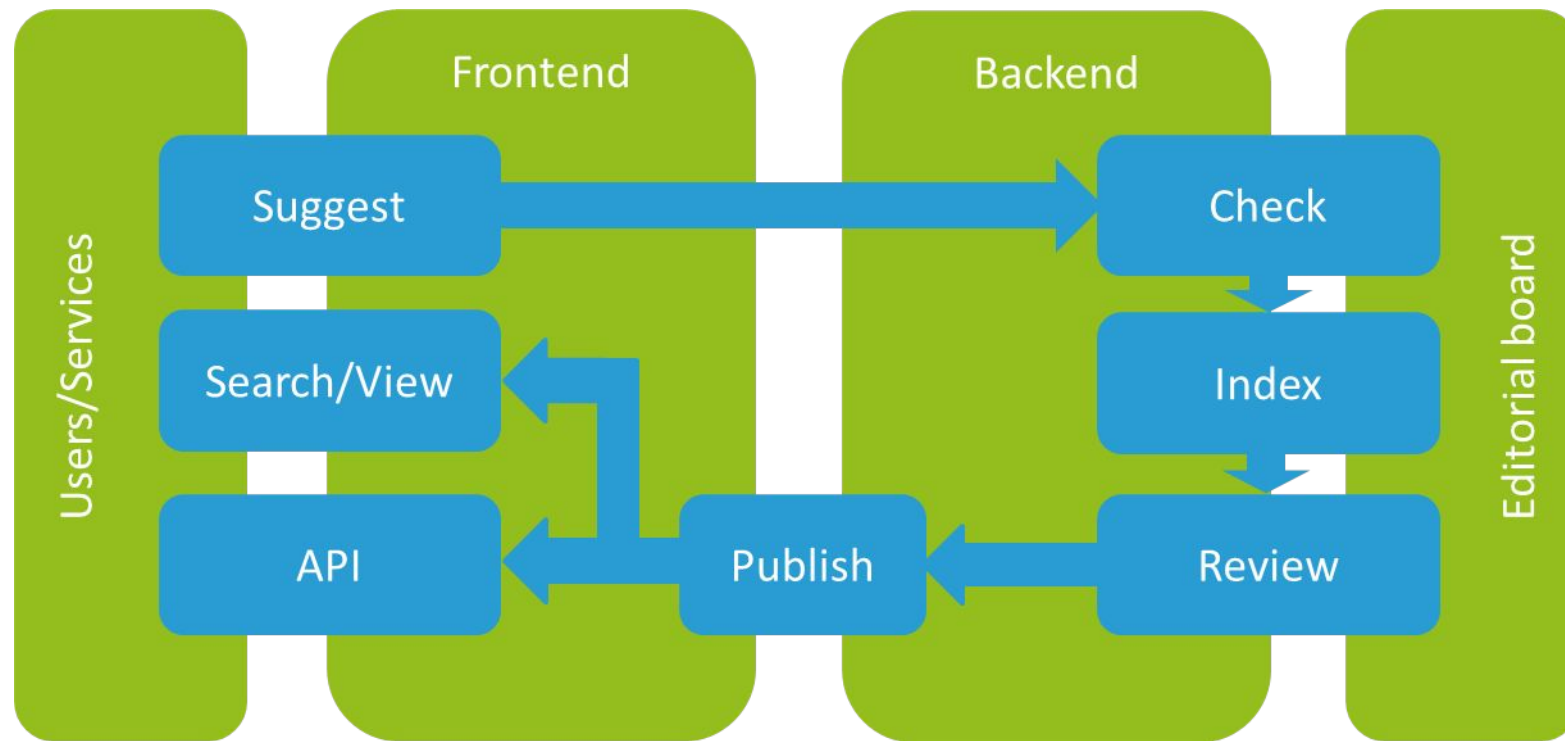
DIF - Directory Interchange Format

DCC

# Icons – facilitating the selection process of appropriate research data repositories



# Workflow



re3data Research Data Repository Registration

# Editorial Board

**Catherine Jones**

Science and Technology Facilities Council (STFC)



**Jiban K. Pal**

Indian Statistical Institute (ISI)



**Iris Lindenmann**

University of Basel

**Rouven Schabinger**

Karlsruhe Institute of Technology (KIT)



**Angelika Semrau**

Karlsruhe Institute of Technology (KIT)

**Edeltraud Schnepf**

Karlsruhe Institute of Technology (KIT)



**Gail Steinhart**

Cornell University

**Hui Wang**

Chinese Academy of Sciences (CAS) / National Science Library (NSL)



**Gabriele Weickert**

Karlsruhe Institute of Technology (KIT)

**Sarah Williams**

University of Illinois



**Michael Witt**

Purdue University

# Keep your record up-to-date

- automated (e.g. CTS)
- Check ups by Editorial Team
- your change request?

