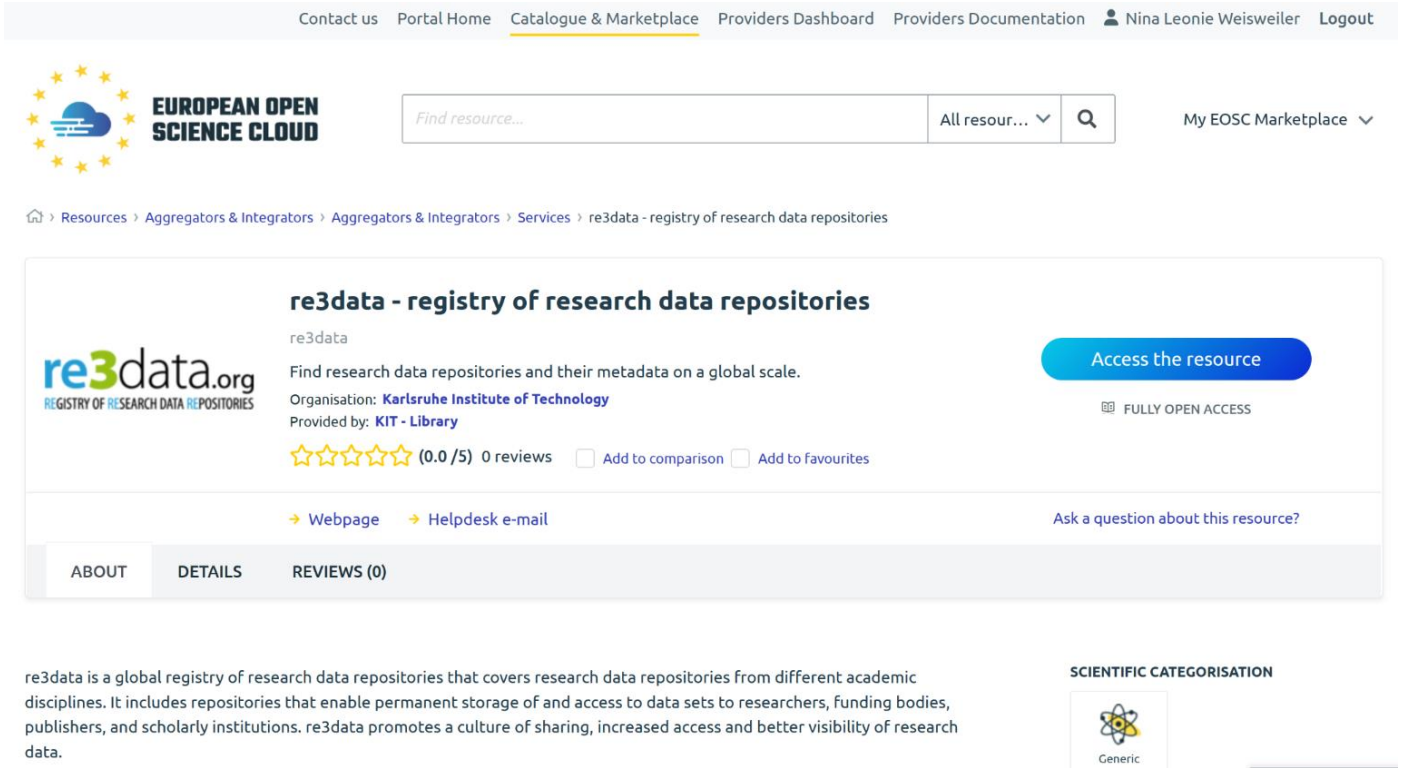
EOSC Symposium 2022

Talk 3: Engaging data repositories in EOSC
Session: Engagement of research communities and service providers in EOSC

Tuesday, November 15, 16:00 - 17:30

Speaker: Nina Leonie Weisweiler
DOI: https://doi.org/10.5281/zenodo.7334302

re3data: Connected via EOSC Marketplace



EU Projects with re3data participation supporting the EOSC

2019-2022:

Tools for finding and selecting FAIR enabling repositories to researchers and other stakeholders

Integration of repository information into DataCite Commons

- re3data as metadata source
- linking with other entities of the PID graph
- providing additional insights

Launched in Summer 2022:

Practical implementation of the FAIR principles across all domains through FAIR-enabling practices, tools and services.

