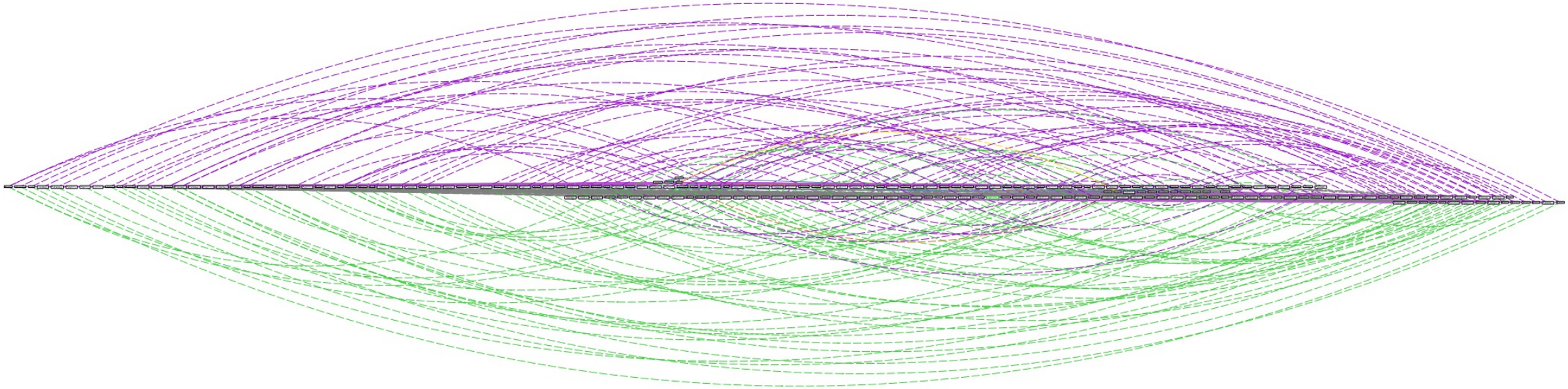
A Short Introduction to the Data Collections Explorer

NFDI4Ing SIG RDM Training & Education – 6. October 2023

Philipp Ost philipp.ost@kit.edu



CC BY-SA 4.0

The big Picture: Data Repositories

What is a data repository?
A *data repository* registers, archives, and provides access to data collections.

- They (usually) provide unique identifiers for data sets
- They help making data FAIR

There are a lot of repositories available – the challenge is finding the right one

- re3data provides a lot of information
- However, there is more…

What is the Data Collections Explorer?

The Data Collections Explorer is an information system for repositories and data collections focussed on the engineering sciences.

Primary use cases include

- Scientists searching for datasets
  - Example: An engineer is interested in material properties
    - Are there other materials that might fit the requirements?

- Scientists looking for repositories where to publish their research data
  - Example: PhD create data sets as part of their research
    - Where to publish these?
    - Are there size limits or costs involved?

