



FAIRSFair

Fostering Fair Data Practices in Europe

Tools to support the journey towards FAIR data

Ilona von Stein, Joy Davidson, Mustapha Mokrane

STM Research Data Webinar

22 March 2021

Outline

The FAIRsFAIR project in a nutshell

**FAIRsFAIR policy enhancement
recommendations**

FAIR data assessment: implementations and tools



FAIRsFAIR in a nutshell

FAIRsFAIR - *Fostering Fair Data Practices in Europe* - aims to supply **practical solutions** for the use of the FAIR data principles throughout the research data life cycle with an emphasis on **fostering FAIR data culture** and **the uptake of good practices** in making data FAIR.

FAIRsFAIR key facts

- Starting date: March 1 2019
- Duration: 36 months
- Budget: 10 million euro
- 22 partners from 8 European Member States
- 6 core partners: DANS (project coordinator), CSC, DCC, TrustIT, STFC, EUA
- Support to the European Open Science Cloud (EOSC) governance

FAIRsFAIR partners

Data Archiving and Networked Services



UNIVERSITEIT VAN AMSTERDAM



HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI



Universidade do Minho



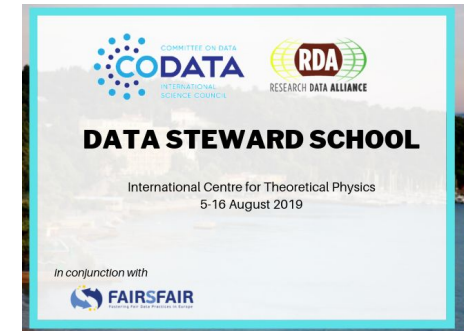
UK Research
and Innovation



GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN



Results in many shapes and forms



Recommendations to increase FAIR practices in academia

ADDRESSING FAIR IN HIGHER EDUCATION

1. Include more specific research data-related skills and competencies throughout the universities' educational portfolio
2. Establish minimum level of awareness of RDM and FAIR
3. Practical guidance on the application of the FAIR principles in different domains and disciplines, and related skills and competences
4. Define training programmes on research data skills and competences for researchers and professional staff,

RESEARCH DATA POLICIES & SUPPORT SERVICES

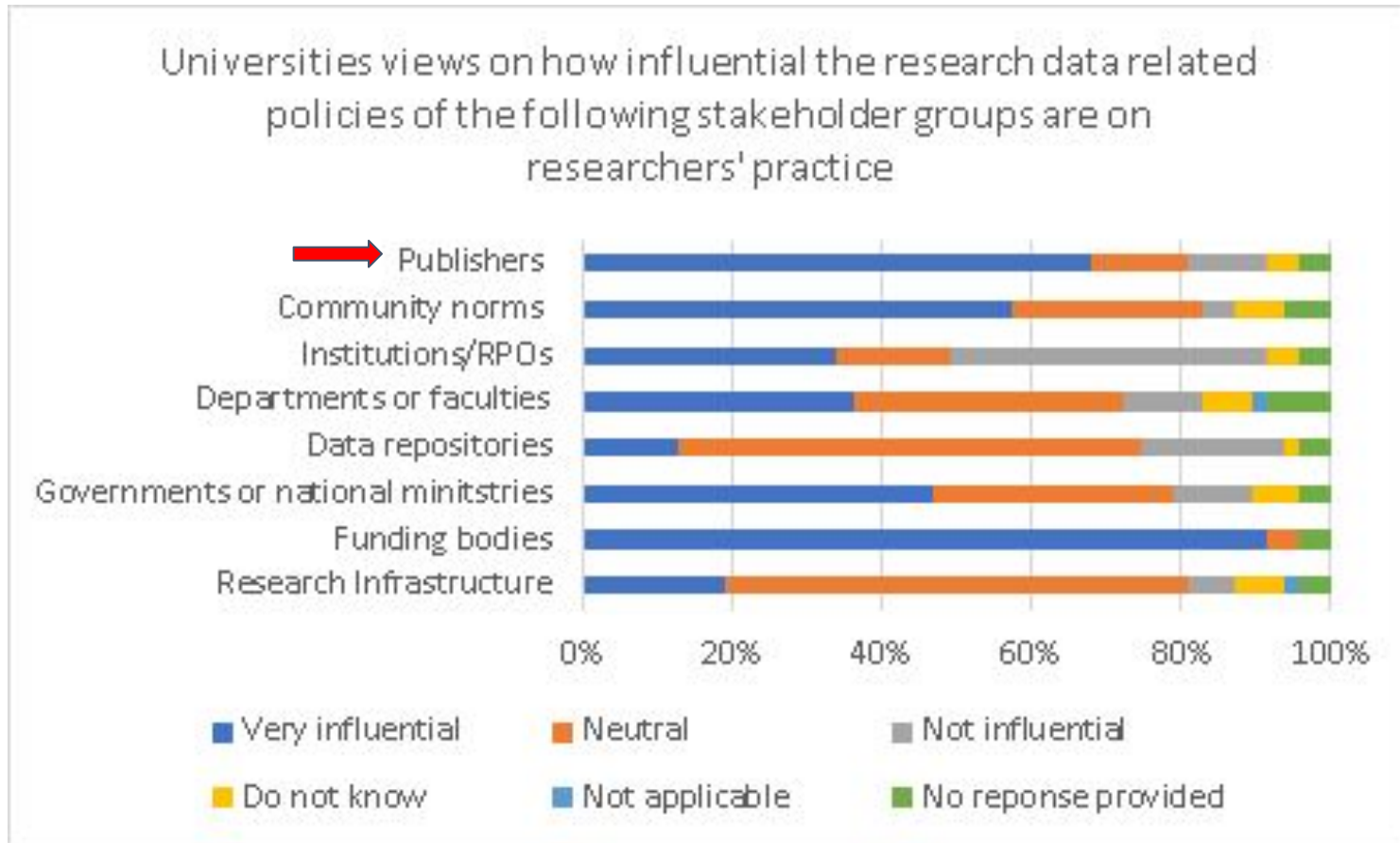
1. Define institutional research data policies together with all relevant units
2. Offer research data support services to researchers and staff
3. Include FAIR data in research data policies
4. Align FAIR policies with funders' requirements
5. Enable time and/or financial resources to train staff

