Towards the scalability of databases’ harmonisation and expansion of RWE networks, Promptly introduces AI-based suggestions and automated vocabulary mapping platform into the ETL pipeline.

Improving vocabulary mapping to OMOP-CDM concepts: an AI-based automated and collaborative mapping platform

**Background** – During the harmonization of 38,162 patients, we identified a major bottleneck: vocabulary mapping of source database concepts to OMOP-CDM concepts, due to the burden of manual assignments requiring highly specialized human resources with a strong clinical background. Considering the limitations identified in USAGI OHSI tool, we are developing an AI-based automated and collaborative mapping platform service – Promptly mapping platform.

> The flow designed for Promptly AI-based automated and collaborative mapping platform was:
> 1. Access data partner’s database to extract descriptions of source concepts to map to OMOP-CDM concepts
> 2. Check if descriptions of source database concepts have been mapped before
> 3. Expose source database concepts and codes in a way that is easily readable by the clinician to map or review them
> 4. Integrate with Athena vocabularies directly, so the clinician can search for concepts
> 5. Suggest the most likely OMOP-CDM concepts for descriptions based on a recommendation system with an AI model
> 6. Select one or more OMOP-CDM concepts for each new description of source database. This mapping is performed by a clinician and helped by the AI suggestion and ability to search in Athena
> 7. Tag source database concepts according to priority and readiness state of the available descriptions, following a collaborative and efficient workflow
> 8. Export OMOP-CDM concepts marked as ready to the database/pipeline that populates the OMOP-CDM version

Intercommunicative service integrated into the ELT pipeline with three main communications:

- **ELT pipeline**: Receive new source codes that need to be mapped and send newly mapped target concept IDs
- **Users**: Alert medical users regarding a new mapping task
- **AI-based suggestion service**: Send source description, domain and metadata to the AI-based suggestion service and receive a suggestion for the target_concept_id.