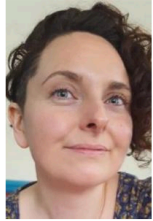


EOSC FAIR Working Group

Dr. Natalie Harrower
Director, Digital Repository of Ireland

*FAIR Forever? FAIRer for Longer: Digital Preservation and the
European Open Science Cloud*
DPC Webinar | 18th March 2021

EOSC FAIR WG Membership



SARAH JONES



FRANÇOISE GENOVA



PALINA SHAUCHUK



CHRISTIAN
KIRCHSTEIGER



MICHAEL BALL



MILAN OJSTERŠEK



NATALIE HARROWER



NEIL CHUE HONG



KOSTAS REPANAS



ANA SLAVEC



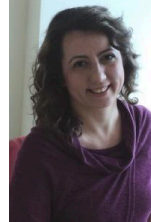
ANDERS CONRAD



ANDRÁS HOLL



OSCAR CORCHO



OYA BEYAN



PEDRO PRÍNCIPE



RACHAEL KOTARSKI



ANDRÉ HEUGHEBAERT



CHRISTINE CHOIRAT



ERNEST CACHIA



HARIS GEORGIADIS



ROB HOOFT



STEFANO COZZINI



JUUSO MARTTILA



KRZYSZTOF KUROWSKI



LESLIE MCINTOSH



LIISI LEMINEN



To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
 - A1.1. the protocol is free, open and universally implementable
 - A1.2. the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

To be Interoperable:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation
- I2. (meta)data uses vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

To be reusable:





- R1. (meta)data are richly described with a plurality of accurate and relevant attributes
 - R1.1. (meta)data are released with a clear and accessible data usage license
 - R1.2. (meta)data are associated with data provenance
 - R1.3. (meta)data meet domain relevant community standards



The FAIR guiding principles from
Wilkinson et al. (2016)



Task Forces

-  PID Policy
-  FAIR Practice
-  Interoperability
-  Metrics & Certification

Implementing the FAIR data principles by defining the corresponding requirements for the development of EOSC services, in order to foster cross-disciplinary interoperability

EOSC INTEROPERABILITY FRAMEWORK

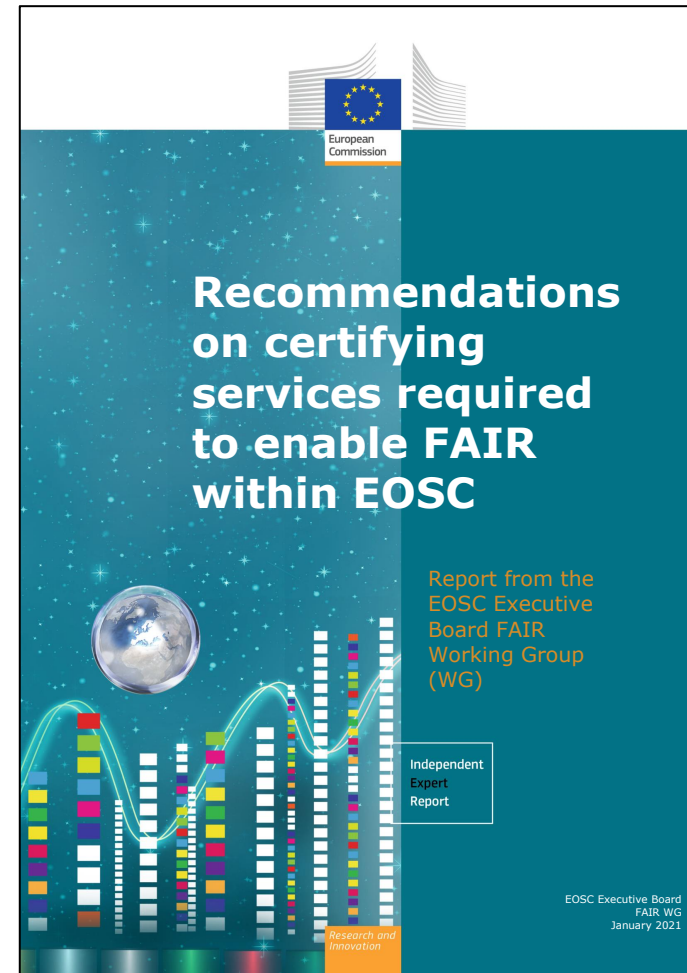
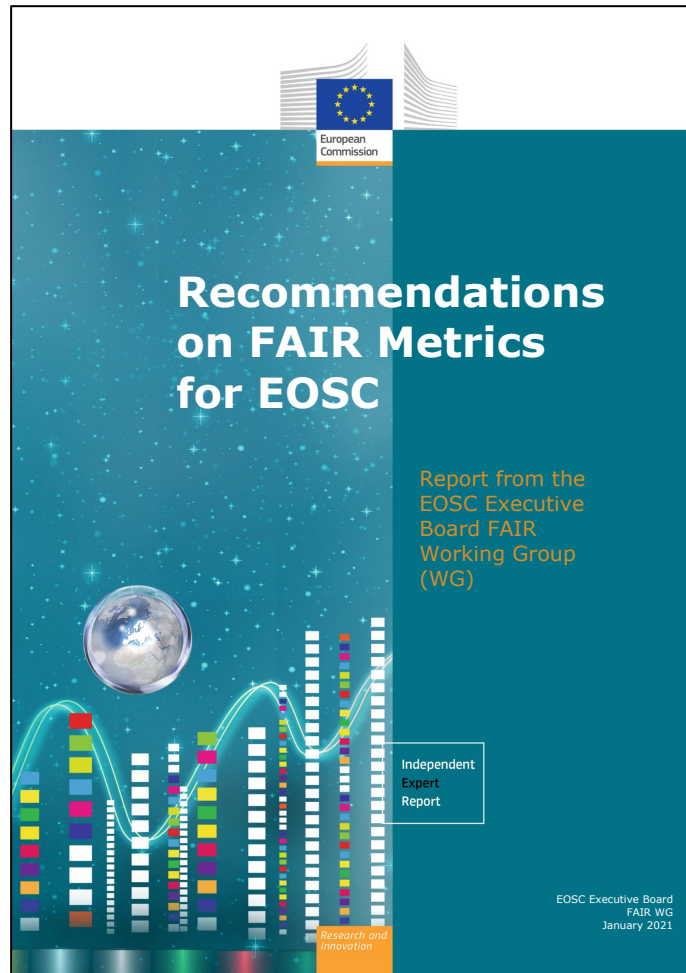
SIX RECOMMENDATIONS FOR IMPLEMENTATION OF FAIR PRACTICE

