EOSC-A Task Force: "PID Policy and Implementation"

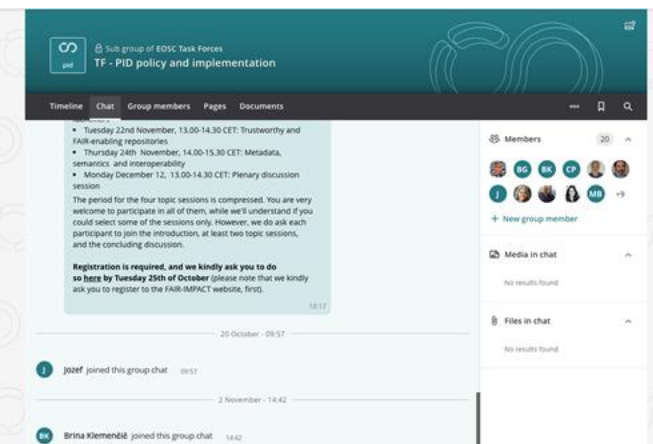**Taking EOSC into the future**

Themis Zamani - GRNET

Tibor Kalman - GWDG

Let's Introduce the TF

!!! Promotion !!!

EOSC PID Policy document

EOSC PID Architecture document

TF Charter (updated 2023)

SRIA

23 Members

12 Service providing organisations

1 Research funding organisation

9 Research performing organisations

1 Research community

A balanced group of PID users , PID experts (including technology and domain experts), and PID Providers.

Task Force Goals

Goals & Core Activities

1. Provide input based on identified gaps in the PID ecosystem (i.e SRIA, MAR);
2. Monitor and provide community feedback
3. Make recommendations to the EC via the EOSC Association for the integration of PID services in the EOSC ecosystem, its implementation and test.

The "core activities"

- Liaison, monitoring and collaboration with the relative initiatives
- Identify, describe emerging and standardised identifiers for resource types
- Global PID resolution
- Review efforts to develop definitions for the most common data formats or building blocks.
- EOSC PID Graph
- Criteria against which PID will be certified eventually to be implemented by tools supporting the certification of PID infrastructure against the EOSC PID Policy.
- Best practices and PID use cases that exemplify FAIR data management

External Collaborations

PID-related projects, communities and providers technology- and domain experts

Other Task Forces

Interact mainly with data, metadata, data lifecycle and data quality naturally have many touchpoints.

- TF FAIR metrics and Data Quality
- TF Semantic Interoperability.
- TF Technical Interoperability of Data and Services.
- TF Long-term Data Preservation.
- TF Rules of Participation (RoP).

PID TF: Work done until today

Group A: Mapping current PID-related activities

- A subset of PID types that could be the primary focus for EOSC
- A list of PID-related projects and providers

Group B: Collecting community-specific use cases and perspectives on the EOSC PID architecture and the EOSC PID policy

- Survey to collect priorities for addressing gaps, as well as use cases
- Validated the results

PID TF: Focus Group C

Harmonization of PID-related Vocabularies

The lack of clearly defined vocabularies covering the PID related ecosystem is the main reason for starting this group. The main areas that the focus group will investigate are:

- Identify work already done
- Analyze the vocabularies needed
- Try to identify the owner of these vocabularies in EOSC (manage/curate).
- Identify who will be the hosting body and the governance body of these vocabularies in EOSC so that they can be used by all related projects and services.

Technology Independence	A-FDO	FDO Requirement Specifiction		FDOs and their key components need to support technology independence, allowing implementations using different technologies.
Versioning	A-FDO	FDO Granularity		There are different policies between communities on how to indicate versioning – some use the PID string, some use the PID record and some use extended metadata for these purposes. The choice widely depends on efficiency criteria and no general recommendation can be made except that data providers need to make clear statements on how versioning is handled.
Attribute	X-EOSC	EOSC PID Architecture	Kernel Attribute, Kernel Information Type	A value that describes a feature of an object or its representation, as part of PID Kernel Information or other metadata.
Digital Entity	X-EOSC	EOSC PID Architecture		A digital Entity denotes any sort of bit sequence that is being stored or transmitted without being registered to enable sharing.
Digital Object	X-EOSC	EOSC PID Architecture		A Digital Object has a bit sequence that can be stored in multiple repositories and is associated with a Persistent Identifier (PID) and metadata.
Digital Object Identifier	X-EOSC	EOSC PID Architecture		A digital object identifier (DOI) is a persistent identifier based on Handle used to identify objects uniquely, standardized by the International Organization for Standardization (ISO).
Granularity	X-EOSC	EOSC PID Architecture		The varying levels of hierachy or constituent parts that may form data or other research outputs. For example, the differing levels of granularity of a research publication, going from a whole Journal issue, the level of detail in a large scientific database. to its constituent articles, to the article constituent sections or figures, the levels in a complex scientific collection or the level of detail in a large scientific database.
				A Handle is a globally resolvable, unique and persistent PID which is defined by RFCs 3650, 3651 and 3652 of the

Expected output(s)

- The need/requirements of the projects (landscape analysis)
- (Linked) vocabularies
- Update the PID Policy with the vocabularies so as to be consistent with the work done

PID TF: Focus Group D

PID Policy in Examples

One of the main questions of the users reading the PID Policy is if there are any examples for the different sections of the document. The Policy is describing the steps the different types of users should follow but what we are looking for is practical implementation and policy examples. This new focus group is about policy and implementation examples regarding PIDs; This FG will collect and describe the activities going on.

Expected output(s)

- Examples would be collected and described, to make something pragmatic, high-level on top of the complex PID Policy. Provide examples on how to better understand the PID Policy
- Collect examples from different areas and domains