

Development of a metadata schema for publication of health-related research data on the German Central Health Study Hub of the NFDI4Health

Introduction: The National Research Data Infrastructure for Personal Health Data (NFDI4Health) aims to improve the FAIRness of health-related data from epidemiological, public health and clinical studies as well as registries and administrative health databases[1]. One key service of NFDI4Health is the German Central Health Study Hub[2] that supports a standardised publication and search of research (meta-)data (i.e. study-level data and related documents[3]) and is based on a metadata schema developed by NFDI4Health[4, 5].

Concept: The NFDI4Health metadata schema allows to collect basic bibliographic information such as title and description as well as information about related persons, organisations and publications. Additionally, details about study design and data accessibility can be provided. The DataCite Metadata Schema[6] was taken as a basis for the schema as it utilises a broadly applied vocabulary and supports publication of research data including DOI assignment. To describe health studies, the schema was extended by attributes from well-established data models in clinical and population-based research[7, 8, 9, 10, 11]. Additionally, a mapping of the schema against the HL7®FHIR® standard was conducted to evaluate its compatibility and to prepare profiling[12].

Conclusion and outlook: The NFDI4Health metadata schema provides a structured way to bundle metadata of various research data types from different health-related domains. The incorporation of a generic metadata standard enables description, publication and comparison of health studies and related resources at a general level. At the same time, the schema is scalable to adjacent research fields such as social sciences. Furthermore, domain-specific attributes captured in the schema for health research allow access to more in-depth information. Currently, the schema is being expanded to meet the needs of specific use cases, i.e. nutritional epidemiology, epidemiology of chronic diseases, and secondary data and record linkage. Mappings to other health metadata schemas such as ECRIN[13] and ERDRI[14] are also intended.

Please assign your poster to one of the following keywords.

Standards

Please assign yourself (presenting author) to one of the stakeholders.

Data Infrastructure Provider

Please specify "other" (stakeholder)

In addition please add keywords.

Metadata model; FAIR principles; Health

Primary author: SHUTSKO, Aliaksandra (ZB MED –Information Centre for Life Sciences)

Co-authors: SCHMIDT, Carsten Oliver (Institute of Community Medicine, University Greifswald); KLOPFEN-STEIN, Sophie Anne Ines (Berlin Institute of Health (BIH) at Charité –Universitätsmedizin Berlin); LÖBE, Matthias (Institute for Medical Informatics (IMISE), University of Leipzig); ABAZA, Haitham (Heidelberg Institute for Theoretical Studies (HITS)); GOLEBIEWSKI, Martin (Heidelberg Institute for Theoretical Studies (HITS)); DARMS, Johannes (ZB MED –Information Centre for Life Sciences); FLUCK, Juliane (ZB MED –Information Centre for Life Sciences)

Presenter: SHUTSKO, Aliaksandra (ZB MED –Information Centre for Life Sciences)

Session Classification: Postersession I

Track Classification: Postersession