



FAIRSFair
Fostering Fair Data Practices in Europe

How repositories can increase their FAIR share

Patricia Herterich | Olivier Rouchon | Maaïke Verburg

Guest speakers: Rudolf Dimper | Peter Sutton-Long

[@FAIRsFAIR_EU](#) | [#OpenRepos2021](#)

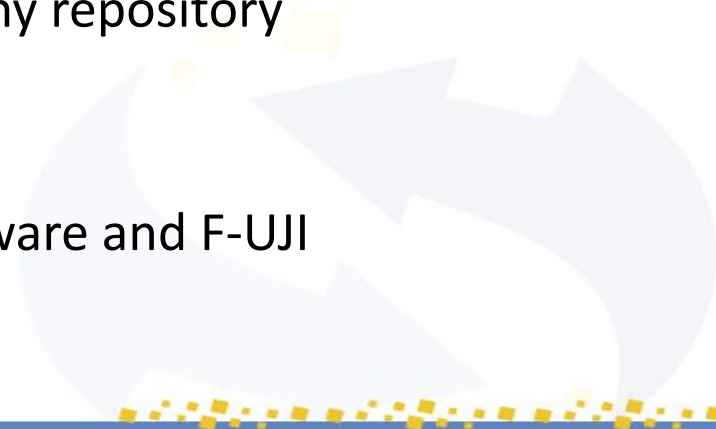
OR2021 Workshop
7 June 2021 - 9:30 (UTC)



Agenda

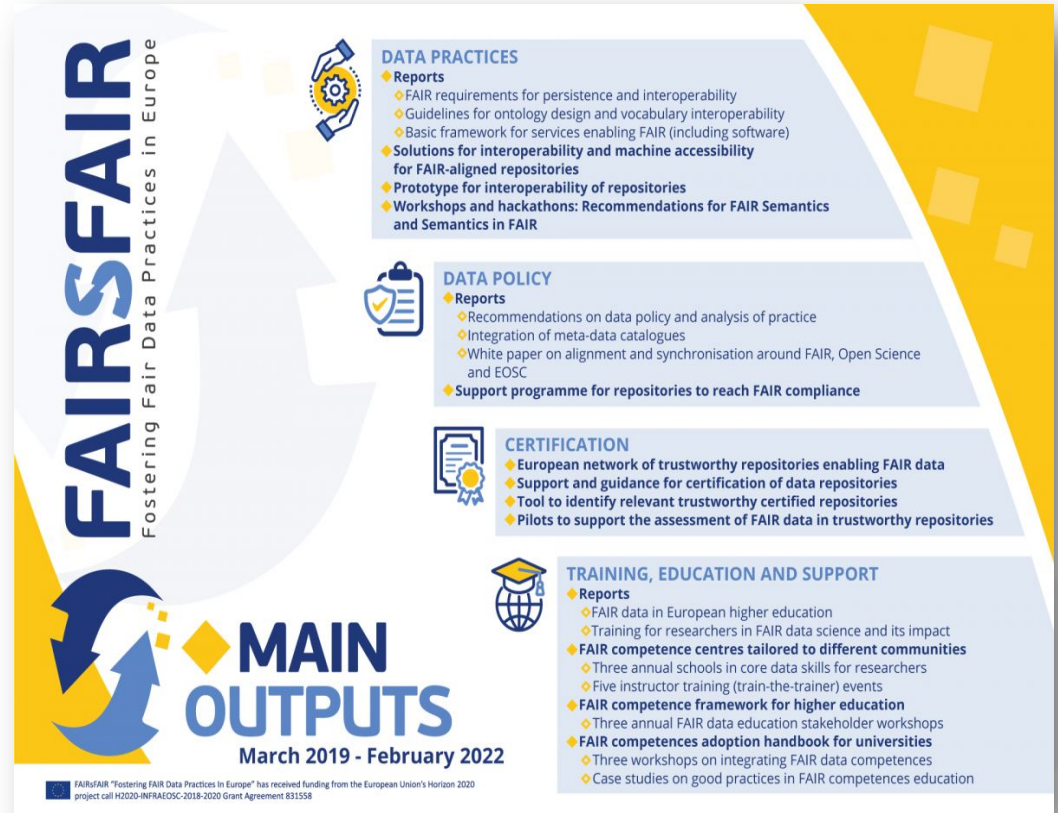
[Note taking doc](#)
(For 100 contributors)

- 9:30 - 9:35 Welcome
- 9:35 - 9:40 FAIRsFAIR repository support program
- 9:40 - 9:55 Lessons learned - repositories and support team
- 9:55 - 10:15 Discussion: FAIRsFAIR tools and other support on your journey to become a trustworthy repository
- 10:15 - 10:25 Break
- 10:25 - 10:35 FAIRsFAIR assessment tools
- 10:35 - 10:50 Hands-on exercise with FAIR-Aware and F-UJI
- 10:50 - 11:00 Wrap-up



FAIRsFAIR - Fostering FAIR data practices in Europe

- Goal: Practical solutions for the use of FAIR principles throughout the research data life cycle
- Budget: €10 million
- Duration: 2019-2022
- Partners: 22 partners from 8 EU member states



FAIRSFair
Fostering Fair Data Practices in Europe

DATA PRACTICES

- ◆ Reports
 - ◆ FAIR requirements for persistence and interoperability
 - ◆ Guidelines for ontology design and vocabulary interoperability
 - ◆ Basic framework for services enabling FAIR (including software)
- ◆ Solutions for interoperability and machine accessibility for FAIR-aligned repositories
- ◆ Prototype for interoperability of repositories
- ◆ Workshops and hackathons: Recommendations for FAIR Semantics and Semantics in FAIR

DATA POLICY

- ◆ Reports
 - ◆ Recommendations on data policy and analysis of practice
 - ◆ Integration of meta-data catalogues
 - ◆ White paper on alignment and synchronisation around FAIR, Open Science and EOSC
- ◆ Support programme for repositories to reach FAIR compliance

CERTIFICATION

- ◆ European network of trustworthy repositories enabling FAIR data
- ◆ Support and guidance for certification of data repositories
- ◆ Tool to identify relevant trustworthy certified repositories
- ◆ Pilots to support the assessment of FAIR data in trustworthy repositories

