

**Title: ANANKE 2.0:**  
*An Update on the Tool  
for Mapping Between  
OHDSI Concept  
Identifiers to Unified  
Medical Language  
System (UMLS) identifiers  
– now with mappings to  
the Human Phenotype  
Ontology (HPO)*

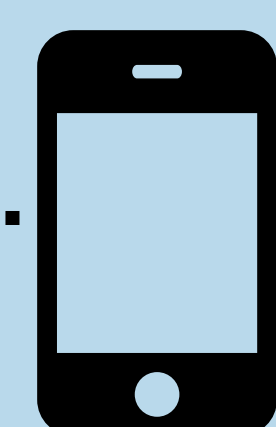
👤 PRESENTER: **Juan M. Banda**

Originally introduced at the 2018 OHDSI symposium, Ananke is a tool that provides automated mapping between UMLS concept unique identifiers (CUIs) and OHDSI concept identifiers. While the first version of Ananke satisfied the basic requirements of mappings between popular vocabularies such as: SNOMED, LOINC, RxNorm, ICD9, and ICD10, among others, coverage was not very comprehensive. In this work we introduce version 2 of the software and mappings that incorporate the additions made to UMLS since version 2017AB to version 2020AA and to the OHDSI Vocabulary up to the end of June 2020. This increases coverage by over 300,000 concepts since the first release. We also add vital mappings between OHDSI concept\_id's and the Human Phenotype Ontology (HPO) as part of ongoing work being performed on the N3C Tools and Resources group.

# The expansion into adding mappings to the HPO opens a very interesting line of work which can allow others to leverage the internal UMLS mappings to link additional resources not found in the OHDSI Vocabulary in an automated way

Table 1. Total number of UMLS concepts mapped into their respective OHDSI concept identifiers between versions of Ananke.

Vocabulary Name	Mapped concepts (Version 1)	Mapped concept (Version 2)
CPT4	39,947	44,083
HCPCS	6,376	7,193
ICD10CM	101,725	104,733
ICD10PCS	180,450	189,499
ICD9CM	16,376	17,553
LOINC	124,813	215,111
MedDRA	52,953	76,241
RxNorm	202,627	198,962
SNOMED	371,013	493,665



Take a picture to visit Ananke's GitHub Repository

## How Ananke mappings work

The process of mapping UMLS concept unique identifiers (CUI) begins with selecting the proper vocabulary in UMLS and on the OHDSI vocabulary. This is done via the SAB field in UMLS and the vocabulary\_id field on the OHDSI Vocabulary. The vital detail in this step is that the versions of the source vocabularies used usually differ between the resources, with OHDSI Vocabulary having newer versions of them. However, these discrepancies are usually resolved, and new mappings can be added when a newer version of UMLS is released (as shown on Table 1).

Once the UMLS and OHDSI vocabulary names are identified, the mapping relies on using the CODE field in UMLS and the concept\_code in OHDSI vocabulary. That is, using their source vocabulary identifiers to get to the CUI and the concept\_id respectively.



👤 Juan M. Banda  
**Georgia State University,  
Atlanta, Georgia, USA**

