

TC-IM 1449 *Research data management and the European Open Science Cloud (EOSC)*

Date: 13 June 2019; 10 a.m. – 3 p.m.

Venue: Chadwick Building, Room 311, Physikalisch-Technische Bundesanstalt, Braunschweig, Germany

Reception and introduction of the participants

Anna Cypionka (head of the Presidential Staff) welcomed the participants to the workshop and introduced the rest of the PTB staff: Sascha Eichstädt (leader of the working group "Digitisation" and initiator of the TC-IM projects 1448 and 1449), Michael Brinkschröder (project manager) and Giacomo Lanza (coordinator for research data management).

The other participants introduced themselves

- Alistair Forbes (NPL)
- Clifford Brown (NPL)
- Federico Grasso Toro (METAS)
- Peter Friis Østergaard (DTI)
- Claudio Francese (INRiM)
- Thomas Wiedenhöfer (PTB)
- Wolfgang Schmid (EURAMET)

Absent: Marina Romanchikova (NPL), Sari Saxholm (VTT), Per Olof Hedekvist (RISE).

The expectations of the participants for the workshop were collected on a flipchart and used as a basis for the further discussion.

Aims of the project and previous work

Giacomo Lanza gave an overview on the motivation for the project, the defined aims of the project and the tasks to be performed within the workshop. All participants reported the current and scheduled tasks of their institutions with regard to the project aims.

Standardisation (PTB – Lanza)

Starting with the project *SmartCom*, PTB is developing metadata schemas for the transmission of metrologic relevant information (numerical values with quantity name, unit and uncertainty). The atomic schema is to be published with the name **D-SI** and can be expanded to cover different application cases, such as the report of calibration values in a *digital calibration certificate (D-CC)* and the ingestion of research data into a data repository. Some machine-readable controlled vocabularies for physical quantities and units are also in development, which can be used for an unambiguous compilation of the metadata. This approach is also being proposed within the framework of the NFDI (German National Research Data Infrastructure).

Networking and infrastructure development (PTB – Lanza)

The PTB supports the development of RDMO, an on-line tool for the composition of data

management plans and for the organisation of data workflows within a project. Its local instance has been used by project leaders and their external partners within the last EMPIR call. There are attempts to make the tool available to all EURAMET members. The PTB is also active within the RDMO community and taking part in the development of the platform.

Information value chain (NPL – Forbes)

NPL understands its role as an institution which conveys information from the real world and delivers it to the decision makers for the increase of efficiency and the development of new products and services.

Therefore, a data workflow should pass through qualified chain elements and its output should be (FAIR) trusted reference data, which are available for primary use as well as secondary reuse. The value of the information is guaranteed by a strict QM framework.

Against this background, NPL is engaged in the following projects:

- Mathmet JNP, WP2 - QM: risk assessment, glossary development;
- Digital calibration certificate;
- Software sustainability initiative.

Metadata standards (NPL – Forbes on behalf of Romanchikova and Brown)

Nowadays reuse of data is difficult due to inconsistency in the data formats and descriptions (black-box), so that data mining is often replaced by data archaeology. This way of going is far from being FAIR!

Because of this NPL is active in:

- creating a minimum viable metadata set, at first in the context of imaging data for life sciences;
- creating a DTF graph ontology for material testing facilities.

MIMA concept (METAS – Grasso Toro)

A model is presented to classify data availability:

1. Level "reader": access is granted only to results (in form of tables and figures)
2. Level "reviewer": access is granted to the input data and the generating algorithm.
3. Level "peer researcher": access is granted to all data and the context.

The model will be implemented at METAS for the internal RDM.

The metadata will be included in XMP format (an ISO standard for metadata as machine-readable text, developed by Adobe, which is included automatically in PDF/A3 files).

European open science cloud (Eichstädt)

The European open science cloud was launched in November 2018 and it is conceived as a network of resources for research data management. It includes for example: computation facilities, tools for data anonymisation, training & education, modules for metadata and PIDs, etc., which can be accessed by browsing on its [website](#).

The network is growing very fast. Unfortunately, there is no action from the side of the metrology institutes and the same EURAMET is not aware of its potential role in the EOSC.

Suggested readings:

- **"Turning FAIR data into reality"** – Final Report and Action Plan from the European Commission Expert Group on FAIR Data, 2018. <https://doi.org/10.2777/1524>
- **"Prompting an EOSC in practice"** – Final report and recommendations of the Commission 2nd High Level Expert Group on the European Open Science Cloud (EOSC), 2018. <https://doi.org/10.2777/112658>

Discussion on the project aims and definition of work packages

During the discussion, the wishes of the participants and some new ideas were grouped and structured. As a result seven work packages were defined.

For each work package a coordinator, a timeline and some milestones were defined. An overview is reported in the following table.

Expected output	Important milestones	Target date	Work package coordinator	Necessary resources
Increase impact of metrology on EOSC	Proposal to BoD for involvement in EOSC	Sep 2019 (BoD meeting)	PTB (Eichstädt)	Materials from member NMIs
	Case submission to EOSC			
	Involvement in "training & education"			
	Link at EURAMET and IMRR sites to EOSC			Materials from EURAMET website
Advice for EURAMET about how to fulfill future EU requirements	Q&A to BoD; draft to workshop participants	Sep 2019 (BoD meeting)	METAS (Grasso Toro) + PTB (Lanza)	
	Baseline survey (users + MSU?); draft	Feb 2020 (TC chairs meeting)		
	Final draft	May 2020 (GA)		
Motivation of the users	Survey under EMPIR coordinators	Dec 2019	NPL	Tools survey (e.g. Zenodo, ORCID)
	Monitoring strategy for data reuse; pilot analysis	2 nd half 2020		
	Training courses for EMPIR projects	1 st quarter 2020		

Expected output	Important milestones	Target date	Work package coordinator	Necessary resources
Case studies + application design guidelines	"getting started" guide; core tools; guidance for MD	2 nd half 2020	NPL	
Development of RDM structures	EURAMET guide on RDM; minimum requirements	2021 (committee draft)	PTB (Lanza)	
Cooperation with stakeholders / scientific community	White paper for data repositories	May 2020	PTB (Lanza)	D-SI
	discussion with big data stakeholders	??? (NPL big data workshop)	NPL	
	presence at CIPM's data WG	Oct 2019	PTB	
Formulation of a vision for metrology in RDM	ideas collection	Nov 2019 (next workshop)	METAS + ideas from everybody	"masters of reproducibility and traceability"
	Vision as first chapter of white paper	May 2020	everybody	

Conclusion

Summary

The workshop was successful in defining concrete work packages and a concrete timeline. The project description will be amended correspondingly (Lanza).

Communication channels

For the exchange of files and material the BSCW server of PTB will be used. Access permissions will be granted to all workshop participants and future partners (Eichstädt).

Scheduling of the next meeting

Next meeting will take place in November 2019 at METAS, possibly on the day after the next meeting of TC-IM 1448 (Grasso Toro).

Tanasko Tasic and Miruna Dobre are the partners in EURAMET who are to be addressed concerning the funding for the organisation of workshops and the travel to workshops.