UNIVERSITIES AND THE EUROPEAN OPEN SCIENCE CLOUD

1. Increase institutional capacity and awareness of EOSC
2. Develop use-cases for EOSC to facilitate collaborative research
3. Ensure that EOSC is a publicly-owned infrastructure that has open principles at its core



Publishers' influence on researchers' practice



From: FAIRsFAIR D3.1 Policy landscape analysis report, <https://doi.org/10.5281/zenodo.3558173>

Publisher involvement in the project



Eefke Smit

Affiliation: International Association of STM Publishers
Country: Netherlands

Eefke Smit is the Director of Standards and Technology of the International **STM Association**, an international trade body for publishing organisations. She has more than 20 years of experience in the academic and professional publishing sector, having become involved in digital publishing as an early pioneer. In the first decade of this century, she held personal responsibility for several large internet platforms and search solutions for researchers that continue to define scholarly communications today. Since 2010, then on behalf of STM, she became actively involved in a whole range of research data initiatives. During her career, especially as a science publisher, Eefke has worked closely with several different research communities: Computer Science and Engineering, Mathematics, Physics and Astronomy, Life Sciences.

► **Listen to the podcast "The role of Publishers in the implementation of FAIR policies and practice for RDM"**



[in LinkedIn](#) |

**FAIR
Champions**

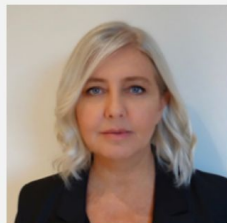


Shelley Stall

Affiliation: American Geophysical Union
Country: United States

Shelley Stall is the Senior Director for the American Geophysical Union's Data Leadership Program. She works with AGU's members, their organizations, and the broader research community to improve data and digital object practices with the ultimate goal of elevating how research data is managed and valued. Better data management results in better science.

[in LinkedIn](#) | [Twitter](#)



Susanna-Assunta Sansone

Affiliation: University of Oxford
Country: United Kingdom

Prof. Sansone has worked since 2001 in the areas of data interoperability and reproducibility, research integrity, and the evolution of scholarly publishing, and she collaborates with researchers, service providers, journal publishers, library science experts, funders and learned societies in academic, commercial and government settings alike. One of the authors of the FAIR Principles and an active contributor to several FAIR-enabling maturity indicators and tools, Prof. Sansone has been a member of the Dryad and RDA Technical Advisory Boards and currently serves as co-chair of numerous working groups.

HOW CAN SUSANNA-ASSUNTA SUPPORT YOU?

