A central ontology platform enables OMOP consortia to pre-coordinate their semantic mapping approach and to maintain necessary granularity until it is possible to formally integrate those concepts into a new standard OMOP vocabulary version.

Title: Construction of a central ontology platform for semantic mapping coordination and vocabulary augmentation across a multi-partner oncology consortium

BACKGROUND: The Integraal Kankercentrum Nederland, or IKNL, collects and catalogues data from individuals diagnosed with cancer across the Netherlands [1]. In an effort to enrich this national dataset, as well as to validate and expand upon ongoing analyses, IKNL has joined efforts with the Blueberry consortium alongside other regional and national cancer registries. Much of the critical information the consortium aims to capture exists as part of the ICDO3 vocabulary, which is not comprehensively represented in the standard Observational Medical Outcomes Partnership (OMOP) vocabulary tables as of spring 2023.

In order to capture sufficient granularity of source concepts across the network, we have constructed a central ontology platform that provides Blueberry users with the following functionalities:

1. Submission of unmapped source codes to a semantic mapping portal for automated standard suggestions (edenceMapper)
2. Creation of semantic mapping sets to be collaboratively reviewed, and subsequently approved or modified, in an online web portal (edenceReviewer)
3. Maintenance and update of the central consortium vocabulary tables using PROSA, an OMOP ontology application,
4. Search and reference of the latest central consortium vocabulary version using a centrally deployed Athena web application
5. Utilities and triggers that help integrate and synchronize each of the independent components

In this work, we describe each of these functionalities and the tools that support them, and we present the benefits and drawbacks of using such an approach to fill gaps (temporarily) in the OMOP vocabularies to support semantic mapping coordination across a targeted federated network.

LIMITATIONS AND DISCUSSION: In comparison with site-specific vocabulary maintenance, the ontology management structure we have established above has several advantages: all partners can implement the same standard vocabulary version in their OMOP CDM instances, all partners can reference the consortium-specific vocabulary through a familiar Athena interface, and all partners can quickly request and update the vocabulary version to incorporate concepts that capture source-specific information, without concern of concept redundancy across the network.

One critical consideration with this type of approach is that data partners need to establish consistent ETL deployment schedules to reference a dynamic vocabulary version. We expect to keep track of source releases and referenced vocabulary versions across the network using the Ares network tool [3, 4], which has already been implemented at IKNL to track OMOP version releases internally.