

# Turbocharging Clinical NLP: Using UMLS Hierarchies to Model Word Cloud Knowledge Base for Faster Insight and Relationship Discovery with AI Assistance

Sina Madani, Asli Weitkamp  
Department of HealthIT, Vanderbilt University Medical Center

## PURPOSE

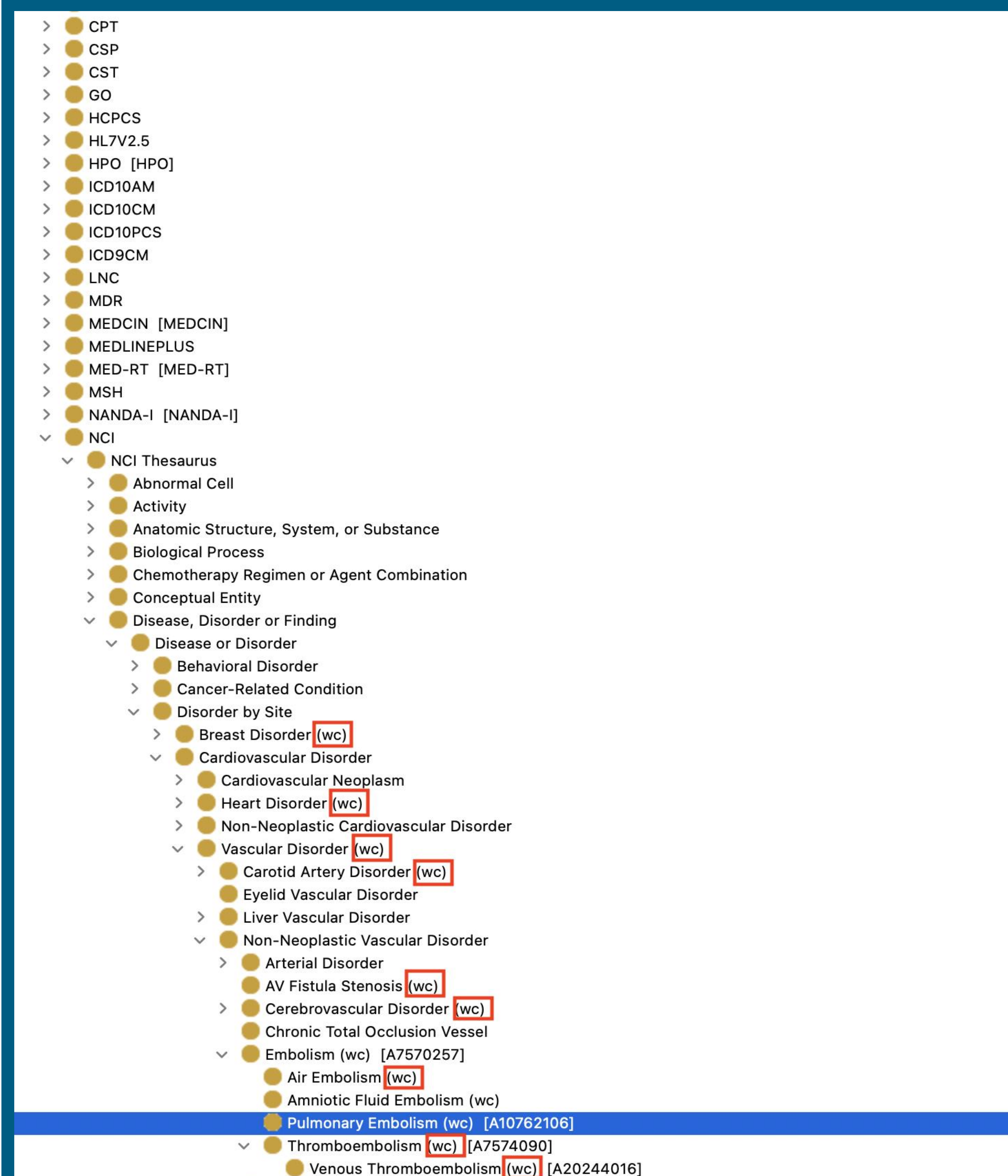
- VUMC Enterprise NLP application (World Cloud) lacks hierarchical structure
- UMLS KB provides >100 terminologies with rich relationships.
- Goal: Embed local CUIs in UMLS hierarchy → scalable, reusable ontology for NLP projects.**