More at: https://fair-impact.eu/

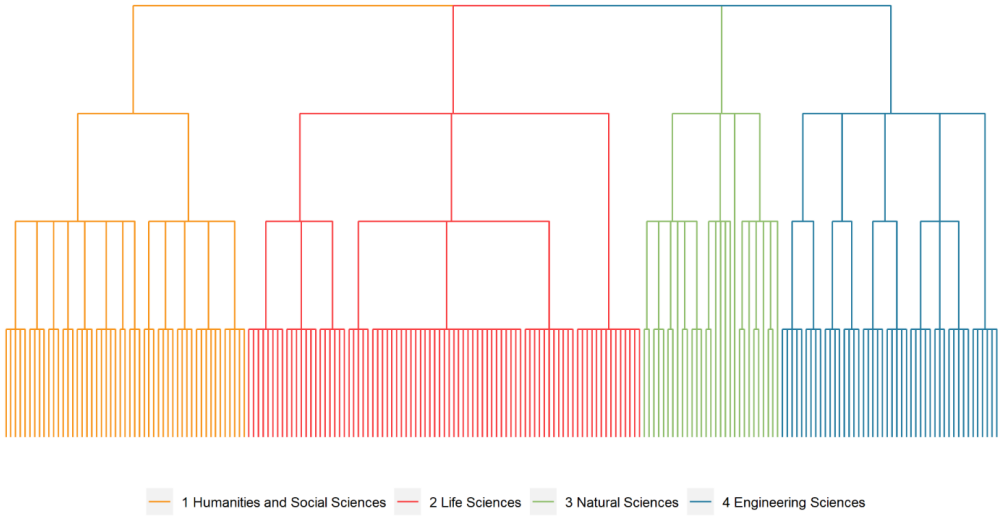
Networks and integrations

re3data is a **central reference point and metadata source** for other service providers on the European and international level:

- OpenAIRE's Open Science Observatory
- European Open Science Monitor
- F-UJI Automated FAIR Data Assessment Tool
- DARIAH-EU Data Deposit Recommendation Service
- DataCite Commons
- and many others…

Collaboration with EUDAT service B2FIND

re3data collaborates with the B2FIND team with the goal of harmonizing their subject classifications.



See also: https://coref.project.re3data.org/blog/reviewing-the-subject-classification-in-re3data

