



Contribution ID: 47

Type: **Talk (15min + 5min)**

How to “unHIDE” and improve the metadata landscape of research software in Helmholtz.

Thursday, March 7, 2024 11:40 AM (20 minutes)

Research across the Helmholtz Association is based on inter- and multidisciplinary collaborations across its 18 centers and beyond. However, the wealth of Helmholtz’s (meta)data and digital assets are stored in a distributed and incoherent manner, with varying quality.

To address this challenge, the Helmholtz Metadata Collaboration (HMC) launched the unified Helmholtz Information and Data Exchange (unHIDE) project in 2022. UnHIDE aggregates metadata harvested from Helmholtz infrastructure in the Helmholtz Knowledge Graph (Helmholtz KG). This serves as a lightweight and sustainable interoperability layer to interlink data infrastructures and increase visibility and access to the Helmholtz Association’s (meta)data and information assets

Version 1.0.0 of the Helmholtz KG was released in October 2023. This includes a comprehensive web front end for manual search of resources [1], a stable and documented [2] backend with a tested data ingestion and integration pipeline, and machine accessible endpoints [3].

In this talk we present an overview of the Helmholtz metadata ecosystem, we describe the semantic and technological architecture of the Helmholtz KG and how it integrates metadata from heterogeneous sources to improve visibility and findability. We will show how code and research software is scattered throughout different platforms (such as institutional gitlab instances), how its metadata is lacking connection to other (research) publications and that only a minority is formally published in central indexes [4]. We will show and discuss some results of our efforts to integrate and improve software metadata in Helmholtz as well as future ways how the Helmholtz KG is envisioned to harmonize and improve quality of metadata at the source: in the respective infrastructures.

[1] <https://search.unhide.helmholtz-metadaten.de/>

[2] <https://docs.unhide.helmholtz-metadaten.de/>

[3] <https://sparql.unhide.helmholtz-metadaten.de/>

[4] e.g. <https://helmholtz.software/>

Slot length

Primary author: BRÖDER, Jens (Forschungszentrum Jülich GmbH (IAS-9))

Co-authors: D’MELLO, Fiona (Forschungszentrum Jülich); PREUSS, Gabriel (HZB); FATHALLA, Said; BUTTIGIEG, Pier (GEOMAR Helmholtz Centre for Ocean Research); MANNIX, Oonagh (HMC matter/HZB); HOFMANN, Volker; SANDFELD, Stefan (Forschungszentrum Jülich)

Presenter: BRÖDER, Jens (Forschungszentrum Jülich GmbH (IAS-9))

Session Classification: Metadata for Research Software

Track Classification: Research Software: Metadata for Research Software