

SUMMARY REPORT OF WORKSHOP ON HARMONISING TRAINING RESOURCE METADATA FOR EOSC COMMUNITIES

Wednesday 14 April, 2021 14:00 -16:30 CEST

Organised by: FAIRsFAIR/ EOSC-5- Task Force Training & Skills

Report prepared by Gabin Kayumbi, Elizabeth Newbold, Angus Whyte and Emma Lazzeri based on the collective notes taken on the day. With thanks to the participants and co-organisers.

About the workshop:

The workshop was co-organised by the [FAIRsFAIR](#) initiative and the EOSC-5 Task Force Training & Skills. It engaged all the InfraEOSC-5 projects (EOSC-Pillar, [ESOC-synergy](#), [EOSC-Nordic](#), [NI4OS-Europe](#), [ExPaNDs](#) and [EOSCSecretariat.eu](#)) as well as representation from related ESFRI Cluster Projects ([PaNOSC](#), [SSHOC](#), [ENVRI-FAIR](#), [EOSC-Life](#) - including [ELIXR TeSS](#)) and other related initiatives ([Terms4FAIRskills](#), [FAIRsharing](#), [RDA-IG ETHRD](#), [OpenAIRE](#), EOSC-FUTURE)

The workshop goals were two fold:

1. Plan joint action by current EOSC projects for a pilot to address recommendations in the report of the EOSC Working Group on Training & Skills, identifying issues and decisions to be taken on 'priority areas' selected (see [notes from previous workshop](#) and charts below)
2. Produce a publicly available summary report from the workshop based on the collaborative notes.

The workshop was a follow-up to an early meeting held in October 2020 where the participants discussed shared priorities.

Agenda:

| | |
|---|--|
| 14:00 Introduction | Emma Lazzeri, EOSC-Pillar, Task Force Chair |
| 14:10 FAIRsFAIR - a proposal on next steps to address the priorities and challenges from the October 29 workshop | Elizabeth Newbold, Angus Whyte FAIRsFAIR |
| 14:20 EOSC Future - Proposed scope of catalogue | Pedro Príncipe, UMinho/ OpenAIRE |
| 14:30 - 14:40 Questions | |

| | |
|---|---|
| 14:40 Break | |
| <p>14:45 Discussion and conclusions - (3 x 25m)</p> <p>Topic 1: 14:45 - 15:10 Minimal metadata for learning resources - comparison of minimal terms draft proposed by RDA ETHRD-IG and those in EOSC WG report (see table below)</p> <p>Topic 2: 15:10 - 15:35 Base standard/schema to use for a metadata application profile</p> <p>Topic 3: 15:35 - 16:00 Sustaining contribution of quality metadata to the catalogue</p> | All |
| 16: 00 Overall conclusions and next steps | Emma Lazzeri, EOSC-Pillar, Task Force Chair |

Introduction:

The workshop started with an introduction from Emma Lazzeri (EOSC-Pillar and Chair of the INFRAEOSC-5 task force on Training & Skills). A brief recap of the recommendations from the EOSC Working Group on Skills and Training as outlined in their report [Digital skills for FAIR and Open Science](#). This report published in Feb 2021 presented a framework for EOSC actors involved in training. The main recommendation from the 4 priority areas was:

- Building a learning and training catalogue.
- User-centric design in building the catalogue.
- Interoperability between catalogues is another priority.
- Quality insurance for the material and sustainability of catalogues too.

FAIRsFAIR Proposal:

Following on from Emma, Elizabeth Newbold and Angus Whyte from the FAIRsFAIR project set out a proposal to address the priorities and challenges from the October 2020 workshop and the recommendations from the EOSC Working Group. In order to progress some of the work already undertaken on minimal metadata it was proposed that a testbed of resources is created drawn from multiple communities that have already established catalogues, building on the recommendations from EOSC to apply the minimal metadata, and the development of an application profile to describe the testbed of materials, test their interoperability between catalogues and identify the curation policy issues that arise from describing them in a standardised way.

