

Modular recommendation system and review application for semantic mapping

Freija Descamps¹, Panagiotis Gialernios¹, Shirah Cashriel¹, Isaac Claessen¹, Mythili Palanisamy¹,
Silvia Jimenez¹, Lars Halvorsen¹, Peter Moorthamer¹, Lore Vermeylen¹
¹edenceHealth NV

Background

Semantic mapping of medical codes and descriptions to standardized vocabularies is a time-consuming task. edenceHealth NV has developed multiple suggestion frameworks that can be used in a modular way depending on the type of input. These suggestions can then be loaded in the edenceReviewer online semantic mapping tool that provides a clean and visual way of validating these mappings.

Methods

The edenceHealth team has developed 3 different mapping suggestion tools that can be used in different situations. LabMapper can be used to map Measurements to the correct LOINC¹ code, making use of the hierarchy and Apache Lucene² text searching techniques. Custom scripts in combination with RxNormBuilder^{3,4,5} will be used to map Drugs to RxNorm⁶ or RxNorm Extension codes by mapping the components and combining everything using the drug hierarchy. Finally, edenceMapper is a more general tool that can be used for mapping any source code to any domain. It includes optional translation to English and is designed to use multiple mapping algorithms like for example simple fuzzy string matching, Apache Lucene search and a text embedding LLM (Multilingual-E5-small architecture⁷). From all models, edenceMapper aims to return best match as well as a ranked list of additional suggestions.

The result of these 3 suggestion systems will be combined into a database that is then used to display the suggestions in edenceReviewer. This is a tool developed by edenceHealth to facilitate the review and validation of the suggested mappings.

Results

Figure 1 shows the flow of the suggestions from the different suggestion systems into a database. The database feeds these suggestions to the edenceReviewer where users can review and change the mappings, where needed. The result will flow back to the database. The resulting mappings can be exported and used in any ETL.

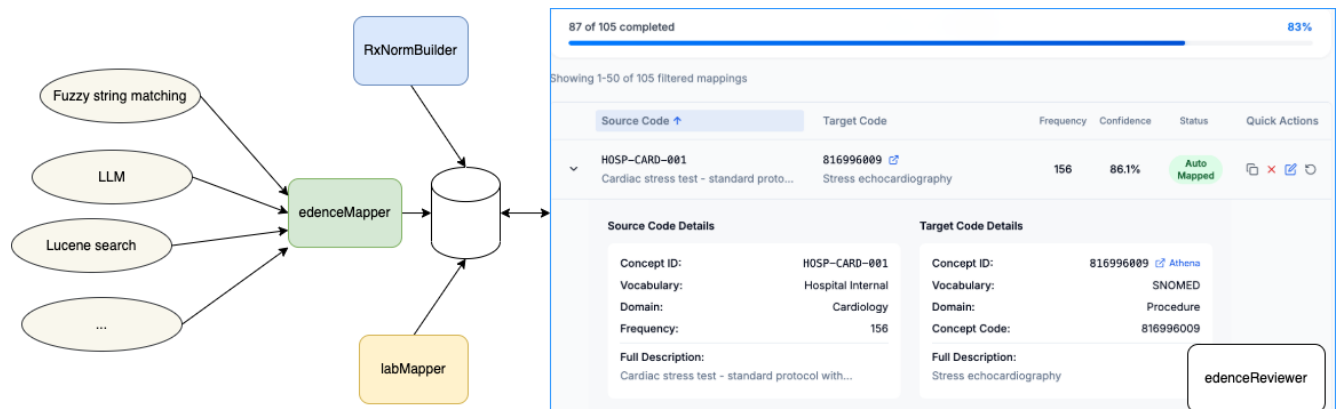


Figure 1. data flow from suggestion systems to edenceReviewer.

Conclusion

The combination of these tools makes the semantic mapping process more streamlined and facilitates the validation of the suggested mappings. It is constructed in a modular way that makes it easy to add more suggestion pipelines for specific mapping purposes or to add extra algorithms in edenceMapper.

References

¹ LOINC from Regenstrief institute. The international standard for identifying health measurements, observations, and document. <http://www.LOINC.org/>.

² Apache Lucene from The Apache Software Foundation. Welcome to Apache Lucene. <https://lucene.apache.org/>.

³ OHDSI, Incorporate your drug vocabulary into the OHDSI Vocabularies: path through RxNorm Extension, [Online]. <https://www.ohdsi.org/wp-content/uploads/2023/12/OHDSI-APAC-Scientific-Forum-20231207.pdf>.

⁴ OHDSI, International Drug Vocabulary Implementation Process - Creation of the input tables, 2024. [Online]. https://github.com/OHDSI/Vocabulary-v5.0/wiki/International-Drug-Vocabulary-Implementation-Process#drug_concept_stage.

⁵ OHDSI, BuildRxE SQL script, https://github.com/OHDSI/Vocabulary-v5.0/blob/master/working/packages/vocabulary_pack/BuildRxE.sql

⁶ U.S. National Library of Medicine, 2025. [Online]. <https://www.nlm.nih.gov/research/umls/rxnorm/index.html>.

⁷ Multilingual-LLM-model, <https://huggingface.co/intfloat/multilingual-e5-small>