

NFDI4DS Shared Tasks

Sonja Schimmler¹² 

¹Fraunhofer FOKUS, Berlin, Germany

²Technische Universität Berlin, Germany

*Correspondence: Sonja Schimmler, sonja.schimmler@tu-berlin.de

Abstract

The NFDI4DS consortium continuously hosts various shared tasks to tackle problems under the umbrella of scholarly information processing.[1] In this abstract, we will describe the overall idea of shared tasks within NFDI4DS and give an overview of the shared tasks that are currently being prepared and executed.

Shared tasks are scientific competitions in which teams attempt to find efficient solutions to a specific problem using shared datasets and evaluation measures. The goal is to objectively and directly compare different methods for tackling the same problem by using gold-standard datasets and common performance measures. Shared tasks are usually organized either at conferences or workshops such as the Workshop on Natural Scientific Language Processing (NSLP)¹ organized by NFDI4DS, or by companies such as Kaggle².

In recent years, shared tasks have been very successful in advancing state-of-the-art methods and standards to solve challenging problems. In most shared tasks, the gold-standard dataset is made publicly available after the competition, thus providing valuable resources for the research community. Recent work has also introduced guidelines to ensure transparency and reproducibility of shared tasks to benefit scientific progress. Using competition between participants, shared tasks have proven to significantly encourage the development of novel and innovative solutions.

NFDI4DS has been quite successful in 2024 and 2025³, with 14 shared tasks having been accepted at international conferences, most of them at ACL⁴, ESWC⁵ and ISWC⁶. In 2024, we hosted the challenge 'KGQA: Knowledge Graph for Question Answering' at ACL, the challenge 'SOTA?: Tracking the State-of-the-Art in AI Scholarly Publications' at CLEF, the challenges 'FoRC: Field of Research Classification of Scholarly Publications' and 'SOMD: Software Mention Detection in Scholarly Publications' at ESWC and the challenges 'LLMsOL: Large Language Models for Ontology Learning' and 'Scholarly QALD: Hybrid Question Answering over Scholarly Knowledge Graphs and Text' at

¹<https://nfdi4ds.github.io/nslp2025/>

²<https://www.kaggle.com>

³<https://www.nfdi4datascience.de/community/shared-tasks>

⁴<https://2025.aclweb.org>

⁵<https://2025.eswc-conferences.org>

⁶<https://iswc2025.semanticweb.org>

ISWC. In 2025, we host the challenges 'FORC: Field of Research Classification (2nd edition)', 'MESD: Metadata Extraction from Scholarly Documents' and 'ReadMe2KG: GitHub ReadMe to Knowledge Graph' at ESWC, the shared tasks 'ClimateCheck: Scientific Fact-checking of Social Media Posts on Climate Change', 'SciVQA: Scientific Visual Question Answering', 'SOMD: Software Mention Detection in Scholarly Publications (2nd edition)' and 'LLMs4Subjects: LLM-based Subject Tagging' at ACL, and the shared task 'LLMs4Subjects: LLM-based Subject Tagging (2nd edition)' at GermEval.

NFDI4DS aims to utilize shared tasks to activate and serve the community, by defining shared tasks, providing gold-standard datasets, and by establishing quality standards for both. The community-driven solutions, many of them currently based on generative AI, have the potential to lay the foundation for novel and innovative NFDI4DS services.

Author contributions

- **Writing - original draft:** Sonja Schimmler
- **Writing - editing & review:** Sonja Schimmler
- **Software:** NFDI4DS Consortium

Funding

This work has received funding through the DFG project NFDI4DS (no. 460234259).

Acknowledgements

We thank the shared task organizers, Tilahun Abedissa (Leuphana University), Ricardo Usbeck (Leuphana University), Jennifer d'Souza (TIB), Raia Abu Ahmad (DFKI), Georg Rehm (DFKI), Frank Krüger (HS Wismar), Stefan Dietze (GESIS), Zeyd Boukhers (Fraunhofer FIT), Genet Gesese (FIZ) and Ekaterina Borisova (DFKI).

References

- [1] R. A. Ahmad, E. Borisova, G. Rehm, *et al.*, "NFDI4DS Shared Tasks," in *INFORMATIK 2023*, Gesellschaft für Informatik, 2023. DOI: [10.18420/INF2023_105](https://doi.org/10.18420/INF2023_105).