Global Registry of Research Data Repositories

re3data.org

REGISTRY OF RESEARCH DATA REPOSITORIES

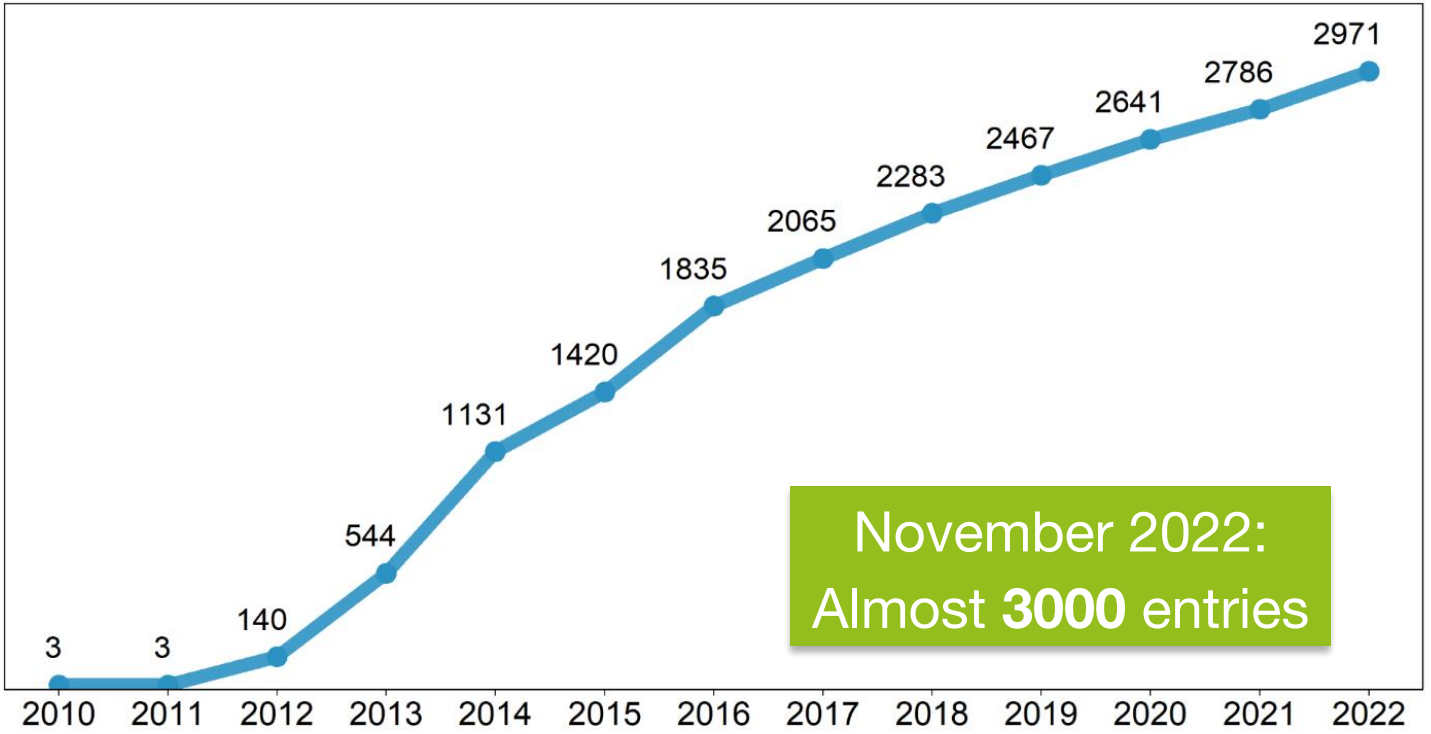
Search... Search

across **all research disciplines**

… presents **repositories** 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

Number of repositories indexed in re3data per year:



re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

Celebrating 10 Years of re3data

Current Projects:

re3data coref

eosc | FAIR-IMPACT
Expanding FAIR solutions across EOSC

2011
Start of the first DFG-funded project

2012
re3data website goes live

2013
Merge with DataBib

2014
Over 1000 indexed repositories

2015
re3data becomes a DataCite Partner Service

2016
Continuation of re3data beyond the DFG-funded project phase

2017
Over 2000 repositories indexed

2019
Start of the FAIRsFAIR project

2020
Start of the re3data COREF project (DataCite, Helmholtz Open Science Office, Humboldt-Universität zu Berlin, Karlsruhe Institute of Technology)

2021
Metadata Schema Version 3.1 & Conceptual Model for User Stories published (re3data COREF)

2022
re3data integration in DataCite Commons;
Start of the FAIR-IMPACT project

Partners:

