



## Incentives and metrics for FAIR data and services - Pillar 5

Session date: 14th of May 2020

Chair: Josefine Nordling (FAIRsFAIR)

Rapporteur: Ilona von Stein (FAIRsFAIR)

Information on the participants, the projects and working groups that they represent, and the spreadsheet used during the workshop can be found in the workshop report: <https://doi.org/10.5281/zenodo.3953979>

All recommendations and the action plan can be found on pp. 59-75 in *Turning FAIR into Reality*: [https://ec.europa.eu/info/sites/info/files/turning\\_fair\\_into\\_reality\\_1.pdf](https://ec.europa.eu/info/sites/info/files/turning_fair_into_reality_1.pdf).

**This session is about recommendations 12, 13 (priority) and 25, 26.**

### Table of contents

Incentives and metrics for FAIR data and services - Pillar 5	1
Whole-Pillar.1 What's missing in the recommendations and actions in this pillar?	2
Whole-Pillar.2 Any recommendations not addressed?	2
Rec. 12: Develop metrics for FAIR Digital Objects	3
12.1 In place	3
12.2 Planned	4
Rec. 13: Develop metrics to certify FAIR services	4
13.1 In place	4
13.2 Planned	5
Rec. 25: Implement and monitor metrics	5
25.1 In place	5
25.2 Planned	5
Rec. 26: Support data citation and next generation metrics	6
26.1 In place	6
26.2 Planned	6





## Whole-Pillar.1 What's missing in the recommendations and actions in this pillar?

*What do projects do - related to implementing FAIR in the context of the EOSC - that is not covered by the original recommendations? Should it be included in an updated action plan and revised set of recommendations? Please focus on this pillar.*

**EOSC FAIR WG:** Monitor the monitoring and the monitoring tools

Addition after the end of the meeting: the maintenance of the FAIR principles themselves is a question

**FAIRsFAIR:** Registries of components of the FAIR ecosystem will be a critical part of the evolving FAIR ecosystem

**FAIRsFAIR:** How can we communicate the results of the assessment based on the metrics to different FAIR stakeholders?

## Whole-Pillar.2 Any recommendations not addressed?

*Are the recommendations being covered enough by these activities? If not: what should be done? And by whom?*

- General discussion points:
  - Ongoing definition on FAIR software: on-going work on the definition of FAIR software (RDA/Force11/ReSA WG being formed). Charter open for comment since yesterday in RDA. Community practices are important. Governance for metrics maintenance will be important.
  - Additional comment: Enabling progress is of course a key objective for metrics. I agree with the fact that measurements should be taken as a help not as punishment but it depends who requires the measurements, eg funders, eg as a yes/no test for allowing a repository to be 'accepted', and if there are thresholds they have to be used cautiously. EOSC-Nordic: evaluation work fits within pillar 5. Reusing existing Wilkinson et al tool. Large sample of repos tested. Proposed to be a guideline for repos to make efforts to become FAIR and how to address aspects of FAIRness. D4.1 (end of August) deliverable output. Improvement and sharing is important, not to exclude or punish.
  - FAIR champion: effectiveness of metrics; improvement of repos through maturity approach might be useful, rather than granting fulfillment of a specific metrics against a rather weak baseline. (There is perceived pressure not to be too harsh with repositories / institutions when it comes to fulfilling metrics, leading to low baselines, which makes adoption easier, people happier in the short run, but lowers the value) Fine balance qualitative/quantitative. clear, quantitative metrics obviously preferred. Subjective measures will exist, but there is a need for making them clearly (reproducibly) measurable if they are to be used for auditing/certifying repositories. If subject experts do not agree on the evaluation of a metric, then either auditors need to be better



trained (i.e. the metric better/more precisely described/defined) or it needs to be changed to make it more consistently evaluable. (Question: in how far have the metrics been evaluated with respect to these criteria? Have there been any studies on how consistently specific repositories have been evaluated by different evaluators?)(EOSC WG FAIR: CoreTrustSeal had two evaluators & the Board is checking the evaluations. They would likely be interested in feedback)

## Rec. 12: Develop metrics for FAIR Digital Objects

### 12.1 In place

*What have the projects **already done** that addresses this recommendation? This should build on the information in the spreadsheet. Please check that there is a link to the concrete deliverable.*

- **FAIRsFAIR:** Object assessment metrics are developed. Pilots will be run to help the assessment of individual datasets within repositories. Focus on two primary use-cases: researchers (manual awareness tool) and repositories (automated FAIR assessment). First deliverable available: D4.1 Draft recommendations on requirements for FAIR datasets in certified repositories.
- **EOSC FAIR WG:** metrics and certification taskforce in EOSC EB FAIR WG. Review what is done by others. Interim report launched on those two topics (will be added to spreadsheet);
- **EOSC-Nordic:** Evaluation of repository FAIRness through testing of selection of datasets within repositories. Project has contributed to the extensive testing and multiple improvements of the FAIR Maturity evaluator tool, including harvester and many of the indicators. [Details to be provided in D4.1 (Aug2020) and D4.3 (Feb2022)]
- **NI4OS:** D4.3 “Mapping of legal, technical and procedural tools”, <http://doi.org/10.5281/zenodo.3820880> has identified studies, frameworks and tools that propose and/ or make use of FAIR metrics.
- **ENVRI-FAIR:** The result of the assessment is currently being implemented to a common system, for more details see D5.1 Requirement analysis, technology review and gap analysis of environmental research infrastructures at: <https://envri.eu/deliverables/>. Heterogeneous datasets makes it somewhat challenging. [Remark from the rapporteur: This is key: we would not want to be data to NOT be FAIR because a data collection is heterogeneous. Also we would not want to leave out older (read: non-machine readable) complex collections.]
- **ESCAPE:** WP4 participated in the tests of the RDA FAIR Data Maturity Model criteria. Astronomy has been developing well established FAIR practices and it was a useful test. The main aims for astronomy are Interoperability and Reuse, Find and Access are mostly in support to I and R. Finding is mostly a dynamic process.