Whether you are a researcher, standard/database developer, curator, funder, journal editor, librarian or data manager, Susanna can help you make the most of FAIRsharing, a widely used informative and educational resource on data and metadata standards, repositories and data policies. An author of the FAIR Principles, and principal investigator of over 50 research and infrastructure projects, she can also assist with data FAIRification experiences and expertise.

SEE SUSANNA-ASSUNTA CONTRIBUTION TO FAIRFAIR

- Susanna co-authored article based on the Open Science Fair 2019 (Porto, Portugal)
- Susanna's interview at Open Science Fair 2019 (Porto, Portugal)

[in LinkedIn](#) | [Twitter](#)





FAIRsFAIR

Fostering Fair Data Practices in Europe

Policy enhancement recommendations

FAIRsFAIR landscape activities



Semantics and
Interoperability

Policies and
Practices

HEIs and FAIR
Data

- **D3.1 FAIR Policy landscape analysis**
<https://zenodo.org/record/3558173>
- **D3.3 Policy Enhancement Recommendations**
<https://zenodo.org/record/3686901>
- **D7.1 FAIR in European Higher Education**
<https://zenodo.org/record/3629683>

FAIRsFAIR recommendations

Collectively agree on a **common set of FAIR policy elements** to describe policies.

FAIRsFAIR Policy Characterisation Data for D3.1

<https://zenodo.org/record/3550544>



SRIA: https://www.eoscsecretariat.eu/sites/default/files/open_consultation_booklet_sria-eosc_20-july-2020.pdf

FAIRsFAIR recommendations

**Accessibility of policies
via web pages**
PIDs should be assigned
to clearly versioned
policies made available
via registries.

Model Policy on Open Access to Publications and Data for Research Performing Organisations (RPOs)

1. Preamble

1. The **[Name of RPO]** commits to the advancement of science and the wide dissemination of knowledge to the benefit of society by adopting practices on open, reproducible and responsible research.
2. The **[Name of the RPO]** recognizes “openness” as one of its guiding principles and commits to promoting it by –among others- encouraging and supporting research processes and tools that encourage collaboration, enabling new working models and new social relationships, stimulating the dissemination of knowledge and the accessibility and re-usability of research outputs, encouraging open access to publications and data and building the necessary infrastructure to support open science.

2. Jurisdiction and Effect of Policy

The Policy applies to all researchers active at **[Name of RPO]**. In cases where research is funded by a third party, any agreement with that party concerning access rights, deposit and storage takes precedence over this Policy.

The Policy has been approved by **[dean/ commission/...]** and takes effect from **[dd/mm/yy]**

3. Rights, Responsibilities, and Duties

3.1 The **[Name of RPO]** is responsible for:

1. Supporting and empowering the transition to Open Access/ Open Science through education, training and awareness-raising actions targeting researchers and other employees, along with the provision of the necessary infrastructure and funding to support this transition. Acquisition of Open Science skills should form an integral part of professional training and career development offered to researchers.
2. Establishing (if one does not yet exist) an **Institutional Open Access Repository** **[Name of the Repository]** according to international standards, containing digital content and providing advanced tools for search, navigation and Open Access to its content.

<https://www.openaire.eu/model-policy-on-open-science-for-research-performing-organisations>

FAIRsFAIR recommendation

FAIR can be referred to both explicitly and implicitly.

IMPLICIT	EXPLICIT
Implicit refers to something that is implied but not said directly	Explicit refers to something is that stated directly
When someone expresses something implicitly, there is room for confusion	When someone expresses something explicitly, there tends to be no confusion

Image from <https://pediaa.com/what-is-the-difference-between-implicit-and-explicit/>