DataCite
CONNECTING RESEARCH, IDENTIFYING KNOWLEDGE

HELMHOLTZ Open Science

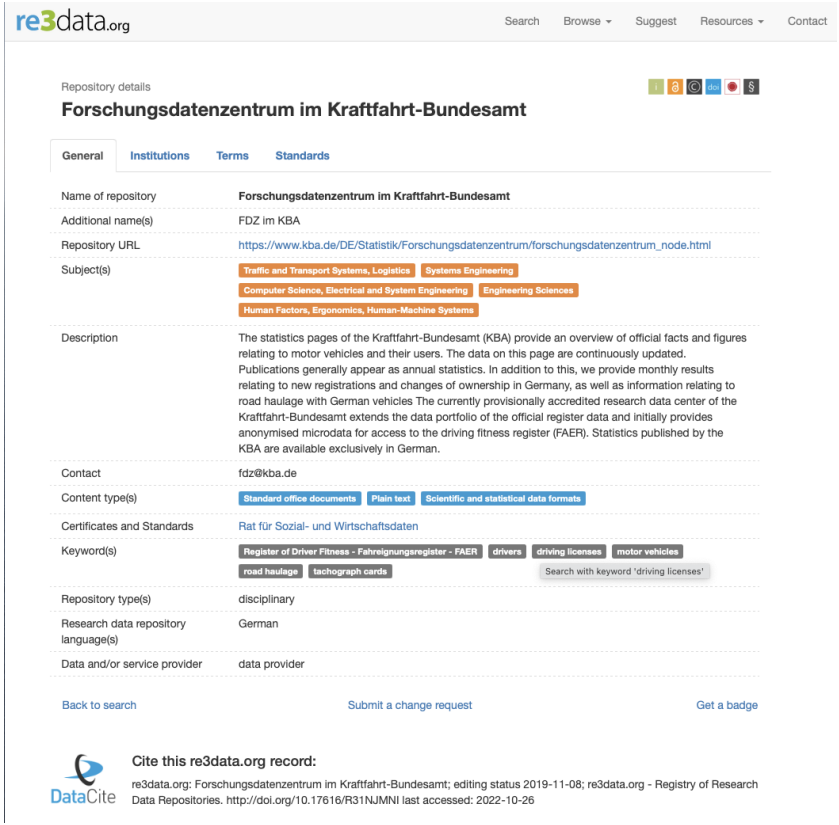
KIT Karlsruher Institut für Technologie

HUMBOLDT-UNIVERSITÄT ZU BERLIN

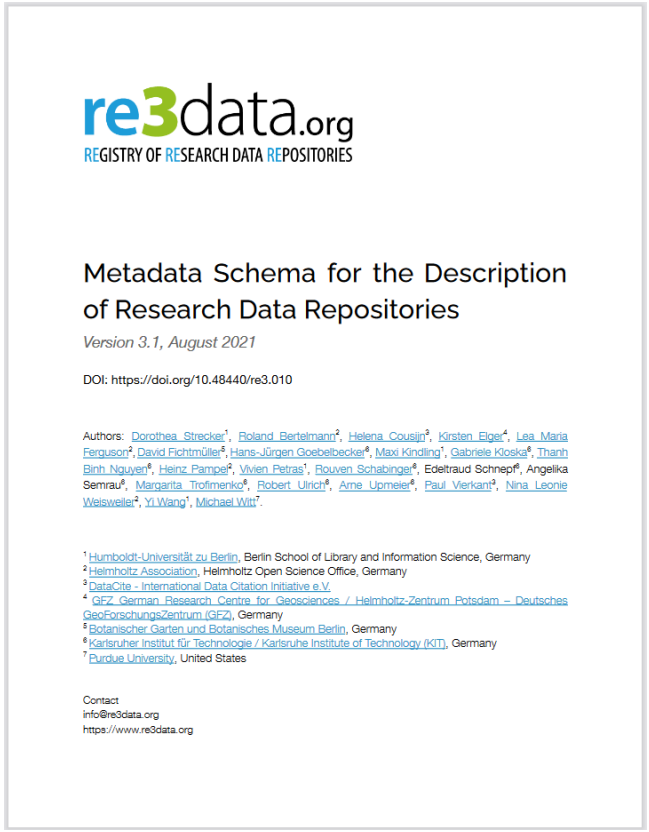
PURDUE UNIVERSITY

What is a “repository” in re3data?