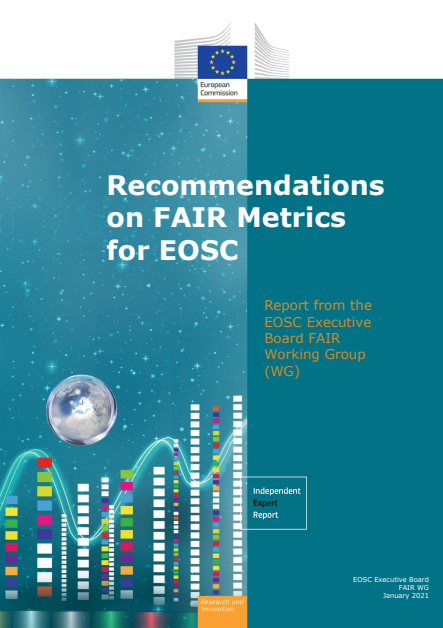
A PERSISTENT IDENTIFIER (PID) POLICY FOR THE EUROPEAN OPEN SCIENCE CLOUD

RECOMMENDATIONS ON FAIR METRICS FOR EOSC

RECOMMENDATIONS ON CERTIFYING SERVICES REQUIRED TO ENABLE FAIR WITHIN EOSC

Metrics & Certification TF





Metrics

Tasked with specifying a set of metrics assessing the FAIRness of datasets and *other digital objects* which should be applied within the European Open Science Cloud (EOSC)

Report builds on work of the RDA FAIR Data Maturity Model Working Group, FAIRsFAIR Data Object Assessment Metrics, and Six Recommendations for FAIR Practice (Another Taskforce in the EOSC FAIR WG)

Understanding of FAIR as an ecosystem; e.g. a dataset is assigned a PID, but then is retrievable via standard communication protocol from a repository.