TRAINING, EDUCATION AND SUPPORT

- ◆ Reports
 - ◆ FAIR data in European higher education
 - ◆ Training for researchers in FAIR data science and its impact
- ◆ FAIR competence centres tailored to different communities
 - ◆ Three annual schools in core data skills for researchers
 - ◆ Five instructor training (train-the-trainer) events
- ◆ FAIR competence framework for higher education
 - ◆ Three annual FAIR data education stakeholder workshops
- ◆ FAIR competences adoption handbook for universities
 - ◆ Three workshops on integrating FAIR data competences
 - ◆ Case studies on good practices in FAIR competences education

MAIN OUTPUTS
March 2019 - February 2022

FAIRsFAIR "Fostering FAIR Data Practices in Europe" has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2019-2020 Grant Agreement 831558

Image source: <https://www.fairsfair.eu>

FAIRsFAIR results



FAIRSFAR

Fostering Fair Data Practices in Europe

DATA PRACTICES

- ◆ Reports
 - ◆ FAIR requirements for persistence and interoperability
 - ◆ Guidelines for ontology design and vocabulary interoperability
 - ◆ Basic framework for services enabling FAIR (including software)
- ◆ Solutions for interoperability and machine accessibility for FAIR-aligned repositories
- ◆ Prototype for interoperability of repositories
- ◆ Workshops and hackathons: Recommendations for FAIR Semantics and Semantics in FAIR

DATA POLICY

- ◆ Reports
 - ◆ Recommendations on data policy and analysis of practice
 - ◆ Integration of meta-data catalogues
 - ◆ White paper on alignment and synchronisation around FAIR, Open Science and EOSC
- ◆ Support programme for repositories to reach FAIR compliance

CERTIFICATION

- ◆ European network of trustworthy repositories enabling FAIR data
- ◆ Support and guidance for certification of data repositories
- ◆ Tool to identify relevant trustworthy certified repositories
- ◆ Pilots to support the assessment of FAIR data in trustworthy repositories

TRAINING, EDUCATION AND SUPPORT

- ◆ Reports
 - ◆ FAIR data in European higher education
 - ◆ Training for researchers in FAIR data science and its impact
- ◆ FAIR competence centres tailored to different communities
 - ◆ Three annual schools in core data skills for researchers
 - ◆ Five instructor training (train-the-trainer) events
- ◆ FAIR competence framework for higher education
 - ◆ Three annual FAIR data education stakeholder workshops
- ◆ FAIR competences adoption handbook for universities
 - ◆ Three workshops on integrating FAIR data competences
 - ◆ Case studies on good practices in FAIR competences education

MAIN OUTPUTS

March 2019 - February 2022

CERTIFICATION

- ◆ European network of trustworthy repositories enabling FAIR data
- ◆ Support and guidance for certification of data repositories
- ◆ Tool to identify relevant trustworthy certified repositories
- ◆ Pilots to support the assessment of FAIR data in trustworthy repositories



FAIR Aware



F-UJI Automated FAIR Data Assessment Tool

Image source: <https://www.fairsfair.eu>

Tools to increase your FAIR share



Researchers & Data stewards

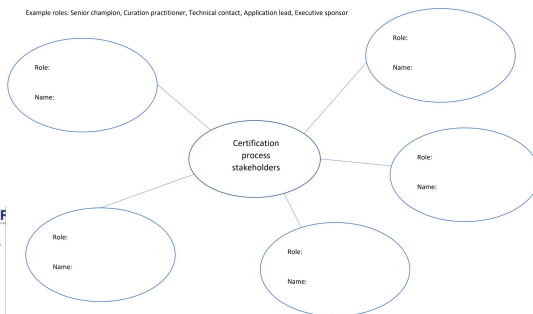
Researchers & Data stewards



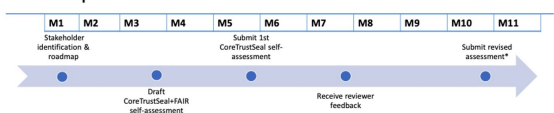
Repository managers

Stakeholder mind map for your institution

Example roles: Senior champion, Curation practitioner, Technical contact, Application lead, Executive sponsor



Roadmap



- Activity/task 1: _____
- Activity/task 2: _____
- Activity/task 3: _____
- Activity/task 4: _____
- Activity/task 5: _____

Roadmap



Risk	Risk level



Repository managers

Preservation Policy Planning - Worksheet



This exercise has been developed to support selected repositories in the FAIR4AIR repository support programme in getting CoreTrustSeal certified. It is part of a variety of resources prepared by FAIR4AIR to support services and individuals in their FAIR data practices. Depending on your situation, it may take around 60 minutes to complete the exercise. The worksheet has been partially adapted from the Digital Preservation Coalition's (DPC) [policy toolkit](#) and the [Conversational CoreTrustSeal working paper](#). This worksheet is not endorsed by the CoreTrustSeal Board and will not guarantee repository certification.

Objectives

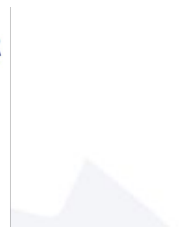
- To identify the status of relevant policy components to have a preservation policy in place.
- To structure and plan the writing process of the preservation policy.

Keep in mind that the construction of a preservation policy content, format and style of your policy accordingly.

Instructions

- To start, indicate the status of each policy component (defined and forms part of an integrated management); the below mentioned components are relevant when do some of the most challenging. You may also leave your workshop.
- Identify the people within your organisation who are so not need to share this information with us). In case you an answer.

Component	Guiding question	Status	Comment/question	Expert	Responsible person for text writing	Duration	Deadline
1. Summary	Is the summary of your policy concise and easily digestible for the reader?						
2. Purpose (RQ)	Why has the policy been developed and what does it aim to achieve?						
3. Organisational Strategy Alignment (RQ)	How does the policy support the overarching purpose and objectives of your organisation?						
4. Mandate (R1)	What are the drivers for the policy and why do they necessitate digital preservation?						
5. Scope (RQ)	What is the scope of the policy in terms of the organisational context and the digital content to be preserved?						
6. Policy Principles	What is the agreed framework and directions for how your organisation approaches digital preservation in a consistent way?						