EOSC Future:

Pedro Principe (University of Minho/OpenAIRE) gave a brief outline of what is within the remit of the EOSC Future project, starting April 2021, is the development of an EOSC knowledge hub which will have both an EOSC learning platform and an EOSC Training Resource Catalogue both of which are 'intended to support FAIR sharing and reuse of training resources in EOSC. The catalogue will be an aggregator, there will be the opportunity to provide direct link to resources and the deposit of content and it is envisaged that one will not need to be in Europe to contribute but will need to comply with any rules of participation.

Training Catalogue Interoperability:

In the October workshop participants were asked about the steps necessary for training catalogue interoperability.

At that time the highest priority and which of the steps are most difficult.

Training catalogue interoperability: which steps are the highest priority?

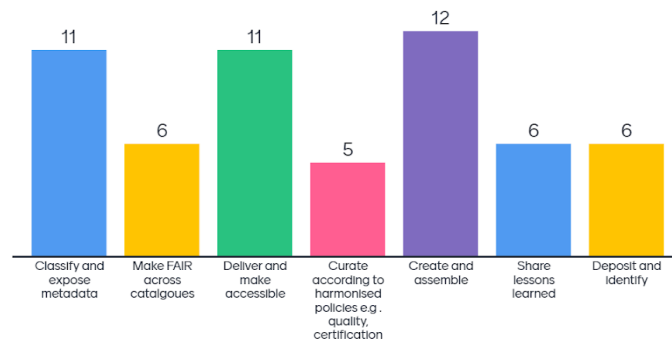


Training catalogue interoperability: which steps are the most challenging?



As a follow-up in April the question was asked again but this time to see which of the steps the projects are actively working on.

Training catalogue interoperability: which steps are you/your project actively working on addressing?



Both the highest priority and most challenging steps identified in the October workshop was 'classify and expose metadata' as we can see from the responses in April, this step along with create and assemble, discover and make accessible are the areas that most projects are working on addressing. Whilst there are different priorities for the projects and other stakeholders all aspects are being addressed albeit at different stages.

Discussion:

The workshop had three topics for discussion using Padlet to gather information and feedback. The topics were selected from those identified in the previous workshop as being priority areas.

Topic 1: Minimal Metadata for Learning Resources (Padlet1)

The wide ranging discussion looked at the draft minimal metadata for learning resources proposed by the RDA-IG ETHRD, and a broader set of metadata terms for this purpose, proposed in the Report from the EOSC Executive Board Skills and Training Working Group¹.

The particular emphasis was on whether the RDA ETHRD proposal would fit for EOSC use cases from the point of view of catalogue users (as learners/researchers, data stewards and training providers).

The padlet records the discussion on individual metadata elements, some additional points were raised in the discussion that weren't recorded on the padlet.

There was a suggestion to clarify the scope and definition of Learning Resource in the context of the remit and scope of the EOSC Training Catalogue e.g. is it for stand-alone training material or material from events? The RDA-ETHRD IG refers to 'Learning Resources' which is the term used in LRMI. The scope of the IG work does include training resources drawn from events, but is not necessarily restricted to these. In the case of EOSC, the scope will be dependent on the selection criteria that are applied to the catalogue.

¹ European Commission DG Research & innovation (2021) 'Digital skills for FAIR and Open Science' <https://op.europa.eu/en/publication-detail/-/publication/af7f7807-6ce1-11eb-aeb5-01aa75ed71a1/language-en/format-PDF/source-190694287>

There is an ongoing debate in the IG between what is required as minimal metadata and what should be included as extended metadata, an example of this is description/abstract which some felt was not necessary as in a minimal set. Again in a similar vein, as to the scope of the catalogue, what counts as minimal will be subjective and context specific as different catalogues may have different use cases that a minimal set of terms need to satisfy. In terms of the future EOSC catalogue these will need to be established building on the work already undertaken in the community.

Topic 2: What 'base' standard/schema to use for an application profile (Padlet 2)

Three questions were discussed:

- What do we mean by an 'application profile' and 'base standard'?
- Which schema is your project using (if any)?
- What issues do we (EOSC projects) need to consider?

