

# GraphLingo: Domain Knowledge Exploration by Synchronizing Knowledge Graphs and Large Language Models



Duy Le, Kris Zhao, Mengying Wang, Yinghui Wu  
Department of Computer and Data Science, Case Western Reserve University, USA

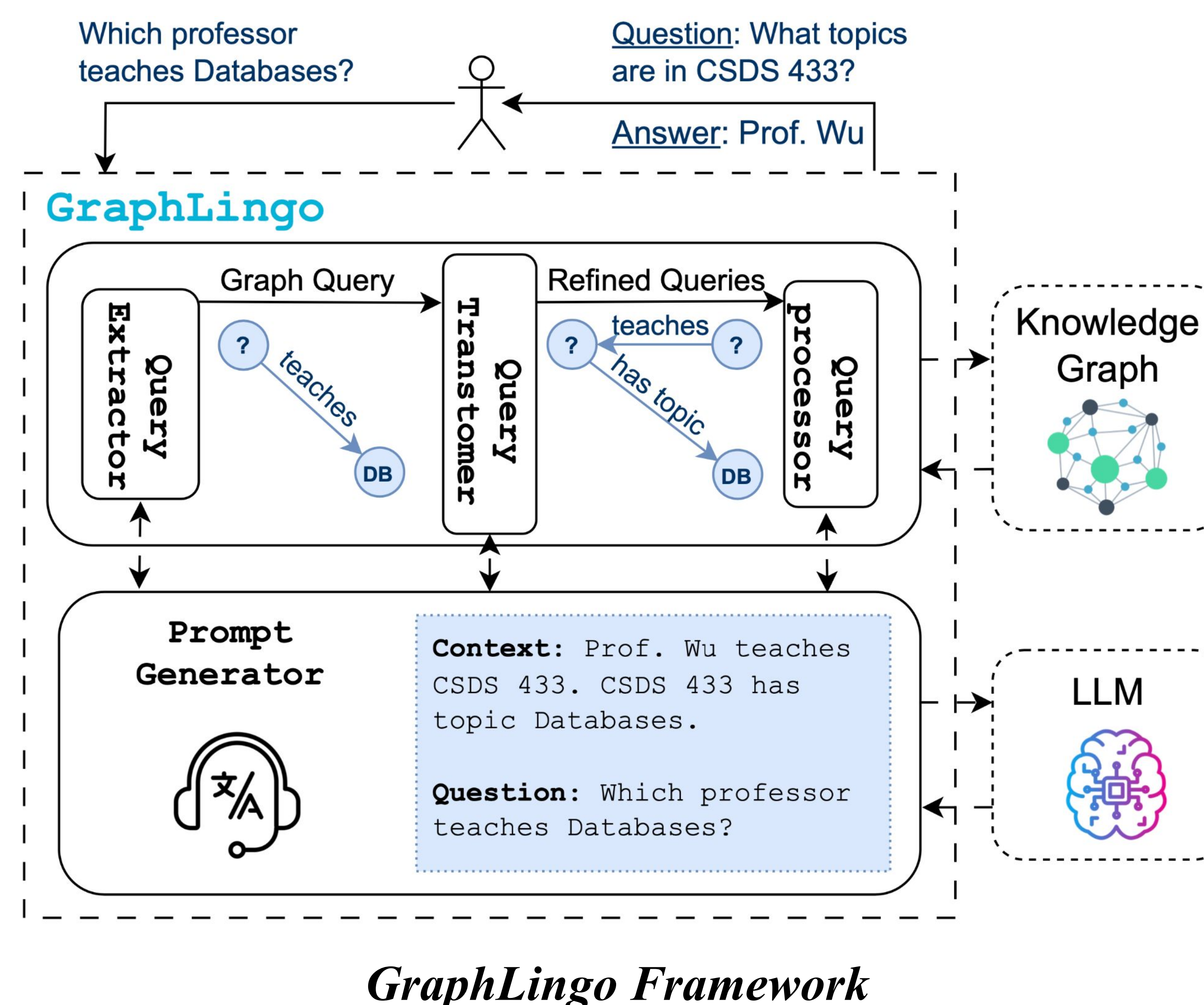
## OVERVIEW

Various domain-specific knowledge graphs (KGs) have been curated to host factual knowledge in e.g., science.