FAIRsFAIR

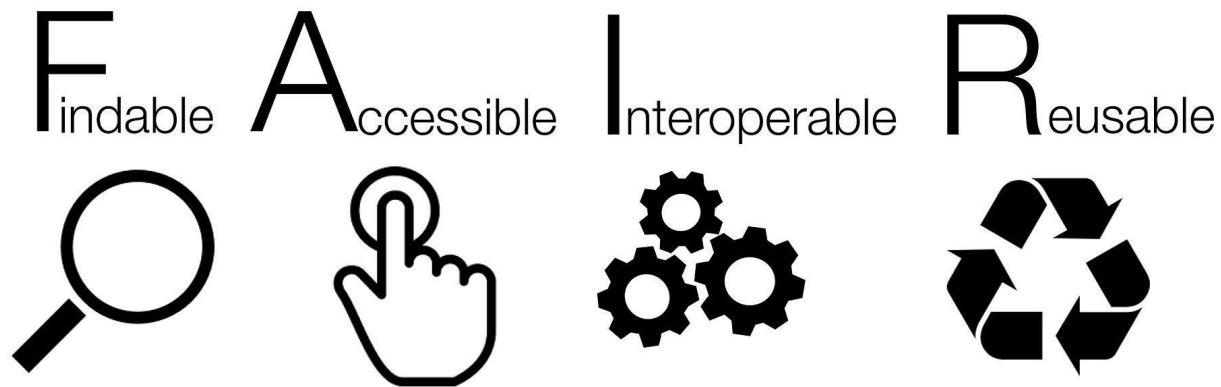
Fostering Fair Data Practices in Europe

FAIR data assessment: implementations and tools



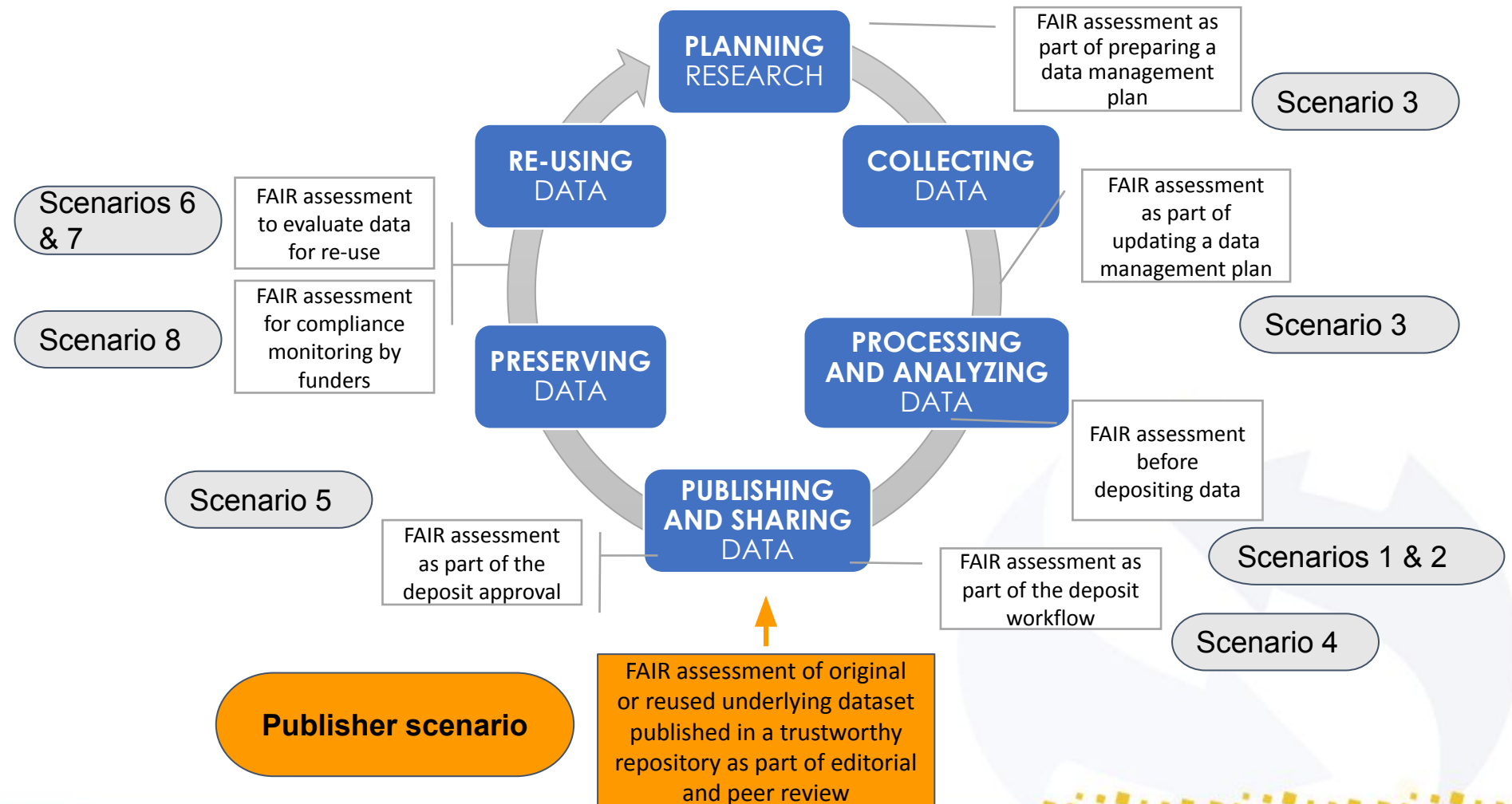
FAIR data

Evaluation and assessment



Publisher endorsement of the [original Force11 FAIR data principles](#) as early as 2016