Underlines importance of disciplinary communities in defining how the implementation of criteria are best suited, and the ‘journey’

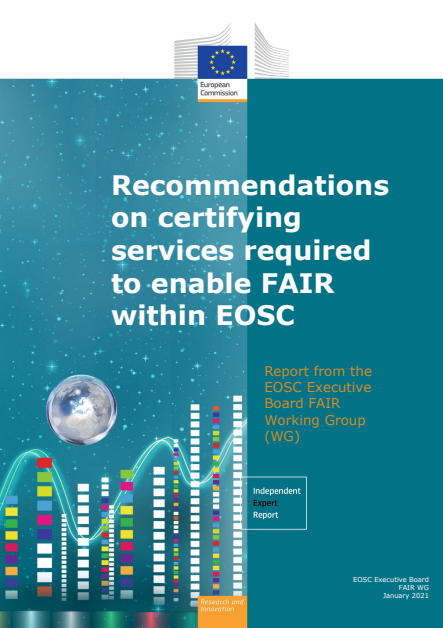
Metrics report Recommendations (redux)

- 1:** Definition of metrics should be a continuous process; regular review
- 2:** Inclusiveness is key: priorities vary by community; FAIR is a journey -- measure progress; build on existing resources
- 3:** Do not reinvent the wheel: use RDA FAIR Data Maturity Model WG
- 4:** Evaluation methods and tools should be thoroughly assessed in a variety of contexts with broad consultation, in particular in different domains to ensure they scale and meet diverse community FAIR practices.
- 5:** FAIR metrics should be developed for digital objects other than data, which may require that the FAIR guidelines be translated to suit these objects, particularly software source code.
- 6:** Guidance should be provided from and to communities for evaluation and implementation.
- 7:** Cross-domain usage should be developed in a pragmatic way based on use-cases, and metrics should be carefully tailored in that respect.

FAIR metrics for EOSC

Proposed list

Name (RDA WG)	RDA WG ID / Priority	EOSC Name	EOSC Timeline			Comments
			2021	2024	2028	
Data is identified by a persistent identifier	RDA-F1-01D/E	EOSC-F1-01D	EOSC-F1-01D	EOSC-F1-01D	EOSC-F1-01D	The expected content of "rich" will evolve with time.
Data is identified by a globally unique identifier	RDA-F1-02D/E	EOSC-F1-02D		EOSC-F1-02D	EOSC-F1-02D	
Rich metadata is provided to allow discovery	RDA-F2-01M/E	EOSC-F2-01M	EOSC-F2-01M	EOSC-F2-02M	EOSC-F2-02M	
Metadata includes the identifier for the data	RDA-F3-01M/E	EOSC-F3-01M	EOSC-F3-01M	EOSC-F3-01M	EOSC-F3-01M	
Metadata is offered in such a way that it can be harvested and indexed	RDA-F4-01M/E	EOSC-F4-01M		EOSC-F4-01M	EOSC-F4-01M	
Metadata contains information to enable the user to get access to the data	RDA-A1-01M/I	EOSC-A1-01M		EOSC-A1-01M	EOSC-A1-01M	This criterion comes later because it is about provenance.
Metadata is guaranteed to remain available, after data is no longer available	RDA-A2-01M/E	EOSC-A2-01M				Progressive implementation when community standards are available.
Metadata uses knowledge representation expressed in standardised format	RDA-I1-01M/I	EOSC-I1-01M				
Metadata is expressed in compliance with a machine understandable community standard	RDA-R1.3-02M/E	EOSC-R1.3-01M				Progressive implementation when community standards are available.



Certification Report Key Findings

At this stage, certification status cannot be a necessary condition for a repository or other key components to be included in EOSC.

May be a prerequisite in future, but must carefully assess certification landscape and possible adverse consequences, such as exclusion of valuable resources used by communities, putting these resources at risk.

Repositories +

Builds on work in
FAIRsFAIR,
CoreTrustSeal,
ELIXIR, TRUST
principles, COAR
framework

Repositories and services wanting to join EOSC should use the certification framework criteria to check and improve their practices, with the aim to progress towards certification. Certified repositories should be clearly identified as such.

CoreTrustSeal, which is a community-driven, international framework used by a large palette of disciplines, is the right level for research data repositories managed in the research environment with respect to DIN 31644 (nestorseal) and ISO 16363:2013. Test model of CTS + FAIR extensively.

Certification report - Priorities

- 1:** Support current efforts to align Certification standards and assessment schemas with FAIR.
- 2:** Test proposed schemas in a variety of communities to gather feedback and update the proposed framework accordingly.
- 3:** Provide support, methodologically as well as financially, to data and service providers to progress towards certification.
- 4:** Monitor the progress of certification, assess the maturity of the certification landscape, and take appropriate action if fields or regions are lagging behind.
- 5:** Support the establishment of core criteria and methodology to certify other key elements of the FAIR ecosystem, in particular in the first instance PID services and vocabulary repositories/metadata registries, and test them extensively.
- 6:** Support the establishment and maintenance of registries of certified components of the ecosystem; if several registries are available for a given component, they should be harvestable and included in registries of registries.
- 7:** Establish a Working Group under the EOSC Stakeholder Forum to ensure the implementation and further development of recommendations in this report.