Contact	<a href="https://nanohub.org/about/contact">https://nanohub.org/about/contact</a>
Content type(s)	<a href="#">Databases</a> <a href="#">Audiovisual data</a> <a href="#">Software applications</a>
Keyword(s)	<a href="#">Nanomaterial Registry</a> <a href="#">nanoBIO</a> <a href="#">nanoscience</a> <a href="#">nan</a>
Persistent identifier(s) of the repository	RRID:SCR_013963 OMICS_27120
Repository size	6.139 resources
Repository type(s)	disciplinary
Mission statement for designated community	<a href="https://nanohub.org/about">https://nanohub.org/about</a>
Research data repository language(s)	English
Data and/or service provider	data provider service provider

[Back to search](#)

[Submit a change request](#)

## Submit a change request

Make changes to the properties that need an update. The editorial board will review the submitted record and put it online

General

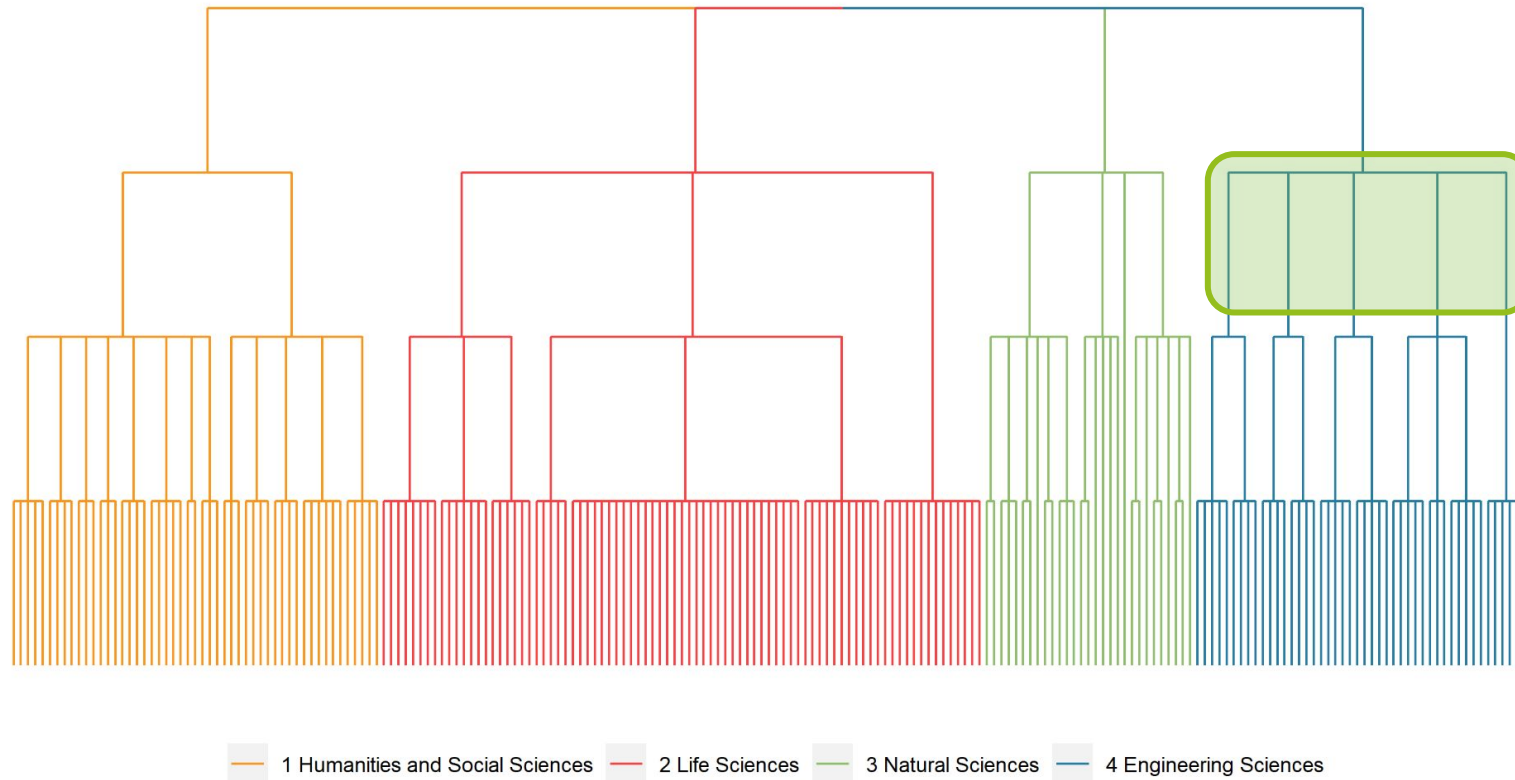
Repository name	<input type="text" value="nanoHUB"/>
Repository name language	<input type="text" value="English"/>
Additional names	<p><b>Text</b> <input type="text" value="nanoHUB.org"/></p>
Language	<input type="text" value="English"/>
Repository url	<input type="text" value="https://nanohub.org/"/>
Subjects	<p><b>Name</b> <input type="text" value="31 Chemistry"/></p> <p><b>Scheme</b> <input type="text" value="DFG"/></p>