FAIR data assessment scenarios



FAIR | Aware

- Online tool to raise awareness and educate on data FAIRness
- For researchers, authors and data stewards
- 10 simple questions with practical tips to improve data FAIRness before data deposit
- Open source. Now available for testing. Continuously improved.
Final release Autumn 2021.

<https://fairaware.dans.knaw.nl/>



The screenshot shows the 'FAIR questions' tool interface. On the left, the word 'FAIR' is written vertically in large blue letters. The main content area is titled 'FAIR questions' and contains a table of 10 questions, grouped into four categories: FINDABLE, ACCESSIBLE, INTERPRETABLE, and REUSABLE. Each question has 'Yes' and 'No' response options.

FINDABLE	
Are you aware that a dataset should be assigned a globally-unique and persistent identifier when deposited with a data repository? (1)	Yes No
Are you aware that when you deposit a dataset with a repository, you will need to provide some details (known as discovery metadata) in order to make the data findable, understandable and reusable to others? (2)	Yes No
Are you aware that the repository providing access to your dataset should make the metadata describing your datasets available in a format readable by machines as well as humans? (3)	Yes No
ACCESSIBLE	
Are you aware that access to your dataset may need to be controlled and that metadata should include license information under which the data can be reused? (4)	Yes No
Are you aware that metadata should remain available over time, even if the data is no longer accessible? (5)	Yes No
INTERPRETABLE	
Are you aware that the metadata describing your datasets should use controlled vocabularies? (6)	Yes No
REUSABLE	
Are you aware that provenance information about the collection and/or generation of data should be included in the metadata? (7)	Yes No
Are you aware that metadata describing your data should follow the specifications of a community-endorsed standard? (8)	Yes No
Are you aware that data should be deposited preferably in a file format that is open - to support reuse - and supported by the repository for long-term preservation? (9)	Yes No
Are you aware that maintaining your dataset FAIR over time requires professional data curation and preservation? (10)	Yes No



F-UJI

Automated FAIR Data Assessment Tool



[About](#) [FAIR Support](#) [FAIR Landscape](#) [Tools & Software](#) [Events](#) [Project Outputs](#) [Outreach](#)

While FAIR principles may apply to any digital objects, we are concerned with the subset of digital objects: research data that are collected, measured, or created for purposes of scientific analysis.

- ✓ **FsF-F1-01D** - Data is assigned a globally unique identifier
- ✓ **FsF-F1-02D** - Data is assigned a persistent identifier
- ✓ **FsF-F2-01M** - Metadata includes descriptive core elements (creator, title, data identifier, publisher, publication date, summary and keywords) to support data findability
- ✓ **FsF-F3-01M** - Metadata includes the identifier of the data it describes
- ✓ **FsF-F4-01M** - Metadata is offered in such a way that it can be retrieved by machines
- ✓ **FsF-A1-01M** - Metadata contains access level and access conditions of the data
- ✓ **FsF-A2-01M** - Metadata remains available, even if the data is no longer available
- ✓ **FsF-I1-01M** - Metadata is represented using a formal knowledge representation language
- ✓ **FsF-I1-02M** - Metadata uses semantic resources
- ✓ **FsF-I3-01M** - Metadata includes links between the data and its related entities
- ✓ **FsF-R1-01MD** - Metadata specifies the content of the data
- ✓ **FsF-R1.1-01M** - Metadata includes license information under which data can be reused
- ✓ **FsF-R1.2-01M** - Metadata includes provenance information about data creation or generation
- ✓ **FsF-R1.3-01M** - Metadata follows a standard recommended by the target research community of the data
- ✓ **FsF-R1.3-02D** - Data is available in a file format recommended by the target research community

Please [login & comment below](#) citing in the subject line the Metric Identifier No. you are referring to - e.g. "FsF-R1.3-01M"