FAIRsFAIR

Fostering Fair Data Practices in Europe



FAIRsFAIR repository support program



FAIRsFAIR repository support program

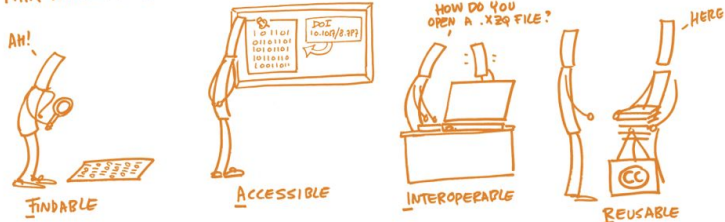
A repository journey towards becoming trustworthy and FAIR-enabling



CoreTrustSeal repository certification



FAIR DATA PRINCIPLES



FAIR data practices

Image: <https://www.openaire.eu/how-to-make-your-data-fair>

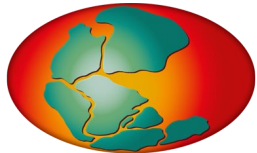
Support team

Data Archiving and Networked Services

DANS



**UK • DATA
ARCHIVE**



Who we work with



→ a country with a selected repository

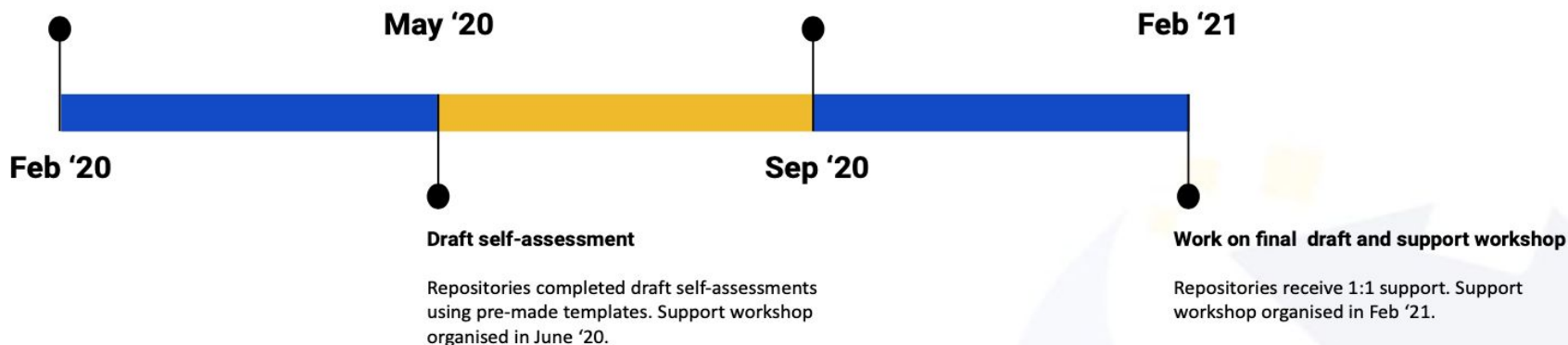
Certification support: stage 1

Programme kick-off

Start of the programme. Representatives from the repositories and FAIRSFAR met at a workshop in The Hague in Feb '20.

First draft submission & peer reviews

Repositories successfully delivered first draft assessments and performed test peer reviews of their submitted work. Support workshop organised in Nov '20



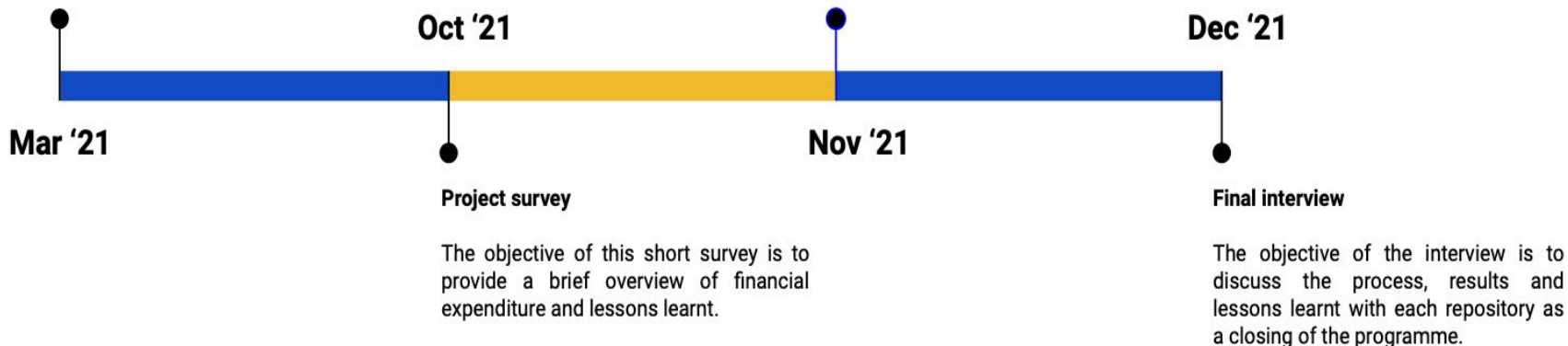
Certification support: stage 2

CoreTrustSeal application submission

The objective is to submit application to CoreTrustSeal.

Final payment & final support workshop

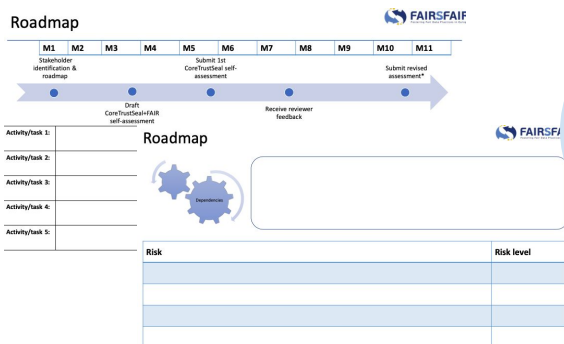
The repositories will receive their second payment. The objective of the workshop is to reflect on key challenges for repositories, lessons learnt and celebrate achievements.



Repository support: services



4 support workshops



Preservation Policy Planning - Worksheet

This exercise has been developed to support selected repositories in the FAIRSFAR repository support programme in getting CoreTrustSeal certified. It is part of a variety of resources prepared by FAIRSFAR to support services and individuals in their FAIR data practices. Depending on your situation, it may take around 60 minutes to complete the exercise. The worksheet has been partially adapted from the Digital Preservation Coalition's (DPC) policy tools; and the Operational CoreTrustSeal workshop assets. This worksheet is not endorsed by the CoreTrustSeal Board and will not guarantee repository certification.