pre-filled  
metadata  
fields

[Get a badge](#)

# Subject classification

Repository subjects in re3data are assigned based on the DFG subject classification.



## NFDI4Ing Community Clusters:

- Mechanical and industrial engineering
- Thermal engineering and process engineering
- Materials science and engineering
- Computer science, systems and electrical engineering
- Construction engineering and architecture

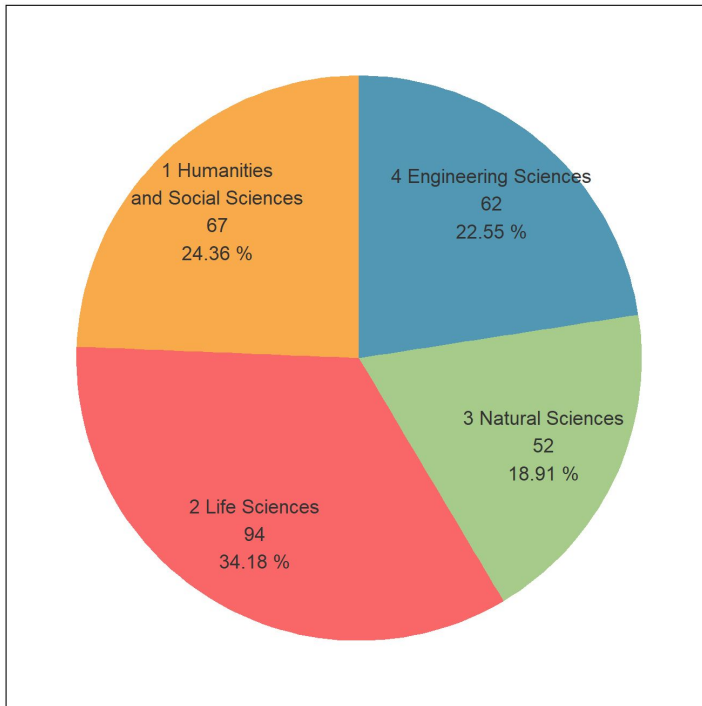
If you are interested in subject classification in re3data, we recommend [this blog post](#). The blog post and these slides are based on re3data metadata extracted in June 2021.

