Contribution ID: **105** Type: **POSTER&PITCH**

The Helmholtz Digitization Ontology: Harmonized semantics for the Helmholtz digital ecosystem

Monday 4 November 2024 14:00 (1 hour)

Abstract
Research in the Helmholtz Association undergoes continuous digitization. The heterogeneity of scientific contexts within Helmholtz leads to ambiguity and conflicts regarding metadata semantics. To ensure semantic interoperability of this decentralized data ecosystem, metadata should be aligned and harmonized with European and global initiatives to ensure an open and interoperable flow of data and information. Thus, HMC provides the Helmholtz Digitization Ontology (HDO), a mid-level ontology that contains concepts and relationships representing digital assets and processes that exist in the Helmholtz digital ecosystem. HDO is developed by contributors from all Helmholtz research fields. The main goal for developing HDO is to serve as a harmonized and machine-actionable institutional reference to represent digital assets and procedures pertinent to their handling and maintenance within Helmholtz.
HDO is aligned to practices and conventions of the Open Biological and Biomedical Ontologies (OBO): we create coherent and precise definitions in the OBO recommended genus-differentia form (i.e. for each term we define a Genus as well as its differentia). Class labels and definitions are developed bilingually in both English and German. Additionally, classes have further information, including synonymy, singular, plural, gloss, comments as well as micro-credits of contributions. To ensure the sustainable development of HDO, we implemented it based on the Ontology Development Kit (ODK).
HDO development is carried out in three phases: 1) Initialization phase: an internal GitLab repository was created to gather a set of core classes and their definitions in per-term YAML files, 2) Implementation phase: YAML files were converted and merged into one OWL file. For this, keys of the template were mapped onto existing and imported annotation properties, and 3) Adoption and Adaption: a phase of continuous iterative development in which the ontology will be used in use cases across the different Helmholtz research fields. This will test the developed ontology against use case-specific requirements and allow further adoption based on iterative exchange. One example we are currently pursuing is the semantic representation of FAIR digital objects.
The current development and the first release can be found in our public git repository [1]. The ontology is made accessible via a persistent identifier [2] and terms are derefrenced via their PIDs. An HTML documentation of HDO is available online [3].

References
[1] https://codebase.helmholtz.cloud/hmc/hmc-public/hob/hdo
[2] https://purls.helmholtz-metadaten.de/hob/hdo.owl
[3] https://purls.helmholtz-metadaten.de/hob/HDO_00000000

Acknowledgements
This work was supported by (1) the Helmholtz Metadata Collaboration (HMC), an incubator-platform of the Helmholtz Association within the framework of the Information and Data Science strategic initiative

Please specify ”other”

In addition, please add 3 to 5 keywords.

Ontology
Semantics
Metadata Management
OWL
FAIR

Please specify ”other”

For whom will your contribution be of most interest?

Data professionals who provide and maintain data infrastructure

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

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

Primary authors: FATHALLA, Said; HOFMANN, Volker

Co-authors: GUENTHER, Gerrit (Helmholtz-Zentrum Berlin); STEINMEIER, Leon (Helmholtz Institute Freiberg); LEMSTER, Christine (Geomar); KOTTMEIER, Dorothee (HMC E&E @PANGAEA/AWI); SIVAPATHAM, Lakxmi; SANDFELD, Stefan

Presenter: FATHALLA, Said

Session Classification: Poster Session A

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