Objectives

- To identify the status of relevant policy components to have a preservation policy in place.
 - To structure and plan the work.
- Keep in mind that the construction content, format and style of your plan will be discussed in the workshop.

Instructions

- To start, indicate the status of each (defined and forms part of an in the below mentioned complete, some of the most challenging, workshop.
- Identify the people within your organisation who will not need to share this information as an answer.

Component	Guiding question	Status	Comment/question	Expert	Responsible person for text writing	Duration	Deadline
1. Summary	Is the summary of your policy concise and easily digestible for the reader?						
2. Purpose (RO)	Why has the policy been developed and what does it aim to achieve?						
3. Organisational Alignment (RO)	How does the policy support the overarching purpose and objectives of your organisation?						
4. Mandate (RI)	What are the drivers for the policy and why do they necessitate digital preservation?						
5. Scope (RO)	What is the scope of the policy in terms of the organisational context and the digital content to be preserved?						
6. Policy Principles	What is the agreed framework and directions for how your organisation approaches digital preservation in a consistent way?						

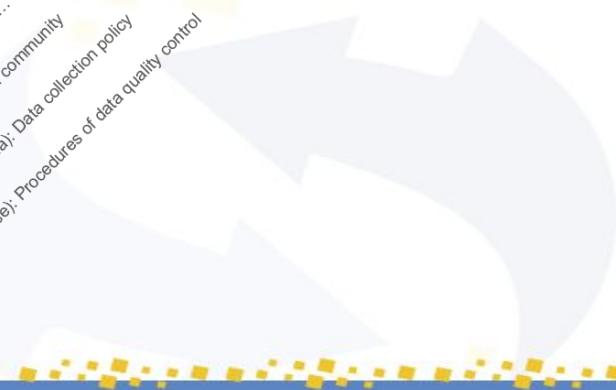
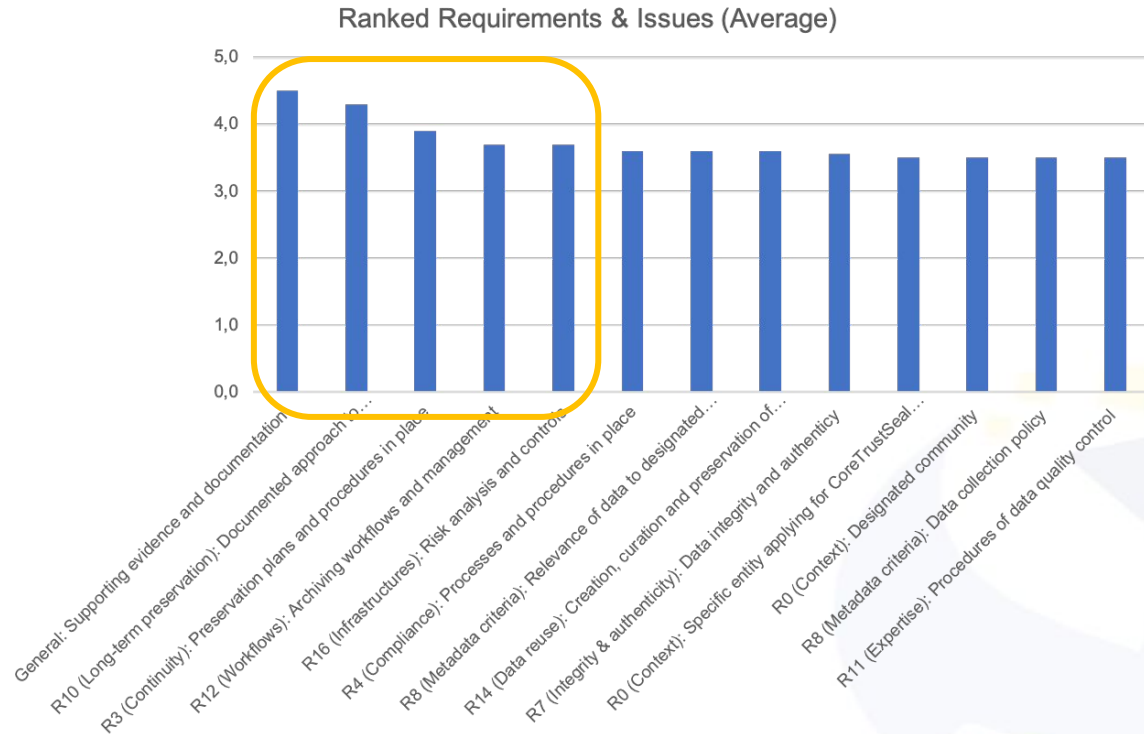
Test reviews

Practical tools



1:1 support and advice

Survey results: top CTS issue areas for repositories

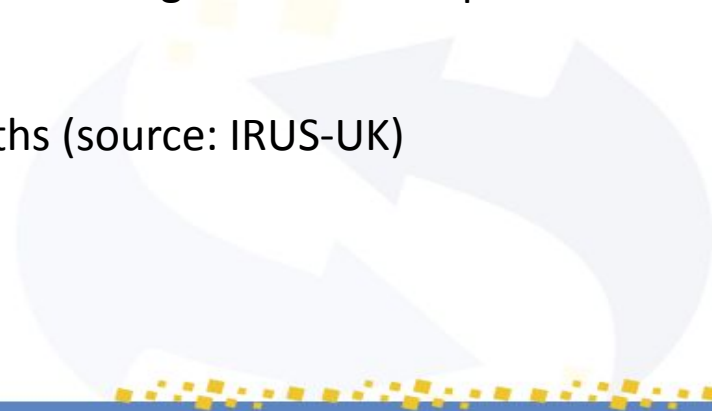


Apollo repository



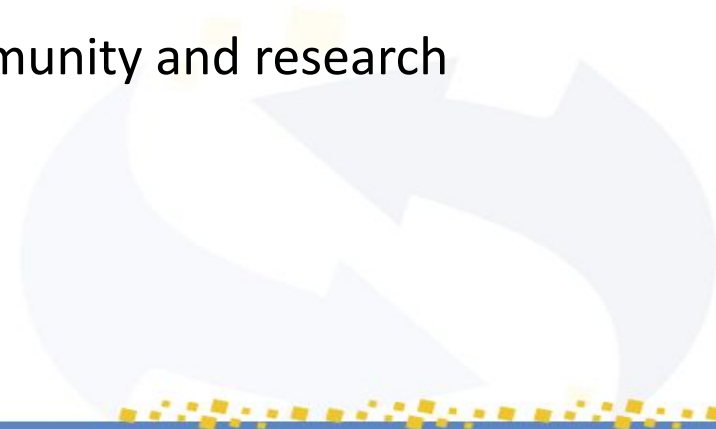
Apollo is the institutional repository of the University of Cambridge:
<https://www.repository.cam.ac.uk/>

- Managed by divisions of Cambridge University Library.
- Central to the University's ambitions to make all Cambridge research outputs available in accordance with the FAIR principles.
- > 250k research outputs
- ~ two million item downloads in the last 12 months (source: IRUS-UK)



Why certify the Apollo repository?

