Lecture#6: Knowledge Graphs with Large Language Models & GraphRAG

Joyce Jiyoung Whang
Department of AI Computing, KAIST

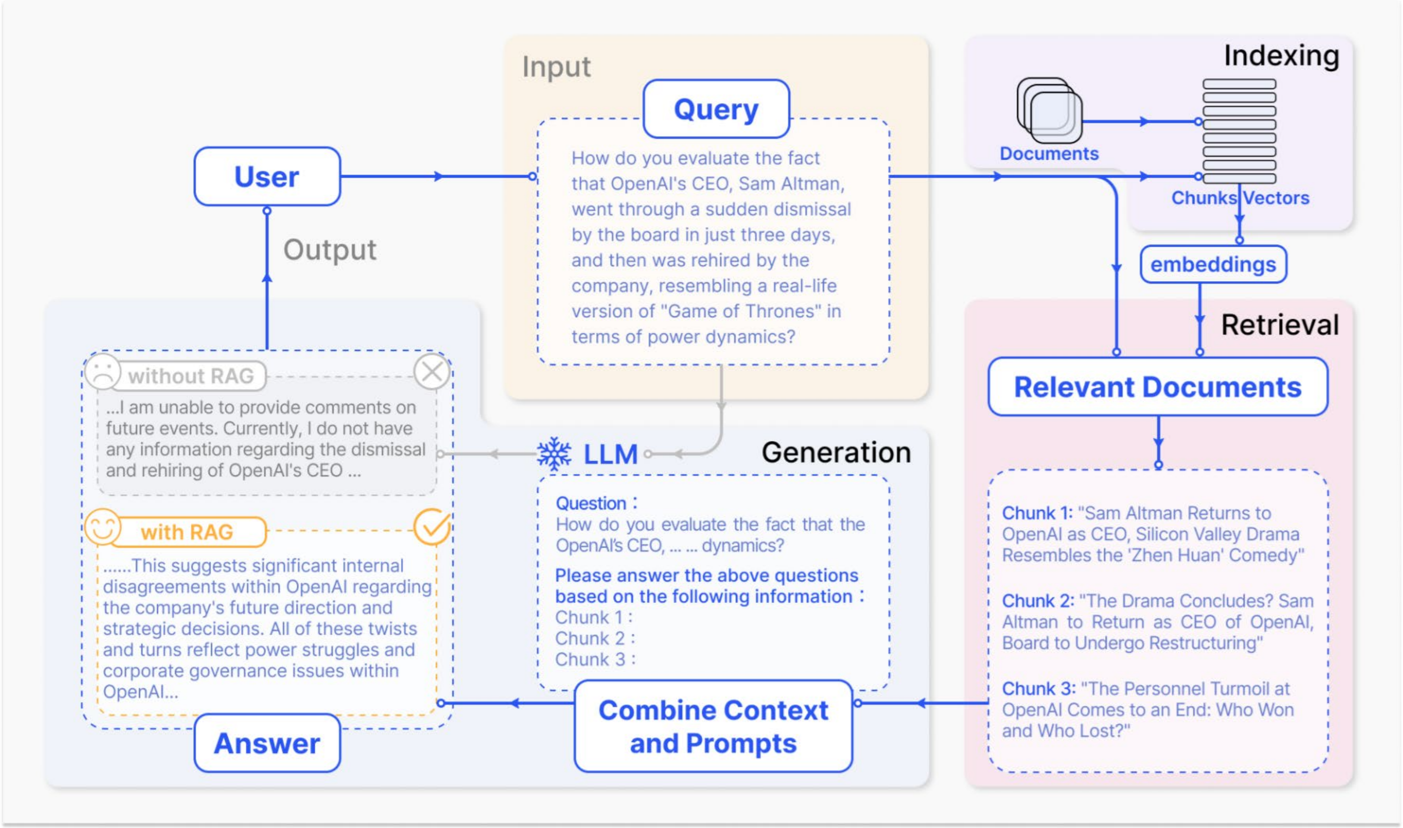
Knowledge Discovery with Knowledge Graphs: From Structural Embeddings to Generative Reasoning (KnowKG), KDD 2026 Tutorial

