Observational Health Data Sciences and Informatics (OHDSI)

Natural Language Processing (NLP) Working Group

03/09/2016
AGENDA

- Note-type mapping Presentation – Karthik
- Share existing ontologies from Vanderbilt (Hua) and Regenstrief (Jon)
- Share strategies for combining data from different searches – Jon
- Improvements to search engine set up using MT samples – Min
AGENDA

- Goals of 2016
- Report on WG for commenting – Hua
- Wrappers for cTAKES and Metamap – Min
- Textual Data Representation – Discussion
Document types from Vanderbilt

- CSV with the typ, styp, and the count of the instances of that document
- The styp is generally what is displayed as the document name
- 380,655 rows