# Subject classification

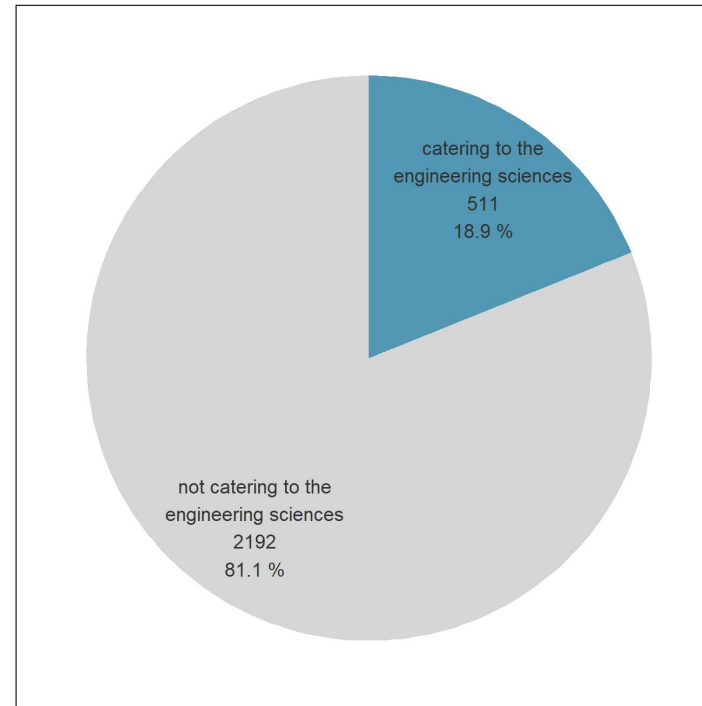
The DFG subject classification was not developed with the description of repositories in mind, its primary purpose is to organize DFG funding. Therefore, if the classification is used for other purposes, the level of detail the DFG subject classification provides for a discipline does not necessarily match the objects being described, for example research data repositories.

18.9 % (511) of all repositories indexed in re3data are assigned **at least one** notation from “Engineering Sciences”.

distribution of notations across subject areas (level 1) in the DFG subject classification



proportion of repositories indexed in re3data catering to the engineering sciences

