



**FAIRsFAIR**  
Fostering Fair Data Practices in Europe

# Supporting data repositories in their FAIR uptake

Josefine Nordling (CSC – IT Center for Science)

18 November 2021  
FAIRsFAIR – Estonian Roadshow



FAIRsFAIR "Fostering FAIR Data Practices in Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 Grant agreement 831558

Twitter: @FAIRsFAIR\_EU

# Repository Support Series webinars



<https://fairsfair.eu/repository-support-webinars>

Some past webinars:

- PID graphs, April 21st
- FAIR-enabling services-Validating the framework, May 20th
- Assessing the FAIRness of data holdings, Sept 29<sup>th</sup>
- Certified Digital Preservation, Nov 4th
- Metadata exchange issues, Oct 21

Possible upcoming webinars:

- Assessment framework for FAIR-enabling services
- DCAT v3
- B2FIND on metadata exchange



# Scope of data repository support

*Focus on the data repository support measures of FAIRsFAIR - both the activities where the project engages directly with repositories, along with the support program being developed. The main goal is to help them in becoming FAIR enabling repositories with FAIR data holdings.*

- Supporting repositories in developing their core services and offering advice on good practice when supporting others in producing and making use of FAIR data.
- More specifically, advising how to become a trustworthy repository supporting FAIR data access, means of increasing the interoperability of datasets, and developing skills to support others in FAIR practice.
- Three types of support:
  - Designing & developing features and tools
  - Hands-on support
  - Recommendations & good practices



# Designing & developing features and tools: **Raising awareness, assessing the FAIRness of data & facilitating the findability of repositories**

- FAIR-Aware: A manual self-assessment tool to educate and raise awareness of researchers on making their data FAIR before depositing the data into the repository. More information: <https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool>
- F-UJI tool: A FAIR metrics developed for automated assessments for published data. More information: <https://fairaware.dans.knaw.nl/>
- Repository finder: <https://repositoryfinder.datacite.org/>
  - Developed by DataCite for the Enabling FAIR Data project
  - Lists repositories that currently are accepting data to support publication, including those that are certified and support the FAIR principles
  - Extended criteria made by the FAIRsFAIR to *filter out FAIR repositories*. The FsF iteration supports three filtering options: (1) Open access data; (2) Persistent Identifiers; (3) Certified repository





Repository Finder, a pilot project of the [Enabling FAIR Data Project](#), in partnership with DataCite and the Earth, space and environmental data community, helps you find an appropriate repository to deposit your research data. The tool is a registry of research data repositories.

As part of the [FAIRsFAIR project](#), which aims to supply practical support throughout the research data life cycle, the Repository Finder is a key component of the FAIRsFAIR Project.

Search [re3data](#) for a repository to upload your data

Type to search...

or

See the [repositories](#) in re3data that meet the criteria of the [Enabling FAIR Data Project](#)

See the [repositories](#) in re3data that meet the criteria of the [FAIRsFAIR Project](#)

<https://repositoryfinder.datacite.org/>

- ☐ Search [re3data](#) for a repository to upload your data
- ☐ See the repositories in re3data that meet the criteria of the [Enabling FAIR Data Project](#)
- ☒ See the repositories in re3data that meet the criteria of the [FAIRsFAIR Project](#) ☒

- ☒ The repository provides open access to its data 

Research data hosted by the repository are accessible without restrictions.

- ☒ The repository uses persistent identifiers 

Persistent identifiers such as DOIs uniquely identifier datasets, enable the linking to publications, and help with discovery.

- ☒ The repository is certified 

Certifications make sure repositories follow community standards and best practices

[Change criteria for this search](#) 

Type to search...

