

EOSC Interoperability Framework Out For Comment

● *By Sarah Jones*

Achieving interoperability within EOSC is essential to federate services and provide added value for users. The draft EOSC Interoperability Framework, which is now open for comment, identifies general principles and organises them into the four layers: technical, semantic, organisational and legal. The framework also contains a proposal for how the management of FAIR Digital Objects should be done in the context of EOSC.

The initial draft has been developed by members of the FAIR and Architecture Working Groups. The authors conducted an extensive review of related literature and interviewed key stakeholders from ERICs, ESFRI projects, service providers and research communities. This helped to identify problems and requirements in each aspect of interoperability to provide recommendations for EOSC. Legal issues will be included in the next version, based on recommendations from a commissioned study.

We welcome community feedback to help us iterate the final version due in October, specifically:

Feedbacks provided by TRIPLE project: <https://www.gotriple.eu/>

Submitted by: Francesca Di Donato

Date: 2020-07-17

Is this what you expected to see or are some things missing?

Overall, it's a good overview of the main interoperability issues and possible solutions, , although the different chapters are often of unequal level of detail.

The 4 layers identification seems efficient to distinguish the actions to take; at the same time it is not always very clear who should take care of which (legal: data stewards? semantics: infra and communities? technical: infra and EOSC?)

Positive aspects:

- Very positive that Semantic Artefacts are mentioned as first-class citizens, but not all problems that relate to their creation and maintenance are mentioned (such as scalability).
- It brings a list of problems that have been gathered from different actors of EOSC - always interesting to know and think about.
- Positive that publications are mentioned as a research output; this is essential for SSH and especially for the Humanities.

Negative aspects:

- Some drawbacks include: the document is in high abstraction, without any guidelines how to design and specify a FAIR digital objects, neither discussions about the feasibility for implementation. It relies on a common referent model/metadata standard to describe and interpret EOSC data and services, however, it would be very hard to create such a model/standard that addresses the requirements of various EOSC stakeholders who come from different domains of science and beyond. It would be even harder to reach the consensus for all EOSC service providers and users to agree on such a model/standard, not even mention the arduous efforts needed for enforcement of adoption and implementation. The concern is although it is a sound approach but infeasible for realisation.
- There's a general lack of detail. Since some interviews have been performed, one could expect a more elaborated approach to various metadata schemes applied by different "research communities". The document notes multiple standards but doesn't mention possible mappings. Perhaps a way to go here would be to work on minimal requirements for metadata interoperability.
- The document lacks solutions in most chapters. And when there are solutions provided, there is no explanation on how to achieve them (e.g. share concepts' definition in EOSC). A more detailed "hands-on" description was expected. Looks like a collection of best practices, not really a (technical) "framework" description.

Are the concepts clear or do some aspects need further clarification?

- Organizational interoperability is focused on as the need for alignment of public organizations dealing with standards and implementation, but for research 'social' interoperability seems more important, which is the need to organize communities to contribute to for instance Semantic Artefacts.
- Very good that Digital Objects (DO) are mentioned, it's very welcome that not everything is a service. This terminology is more close to what the SSH community is familiar with. Nevertheless, it should include a clear interpretation of the concept of FAIR Digital Object in the context of the proposed EOSC Interoperability Framework. Does EOSC implies/equals FAIRfulness? If not, what is missing/constrained?
- Legal interoperability is not defined. If this is part of the overall interoperability model, it is difficult to evaluate the completeness of the current proposal.

- Maybe this is EOSC lingo but we find difficulties in grasping the difference between research communities and disciplines in the context of interoperability. The language should be more precise with regards to the target groups and their problems, or, at least, link to another document with a more precise definition of the target communities.
- The “Scientific workflows” described in 2.1 are not clear, and haven’t found a correct explanation of it in 3.1.

Are the minimum requirements and recommendations appropriate?

- We repeat here that they’re quite broadly sketched, thus the document is too generic. It is understandable that research communities would use different metadata formats, but there are no minimum requirements linked to them explained in the document.
- Recommendations exist, yes and seem appropriate.

Is it clear who is responsible for what and how this should be followed?

- Not always: what is concretely intended by « PID infrastructure » and who is responsible for it? Who will take in charge the definition of a “simple vocabulary ... allowing discovery over existing federated research data and metadata (extension of DCAT-AP, DDI 4 Core, or DataCite core schema)”, knowing that, for instance, the SSHOC project is working on the same topic?

As a service provider, could you conform with / implement the framework?

- At least in some aspects, but the framework at this stage is still too generic.
- It’s not clear what role to play in each case (e.g. on legal interoperability between various sources of various countries: should we provide a policy, recommendations, share information?).
- A “use cases” companion document should be produced. To ensure interoperability you cannot remain at such a high-level description of the technical solutions to apply. There should be at least a reference implementation of this framework (maybe it’s just me that got confused with the term “framework”...).

Is the model for FAIR Digital Objects sound?

- TRIPLE’s partners perspectives are different to this respect: CLARIN ones welcome the DFOs model, while for other it is a bit confusing that it’s not totally similar to the model developed within GOFAIR; various FDOs flavors are possible, however the GOFAIR model only envisioned ID+resolvable link+type, while the model here described, with all semantics, is much more complex and therefore difficult to reach.
- In general, within the various SSH communities there are different levels of readiness regarding FDOs (sometimes far from using more than a generic metadata standard, to not talk about semantics).

What other feedback and comments would you like to offer?

- We'd suggest to work on a reference implementation and/or a use cases document for this framework, at least for the technical and semantic interoperability aspects, with a specific and detailed description of the technical solutions to apply.
- It would be great to see slightly more information about the interviews with stakeholders - the disciplines are listed but countries, career stage and the number of interviews would be helpful in understanding who guided the recommendations.
- Short comment: in the doc you point to "interview of CESSDA CTO", is this a public document? If yes, a link would be helpful - if not, better to delete the footnote.