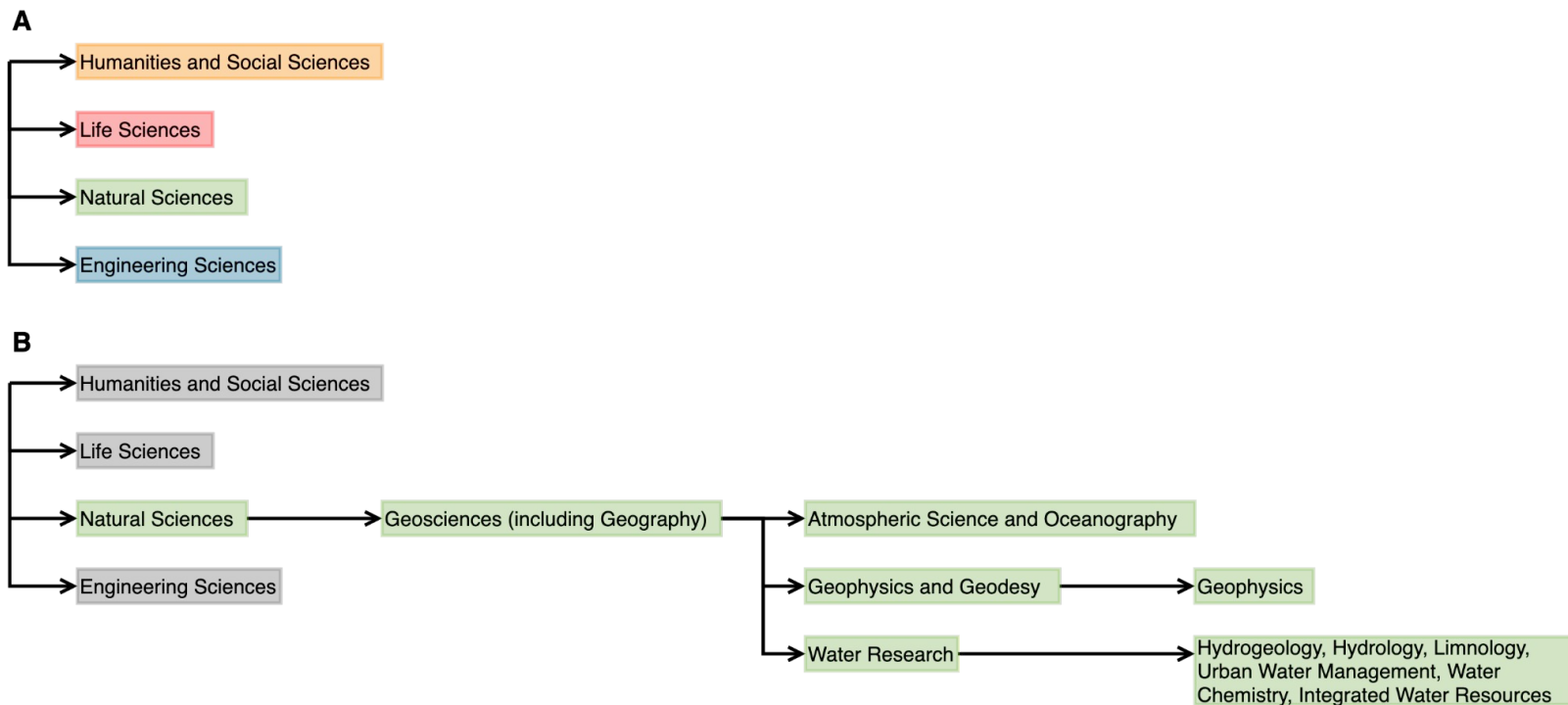




# Subject classification

Of these, only 7.24 % (37) are **exclusively** assigned notations from “Engineering Sciences”; 60.47 % (309) are assigned notations from **all four** disciplines.