Building on prior global indicators & initiatives

Refine and revise metrics through iterations (v0.4)

<https://fairsfair.eu/fairsfair-data-object-assessment-metrics-request-comments>



FAIR

Data Object Assessment Metrics

FAIRsFAIR Data Object Assessment Metrics:

<https://doi.org/10.5281/zenodo.4081213>

Practical tests
implemented against
metrics

Programmatic
assessment of
published
datasets



<https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool>


Repositories

Existing web standards and
external registries and
resources

Machine actionable assessment tool

F-UJI 1.0.0 OAS3
fujii@pangaea.de

A Service for Evaluating Research Data Objects Based on FAIRFAIR Metrics.
This work was supported by the FAIRFAIR project (H2020-INFRAEOSC-2019-2020 Grant Agreement 831558).
Contact the developer
MIT License
Find out more about F-UJI

Servers
./fujii/api/v1 Authorize 

FAIR object FAIRness assessment of a data object ▼

POST /evaluate lock icon

FAIR metric FAIRFAIR assessment metrics ▼

GET /metrics Return all metrics and their definitions lock icon

Response body


```

{
  "metric_identifier": "FsF-F1-02D",
  "metric_name": "Persistent identifier",
  "output": {
    "pid": "https://doi.org/10.1594/PANGAEA.902845",
    "pid_scheme": "doi",
    "resolvable_status": true,
    "resolved_url": "https://doi.pangaea.de/10.1594/PANGAEA.902845"
  }
}

```

<https://github.com/pangaea-data-publisher/fuji>

Graphical user interface / www.f-uji.net



Automated FAIR Data Assessment Tool

Disclaimer: The test results shown here are based on preliminary data and code which still is under development. F-UJI is rapidly evolving and not yet available in a productive environment.

Research Data Object (URL/PID):* OAI-PMH:

(Optional) Enter the URL of an OAI-PMH endpoint run test


☐ Enable caching? ⓘ ☒ Use DataCite? ⓘ


Assessment Results:


Evaluated Resource:


Data Discovery and Reuse Practices in Research	
Resource PID/URL:	https://doi.org/10.17026/dans-xsw-kkeq
Metric Version:	metrics_v0.4
Metric Specification:	https://doi.org/10.5281/zenodo.4081213
Software version:	v1.0.3


Summary:



