A Pan-European and Cross-Discipline Metadata Portal

Heinrich Widmann, Hannes Thiemann, and Michael Lautenschlager
The German Climate Computing Center (DKRZ), Hamburg, Germany

In recent years, significant investments have been made to create a pan-European e-infrastructure supporting multiple and diverse research communities. This led to the establishment of the community-driven European Data Infrastructure (EUDAT) project that implements services to tackle the specific challenges of international and interdisciplinary research data management.

The EUDAT metadata service B2FIND plays a central role in this context as a repository and a search portal for the diverse metadata collected from heterogeneous sources. For this we built up a comprehensive joint metadata catalogue and an open data portal and offer support for new communities interested in publishing their data within EUDAT.

The implemented metadata ingestion workflow consists in three steps. First the metadata records - provided either by various research communities or via other EUDAT services - are harvested. Afterwards the raw metadata records are converted and mapped to unified key-value dictionaries. The semantic mapping of the non-uniform community specific metadata to homogenous structured datasets is hereby the most subtle and challenging task. Finally the mapped records are uploaded as datasets to the catalogue and displayed in the portal.

The homogenisation of the different community specific data models and vocabularies enables not only the unique presentation of these datasets as tables of field-value pairs but also the faceted, spatial and temporal search in the B2FIND metadata portal. Furthermore the service provides transparent access to the scientific data objects through the given references in the metadata.

We present here the functionality and the features of the B2FIND service and give an outlook of further developments.