- Institutional level support for best practice in research data management
 - Research Data Management Policy <https://doi.org/10.17863/CAM.10528>
 - Open Research Position Statement
(<https://osc.cam.ac.uk/open-research/open-research-position-statement>)
- Response to requests from research community and research partners



Lessons learned

- Self-evaluation is a challenge and is an iterative process.
- Engage key stakeholders early.
- Gain a thorough understanding of the policy landscape
- Appreciate the importance of providing context (R0)



Lessons learned - ESRF



The European Synchrotron Radiation Facility (ESRF) is a 4th generation synchrotron to produce extremely intense light (X-rays) for studying the structure and functioning of matter.

Key figures:

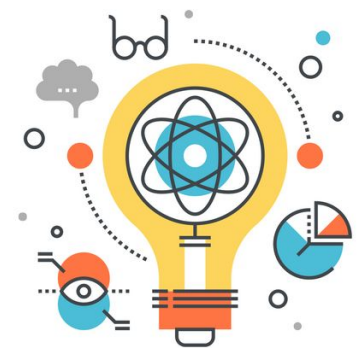
- 13 member countries
- 9 associate countries
- 100 M€/y operation budget
- 700 staff
- 10 000 user visits/y
- 44 beamlines/instruments
- ~2 000 publications/y
- ~10PB/y of data



Why certifying the ESRF scientific data repository

- The ESRF has adopted a research data policy (RDP) since 2015:
 - The RDP is in its final stages of implementation (rich metadata capture, e-logbook, harmonised data format, ...)
 - We commit preserving all data produced at the ESRF for ten years with an embargo period of three years for the initial research team
 - The embargo period can exceptionally be extended
- Experiments at the ESRF produce more and more complex data; some data sets are difficult to transport/transfer
- The visiting scientists, but also the journals in which they publish, must be able to trust in our capacity to manage the data. Trust is essential!
- Reproducibility and re-use important for the future
- ESRF is committed to connect to the EOSC





Key lessons from the ESRF FAIR journey (1/3)

Adopting and Implementing the RDP

Deciding and implementing a RDP is the most important foundation block, it requires:

- careful preparatory work to get maximum adherence from key stakeholders, such as management and in-house scientists
- full involvement of the lab governance to get it adopted and to make the RDP part of the governance papers
- dedicated human and financial resources for its implementation
- not to underestimate the time it takes to define rich metadata and to develop the necessary tools for FAIR data management
- → we were fortunate to have the EU funded project PaNdata-Europe (2010) which allowed to have a common RDP framework for the Photon and Neutron community.

Key lessons from the ESRF FAIR journey (2/3)

The RDP is just the beginning...

FAIR data management has to become a company culture and this requires:

- information and training sessions for scientific and technical staff
- a working group with the main stakeholders to discuss regularly implications of FAIR data management
- fostering community standards such as AAI, data formats, data transfer, data analysis services
- preparing the road from FAIR to Open Science

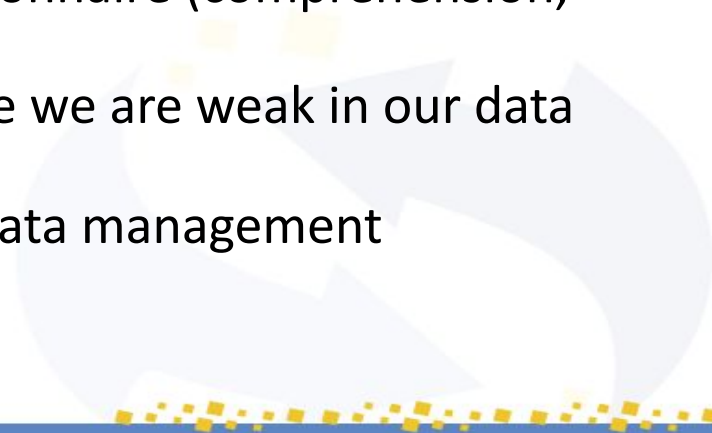


Key lessons from the ESRF FAIR journey (3/3)

The CTS certification process:

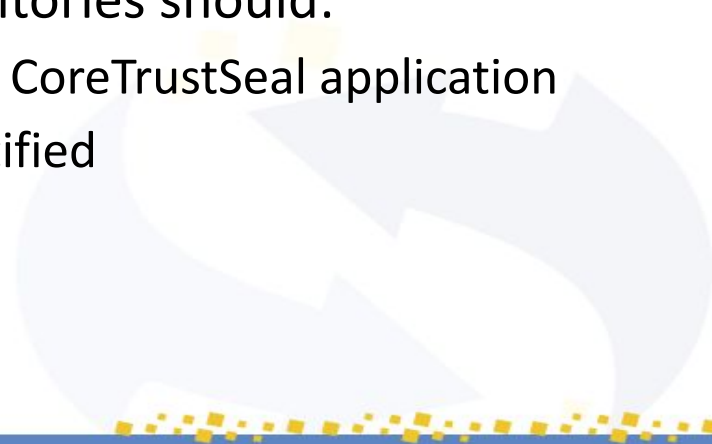
Assistance from FAIRsFAIR to prepare the CTS certification was extremely useful and enriching to:

- clarify and adjust our replies to the questionnaire (comprehension, acronyms, missing parts, length)
- understand where we are good and where we are weak in our data management practices
- measure possible future progress in our data management



Challenges

- Planning is key. Reviewers need to be available for self-assessments.
- Tailored support is needed. Repositories have different starting points on their journey towards trust and FAIR.
- Before the end of FAIRsFAIR all ten repositories should:
 - Have a revised up-to-date version of their CoreTrustSeal application
 - Be as much as possible CoreTrustSeal certified



Discussion

- Short introduction of yourself
- Feedback on FAIRsFAIR tools and other support on your journey to become a trustworthy repository
- Open space to suggest other ways of support



Discussion

Please go to www.menti.com

Enter code: 9614 5710






FAIRSFair

Fostering Fair Data Practices in Europe

Break

See you back in 10 minutes (10:25 UTC)

FAIRsFAIR results - FAIR data assessment



CERTIFICATION

- ◆ European network of trustworthy repositories enabling FAIR data
- ◆ Support and guidance for certification of data repositories
- ◆ Tool to identify relevant trustworthy certified repositories
- ◆ Pilots to support the assessment of FAIR data in trustworthy repositories



How can you make new data more FAIR?