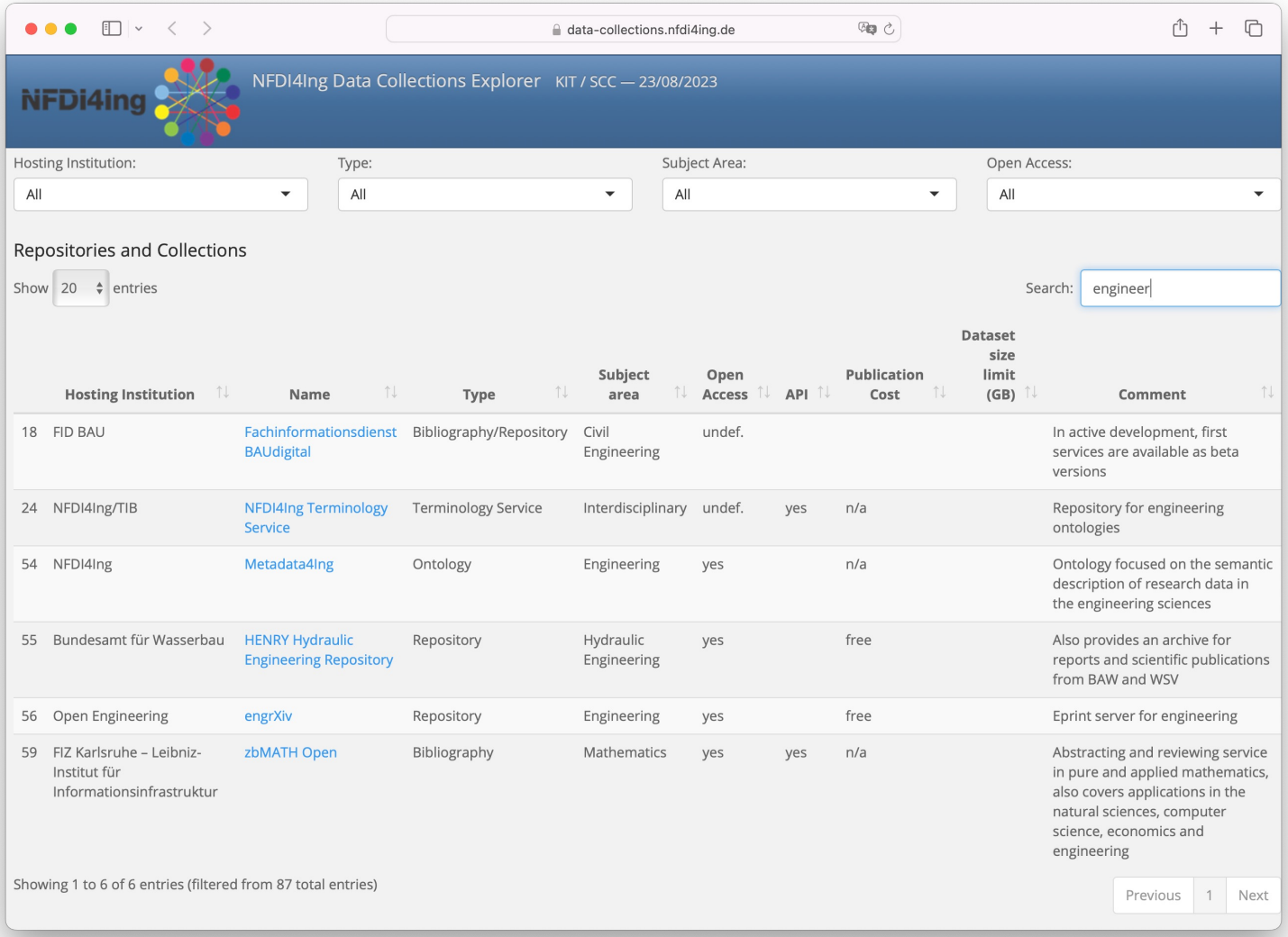
What is the Data Collections Explorer?

The Data Collections Explorer

- Allows filtering and searching for entries
- Complements re3data
  - Individual data sets (e.g., The KITTI Vision Benchmark)
  - Privately or community curated repositories and databases (e.g., The Wood Database, The Phase Field Hub)
  - Repositories not yet listed in re3data
- Receives continuous updates based on community feedback

What Does the Data Collections Explorer Offer?

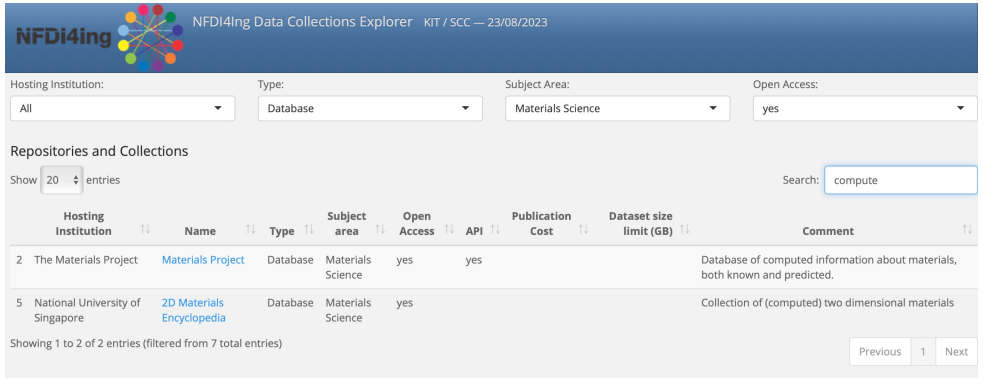
- Currently 87 repositories and data sets listed
- Access it here: data-collections.nfdi4ing.de