Search

## 111 Repositories

Sort by Name

### 4TU.Centre for Research Data

[4TU.ResearchData](#)

4TU.ResearchData, previously known as 3TU.Datacentrum, is an archive for research data. It offers the knowledge, experience and the tools to share and safely store scientific research data in a standardized, secure and well-documented manner. 4TU.Centre for Research Data provides the research community with: Advice and support on data management; A long-term archive for scientific research data; Support for current research projects; Tools for reusing research data.

life sciences

basic biological and medical research

biophysics

bioinformatics and theoretical biology

agriculture, forestry, horticulture and veterinary medicine

soil sciences

biology

agriculture, forestry, horticulture and veterinary medicine



## Hands-on support: **Facilitating interoperability between repositories**

- Working with 12 repositories on designing implementation features to increase repository interoperability: <https://fairsfair.eu/application-results-open-call-data-repositories>
- Set of FAIR data repositories features: <https://doi.org/10.5281/zenodo.3631528>
- Create the bases for continued work on *sustainable technical implementation* of the FAIR principles on a broad level
- *Improve understanding* of the current variety in *FAIR data standards*
- *Improve the semantic interoperability* of research resources by specifying FAIR metadata schemas, vocabularies, protocols, and ontologies
- *Provide solutions for interoperability requirements and machine-accessibility* for FAIR enabling repositories