- We refer to all re3data entries as “repositories”.
  - Repositories in re3data are defined as “[...] a subtype of a sustainable information infrastructure providing long-term storage of and access to research data.”(Metadata Schema for the Description of Research Data Repositories, Version 3.1)
- The re3data registration policy is intentionally inclusive to cover a broad range of infrastructures for research data.



re3data Metadata Schema



re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

Metadata Schema for the Description of Research Data Repositories

Version 3.1, August 2021

DOI: https://doi.org/10.48440/re3.010

Authors: Dorothea Strecker[1], Roland Bertelmann[2], Helena Cousijn[3], Kirsten Elger[4], Lea Maria Ferguson[2], David Fichtmüller[5], Hans-Jürgen Goebelbecker[6], Maxi Kindling[1], Gabriele Kloska[6], Thanh Binh Nguyen[6], Heinz Pampel[2], Vivien Petras[1], Rouven Schabinger[6], Edeltraud Schnepf[6], Angelika Semrau[6], Margarita Trofimenko[6], Robert Ulrich[6], Arne Upmeier[6], Paul Vierkant[3], Nina Leonie Weisweiler[2], Yi Wang[1], Michael Witt[7].

[1] Humboldt-Universität zu Berlin, Berlin School of Library and Information Science, Germany
[2] Helmholtz Association, Helmholtz Open Science Office, Germany
[3] DataCite - International Data Citation Initiative e.V.
[4] GFZ German Research Centre for Geosciences / Helmholtz-Zentrum Potsdam – Deutsches GeoForschungsZentrum (GFZ), Germany
[5] Botanischer Garten und Botanisches Museum Berlin, Germany
[6] Karlsruher Institut für Technologie / Karlsruhe Institute of Technology (KIT), Germany
[7] Purdue University, United States

Contact
info@re3data.org
https://www.re3data.org

Strecker S., et al. (2021): Metadata Schema for the Description of Research Data Repositories: version 3.1, https://doi.org/10.48440/re3.010