<https://www.fairsfair.eu/fair-aware>



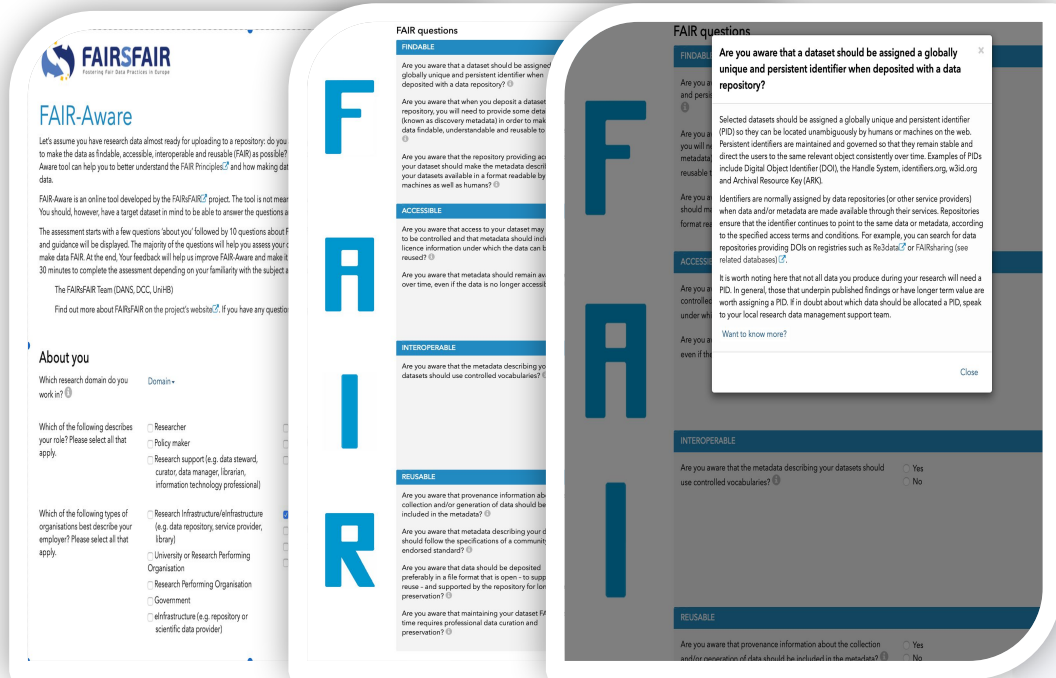
How FAIR are existing data in repositories?

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



FAIR-Aware

- For researchers and data stewards
- Online tool to raise awareness and educate on the FAIR data principles
- 10 simple questions with practical tips to improve data FAIRness before deposit



➔ <https://fairaware.dans.knaw.nl/> | <https://youtu.be/TkxSXxkqhfy>

FAIR-Aware

- Use the resources presented in the tool to increase your knowledge and develop FAIR skills
- Leave feedback about your experience and how the tool can improve
- Spread the word!

Feedback

Please answer the four questions below to help us improve the tool and make it even more relevant for the community.

- Which of the following issues do you find hard to understand and/or to answer?

1. Globally unique persistent identifier (PID)



2. Metadata for citation and discovery

Thank you for your participation!



Awareness:

High (8/10)



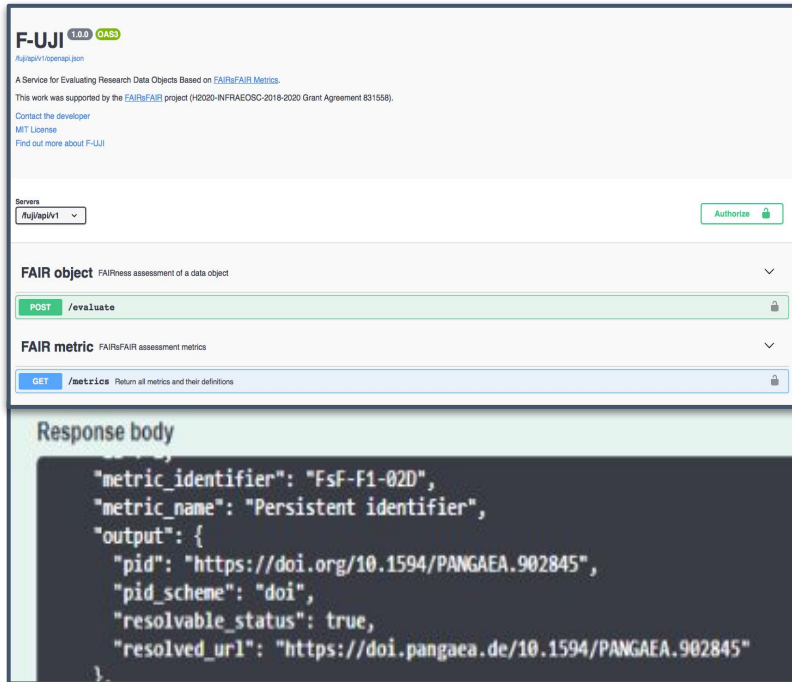
Willingness to comply:

Low (27/50)

→ <https://fairaware.dans.knaw.nl/> | <https://youtu.be/TkxSXxkqhFY>

F-UJI

Machine-actionable assessment tool



F-UJI 1.0.0 OASIS
 fujijpvi@openapi.json

A Service for Evaluating Research Data Objects Based on [FAIRFAIR Metrics](#).
 This work was supported by the [FAIRFAIR](#) project (H2020-INFRAEOSC-2018-2020 Grant Agreement 831558).