## METHODS

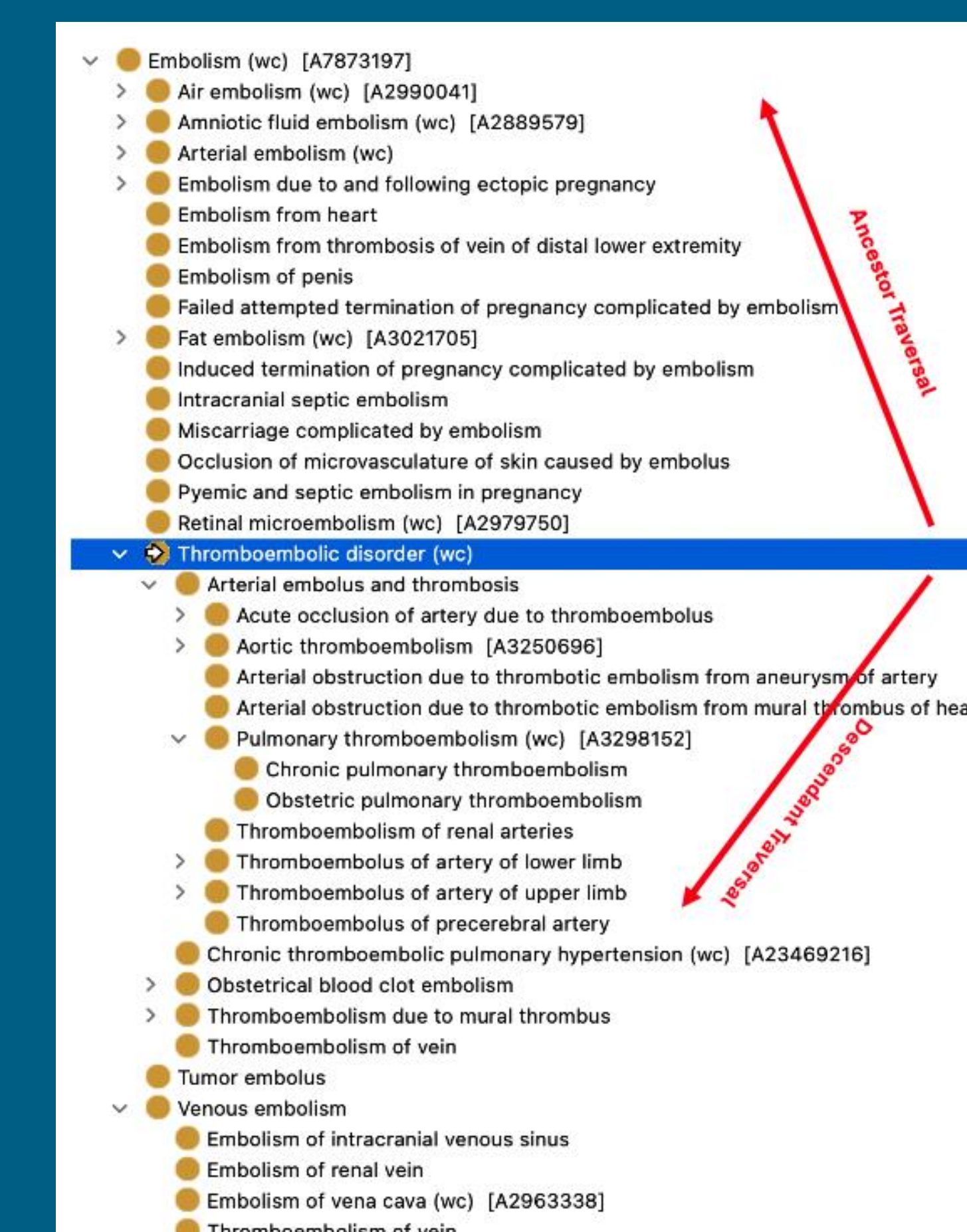
- Queried UMLS MRCONSO & MRHIER tables in Databricks to filter CUIs, AUIs, and hierarchical paths.
- Used AI tools for schema interpretation, Spark SQL design, and RDF modeling
- Upward & downward hierarchical traversal for top-level and granular concept retrieval
- Graph Build Ontology & integration into TopBraid / NCBO Bioportal

## Ontology enables instant expansion of clinical concepts within Word Cloud KB

*Word Cloud KB Concepts like "heart valve disease", "thromboembolism" can be quickly expanded into narrower/broader concepts across terminologies*



Concepts labeled (wc) belong to the Word Cloud Knowledge Base



AI tools like aiChat, ChatGPT, and Databricks Assistant reduced ontology build and NLP concept analysis time from days to hours

## Results

- Ontology coverage:** > 15,000 CUIs classified
- Efficiency:** Ontology generated from multi-million paths in < 2 hours
- Utility:** Supports instant terminology expansion in NLP pipelines

## Conclusion

- Faster development:** AI reduced manual coding/debugging time.
- More accurate retrieval:** Consistent terminology expansion and relationship discovery
- Clinical Impact:** Improves speed and precision of concept identification for enterprise NLP projects

Presenter: Sina Madani

 sina.madani@vumc.org

VANDERBILT  UNIVERSITY  
MEDICAL CENTER

Department of HealthIT