A terminology service via TS4NFDI for BERD@NFDI knowledge graph infrastructure

Authors:
Renat Shigapov[1]*, Oliver Koepler[2], and Roman Baum[3]
*Lead presenter
[1] renat.shigapov@uni-mannheim.de, University of Mannheim, Mannheim, Germany
[2] TIB - Leibniz Information Centre for Science and Technology, Germany, Hannover
[3] ZB MED - Information Centre for Life Sciences, Cologne, Germany

Abstract:
The use of standardised terminologies facilitates interoperability and data integration, making research data FAIR (Findable, Accessible, Interoperable, Reusable). Terminologies are often made available via terminology services through graphical user interfaces (GUIs) and APIs. Several domain-specific terminology services already exist within the NFDI consortia such as SemLookP [1, 2], BiodivPortal [3], and NFDI4Chem Terminology Service [4, 5]. In the social sciences, services such as the GESIS Controlled Vocabulary Service [6] which is based on SKOSMOS and the STW Thesaurus for Economics [7] hosted by the ZBW are notable examples. However, a common terminology service for BERD@NFDI does not yet exist. BERD@NFDI is developing a knowledge graph infrastructure (KGI) for German company data, transforming analogue books into multiple knowledge graphs through optical character recognition, automatic structuring, and semantification [8, 9, 10]. These knowledge graphs are a part of distributed KGI4NFDI [11, 12]. A common terminology service for BERD@NFDI would significantly enhance the efficiency of semantification and linking processes.

One of the aims of BERD@NFDI is to ensure that data management adheres to the highest standards of integrity and accessibility. To achieve this, we are setting up a terminology service for BERD@NFDI using the Terminology Service Suite (TSS) [13]. The widgets of the TSS could handle multiple backends such as SKOSMOS, OLS and OntoPortal. This ensures that the newly developed terminology service for the BERD community meets the standards of other terminology services, promoting interoperability with services such as the GESIS Controlled Vocabulary Service, which utilises SKOSMOS as a backend system. The new terminology service can be seamlessly integrated into the BERD knowledge graph infrastructure. The GUI of the terminology service, constructed using TSS widgets, allows for straightforward embedding of terminology service data into other applications or services within the Business, Economic, and Related Data domains. Additionally, terminologies can be accessed via the provided APIs, ensuring wide usability and integration.

By implementing a standardised terminology service, BERD@NFDI not only enhances the semantification and linking of knowledge graphs but also ensures that these processes are aligned with the broader NFDI goals. The terminology service supports the overall mission of the NFDI by promoting data interoperability, thereby advancing research capabilities across various domains.

Keywords: terminology service, knowledge graph, FAIR research data

[1] Baum, R. (2024, Januar 18). SemLookP widgets - Embed ontologies in service applications. Zenodo. doi.org/10.5281/zenodo.10529181
[2] https://semanticlookup.zbmed.de/
[3] https://biodivportal.gfbio.org/
[4] Strömert, P., Limbachia, V., Oladazimi, P., Hunold, J., Koepler, O., Towards a versatile Terminology Service for empowering FAIR research data: Enabling ontology discovery,

design, curation, and utilization across scientific communities. Studies on the Semantic Web, Vol. 56 Knowledge Graphs: Semantics, Machine Learning, and Languages. IOS Press; 2023. doi:10.3233/ssw230005
[5] https://terminology.nfdi4chem.de
[6] https://lod.gesis.org/en/
[7] https://zbw.eu/stw
[8] Shigapov, R. (2022, November 28). Knowledge graphs in BERD and in NFDI. Focused Tutorial on Capturing, Enriching, Disseminating Research Data Objects. Use Cases from Text+, NFDI4Culture and BERD@NFDI, Mannheim and online. Zenodo. doi.org/10.5281/zenodo.7373258
[9] Kamlah, J., Schmidt, T., & Shigapov, R. (2022, November 28). Extracting research data from historical documents with eScriptorium and Python. Focused Tutorial on Capturing, Enriching, Disseminating Research Data Objects. Use Cases from Text+, NFDI4Culture and BERD@NFDI, Mannheim and online. Zenodo. https://doi.org/10.5281/zenodo.7373135
[10] Kamlah, J., & Shigapov, R. (2023, Mai 5). The German Production Pipeline: Mannheim - OCR & Knowledge Graphs. Zenodo. https://doi.org/10.5281/zenodo.7900133
[11] Rossenova, L., Schubotz, M., & Shigapov, R. (2023). The Case for a Common, Reusable Knowledge Graph Infrastructure for NFDI. Proceedings of the Conference on Research Data Infrastructure , 1. https://doi.org/10.52825/cordi.v1i.266
[12] Rossenova, L., Shigapov, R., Schubotz, M., Limani, F., Zapilko, B., & Förstner, K. U. (2024). KGI4NFDI - Knowledge Graph Infrastructure for the German National Research Data Infrastructure. Zenodo. https://doi.org/10.5281/zenodo.13118749
[13] https://github.com/ts4nfdi/terminology-service-suite