In re3data, notations from “Engineering Sciences” are often used to indicate multidisciplinary:

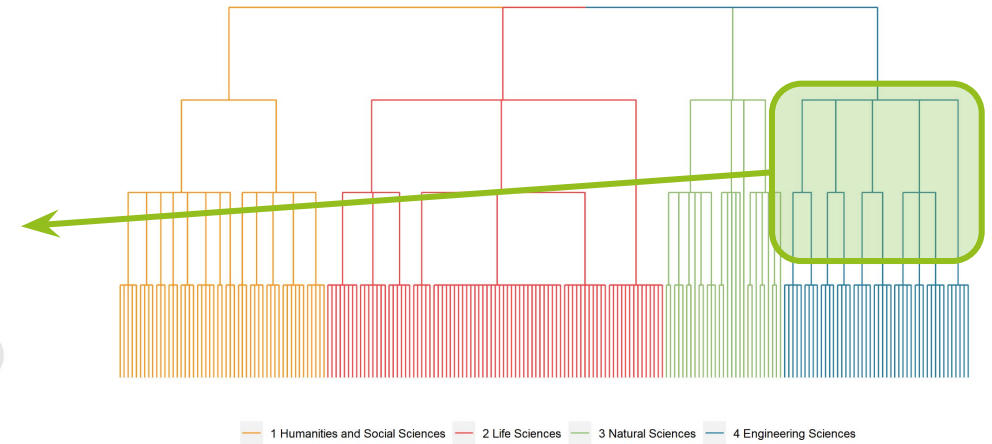
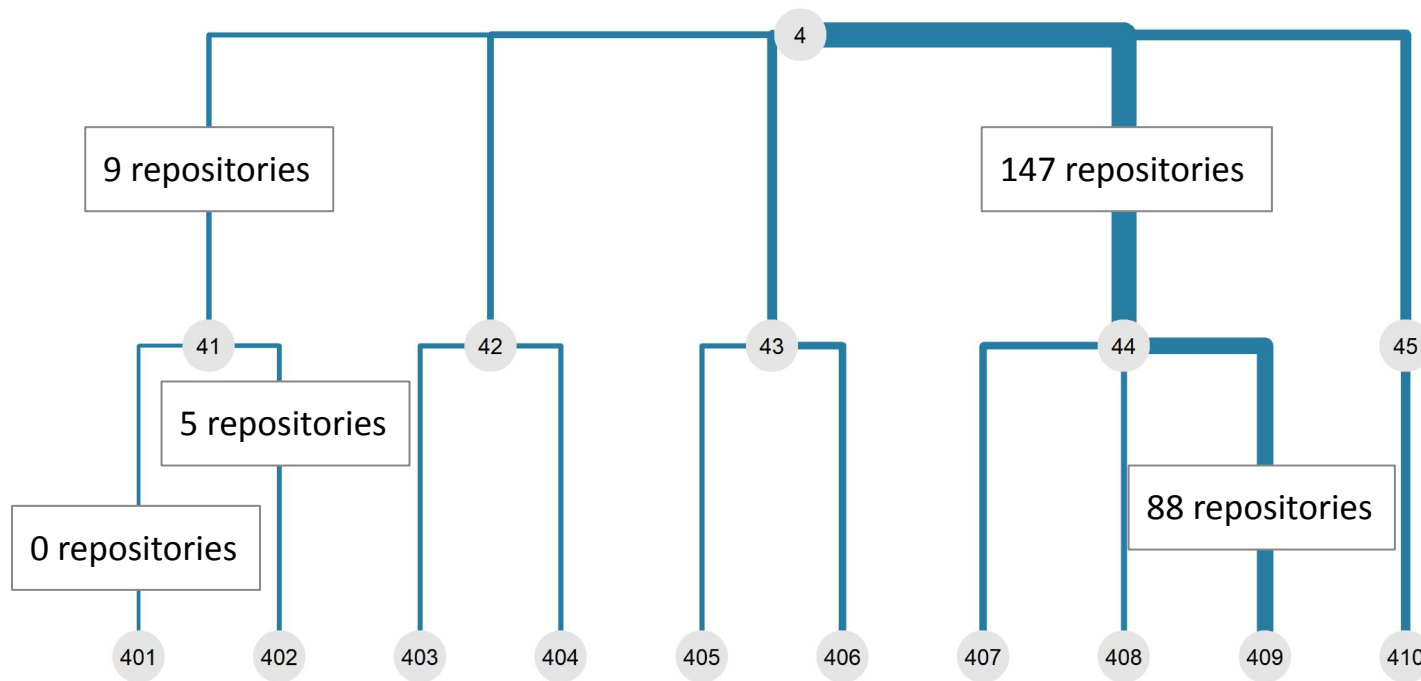


