NFDI4Cat Web Application for User-Friendly Generation and Browsing of Catalysis Metadata within a Semantically Rich RDF Framework

Authors:
Taras Petrenko[1]*, Yuliia Dikova[1], Nadiia Huskova[1], Volodymyr Kushnarenko[1],Preston Rodrigues[1], and Thomas Bönisch[1]
*taras.petrenko@hlrs.de
[1]High Performance Computing Center Stuttgart (HLRS), Stuttgart, Germany

Abstract:
The key to the success of the NFDI initiative lies in the creation of a common data instructure, the development of standardized ontologies, efficient data integration techniques, user-friendly generation of RDF data, and the use of advanced tools and technologies, including AI. Generating RDF data based on a common ontology for a semantic metadata framework, particularly in a specialized field like catalysis, involves several significant challenges::

- Representation of complex concepts and relationships for the domain of catalysis
- Lack of standardized ontologies which hinders interoperability of data
- Heterogeneous data sources
- Ensuring consistency, accuracy and completeness of data
- Generation and collection of RDF data is too comlpicated for the most researchers to be done manually

The latter aspect poses a major barrier to the wide adoption of the semantic methodologies in the research community. In addition, the exploration of RDF data by means of conventional SPARQL queries is also not very user-friendly. To address the problems systematically, we have developed the following approach. The reference Metadata4Cat ontology is based on the middle-level Metadata4Ing ontology, which provides a reasonable framework for the semantic description of research data. Within the Metadata4Cat ontology, the required metada fields are indicated by appropriate restriction classes. Such constructs are then used by the questionnaire program TRIQ, developed in our group, for reasoning and querying the corresponding metadata via user-friendly dynamic web forms. This leads to a chain-like instantiation of various resources representing a data context which is required to make data FAIR. The generated RDF metadata are immediately available in a browseable HTML form. In addition, the TRIQ program will be coupled with a triplestore, PID handle service, and the terminology services of NFDI. The proposed methodology can be easily extended to other domains of knowledge. We believe that our application can play an important role in the broad adoption of the semantic data technologies within NFDI.

Keywords: NFDI4Cat, TRIQ, generation of RDF data, semantically rich metadata metadata enrichment, Metadata4Cat ontology, Metadata4Ing, catalysis