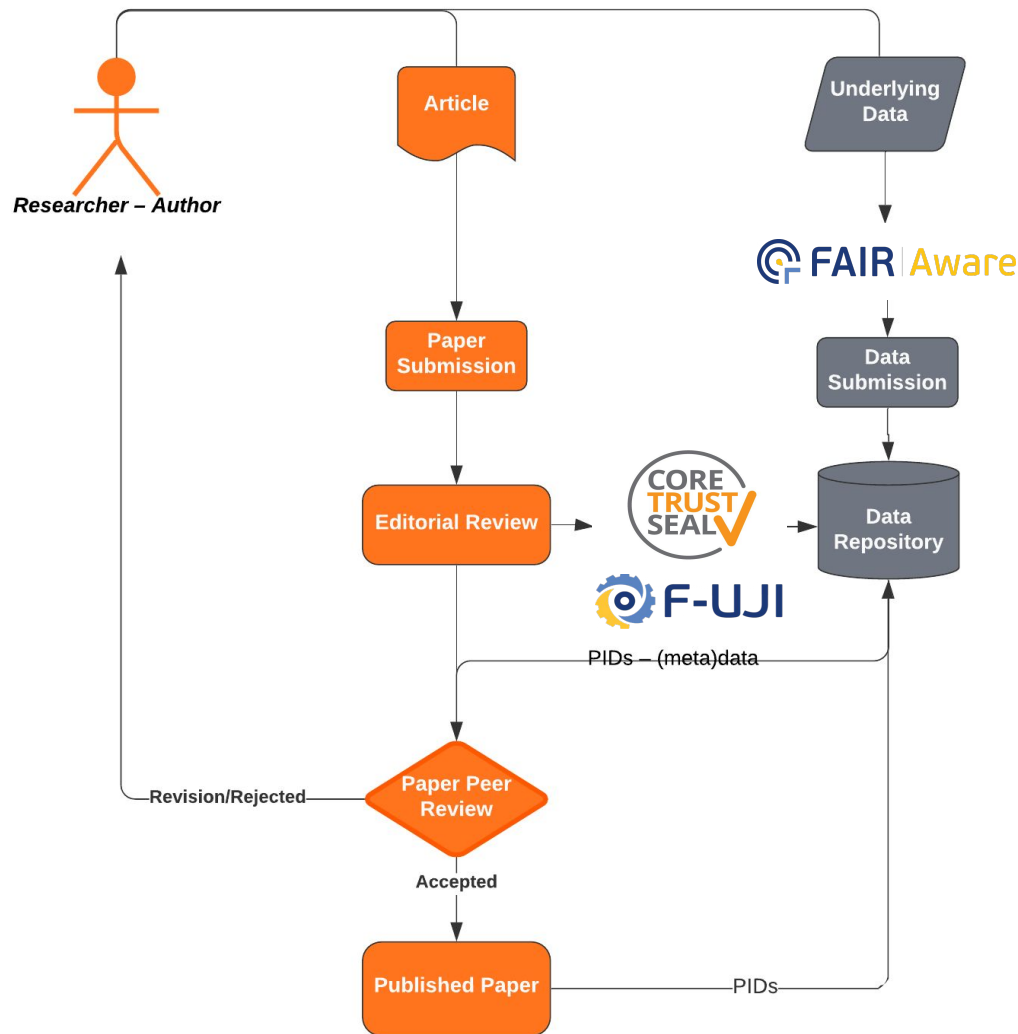
Findable: 6 of 7 

Accessible: 2 of 3 

Interoperable: 1 of 4 

Reusable: 5 of 10 

FAIRsFAIR tools for publishers



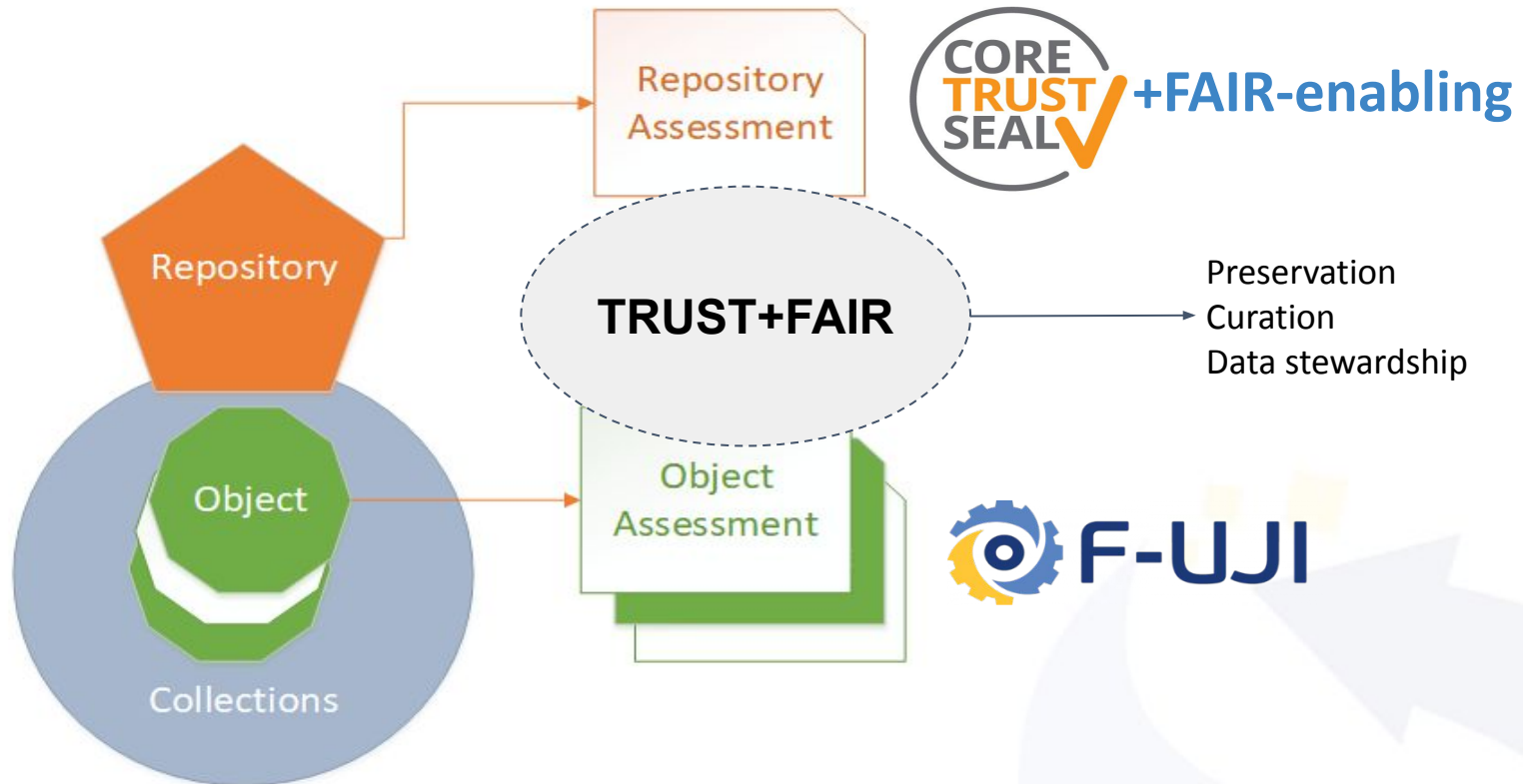


FAIRSFair

Fostering Fair Data Practices in Europe

And for dessert...

FAIR data over time

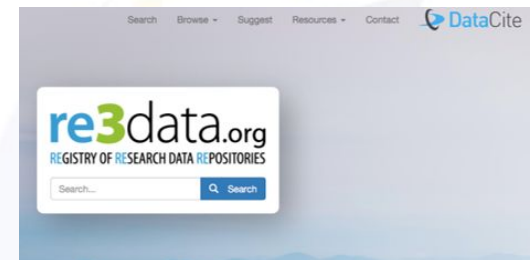


See also: FAIRsFAIR M4.2 Draft Maturity Model Based on Extensions and-or Additions to CoreTrustSeal Requirements,
<https://doi.org/10.5281/zenodo.4003598>

Repository finder tool

developing a tool to identify FAIR-enabling and certified repositories

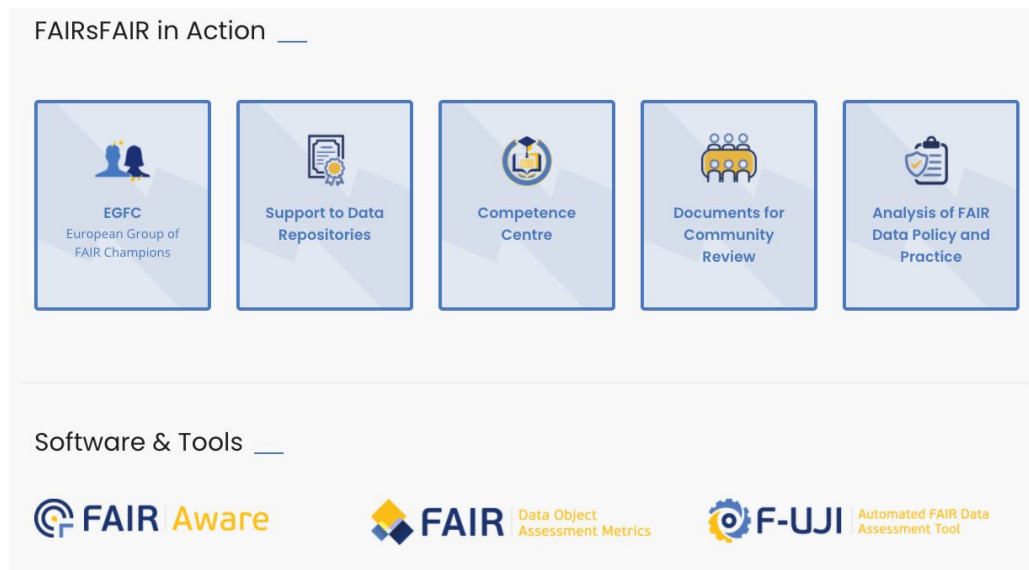
- make FAIR-enabling and certified repositories findable
- revised and improved re3data metadata schema
- integration of repository finder into DataCite Commons



See also: FAIRsFAIR M4.8 Introduce additional components and practices to metadata schema and align them with FAIR data practices, <https://doi.org/10.5281/zenodo.4590298>

What could be next?

- meetings on specific relevant topics, like e.g. publisher-related policy enhancement
- demo of the FAIR assessment tools FAIR-Aware and F-UJI



Thank you for your attention!



www.fairsfair.eu

 FAIRsFAIR_EU

 In/fairsfair

info@fairsfair.eu