FAIR train the trainers
Methodology section
FAIR tools

February 20, 2019

Judit Fazekas-Paragh
Edit Görögh
Ádám Száldobágyi

Map of the presentation

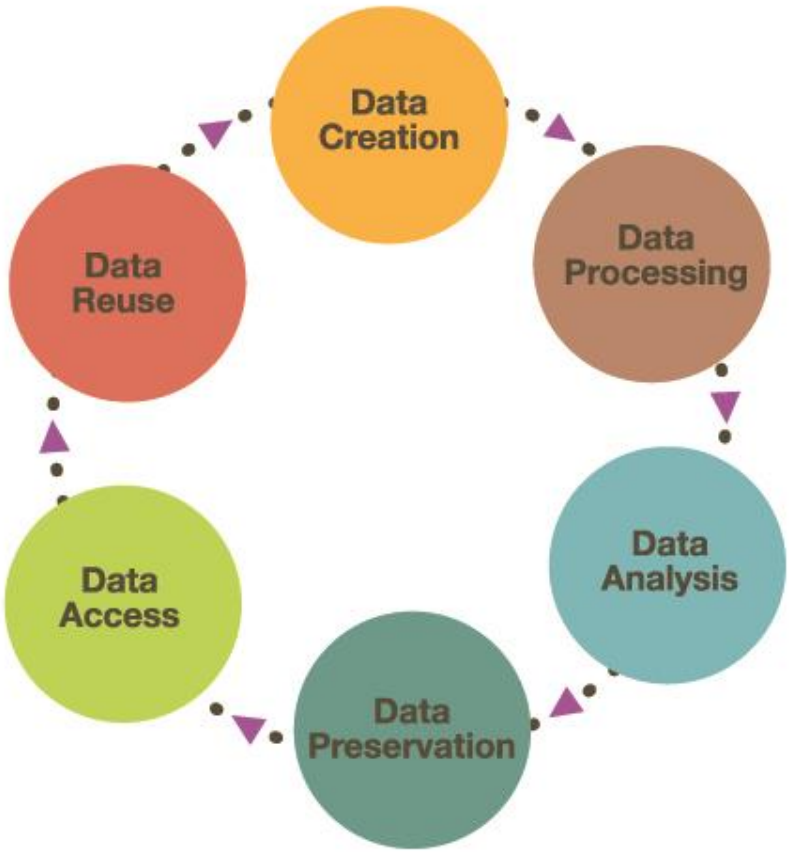
- FAIR principles
- How and where to choose the right FAIR tools
- Tools for all disciplines:
  - FAIRsharing
  - RDA Metadata Standards Directory
  - FAIR Evaluation Services
  - ARGOS
  - AMNESIA
  - B2SHARE, Zenodo
  - Checklist to evaluate FAIRness of data

FAIR principles I

FAIR principles II

What aspects to consider when looking for the right tool?

- Research data life cycle
- Disciplines
- The possibilities and expectations of the given institution
- Funder and journal expectations



