

Building Bridges through Metadata within the NFDI

Andreas Czerniak¹  <https://orcid.org/0000-0003-3883-4169>, Leyla Jael Castro²  <https://orcid.org/0000-0003-3986-0510>, Heike Fliegl³  <https://orcid.org/0000-0002-7541-115X>, Jonas Grieb⁴  <https://orcid.org/0000-0002-8876-1722>,
Christin Henzen⁵  <https://orcid.org/0000-0002-5181-4368>, Oliver Koepler⁶  <https://orcid.org/0000-0003-3385-4232>,
Josh Moore⁷  <https://orcid.org/0000-0003-4028-811X>, Emanuel Söding⁸  <https://orcid.org/0000-0002-4467-642X>,
Ulrik Stervbo⁹  <https://orcid.org/0000-0002-2831-8868>, Thorsten Trippel¹⁰  <https://orcid.org/0000-0002-7211-7393>,
and Cord Wiljes¹¹  <https://orcid.org/0000-0003-2528-5391>

¹ Universitätsbibliothek Bielefeld, Germany

² ZB Med Information Centre for Life Sciences, Germany

³ FIZ Karlsruhe – Leibniz Institut für Informationsinfrastruktur GmbH, Germany

⁴ Senckenberg Gesellschaft für Naturforschung, Germany

⁵ TU Dresden, Germany

⁶ TIB - Leibniz Information Centre for Science and Technology, Germany

⁷ German BiImaging - Gesellschaft für Mikroskopie und Bildanalyse e.V., Germany

⁸ GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel, Germany

⁹ Ruhr University Bochum, Germany

¹⁰ Leibniz-Institute for the German Language and University of Tübingen, Germany

¹¹ Nationale Forschungsdateninfrastruktur (NFDI) e.V.

*Correspondence: Thorsten Trippel, trippel@ids-mannheim.de

Abstract

The accessibility of data for interdisciplinary research requires overarching metadata standards implemented across scientific fields. The Taskforce Metadata within the Section Metadata, Terminologies, Provenance of the German National Research Infrastructure (NFDI) aims to advance metadata standardization across the diverse scientific disciplines represented within NFDI. This effort aligns with the goals set by the “Bund-Länder Vereinbarung (BLV)” of the German federal and state governments, which mandate the NFDI consortia to work towards agreed upon metadata standards [1]. The primary objective of the Taskforce is to encourage the adoption of existing (cross-disciplinary) terminologies and metadata schemas to enhance interoperability among the consortia within the NFDI. To support this goal, the Taskforce will create recommendations and guidelines to help consortia integrate agreed-upon metadata schemas into their repository infrastructure. The Taskforce has outlined a roadmap to achieve this objective, starting with the identification and evaluation of generic metadata schemas currently used for research data. Based on this evaluation, the Taskforce will develop guidelines to: drive schema adoption and harmonization, provide motivation and assistance, and facilitate the integration of modular metadata schemas, initially focusing on more closely related domain clusters. Additionally, the Taskforce plans to create guidelines for the

development and integration of modular metadata schemas, with an initial focus on consortia from related scientific domains. Through this, the Taskforce will deliver a comprehensive set of resources to support standardized metadata practices.

The Taskforce fosters dialogue with the NFDI consortia by organizing a series of workshops. In preparation for the kick-off workshop, the Taskforce identified the Data Catalog Vocabulary (DCAT), DataCite Metadata Schema, and Schema.org as the most suitable generic metadata standards, which allow extensions or mapping to individual domains. These standards were then discussed during the 1st NFDI Metadata Workshop in Dresden in January 2025 with about 50 metadata experts, representing all 26 NFDI consortia and further stakeholders. The focus was on the applicability of these schemas from the perspective of the individual consortia, exploring which metadata fields should be mandatory, recommended, or optional. Early in the workshop, it was agreed that a set of core metadata fields across domains could facilitate interoperability. In fact, many consortia already support general information in mandatory fields, such as title, identifier, creator, and publisher. However, it was also apparent that some fields needed further clarification, like defining the creator for sensor-generated data. To address these challenges and ensure consistency, semantic integration is crucial, especially when discussing keywords, subjects, and topic categories based on controlled vocabularies. The workshop concluded that a single, uniform metadata schema is impractical due to the wide range of domain-specific data characteristics, funding organization requirements, and legal aspects, each imposing unique constraints that necessitate the coexistence of multiple metadata schemas. The 2nd workshop in June 2025 will continue the community process to create recommendations and guidelines for integration of standardized metadata schemas. During this workshop, we will gather best practices for implementations of the identified generic metadata schemas and mapping between the generic and the more expressive discipline-specific schemas. Results of the workshop will be presented in this contribution.

The Taskforce invites participation in this process, contributing to the NFDI strategy on metadata harmonization and standardization framed within the agreement by the federal government and the states in Germany.

Keywords: Metadata, Standardization, NFDI

Author contributions

All authors equally contributed to conceptualization, writing, editing and review.

The authors acknowledge the use of Large Language Models (LLMs) as writing aids in phrasing this abstract, based on the authors' notes, ideas, and concepts. The authors retain full responsibility for the content.

Competing interests

The authors declare no competing interests.

Acknowledgement

We would like to thank all members of the TF Metadata and participants of the NFDI Metadata Workshops for the valuable contributions.

References

- [1] GWK, "Bund-Länder-Vereinbarung zu Aufbau und Förderung einer Nationalen Forschungsdateninfrastruktur (NFDI)", <https://www.gwk-bonn.de/fileadmin/Redaktion/Dokumente/Papers/NFDI.pdf> (last access April 10, 2025)
- [2] Henzen, C., Neidiger, C., Söding, E., & Trippel, T. (2025). Report on the first NFDI Metadata Workshop. With contributions from Leyla Jael Castro, Andreas Czerniak, Heike Fliegl, Jonas Grieb, Oliver Koepler, Josh Moore, and Cord Wiljes. Zenodo. DOI: [10.5281/zenodo.1522723](https://doi.org/10.5281/zenodo.1522723)