## Task 2.3

# Co-development from Open Call

Developer repositories

Tester repositories



**DRYAD**



 **Data INRAE**



**KNOWLEDGE REPOSITORY**  
Central University of Punjab, Bathinda

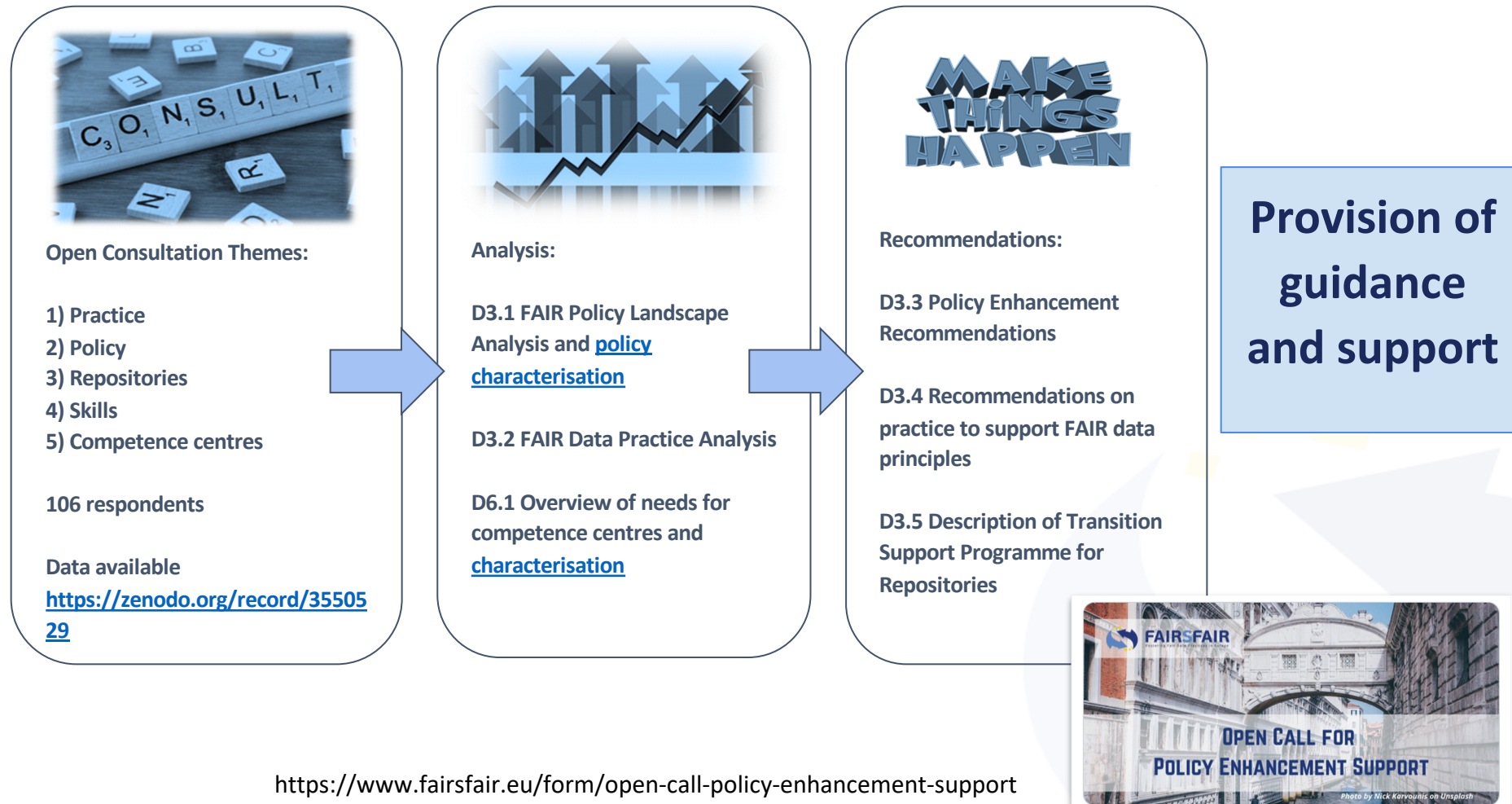


**SciencesPo**

 **DataverseNO**



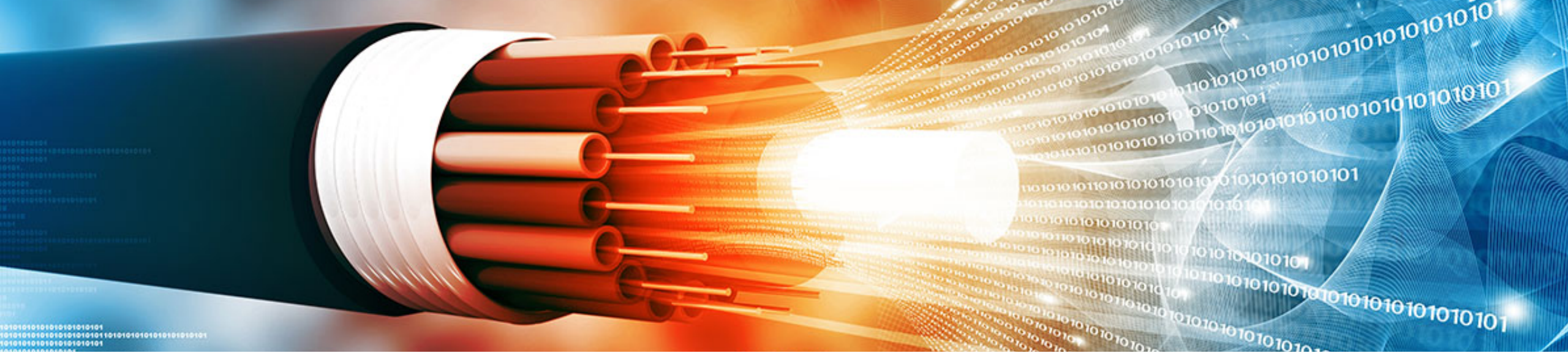
# Recommendations & good practices: **FAIR** policies & practices





## Recommendations & good practices: **FAIR Semantics**

- Recommendations for creating FAIR semantic artefacts
- Collaboration with the RDA Vocabulary Service Interest group
- Webinar recordings on PIDs, metadata and semantic interoperability:
  - For researchers: <https://www.fairsfair.eu/events/clearing-some-highest-fair-hurdles-pids-metadata-and-semantic-interoperability-researchers>
  - For data stewards and service providers: <https://www.fairsfair.eu/events/clearing-some-highest-fair-hurdles-pids-metadata-and-semantic-interoperability-data-stewards>
- Semantic artefacts as FAIR data objects:
  - Implementing machine-actionability
  - Defining metadata



# Supporting data repositories and researchers in their FAIR uptake

Josefine Nordling, CSC – IT Center for Science  
(EOSC-Nordic WP4, FAIR incentives task lead)