Picture: https://blogs.ntu.edu.sg/lib-datamanagement/data-lifecycle/

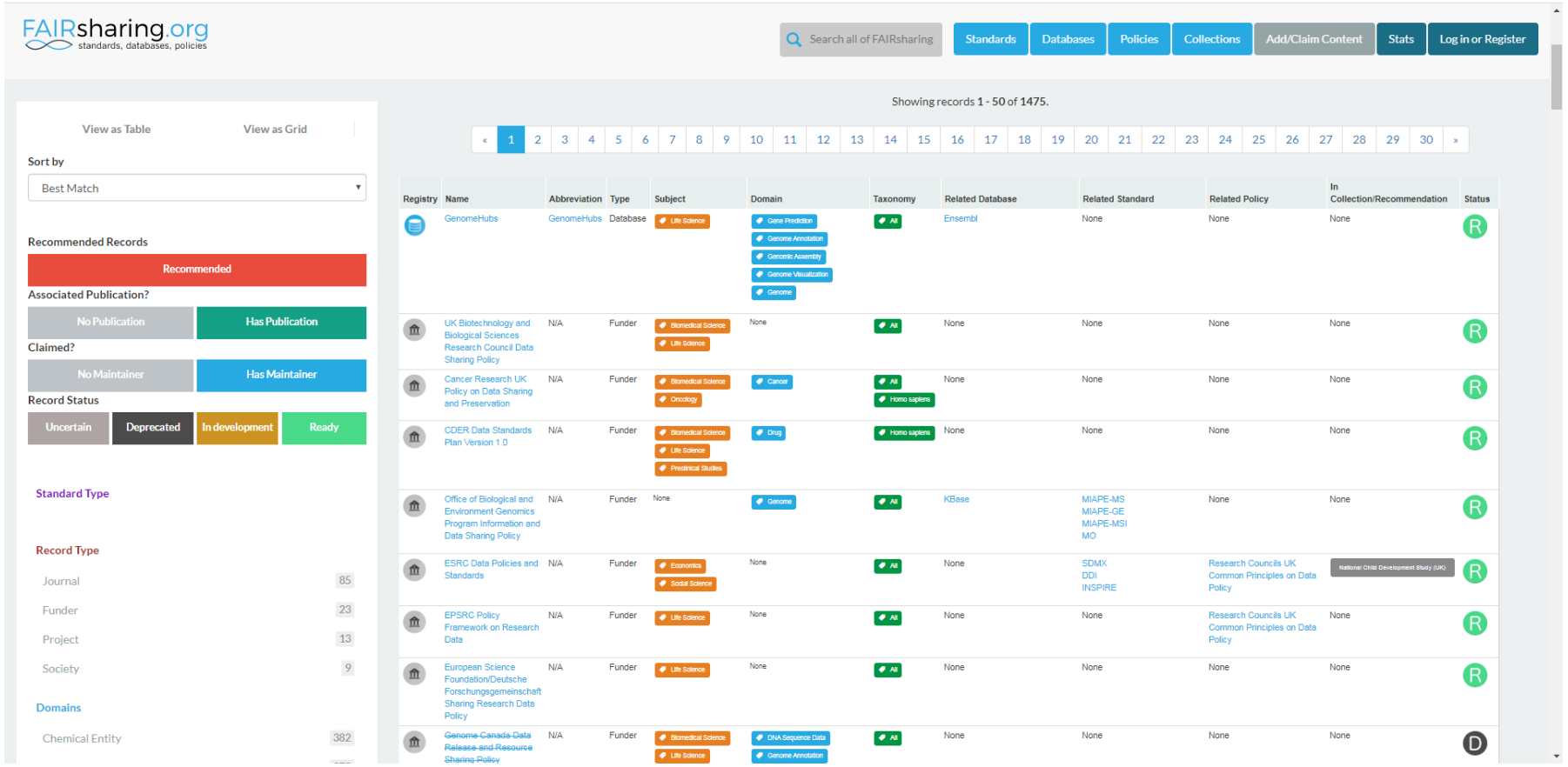
How and where to choose the right FAIR tools

Databeses:

- For everyone:
  - Open Science MOOC
  - Research Data Alliance
  - 400+ Tools and innovations in scholarly communication
- NI4OS partners:
  - NI4OS Box tools
  - NI4OS Box ORDM