### Challenges:

- Domain experts are still expected to write complex queries, e.g. SPARQL, to access the KGs.
- LLMs are good at Q&A, but lacks factual, domain knowledge leading to problems like hallucination.
- Can we have a KG Q&A system that marries the merits of both graph search and LLM?

**GraphLingo:** A KG Q&A tool that *synchronizes domain-specific KGs* and *LLMs* to *guide users in exploring domain knowledge* with natural language.



## Query Extractor

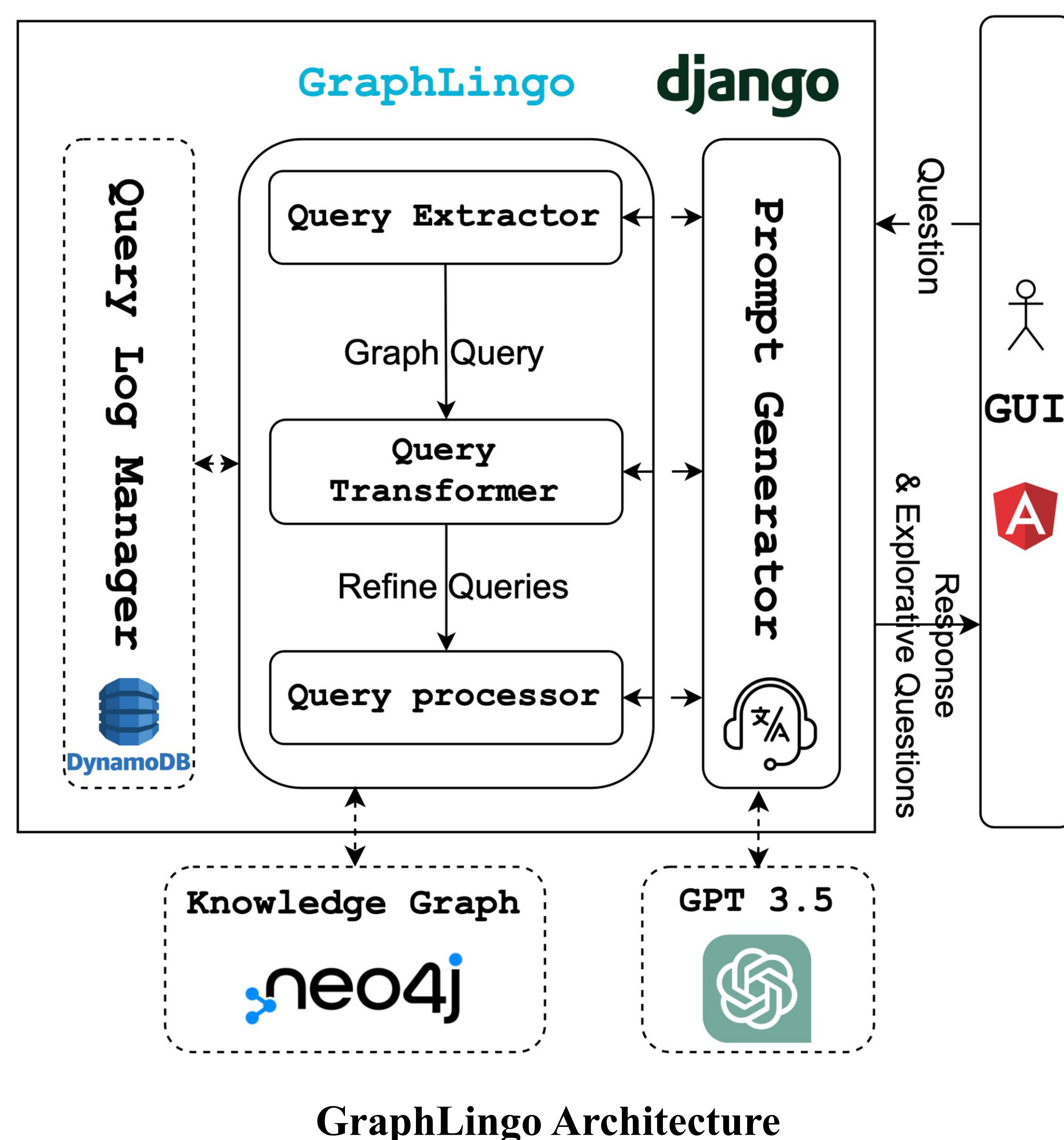
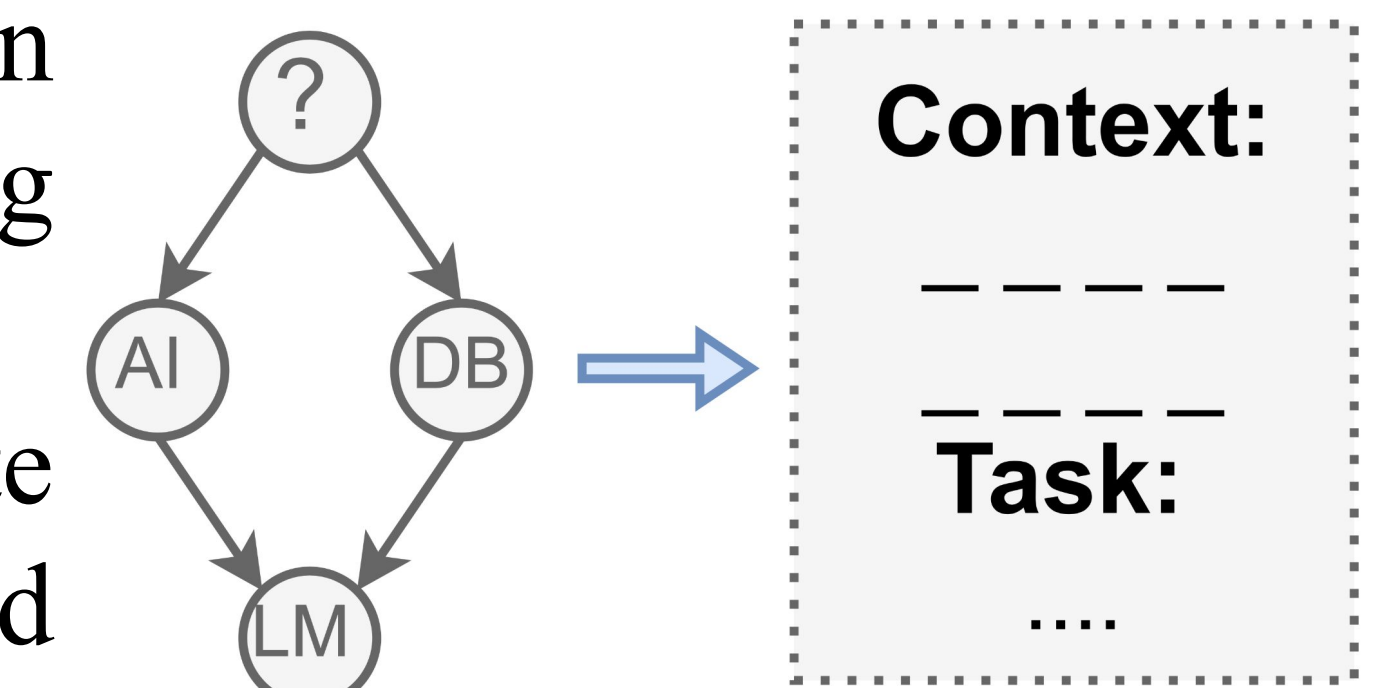
- Translate received natural language question into graph pattern query,  $Q$ .
- By prompting LLMs through prompt generator module.