Contact the developer
 MIT License
 Find out more about F-UJI

Servers
 /fujijpvi/v1 Authorize

FAIR object FAIRness assessment of a data object
POST /evaluate

FAIR metric FAIRFAIR assessment metrics
GET /metrics Return all metrics and their definitions

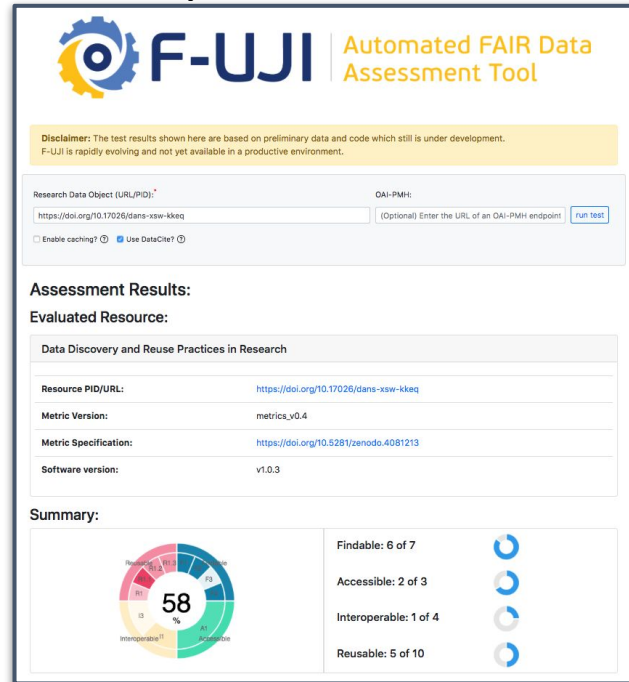
Response body

```

{
  "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



F-UJI | 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.

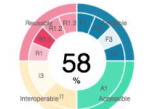
Research Data Object (URL/PID):
 OAI-PMH: run test

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:


 <p>58%</p>	Findable: 6 of 7 🔵
	Accessible: 2 of 3 🔵
	Interoperable: 1 of 4 🔵
	Reusable: 5 of 10 🔵

➔ <https://www.f-uji.net/>

F-UJI

- Provide a PID/URL
- F-UJI test is run
- Output is a FAIRness score + feedback on how to improve the FAIRness of the dataset


FsF-I1-02M - Metadata uses semantic resources 

Status: fail 

Score: 0 of 1

Output: []

Metric tests:

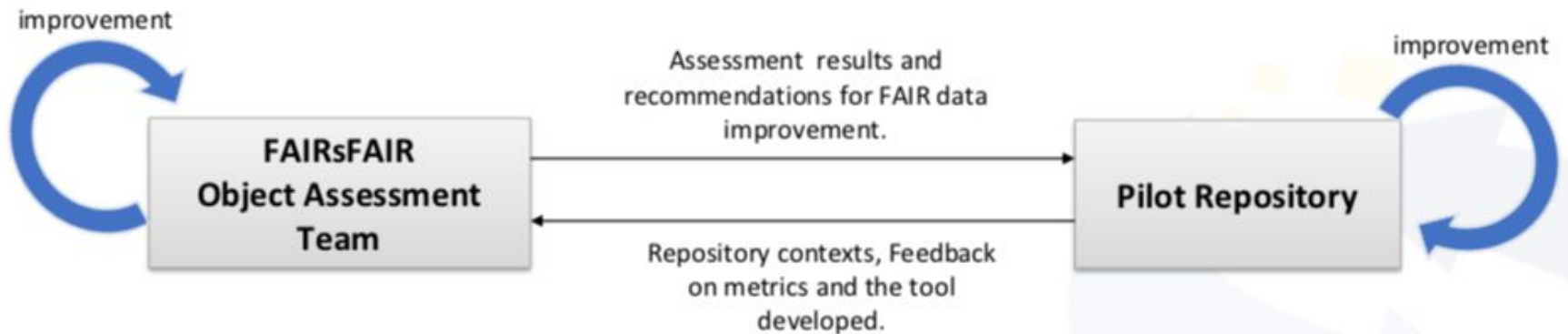
Test:	Test name:	Result:
FsF-I1-02M-1	Namespaces of known semantic resources can be identified in metadata	

Debug:

Level:	Message:
INFO	Number of vocabulary namespaces extracted from all RDF-based metadata -: 1
INFO	Default vocabulary namespace(s) excluded -: ['http://schema.org/']
WARNING	NO namespaces of semantic vocabularies found in the metadata

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

F-UJI - iteration and improvement



Hands-on exercises



- <https://fairaware.dans.knaw.nl/>
- Take a look at the tool (questions, guidance texts, feedback)



- <https://www.f-uji.net/>
- Web-based GUI, dataset level

Use a dataset:

- One of your own, or
- <https://bit.ly/CycloneData>
- <https://bit.ly/BirdsData>

Questions?

Outlook and wrap-up



Tools to increase your FAIR share



Researchers & Data stewards

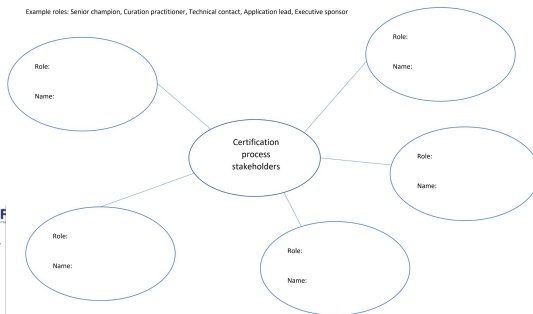
Researchers & Data stewards



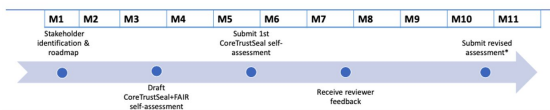
Repository managers

Stakeholder mind map for your institution

Example roles: Senior champion, Curation practitioner, Technical contact, Application lead, Executive sponsor



Roadmap



- Activity/task 1: _____
- Activity/task 2: _____
- Activity/task 3: _____
- Activity/task 4: _____
- Activity/task 5: _____

Roadmap



Risk	Risk level



Repository managers

Preservation Policy Planning - Worksheet



This exercise has been developed to support selected repositories in the FAIR4AIR repository support programme in getting CoreTrustSeal certified. It is part of a variety of resources prepared by FAIR4AIR to support services and individuals in their FAIR data practices. Depending on your situation, it may take around 60 minutes to complete the exercise. The worksheet has been partially adapted from the Digital Preservation Coalition's (DPC) [policy toolkit](#) and the [Conversational CoreTrustSeal working paper](#). This worksheet is not endorsed by the CoreTrustSeal Board and will not guarantee repository certification.

Objectives

- To identify the status of relevant policy components to have a preservation policy in place.
- To structure and plan the writing process of the preservation policy.

