////////////////////////////////////////////////////////////////////////////////////////////////////////////////////

**ONTWERPVERSLAG**

FOSB WG Metadata & Standardisation

////////////////////////////////////////////////////////////////////////////////////////////////////////////////////

Datum: 25/05/2021

* Update Roadmap
* State of affairs TF Metrics
* State of affairs TF DMP
* Questions on implementation of FOSB metadata model

**Aanwezig**

Pascale Dengis Departement EWI

Jochen Deferme Departement EWI

Sadia Vancauwenbergh ECOOM-Hasselt

Evy Neyens ECOOM-Hasselt

Jolien Berckmans UHasselt

Joke Meeus FWO

Helga Deschrijver IMEC

Katrien Desmet VLOHRA

Rinaldo Beck VIB

Alexander Botzki VIB

Niek Van Wettere VUB

Rudi Baccarne UAntwerpen

Roxanne Wyns KU Leuven

Sara Decoster KU Leuven

Dries Moreels UGent

**Verslag Vergadering 25/05/2021**

/////////////////////////////////////////////////////////

Update Roadmap

The roadmap Open Science Policy in Flanders: roadmap, KPIs, and multi-year plan (VR 2020 0412 MED.0379/1) needs to be adapted according to open agenda items and their timeline by June 15, 2021. ECOOM-Hasselt presented an outline of the open roadmap items and the proposed timeline.

From September 2021, the WG M&S will start with the development of disciplinary metadata models for research data. It was suggested to recommend ‘best practices’ per research discipline.

Remarks of the WG members

* Does this overlap with the TF Use cases? They focus on domain-specific repositories.
* Too early to do this theoretical exercise: risk of losing work.
* Postpone target to 2022?
* Which data repositories would we recommend to specific users and which metadata models do they use?
* Suggestion to take one test case as a midterm solution: keep it specific.
* Our minimum model is more elaborate than those of generic repositories such as Zenodo.
* Have a look at EOSC and OpenAIRE.
* Follow existing standards as much as possible.
* Link to mission/vision text.

State of affairs TF Metrics

ECOOM-Hasselt outlined the current state of affairs regarding the TF Metrics. This TF has been tasked to develop a FAIR data label and an Open data label to be able to monitor these KPIs via FRIS in an automated manner. The FAIR data label refers to assigning and completing a FAIR Data label to each relevant dataset that reflects the degree of FAIRness. Currently, three possible pathways have been discussed in the TF Metrics:

1. Automated F-UJI tool: a third party open source tool that is part of the FAIRsFAIR project.   
   F-UJI is an automated assessment of datasets in repositories. An API applies metrics to published metadata of datasets in repositories. Easy to use via a DOI/URL, or DOI/URL + provider info OAI-PMH- + DataCite integration. https://www.f-uji.net/
2. FAIR self-assessment tool for researchers: to be used to stimulate FAIR awareness among researchers to increase the FAIRness of datasets indirectly. Researchers can be asked to complete such an evaluation before depositing their data in a repository. An example of such a tool is Satifyd by DANS. It is explicitly not the purpose to use these scores for KPI measurements.
3. Minimal measurement of objective indicators of the F & A of FAIR using the metadata fields of the FOSB metadata model. e.g. "Is there a DOI present? Yes/No". Objective FAIR elements can be measured automatically in FRIS based on the fields of the metadata model for research datasets.

The TF Metrics will prepare an advisory note describing the options with their respective advantages and disadvantages and will put forward a preferred option.

Remarks of the WG members

F-UJI: "In addition to repository and services requirements, automated testing depends on clear machine assessable criteria. Some aspects (rich, plurality, accurate, relevant) specified in FAIR principles still require human mediation and interpretation." The limits of F-UJI must be clearly stated.

State of affairs TF DMP

ECOOM-Hasselt outlined the current state of affairs regarding the TF DMP. It was agreed to work in two steps: First, construct a list of core (sub)topics that should be part of a minimal Flemish DMP based on the FWO DMP. Second, outline specific questions, their wording, answering options & guidance. The RDM stewards of the TF DMP provided feedback on the current FWO DMP from a user perspective. These insights will be considered when drafting the Flemish minimal DMP standard. Next steps are: create a first draft of the Flemish DMP (Q3), define a process flow for collaboration (Q3 2021), and propose a standard for machine-actionable DMPs in the future (2022).

Remarks of the WG members

* Is it possible to digitally share DMPs?
* In DMPOnline there currently is no milestone version, hence you don’t know which version to send.
* At the European level DMPs are deposited via the application portal, not digitally.

Questions on implementation of FOSB metadata model

The Department EWI issued some open questions regarding the FOSB metadata model. These questions mostly referred to technical issues, CERIF and the connection to OpenAIRE.

Version 2020-11-20\_FOSB\_metadatamodel\_V1.7ter

Concretisations:

* Column with multiplicity (1-n relations)
* Identifiers:
  + Multiplicity: max 1 DOI; other identifier types?
  + URL:
    - Specific enough to identify a dataset?
    - Can the identifier type in FRIS “Institutional Repository URL” be used for URL?
    - Should URL be excluded from the aliasing of identifiers in FRIS?
  + Main identifier: will be the first of the identifiers in the following order:

DOI>Handle>ARK>PURL>(Institutional Repository) URL>URN

Can the same dataset have two DOIs? For example, if two different researchers deposit the same dataset on different platforms?

Remarks of the WG Members

* It was never agreed before that FRIS will deliver metadata to OpenAIRE. Currently, the institutions do this themselves.
* The Flemish government has requested that FRIS acts as a hub for metadata on research data towards EOSC. The institutions are also free to connect directly to OpenAIRE/EOSC.
* Requirement to deliver minimum 1 PID: this will exclude repositories that don’t work with PIDS
* In re3data there are still many repositories that don’t work with PIDS:

[*https://www.re3data.org/search?query=&pidSystems%5B%5D=none*](https://www.re3data.org/search?query=&pidSystems%5B%5D=none)

* It is also possible to have resolvable URLs
* But then you lose the original semantics of the data, for delivery to EOSC
* What if different RPOs deliver different identifiers for the same dataset?
* Merge records?
* How to visualize in FRIS? show all identifiers?
* Aliasing as main identifier?

|  |
| --- |
| 1. Next meeting 31/05/21 we will proceed with discussing the open questions regarding the metadata model. 2. A document with comments to adapt the roadmap will be shared. The WG members will provide their feedback by 10/06/21 (to be confirmed). |