subjects assigned to Zenodo (A) and National Center for Atmospheric Research (B)

# Subject classification

re3data editors generally select the most specific notations applicable, and also add notations higher up in the hierarchy. This results in notations "flowing upwards".

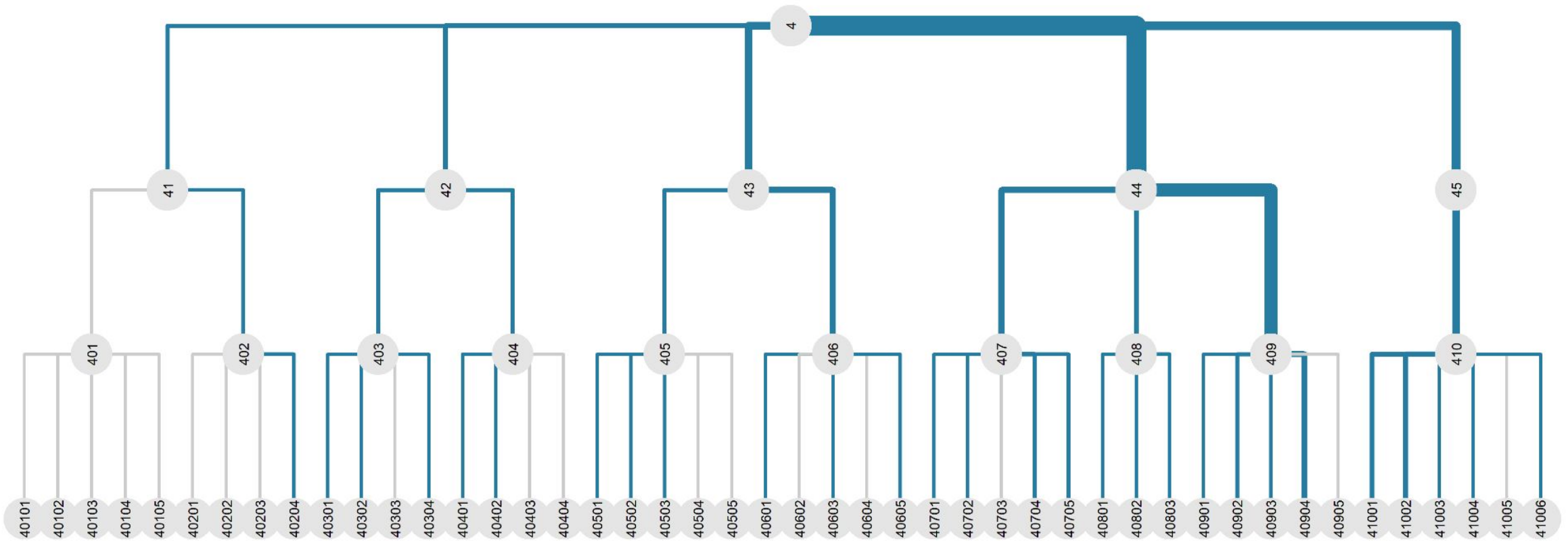
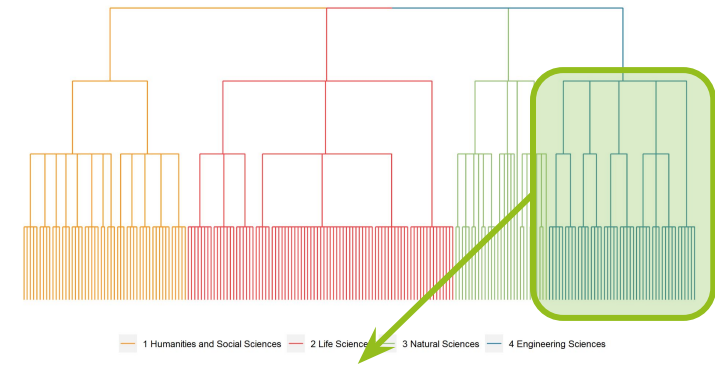
Notations from certain areas tend to be used more frequently, for example 409 (Computer Science). As a result, 44 (Computer science, systems and electrical engineering) directly above it is also more common.



