



Contribution ID: 117

Type: POSTER&PITCH

Harmonizing the Implementations of PIDs across Repositories

Monday 4 November 2024 16:00 (1 hour)

In our increasingly digital and interconnected world, the integration of Persistent Identifiers (PIDs) in metadata are essential for machine-readable and -understandable metadata as also described in the FAIR Guiding Principles for research data management. PIDs provide unique, permanent and machine-readable references to various types of digital objects, including publications, datasets, scientific software, individuals, organizations, samples that together represent the broad range of research outcomes.

Within the AK Metadata-PIDs working group (a joint initiative between the HMC Hub Earth and Environment and the Helmholtz DataHub Earth and Environment), we discussed several PID systems and reached a consensus on recommending specific systems for different purposes: "ORCID" for identifying individuals, "ROR" for organizations, and the "PIDINST" PID for instruments.

For the full integration and sustainability of individual PIDs, different players are involved: these range from the consortium developing a PID to the research institution supporting (and ideally enforcing) its implementation for their employees (e.g., ORCID) to the individual research infrastructures and repositories where the required information is collected and the provision of the PID within the metadata is ensured.

Our working group has focused on supporting the ongoing PID implementation in research infrastructures by conserving existing, well-established PID implementations (best practices) and promoting their integration in future systems. We further aim to provide support and guidance for new implementations.

We observe differences in the metadata content, even between two DOI-registering data repositories that use the DataCite Schema. What is the reason for this? Would we need specific mapping tables to harmonise cross-repository metadata? Could a stronger guidance and definition of metadata properties (note: metadata schemas are intended to be very generic) achieve the envisioned higher grade of harmonization?

This poster highlights and discusses the differences in PID implementations across various exchange formats, such as DataCite, ISO 19115/19139, and schema.org. The goal is to encourage and support scientists and IT specialists who maintain and develop research infrastructure to foster metadata harmonization for ensuring interoperable metadata exchange across repositories.

Please specify "other"

HMC staff

In addition, please add 3 to 5 keywords.

PID, DataCite, ISO, Schema, ORCID, ROR

Please specify "other"

For whom will your contribution be of most interest?

Scientists and technicians who maintain and operate research infrastructure for data generation

Please assign yourself (presenting author) to one of the following groups.

other (please specify)

Primary authors: PÖRSCH, Andrea (HMC Hub EE at GFZ); SÖDING, Emanuel (GEOMAR)

Co-authors: KOTTMEIER, Dorothee (HMC E&E @PANGAEA/AWI); MALINOVSKII, Stanislav (HMC); LORENZ, Sören; RAZEGHI, Yousef

Presenter: PÖRSCH, Andrea (HMC Hub EE at GFZ)

Session Classification: Poster Session C

Track Classification: Connecting research data: 6. Interoperable semantics at domain and application level