https://bdi-lab.kaist.ac.kr

KAIST
BDILab
BIG DATA INTELLIGENCE

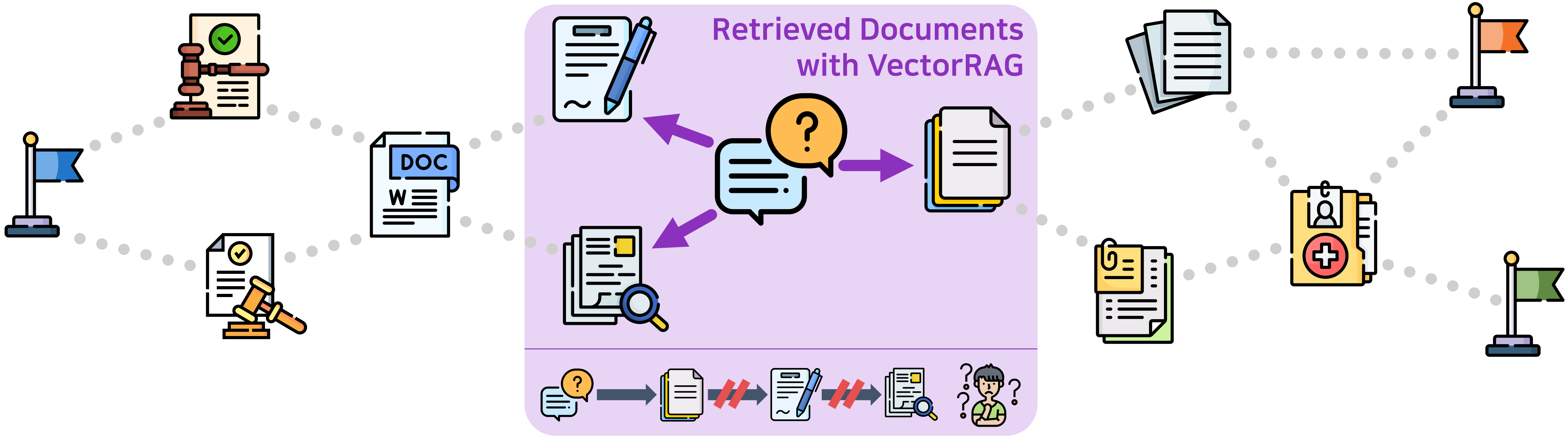
01 Retrieval-Augmented Generation (RAG)

- Enhancing Large Language Models (LLMs) by **providing relevant information retrieved from corpus**
- Allow LLMs to answer **queries based on massive data** that exceeds the models' context window
- **VectorRAG**: Standard RAG setup that retrieves chunks using **embedding-based similarity search**



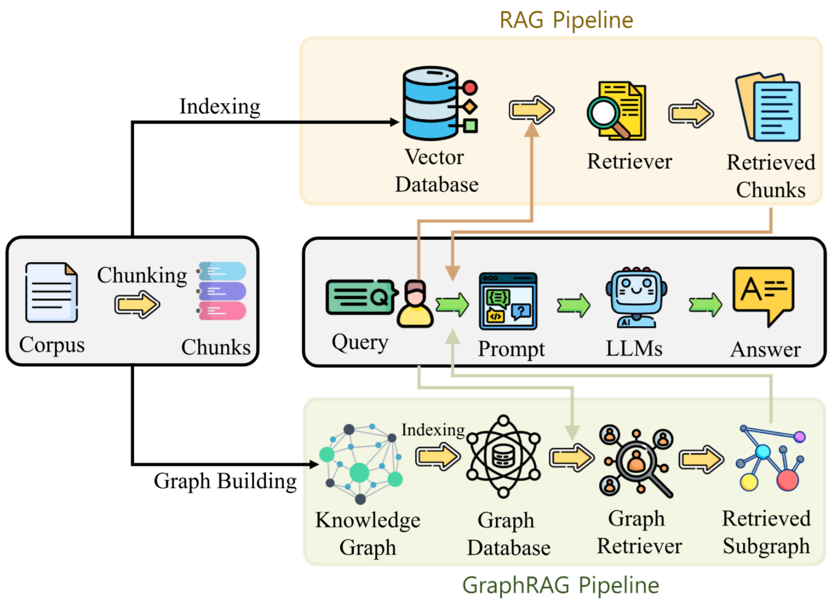
01 Limitations of VectorRAG

- **Lack of Global Understanding**
  - While VectorRAG approaches work well for localized queries that can be answered with a small set of records, they struggle with sensemaking queries that require a global, holistic understanding of the entire data.
- **Fragmented Information**
  - Since VectorRAG approaches treat each text chunk in isolation, they fail to integrate knowledge across different sources and connect related concepts, often resulting in fragmented and disjointed answers.