## 12.2 Planned

What are the projects represented **developing or planning** to do? Again, this should build on the information in the spreadsheet: information about a planned deliverable, i.e. title, due date, short description

- **FAIRsFAIR:** D4.5 (Aug 2021) + Milestone report 4.9 (Aug 2020): Report on FAIR data assessment toolset and badging scheme will contain the **final version** of metrics defined in D4.1 (Feb 2020).
- **EOSC FAIR WG:** metrics and certification taskforce in EOSC EB FAIR WG. Review what is done by others. Around summer hopefully next version of report.
- **EOSC-Nordic:** Evaluation of repository FAIRness through testing of selection of datasets within repositories. Project has contributed to the extensive testing and multiple improvements of the FAIR Maturity evaluator tool, including harvester and many of the indicators. [Details to be provided in D4.1 (Aug2020) and D4.3 (Feb2022)]
- **NI4OS:** FAIR Maturity Model is also guiding practices in that area and service providers will be informed about all possible solutions for assessing FAIRness and will be provided with support within WP3, 4 and 5 should they are willing to implement them
- **ExPaNDs:** develops draft policy data framework. Also in collaboration with PANOSC: working on RDA FAIR data maturity WG. Possible use case approach to be developed - work in progress

## Rec. 13: Develop metrics to certify FAIR services

### 13.1 In place

What have the projects **already done** that addresses this recommendation? This should build on the information in the spreadsheet. Please check that there is a link to the concrete deliverable.

- **FAIRsFAIR:** Assessment report on FAIRness of services. Services enabling FAIR.
- **EOSC FAIR WG:** Metrics & Certification TF works on a recommendation on certification for FAIR-enabling services. Liaison established with FsF WP2 & 4. metrics to be used cautiously for evaluation. For repositories alignment of CoreTrustSeal with FAIR data principles; CTS has cross-disciplinary capacities. Some communities may want to develop their own certification system eg Elixir/EOSC-Life. Supporting activities (attending meetings organised by FsF, including the one foreseen next week during the EOSC week), Session on FAIR Metrics organised by the FAIR WG in Budapest)
- **EOSC-Nordic:** surface maturity framework: work on how to evaluate and certificate services. Services not being repos.





- **ENVRI-FAIR:** sub-domains within ENVRI-FAIR work together on D5.1, on questions around certification of services.
- **SSHOC:** D8.2 Certification plan for SSHOC repositories (<https://sshopencloud.eu/d82-certification-plan-sshoc-repositories>)

## 13.2 Planned

What are the projects represented **developing or planning** to do? Again, this should build on the information in the [spreadsheet](#): information about a planned deliverable, i.e. title, due date, short description

- **FAIRsFAIR:** Updated version of Assessment report on FAIRness of services. Services enabling FAIR.
- **EXpaNDs:** work mostly around repos. Work to be started: certification schemes incl CoreTrustSeal and other: assess them and profile them for the Expands communities. Different schemes that are emerging, how to be applied to their communities
- **SSHOC:** also work with CoreTrustSeal and how to be used for their communities (2021). Work on how should or could CoreTrustSeal criteria evolve.

## Rec. 25: Implement and monitor metrics

### 25.1 In place

What have the projects **already done** that addresses this recommendation? This should build on the information in the spreadsheet. Please check that there is a link to the concrete deliverable.

- **FAIRsFAIR:** pilots of two use cases (due on D4.5, Aug 2021) 1) raise awareness about FAIR through a manual self-assessment tool and 2) automated assessment tool for evaluating publish datasets
- **EOSC FAIR WG:** early to decide on tools, metrics, badges. Governance set up is key. Need to test the tools
- **EOSC-Nordic:** intention to monitor repos through time. challenge to measure with evolving tools. Frequency of planned testing.
- **NI4OS-Europe:** specific aspects of FAIR-aligned repositories will be made possible for repositories using the OpenAIRE Repository Dashboard (eg metadata enhancement); overall, for monitoring specific aspects of FAIRaligned services the use of tools such as Argo has been discussed.

### 25.2 Planned

What are the projects represented **developing or planning** to do? Again, this should build on the information in the spreadsheet: information about a planned deliverable, i.e. title, due date, short description



- NI4OS: Metrics derived from onboarded services and repositories will be monitored at the pre-production platform before fed to EOSC.

## Rec. 26: Support data citation and next generation metrics

### 26.1 In place

What have the projects **already done** that addresses this recommendation? This should build on the information in the spreadsheet. Please check that there is a link to the concrete deliverable.

- **SSHOC**: SSH citation practices deliverable out, D3.2 Inventory of SSH citation practices, and choice for SSHOC citation formats and implementation planning
- RDA dynamic data citation: should be easy for researchers; also will repos do it if lets say no researchers use it

### 26.2 Planned

What are the projects represented **developing or planning** to do? Again, this should build on the information in the spreadsheet: information about a planned deliverable, i.e. title, due date, short description

- **SSHOC**: D3.5 Report on integration and exploitation of citation and semantic annotation in SSH catalogues (August 2021)