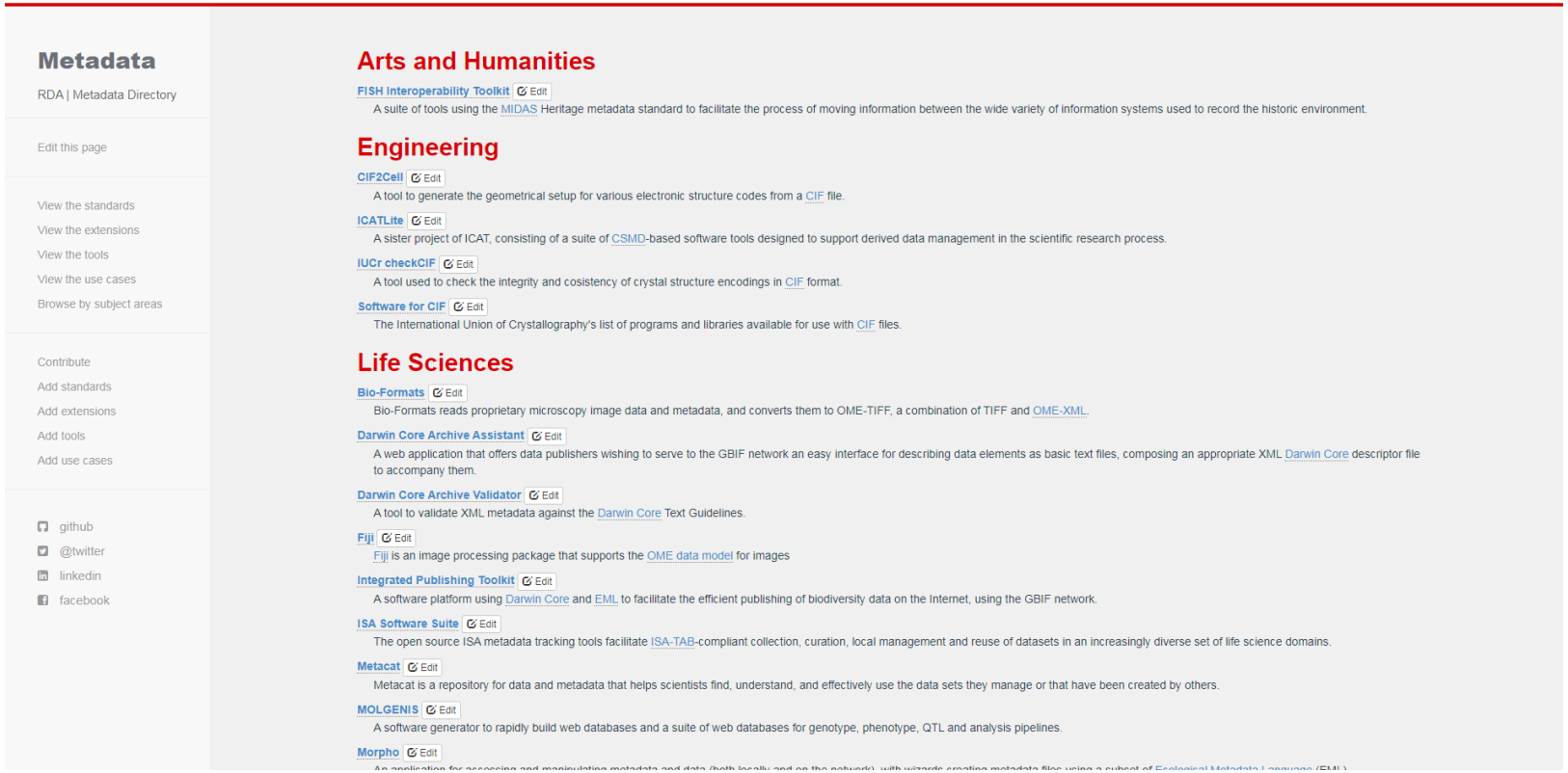
FAIRsharing

A curated, informative and educational resource on data and metadata standards, inter-related to databases and data policies.