## Query Transformer

- Transform initial graph query  $Q$  to be more semantically relevant to the KG.
- With Pattern-of-Thought (PoT) prompting technique.

## Prompt Generator

- Bridge the communication between graph processing modules and the LLM.
- Automatically generate prompts based on induced sub-patterns.



## LLM-enhanced Query Processor

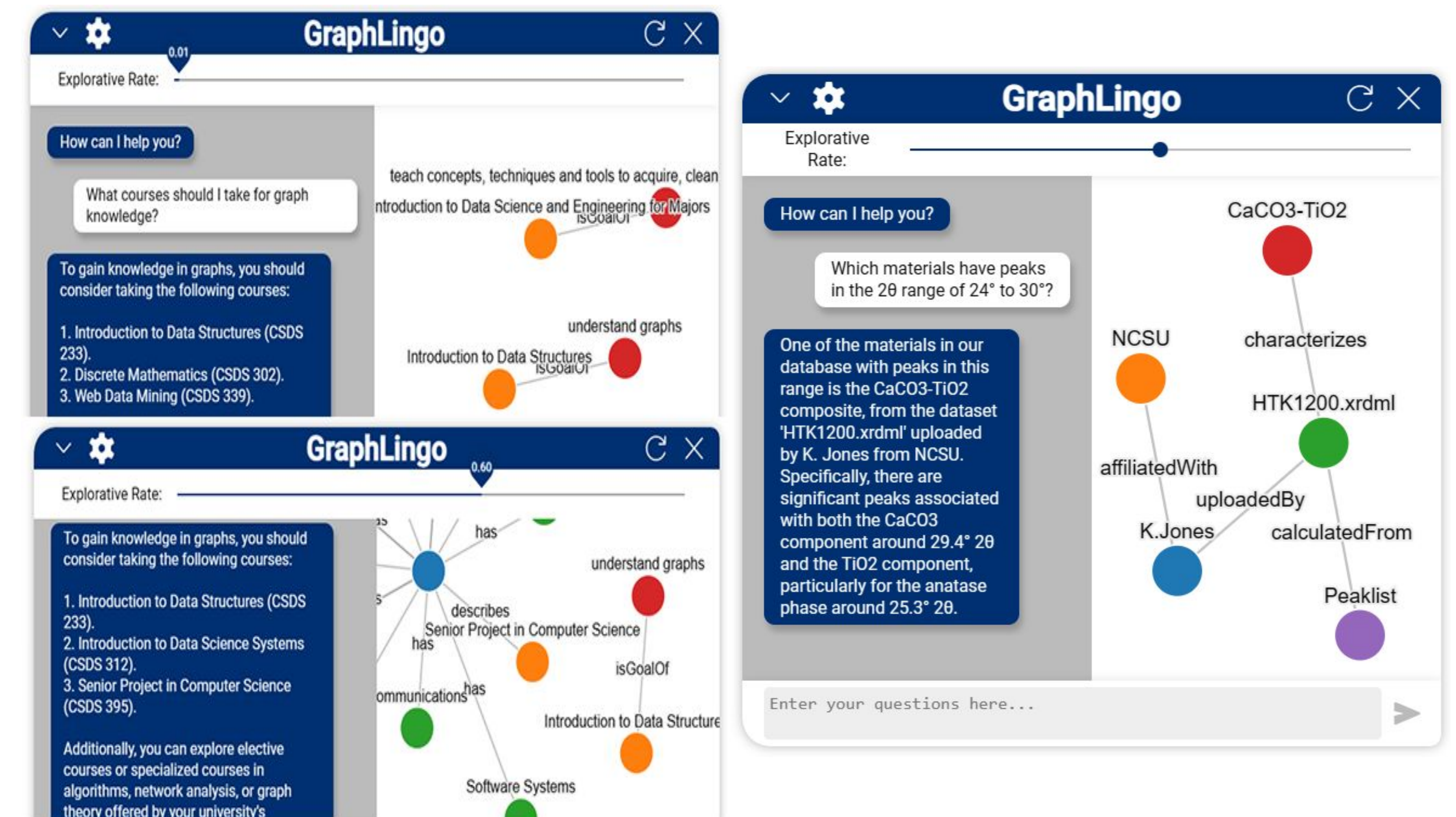
### Graph Query Engine

### LLM Query Processor

### Preference-aware Exploration:

- Allow users to pose ad-hocly a tunable preference, to explore knowledge in a more "explorative" or "conservative" manner.
- Diversely selecting answers from the union set of result triples from the KG and LLM:

$$(1 - \lambda) \sum_{e \in Q(G)} \text{rev}(e, Q) + \lambda \sum_{e \in \mathcal{L}(\mathcal{E}_Q)} \text{sim}(e, Q)$$



Demo Interface

## ACKNOWLEDGEMENT

This work is supported by NSF under OAC-2104007.

## GraphLingo TEAM