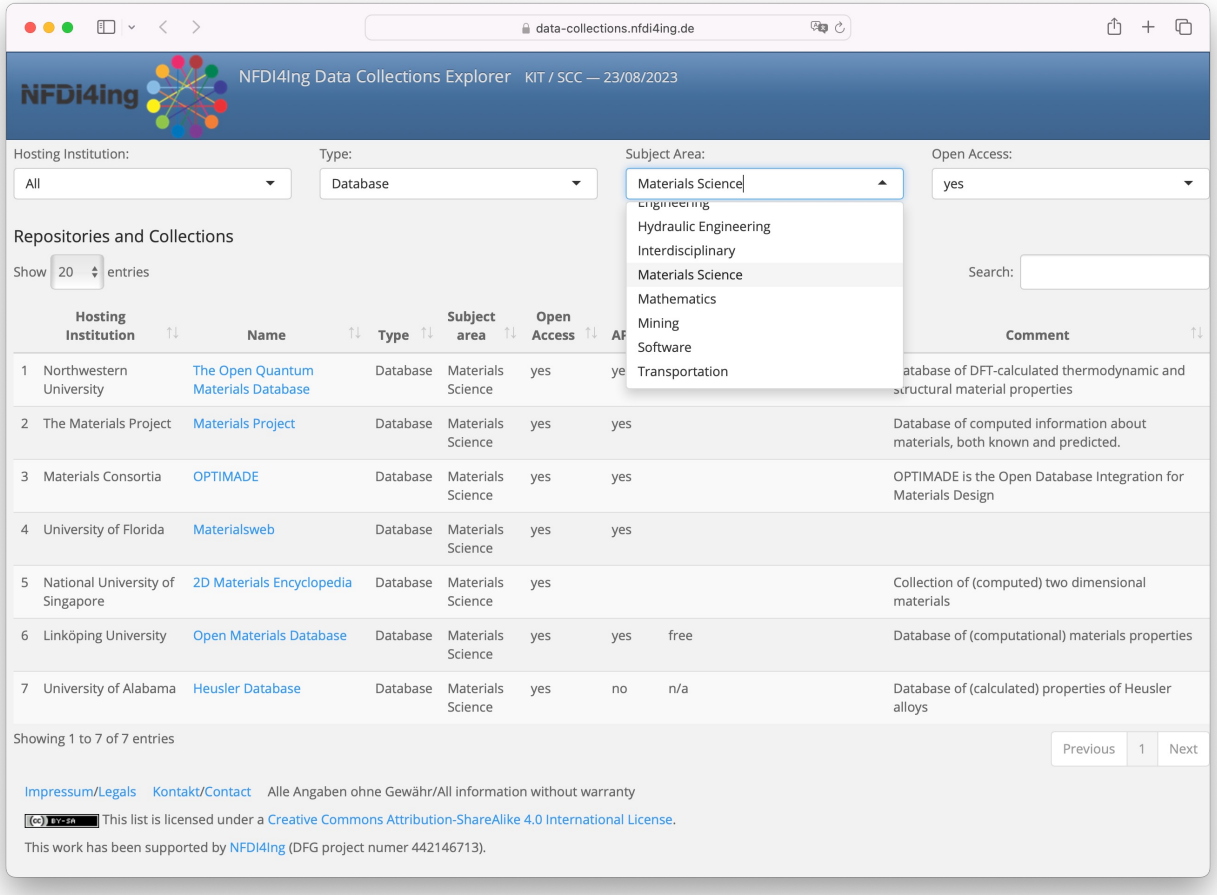


	Hosting Institution	Name	Type	Subject area	Open Access	API	Publication Cost	Dataset size limit (GB)	Comment
18	FID BAU	Fachinformationsdienst BAUdigital	Bibliography/Repository	Civil Engineering	undef.				In active development, first services are available as beta versions
24	NFDI4Ing/TIB	NFDI4Ing Terminology Service	Terminology Service	Interdisciplinary	undef.	yes	n/a		Repository for engineering ontologies
54	NFDI4Ing	Metadata4Ing	Ontology	Engineering	yes		n/a		Ontology focused on the semantic description of research data in the engineering sciences
55	Bundesamt für Wasserbau	HENRY Hydraulic Engineering Repository	Repository	Hydraulic Engineering	yes		free		Also provides an archive for reports and scientific publications from BAW and WSV
56	Open Engineering	engrXiv	Repository	Engineering	yes		free		Eprint server for engineering
59	FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur	zbMATH Open	Bibliography	Mathematics	yes	yes	n/a		Abstracting and reviewing service in pure and applied mathematics, also covers applications in the natural sciences, computer science, economics and engineering

What Does the Data Collections Explorer Offer?

- The drop down menus allow filtering for entries
  - Hosting Institution
  - Type
  - Subject Area
  - Open Access
- The full text search allows to refine results further:





Conclusions

- The Data Collections Explorer is an easy-to-use and flexible tool
  - Share and find data from the engineering domain
  - Low barrier of entry

- The Data Collections Explorer is a service for the community
  - Open for contributions – please contact us!
    - You can also create at pull request at https://github.com/kit-data-manager/Data-Collections-Explorer
  - Continuously updated based on community feedback