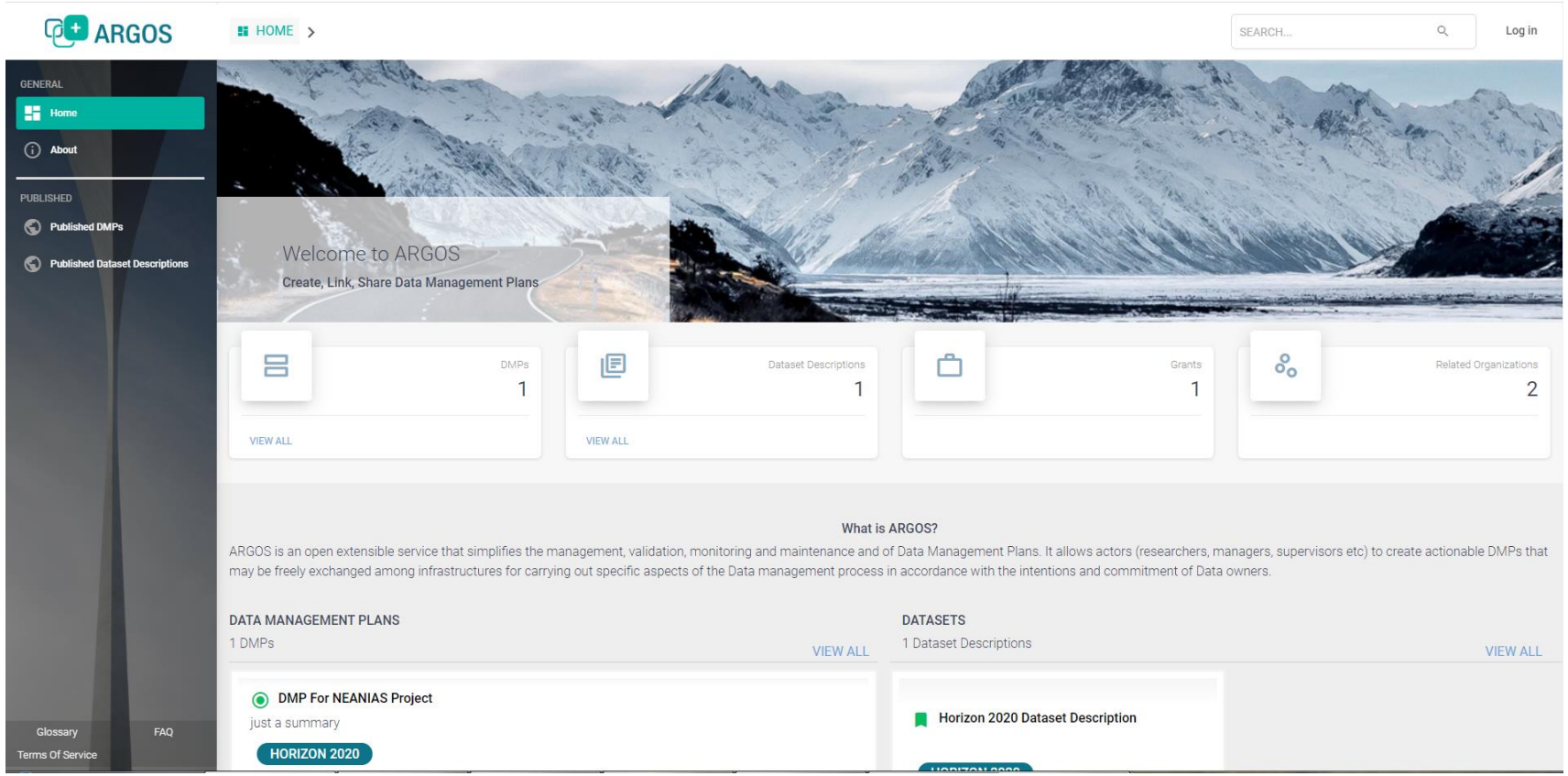
RDA Metadata Standards Directory

The overriding goal is to develop a collaborative, open directory of metadata standards applicable to scientific data can help address infrastructure challenges.



ARGOS

ARGOS is an online tool in support of automated processes to creating, managing, sharing and linking DMPs with research artifacts they correspond to.