# FAIR Data



Survey

FAIR Maturity  
evaluation

GAP  
analysis

FAIR uptake

FAIR-uptake

Nordic region

Repository  
providers

Data standard

Certification

Implementation

Selected  
communities

Map  
incentives

Impact  
analysis

Stakeholders  
engagement

Policies

Policy-  
makers

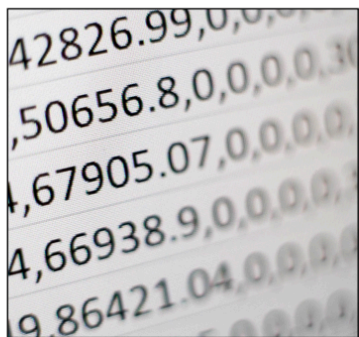
Communities

Funders

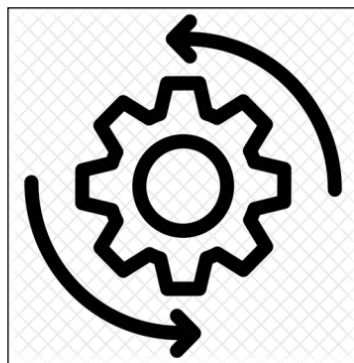
Activities

Stakeholders

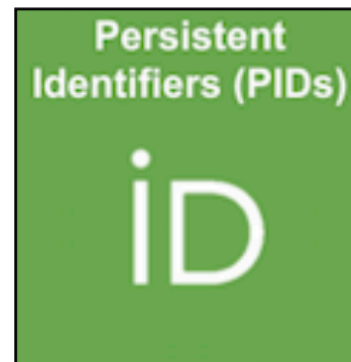
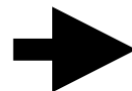
# Measuring FAIR uptake



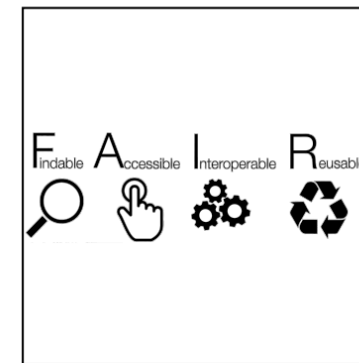
Collected ~100  
research data  
repositories  
in the Nordic+Baltic



FAIR assessments of  
all datasets using fully  
machine-actionable  
metadata



24% of repos do  
NOT provide  
PIDs or GUIDs  
(identifiers)



Average  
FAIR score 24%,  
higher for  
mature / certified  
repositories

Credit: Andreas Jaunsen

<https://docs.google.com/spreadsheets/d/1MBTMXb5SleaBKiyEzImgJtQtxNS1zZTgTvTk9A2vgdc>

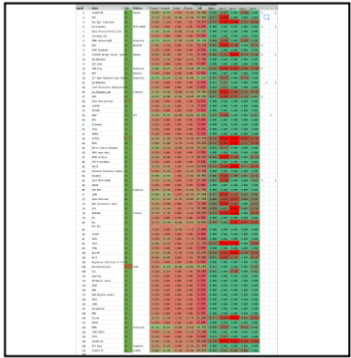
Expose the results to repos/communities, so that they can use the feedback from the tool as a guideline to improve their FAIRness