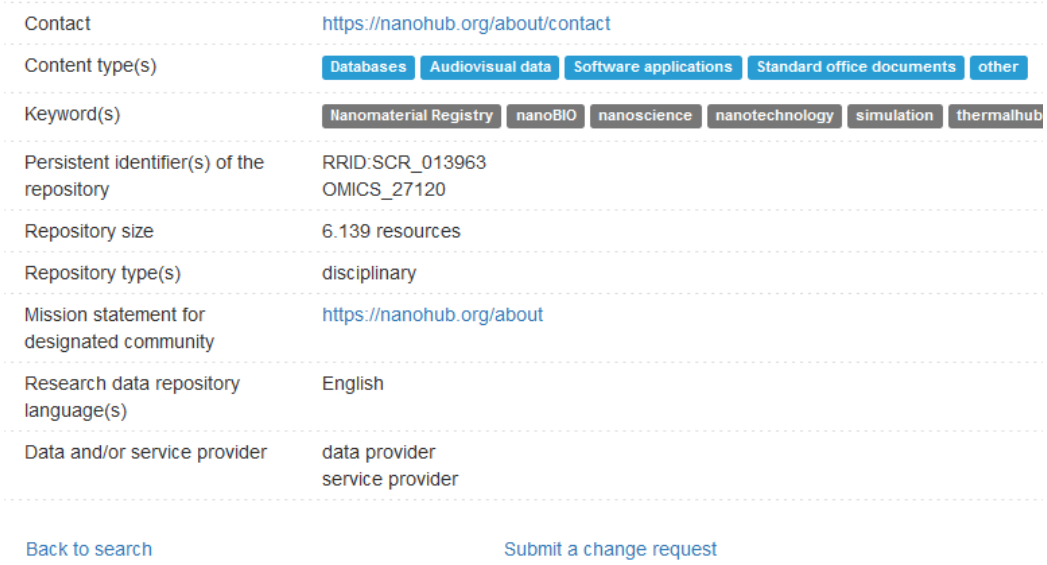
42 Properties on…

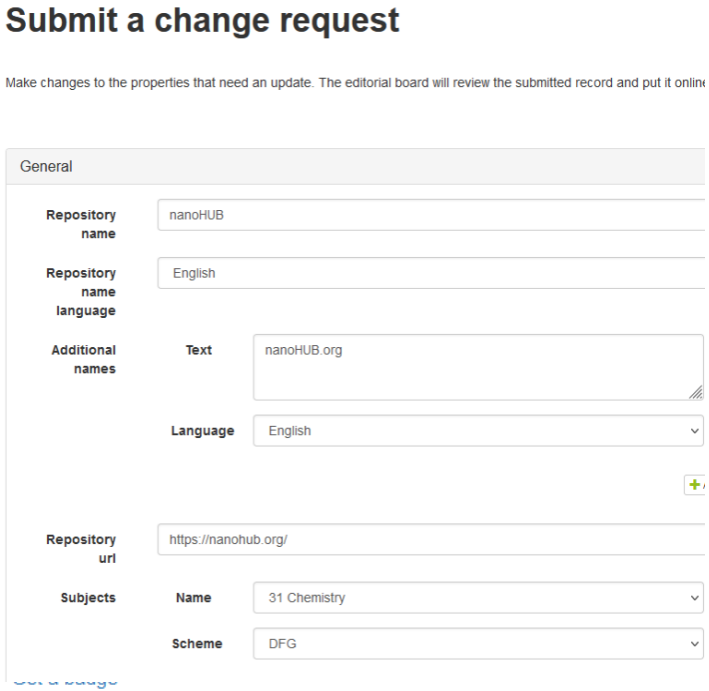
- General information
- Responsibilities
- Policies
- Legal aspects
- Technical standards
- Quality standards
- DOI for repository descriptions
- Since 2020: ROR IDs for institutions

Version 3.1 of the Metadata Schema:
https://www.re3data.org/schema

Input: Keeping records up-to-date

- Automated updates (e.g. CoreTrustSeal)
- Reviews by the editorial team
- Change requests:





Output: Querying the re3data API with R

https://github.com/re3data/using_the_re3data_API

There are many conceivable use cases for re3data metadata.

In this GitHub repository, we provide examples for using the re3data API that can be adapted to fit other use cases.

The example use cases are implemented in R and Python using Jupyter Notebooks.

See also:
https://coref.project.re3data.org/blog/using_the_re3data_api

Develop & enhance re3data: Current project

re3data coref
COMMUNITY DRIVEN OPEN REFERENCE FOR RESEARCH DATA REPOSITORIES

Conceptual Service Model

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

DataCite

Metadata Schema

- Revision and enhancement of the re3data metadata schema for the description of research data repositories.

HELMHOLTZ Open Science

Technical Infrastructure

- Implement new features and the updated metadata schema.
- Provide the necessary technical infrastructure and services.

HUMBOLDT-UNIVERSITÄT ZU BERLIN

Quality & Transparency

- Conduct a comprehensive study on quality assurances at research data repositories to improve re3data metadata information.
- Improve re3data metadata quality, trust and transparency.

KIT Karlsruher Institut für Technologie

Community Building

- Extensive outreach activities, networking and policy development.

Community engagement in re3data COREF

Conceptual Service Model

- Stakeholder survey & workshop series: Read more

Metadata Schema

- Public RFC process: Read more

Technical Infrastructure

- Improvements & development in consultation with the community
- API use cases examples: Read more

Quality & Transparency

- Survey on quality management at research data repositories & related workshop in collaboration with CoreTrustSeal: Check the project blog for updates!

Community Building

- Extensive outreach activities: Blog, social media, mailings, events & workshops…

New study finds that the share of researchers who store data in research data repositories is still low…

re3data is here to support change!

See tomorrow's Session "European Research Data Landscape (2021-2022)"

Wednesday, Nov 16, 14:30 - 16:00 CET

Join in!

Become a part of the EOSC network by registering your repository in re3data:

https://www.re3data.org/suggest

re3data.org

REGISTRY OF RESEARCH DATA REPOSITORIES

re3data.org

REGISTRY OF RESEARCH DATA REPOSITORIES

Any questions?

Stay in touch!

Web: **www.re3data.org**

Mail: **info@re3data.org**

Twitter: **@re3data**

DOI: https://doi.org/10.5281/zenodo.7334302

https://creativecommons.org/licenses/by/4.0/