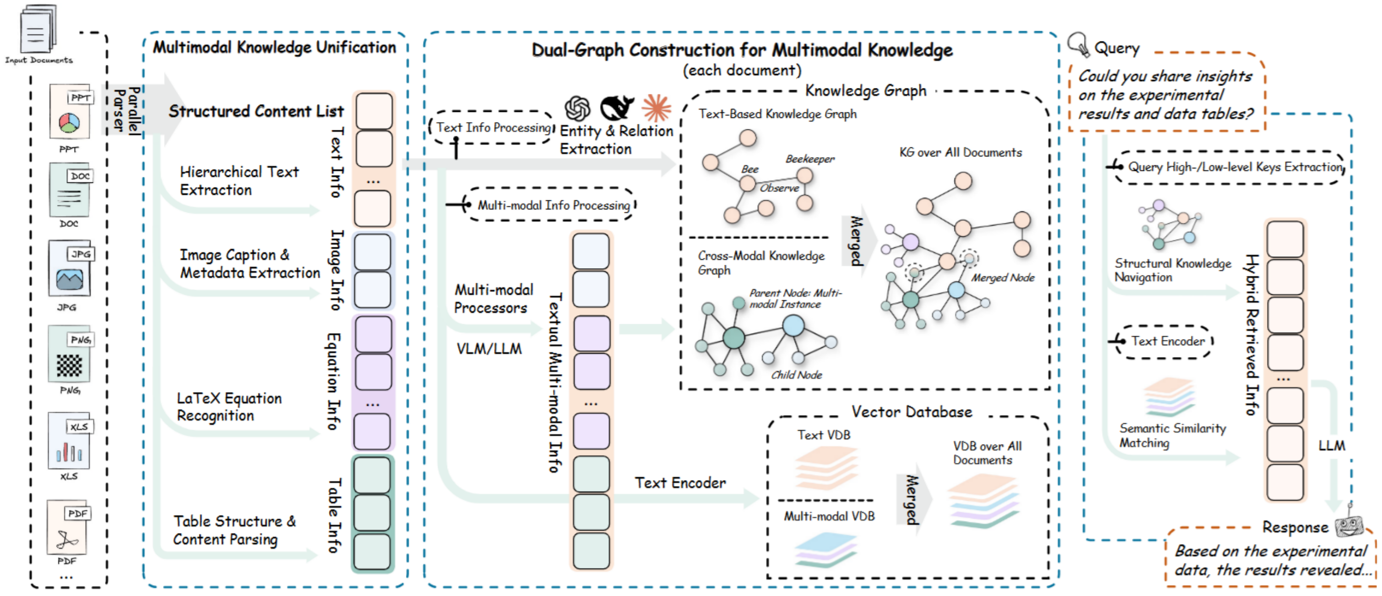


01 GraphRAG

- RAG approaches that leverage graphs as a data source
  - Constructing knowledge graphs from source text to enable reasoning over structured entities and relations
  - Linking isolated data segments to integrate fragmented information across diverse sources
- GraphRAG enhances factual accuracy and multi-hop reasoning of LLMs