AMNESIA

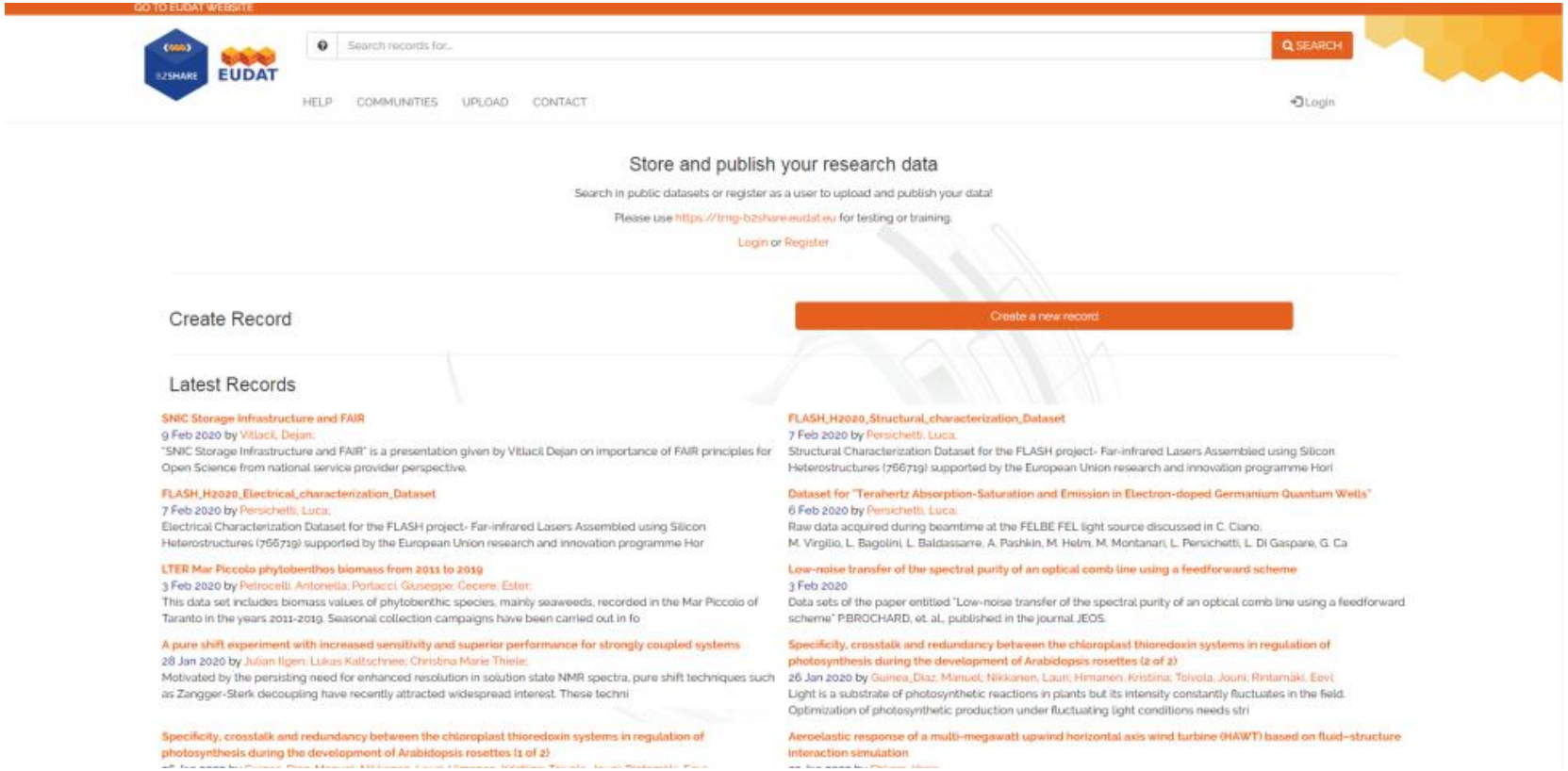
Amnesia is a data anonymization tool, that allows to remove identifying information from data. Amnesia not only removes direct identifiers like names, SSNs etc but also transforms secondary identifiers like birth date and zip code so that individuals cannot be identified in the data.

ID	Age	Zipcode	Diagnosis
1	28	13053	Heart Disease
2	29	13068	Heart Disease
3	21	13068	Viral Infection
4	23	13053	Viral Infection
5	50	14853	Cancer
6	55	14853	Heart Disease
7	47	14850	Viral Infection
8	49	14850	Viral Infection
9	31	13053	Cancer
10	37	13053	Cancer
11	36	13222	Cancer
12	35	13068	Cancer

ID	Age	Zipcode	Diagnosis
1	[20-30]	130**	Heart Disease
2	[20-30]	130**	Heart Disease
3	[20-30]	130**	Viral Infection
4	[20-30]	130**	Viral Infection
5	[40-60]	148**	Cancer
6	[40-60]	148**	Heart Disease
7	[40-60]	148**	Viral Infection
8	[40-60]	148**	Viral Infection
9	[30-40]	13***	Cancer
10	[30-40]	13***	Cancer
11	[30-40]	13***	Cancer
12	[30-40]	13***	Cancer