Keep in mind that the construction of a preservation policy content, format and style of your policy accordingly.

Instructions

- To start, indicate the status of each policy component (defined and forms part of an integrated management) the below mentioned components are relevant when do some of the most challenging. You may also leave your workshop.
- Identify the people within your organisation who are so not need to share this information with us). In case you an answer.

Component	Guiding question	Status	Comment/question	Expert	Responsible person for text writing	Duration	Deadline
1. Summary	Is the summary of your policy concise and easily digestible for the reader?						
2. Purpose (RQ)	Why has the policy been developed and what does it aim to achieve?						
3. Organisational Strategy Alignment (RQ)	How does the policy support the overarching purpose and objectives of your organisation?						
4. Mandate (R1)	What are the drivers for the policy and why do they necessitate digital preservation?						
5. Scope (RQ)	What is the scope of the policy in terms of the organisational context and the digital content to be preserved?						
6. Policy Principles	What is the agreed framework and directions for how your organisation approaches digital preservation in a consistent way?						

DataCite repository finder - current status

- Filter for re3data.org
<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 [FAIRSFAR 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

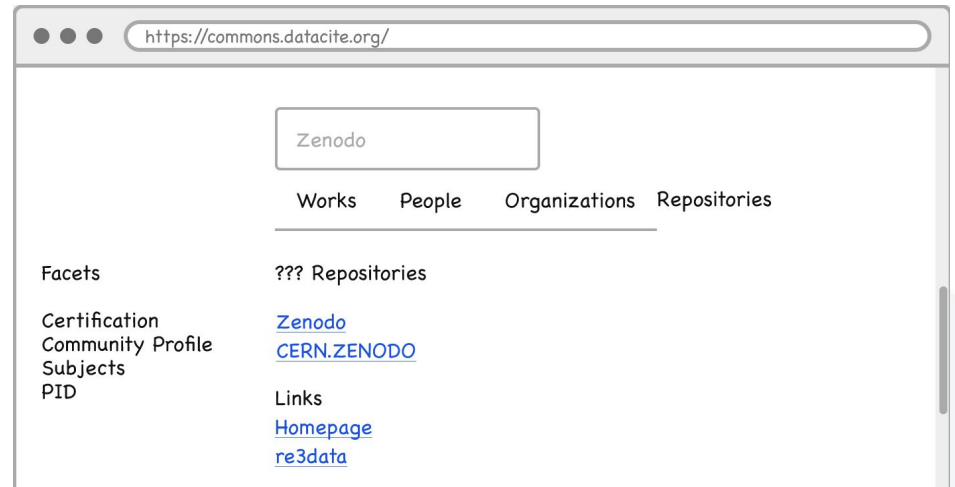
soil sciences

biology

agriculture, forestry, horticulture and veterinary medicine

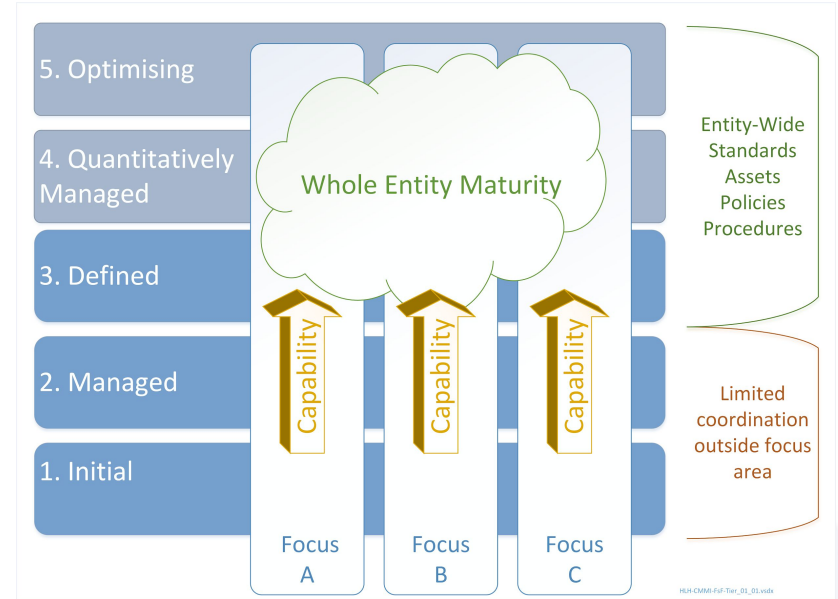
DataCite repository finder - outlook

- Revised and improved re3data metadata schema
- Integration into DataCite Commons:
<https://commons.datacite.org/>



CoreTrustSeal + FAIR - Outlook

- Enhancing CoreTrustSeal with a capability maturity models towards FAIR certification of repositories
- Recommendations for amendments to the next version of CoreTrustSeal



FAIR + Time: Preservation for a Designated Community

- Common discussion paper with EOSC Nordic and SSHOC
- Discusses two key concepts of CoreTrustSeal: designated community and preservation
- Specifies the standards and assessments required for an interoperable ecosystem of FAIR data preserved for the long term
- Intended to help clarify which organisations are potential candidates to receive CoreTrustSeal
- Read it at <http://doi.org/10.5281/zenodo.4783116>

Repository support webinar series

Planned upcoming topics:

- Using Registries to make repository services more visible
- FAIR data over time - developing a Preservation Plan

Upcoming events and sign up for notifications:

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



What else can you do?

- Invite selected repositories to share their FAIR practices widely linking to FAIRsFAIR
- Promote informal networking/support for trustworthy digital repositories (e.g. via [FAIR Data Forum](#)).





FAIRsFAIR

Fostering Fair Data Practices in Europe

Thank you!

info@fairsfair.eu | [@FAIRsFAIR_EU](https://www.fairsfair.eu)