repoID	Name	Data	Platform	F-score	A-score	I-score	R-score	FAIR	Sigma	sigma (F)	sigma (A)	sigma (I)	sigma (R)			
2	CLARIN-DK	48	Dspace	50.00%	66.67%	9.72%	22.22%	34.30%	0.036	0.072	0.000	0.124	0.000	X		X
3	DDA	20		57.14%	33.33%	50.00%	33.33%	43.48%	0.073	0.293	0.000	0.000	0.000			
4	Det Kgl. bibliotek	20		30.00%	13.33%	5.00%	7.22%	14.57%	0.118	0.307	0.294	0.131	0.154			
6	Kielipankki	10	META-SHARE	14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000	X		X
7	Data Service Portal Aila	10		71.43%	66.67%	25.00%	33.33%	47.83%	0.000	0.000	0.000	0.000	0.000			
8	Fairdata IDA	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
9	NMBU dataverseNO	16	Dataverse	85.72%	66.67%	37.50%	43.05%	58.15%	0.050	0.148	0.000	0.129	0.038			
10	NSD	20	NESSTAR	37.86%	13.33%	0.00%	8.89%	16.74%	0.087	0.283	0.168	0.000	0.112	X		
11	HUNT Databank	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
13	CLARINO Bergen Center repo	20	Dspace	30.00%	20.00%	7.50%	6.67%	15.65%	0.096	0.153	0.313	0.118	0.104	X		X
16	Språkbanken	10		14.29%	33.33%	0.00%	22.22%	17.39%	0.000	0.000	0.000	0.000	0.000			
17	ESS Data	9		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
18	TROLLing	22	Dataverse	75.00%	56.67%	27.50%	35.55%	48.91%	0.103	0.240	0.244	0.160	0.160	X		
19	EED	10	Nesstar	14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
20	UIT Open Research Data Dat	20	Dataverse	85.72%	66.67%	32.50%	44.44%	57.83%	0.047	0.147	0.000	0.118	0.000			
24	Språkbanken	13		17.59%	0.00%	0.00%	0.00%	5.35%	0.016	0.063	0.000	0.000	0.000		X	X
25	Lund University Humanities	20		21.43%	0.00%	0.00%	0.00%	6.52%	0.018	0.073	0.000	0.000	0.000			
26	su.figshare.com	20	Figshare	51.43%	45.00%	10.00%	29.44%	34.78%	0.098	0.287	0.196	0.126	0.136			
27	SND	20		64.29%	50.00%	42.50%	38.89%	48.69%	0.077	0.220	0.171	0.118	0.057	X		
28	ICES data portals	8		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
29	JASPAR	10		14.29%	33.33%	0.00%	11.11%	13.04%	0.000	0.000	0.000	0.000	0.000			
30	STRING	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
32	GBIF	22	IPT	59.74%	66.67%	25.00%	38.89%	46.44%	0.047	0.180	0.000	0.000	0.057		X	
39	HPA	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
41	Fishbase	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
45	ISIG	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
47	GERDA	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
49	ACTRIS	8		46.43%	33.34%	12.50%	16.67%	27.17%	0.139	0.312	0.356	0.231	0.178			
52	NPDC	20		37.86%	35.00%	13.75%	17.22%	25.22%	0.127	0.259	0.366	0.151	0.178			
54	Bolin Centre Database	12		61.90%	33.33%	25.00%	22.22%	36.23%	0.028	0.111	0.000	0.000	0.000			
55	SMHI open data	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
57	NIRD Archive	20		44.29%	20.00%	0.00%	11.11%	20.43%	0.093	0.311	0.168	0.000	0.114			
60	GTN-P Database	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
62	UNITE	20		35.72%	33.34%	6.25%	16.67%	22.83%	0.114	0.220	0.342	0.111	0.171			
63	Estonian Biocentre Public	10		14.29%	33.33%	0.00%	22.22%	17.39%	0.000	0.000	0.000	0.000	0.000			
64	DataDOI	18		61.90%	62.97%	2.78%	27.78%	38.40%	0.068	0.208	0.108	0.081	0.116			
65	CELR META-SHARE	20		35.72%	33.33%	0.00%	5.56%	17.39%	0.057	0.220	0.000	0.000	0.057	X		X
66	AHEAD	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
68	USN RDA	10	Figshare	28.57%	33.33%	0.00%	22.22%	21.74%	0.000	0.000	0.000	0.000	0.000			
71	LOAR	19		61.65%	63.16%	2.63%	26.31%	37.76%	0.077	0.243	0.153	0.079	0.085			
72	AIDA Data Hub	20		65.71%	43.33%	23.75%	30.00%	41.52%	0.080	0.188	0.219	0.099	0.103			
73	QoG Institute's data	10		31.43%	10.00%	0.00%	6.67%	13.48%	0.084	0.276	0.161	0.000	0.107			
76	JYX	20		35.72%	33.34%	1.25%	11.11%	19.78%	0.093	0.073	0.342	0.056	0.114			
78	B2SHARE	12	Invenio	38.10%	27.78%	0.00%	13.89%	20.65%	0.118	0.275	0.343	0.000	0.172			
79	DH	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
80	NLL	10		25.71%	53.34%	40.00%	35.55%	35.65%	0.101	0.060	0.281	0.211	0.187			
	RTU RIS	10		14.29%	0.00%	25.00%	11.11%	13.04%	0.000	0.000	0.000	0.000	0.000			
85	FinBIF	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
87	SARV	10		14.29%	33.33%	0.00%	11.11%	13.04%	0.000	0.000	0.000	0.000	0.000			
92	SSRI	15		42.86%	26.67%	0.00%	13.33%	21.74%	0.131	0.382	0.338	0.000	0.169			
94	IINH	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
100	QsarDB	20		35.72%	30.00%	11.25%	15.55%	22.83%	0.110	0.220	0.323	0.128	0.163			
104	Bird	20		30.00%	40.00%	0.00%	6.67%	16.96%	0.074	0.222	0.137	0.000	0.137			
106	Migration Institute of Fin	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
108	Musiikkiarkisto	4	CKAN	28.57%	66.67%	50.00%	50.00%	45.65%	0.016	0.000	0.000	0.000	0.064			
109	SLS	14		14.29%	23.81%	0.00%	7.94%	10.56%	0.041	0.000	0.156	0.000	0.052			
113	SweFreq	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
114	Metabolic Atlas	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
115	SEAD	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
116	NOW	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
117	SNM Digital Assets	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
120	GEUS	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
123	LARM	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
127	Garamantas	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
129	MMB	10		14.29%	0.00%	0.00%	0.00%	4.35%	0.000	0.000	0.000	0.000	0.000			
130	PlutoF	20		35.72%	31.67%	3.75%	16.11%	21.96%	0.111	0.220	0.333	0.092	0.167			