The rationale for a Metadata Application Profile (MAP) is to enable harmonisation across existing standards, to fulfil a more specific subset of requirements. A MAP is created using elements from other schemas; "schemas which consist of data elements drawn from one or more namespaces, combined together by implementers, and optimised for a particular local application" (Heery and Patel, 2000). It was felt that an application profile represents a sound approach, provided that stakeholders are willing to use and combine fields from different schema/metadata profiles. It remains to be seen whether a single application profile is sufficient.

The 'base' standard mentioned above refers to the standard used in the application profile, and which others are mapped onto. The base standard does not replace the variety of standards in use, but facilitates aggregation of items represented according to them and mapped to the 'base'. Currently there are various generic standards, e.g. those considered by the RDA ETHRD-IG task group and there is a need to translate between those.

One approach to harmonisation is to directly map from one standard to another, and do this for each pair of standards. The FAIRsharing collection provides examples of this approach, listing crosswalks of the most used metadata schemes, and guidelines for the description of digital objects in open science². A common profile may provide a more effective and efficient route for harmonisation to catalogues that currently adopt any of the standards, by offering a single schema that these may be mapped to, thereby allowing their resources to be aggregated within the EOSC ecosystem.

It was agreed that it is not realistic to expect an overall agreement on describing metadata for resources so that ALL the providers of training material adhere to the same standard. But that it would not be unreasonable for one to be described within EOSC for those that want to provide material to an EOSC training catalogue. For stakeholders that are in relatively early stages of developing catalogues it could be helpful to have a specification to work to.

Looking at the schemas that are already in use within the represented projects and initiatives, there were a broad range of responses indicating the different approaches, domains that the catalogues serve and the maturity of the catalogue. The following schemes and approaches were highlighted:

- IEEE Learning Object Model
- Bioschemas.org (either already in use or working towards implementation)
- Schema.org (either already in use or planning to use)

² FAIRsharing Collection: Crosswalk of most used metadata schemes and guidelines for metadata interoperability

- o Elements relating to DCAT
- o Self-built
- o None at present.

Finally in this session we briefly discussed what EOSC projects need to consider. Questions posed included :

- How important is it to represent the profile in RDF e.g. for sustainability?
- Inclusiveness - use standards adopted by providers of training/learning resources?
- Feasibility - how does choice constrain the tools used to construct the profile?

Whilst there was not specific or conclusive agreement on this there was broad support for the use of RDF to represent the profile and that using a controlled vocabulary should be a minimal requirement.

Topic 3: Which projects could contribute resources to a testbed (Padlet 3)

Finally we went back to the proposal to develop a testbed of materials with projects suggesting what they could contribute. Interest was noted from ELIXIR, ENVR-FAIR, EOSC-Pillar, NI4OS, SSHOC, DMT Clearing House. Suggestions of existing material to contribute also included MOOC in Delivering RDM Services, OpenAIRE RDM guides on Zenodo

Next steps:

Following the workshop there are a number of next steps:

- It was proposed that a further workshop should be planned for the autumn, to look at a topic (to be determined) in more depth that will also consider the initiatives and projects that are now starting their work on similar/connected topics (EOSC Association Advisory groups and EOSC related projects)
- Continued engagement and participation with the RDA. The output from this workshop was incorporated into the RDA Plenary 17 session of ETHRD-IG, on minimal metadata for training resources, that took place the following week. This IG provides a platform for establishing broad community consensus, and includes participants from FAIRsFAIR, ENVRI-FAIR, SSHOC and ELIXIR.
- FAIRsFAIR will continue to work with colleagues in EOSC Futures and across the projects to progress the following steps: -
 1. Compile a testbed of learning resource metadata drawn from multiple communities that have already established catalogues of these resources.
 2. Establish a metadata application profile in collaboration with the RDA Education and Training Interest Group, drawing on currently used standards for classifying and exposing learning resource metadata.
 3. Assess schema options for representing and aggregating the metadata described according to the profile.
 4. Refine the resource descriptions in the testbed using the application profile definitions and mappings to relevant standards.
 5. Pilot the exchange of the metadata testbed between several catalogues that use different metadata standards and curation policies/ rules of participation, and report on the harmonisation successes and challenges.