

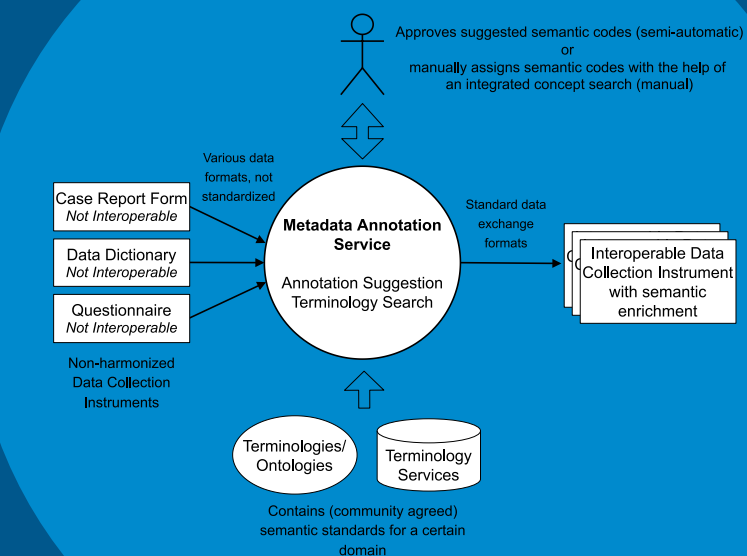
Background: The German project NFDI4Health is dedicated to the development of national research data infrastructure for personal health data, and as such searches for ways to enhance semantic interoperability.

Goal: Interoperability of data dictionaries, questionnaires and data collection tools is key to reusing and combining results from independent research studies. This can be reached by integration of semantic codes into study metadata.

Challenge: The researcher is confronted with terminologies and ontologies and plenty of variables to annotate in standard and non-standard formats. Terminologies, ontologies and standard formats differ depending on the domain. Thus, the process of semantic metadata annotation is labor intensive. While some metadata annotation services already exist [1], they are in many cases tailored to specific data formats and terminologies or they do not meet all basic prerequisites.

Requirements: The following basic requirements for semantic metadata annotation software in the biomedical field and in the context of the FAIR principles were identified:

- 1 **Open accessible and/or open code:** Open-source licence or free to use.
- 2 **Support of common data formats:** Support of common and domain specific import and export data formats (currently Excel Spreadsheet questionnaires).
- 3 **Terminology search and suggestion:** Search for concepts/classes within integrated terminologies/ontologies and suggestions for semantic concept annotations for each data element (semi-autonomous process).
- 4 **Interface to external terminology/ontology services:** Currently: Terminology Service SemLookP.
- 5 **Interactive user interface:** Ease annotation.



Terminology Service SemLookP

<https://semanticlookup.zbmed.de/ols/docs/index>

Based on the Ontology Lookup Service (OLS) [2]

Suggestions

Currently by string matching

Generation of suggestions by other methods is planned e.g.

- Text Mining approach
- Suggesting existing annotations

Conclusion and Outlook

- The first prototype of a metadata annotation service meets the basic requirements and can be adapted to specific use cases.
- Currently, the annotation service is being further adapted for a use case in the medical domain.
- The SNOMED terminology and common standards such as HL7 FHIR are therefore additional requirements that we are currently working on.

Acknowledgement: We gratefully acknowledge the financial support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 442326535. We would like to thank Carina Vorisek (Berlin Institute of Health (BIH)), Sophie Klopfenstein (Berlin Institute of Health (BIH)), Matthias Löbe (Inst. for Medical Informatics (IMISE)) and Carsten Oliver Schmidt (UM Greifswald) for their contributions during the requirement specification.

[1] <https://doi.org/10.3390/app12020796> [2] <https://www.ebi.ac.uk/ols/index>

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).