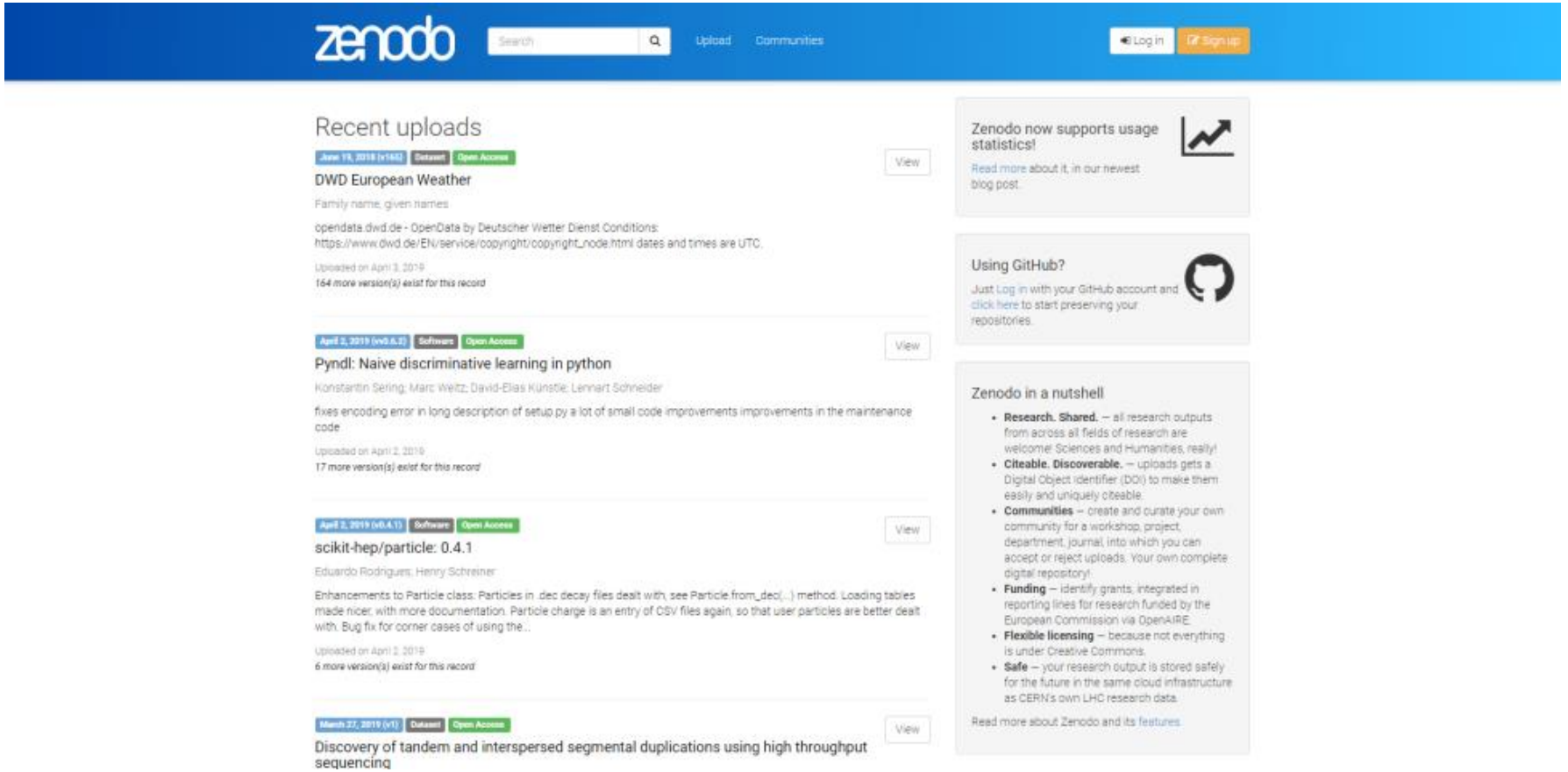
B2SHARE

B2SHARE is a user-friendly, reliable and trustworthy way for researchers, scientific communities and citizen scientists to store and publish small-scale research data from diverse contexts.