depth of notations from "Engineering Sciences"; level 2 and 3 of the DFG subject classification

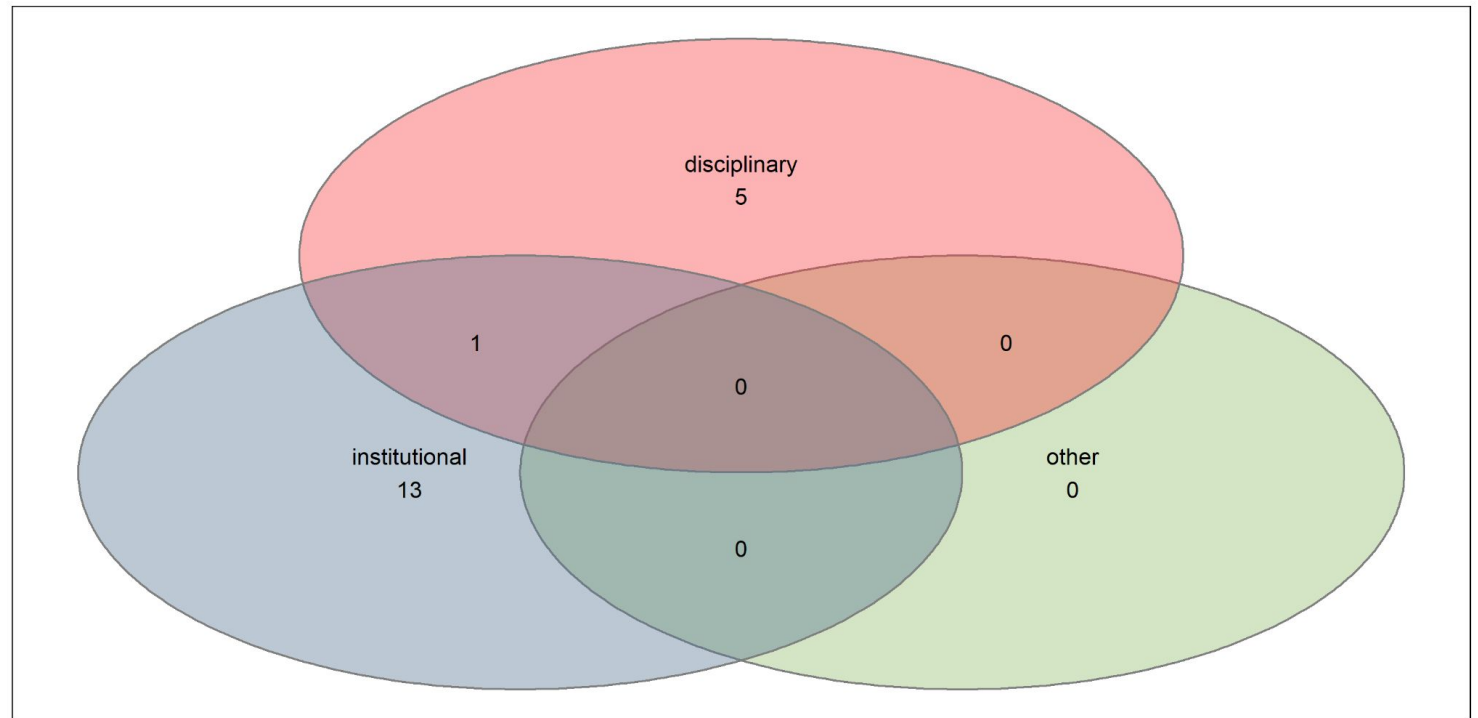
# Subject classification

depth of “Engineering Sciences” notations



depth of notations from “Engineering Sciences”; level 2, 3 and 4 of the DFG subject classification; nodes with grey incoming edges are not used in re3data

# Repositories of NFDI4Ing institutions



As of February 2022, 18 repositories indexed in re3data are affiliated with NFDI4Ing (co-)applicant and participant institutions and also cater to the engineering sciences.

Most of them are institutional repositories.

- maintains repository catering to the engineering sciences
- does not maintain repository catering to the engineering sciences

# Repositories of NFDI4Ing institutions

## How can the representation of the engineering sciences in re3data be improved?

### re3data suggest form:

- Suggest repositories relevant to you that are not yet indexed in re3data
- Update existing re3data entries

### feedback:

- Contact us for improvements and requirements, e.g. search, suggest, schema, API etc:  
[info@re3data.org](mailto:info@re3data.org)

### research data management and key features of repositories in the engineering sciences?

- How is research data management in the engineering sciences organized? Where do you store your data?
- For key features of repositories, check the re3data icon system. What would you consider a “key feature”?

### models for continuous collaboration?

- New re3data metadata element “*profile*” (The selection of repositories based on a set of community-developed criteria.)



## Contact us!

Web: [www.re3data.org](http://www.re3data.org)

Mail: [info@re3data.org](mailto:info@re3data.org)

Twitter: [@re3data](https://twitter.com/re3data)



re3data.org  
REGISTRY OF RESEARCH DATA REPOSITORIES

Thank you!



Visit the project blog:  
[coref.project.re3data.org](http://coref.project.re3data.org)

## COREF Team:

[Roland Bertelmann](#) (Helmholtz-Gemeinschaft, Helmholtz Open Science Office)

[Helena Cousijn](#) (DataCite - International Data Citation Initiative e.V.)

[Kirsten Elger](#) (Deutsches GeoForschungsZentrum GFZ)

[Lea Maria Ferguson](#) (Helmholtz-Gemeinschaft, Helmholtz Open Science Office)

[Maxi Kindling](#) (Humboldt-Universität zu Berlin)

[Thanh Binh Nguyen](#) (Karlsruher Institut für Technologie (KIT))

[Heinz Pampel](#) (Helmholtz-Gemeinschaft, Helmholtz Open Science Office)

[Vivien Petras](#) (Humboldt-Universität zu Berlin)

[Rouven Schabinger](#) (Karlsruher Institut für Technologie (KIT))

Edeltraut Schnepf (Karlsruher Institut für Technologie (KIT))

Angelika Semrau (Karlsruher Institut für Technologie (KIT))

[Dorothea Strecker](#) (Humboldt-Universität zu Berlin)

[Margarita Trofimenko](#) (Karlsruher Institut für Technologie (KIT))

[Robert Ulrich](#) (Karlsruher Institut für Technologie (KIT))

[Arne Upmeyer](#) (Karlsruher Institut für Technologie (KIT))

[Paul Vierkant](#) (DataCite - International Data Citation Initiative e.V.)

[Yi Wang](#) (Karlsruher Institut für Technologie (KIT))

[Nina Weisweiler](#) (Helmholtz-Gemeinschaft, Helmholtz Open Science Office)

[Michael Witt](#) (Purdue University, United States)