# Supporting FAIR uptake



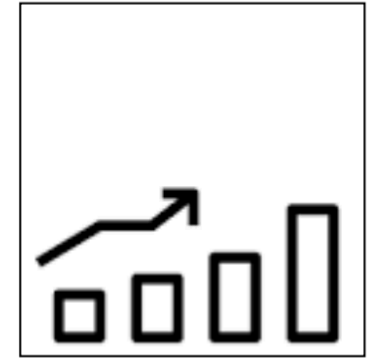
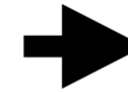
Prepared results  
and shared with  
the community



Offered community  
recommendations  
based on  
results

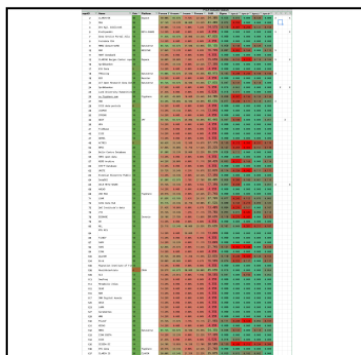


Community  
workshops  
held with >300  
attendees in total



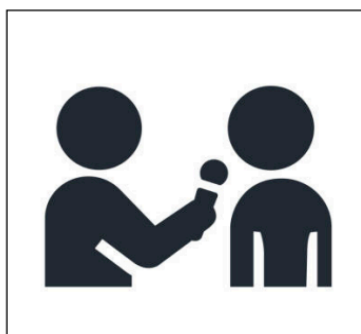
Directly gauge  
uptake by repositories  
by measuring their  
FAIR score

**Certification + FAIR**



Measured  
FAIR maturity

**FAIR  
incentives  
study**

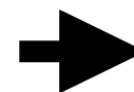


Conducted  
interviews on  
policies FAIR  
incentives

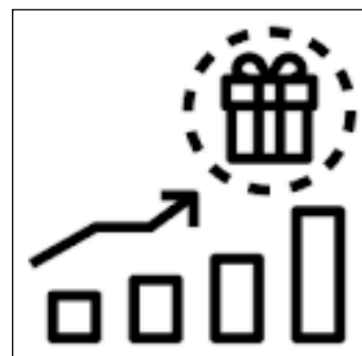
# Additional activities



Certification status of  
repositories



**Test proposed  
FAIR certification  
scheme (by  
FAIRsFAIR)**



Result  
dissemination  
and stakeholder  
liaising



**Aim to bring  
stakeholders inline  
with OS**



# FAIRSFair

Fostering Fair Data Practices in Europe



# EOSC



## Thank you!