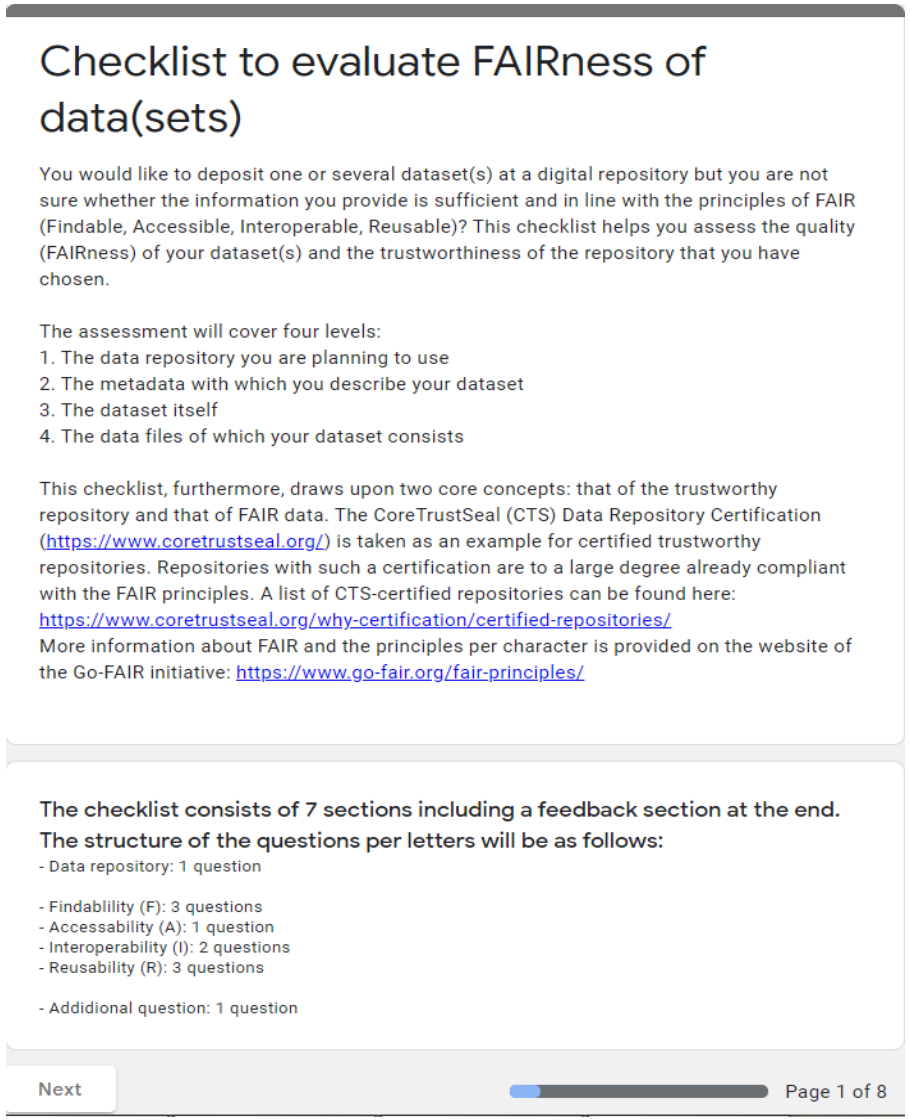
Zenodo

Zenodo is an open repository for all scholarship, enabling researchers from all disciplines to share and preserve their research outputs.



Checklist to evaluate FAIRness of data

This checklist helps you assess the quality (FAIRness) of your dataset(s) and the trustworthiness of the repository that you have chosen.



Checklist to evaluate FAIRness of data(sets)

You would like to deposit one or several dataset(s) at a digital repository but you are not sure whether the information you provide is sufficient and in line with the principles of FAIR (Findable, Accessible, Interoperable, Reusable)? This checklist helps you assess the quality (FAIRness) of your dataset(s) and the trustworthiness of the repository that you have chosen.

The assessment will cover four levels:
1. The data repository you are planning to use
2. The metadata with which you describe your dataset
3. The dataset itself
4. The data files of which your dataset consists

This checklist, furthermore, draws upon two core concepts: that of the trustworthy repository and that of FAIR data. The CoreTrustSeal (CTS) Data Repository Certification (https://www.coretrustseal.org/) is taken as an example for certified trustworthy repositories. Repositories with such a certification are to a large degree already compliant with the FAIR principles. A list of CTS-certified repositories can be found here: https://www.coretrustseal.org/why-certification/certified-repositories/
More information about FAIR and the principles per character is provided on the website of the Go-FAIR initiative: https://www.go-fair.org/fair-principles/

The checklist consists of 7 sections including a feedback section at the end.
The structure of the questions per letters will be as follows:
- Data repository: 1 question

- Findablility (F): 3 questions
- Accessability (A): 1 question
- Interoperability (I): 2 questions
- Reusability (R): 3 questions

- Addidional question: 1 question

Next

Page 1 of 8

Thank you for your attention!