KN-RAG: A Novel Approach to Retrieval-Augmented Generation with Knowledge Graphs and Self-Contextualization

Northeastern University
Khoury College of Computer Sciences

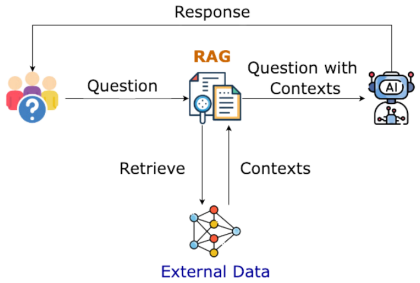
Zhixiao Wang, Haoning Wang, Wenqian Xie, Mingyi Qiu

Supervised by Prof Ryan Rad
Mentored by Hayden Barker at Deloitte

Deloitte.

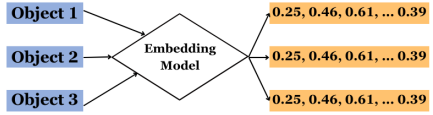
1. Background

- Large language models (LLMs) can generate incorrect answers and need costly updates.
- Retrieval-Augmented Generation (RAG) improves response accuracy without updating LLMs.

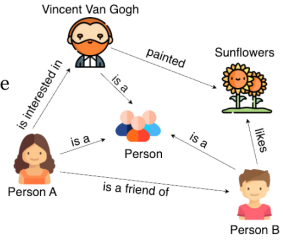


2. Motivations & Objectives

Motivation



- Traditional RAG vector embedding loses information about connections and relationships.
- Knowledge graphs map entities and connections, but are time-consuming to create manually.



Objectives

1. Integrate and automate knowledge graph generation.
2. Introduce a self-contextualization mechanism.

LinkedIn

Zhixiao Wang | Haoning Wang | Wenqian Xie | Mingyi Qiu

3. Methodology

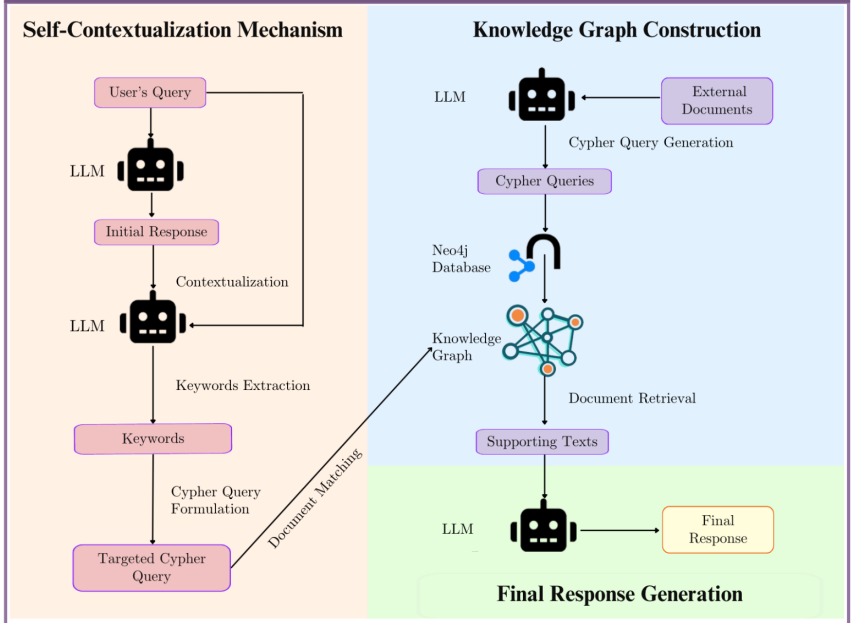
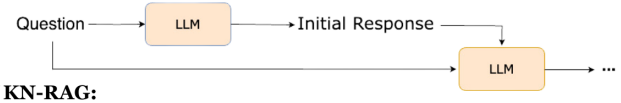


Figure 1. The pipeline of Knowledge-Nexus RAG (KN-RAG)

Differences from other RAG models:

- Automated knowledge graph generation
- Self-contextualization mechanism



Benefits of KN-RAG:

- Enhanced logical association
- Optimized retrieval process

Figure 2. Part of Knowledge Graph for Violent Incident Information from News Articles (VIINA) dataset

4. Experimental Result

- **Models**
  - SR: Standard RAG, which serves as our baseline model.
  - KN1: KN-RAG using direct keyword search.
  - KN2: KN-RAG using semantic similarity search.
- **RAGAS Evaluation - Multihop Dataset**

COMPARATIVE RAGAS EVALUATION RESULTS

	Faith.	Ans. Rel.	Cont. Rec.	Cont. Prec.
SR	0.452	0.423	0.535	0.517
KN1	0.584	0.571	**0.724**	0.703
KN2	**0.602**	**0.586**	0.709	**0.718**

- **Head-to-head Comparison - VIINA Dataset**

New articles

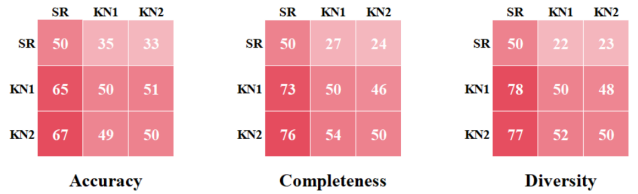
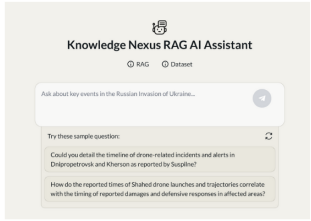


Figure 3. Win rate percentages in head-to-head comparisons between KN-RAG models (KN1 and KN2) and the standard RAG model across three metrics

5. Conclusion & Exhibition

- **Conclusion & Future Work**
  - KN-RAG outperforms the standard RAG model in response accuracy, completeness, and diversity.
  - In the future, we plan to test whether the KN-RAG model can be effectively applied to various domain-specific datasets.
- **Web Application Demo: KN-RAG AI Assistant**



Screenshot of the Web App

QR Code of the Web App