PID INFORMATION TYPES

CONTACT DETAILS:
RDA WG: PID INFORMATION TYPES WORKING GROUP
NAME: TIMOTHY DILAURO & TOBIAS WEIGEL
ORGANISATION: DATA CONSERVANCY / JOHN HOPKINS UNIVERSITY **&** GERMAN CLIMATE COMPUTING CENTER (DKRZ) / UNIVERSITY OF HAMBURG
EMAIL: TIMMO [AT] JHU.EDU, WEIGEL [AT]DKRZ.DE

Objectives:

In complex data domains, unique and persistent identifiers (PIDs) associated with specific information are the core of proper data management and access. They can be used to give every data object (including collection objects) an identity that enables referring to the data resources and metadata and, additionally, to prove integrity, authenticity and other attributes. But this requires a PID to be uniquely associated with specific types of information, and those types and their association with PIDs must be well managed. Automated data management across disciplines and repositories can highly benefit from standardized types.

The PID Information Types WG will provide a first solution for types. Based on practical use cases taken from the interested communities and infrastructures, the group will derive an initial set of types and propose an overarching framework which can be used to incorporate new types in the future. To facilitate practical adoption, the group will also design and prototypically implement an API to interact with types and types information. The API and other group activities and deliverables focus on cross-cutting conern and abstract from particular PID infrastructures or systems.

On-going activities and first Results:

The WG has collected practical use cases from the various communities already working with PIDs or planning to do so in the future. The use cases come in diverse formats and stages of maturity, and the current work of the WG focuses on providing a more streamlined technical form of the use cases that can be used to continue with proposing types and designing the API. Evidently, not all use cases and not their full scope can and should be enabled by the WG; instead, the goal is to identify recurring themes in the use cases and to extract building blocks that can be supported by a types API.

A first proposal for a list of types has been drafted and will be put to discussion at the second plenary. To sort out the role of infrastructures, client users and other service providers, a high-level architecture design for the API implementation has been sketched. The architecture draft also hints at a long-term adoption strategy which places the API closer to the infrastructure provider side.

The second RDA plenary will end with a first set of use cases and types, based on subthemes identified in the use cases. Discussions on use cases and potential types will continue, but in the meantime the group activities will move on towards developing the types framework and the API. One particular discussion point that may be approached in coordination with the DFT WG is how to define types and how to separate them from other concepts such as metadata using for example practical metaphors.

URL: https://rd-alliance.org/working-groups/pid-information-types-wg.html