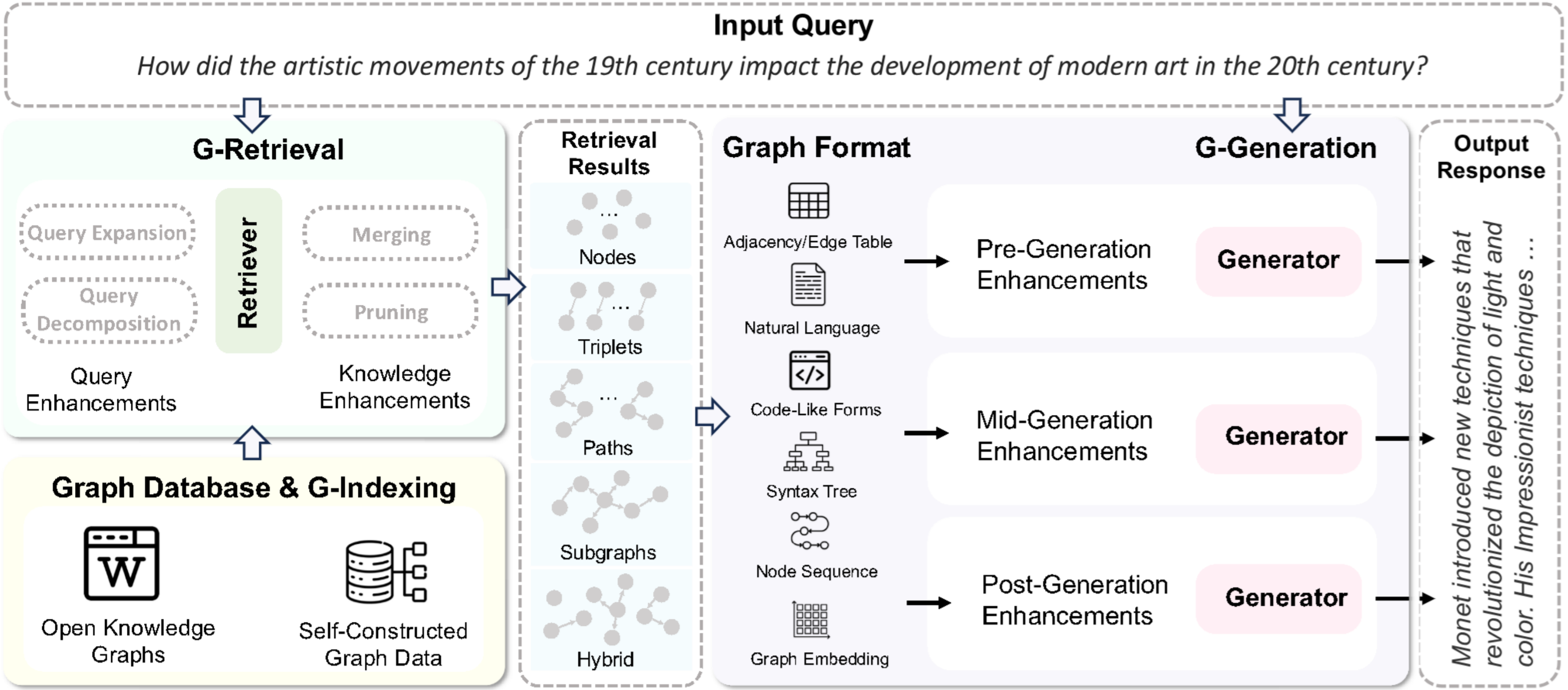
VectorRAG and GraphRAG



RAG-Anything (arXiv 2025)

01 Common Pipeline of GraphRAG

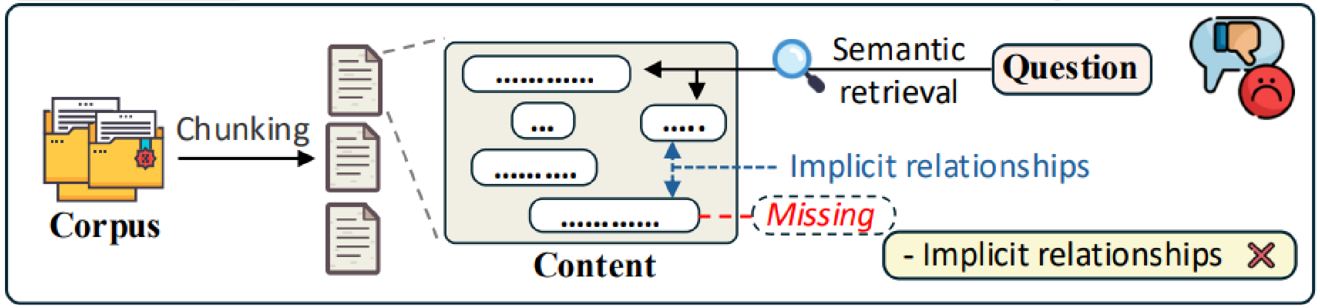
- **Indexing**: Construct a **structured representation** of the input corpus
- **Retrieval**: Retrieve graph elements **relevant to the input query**
- **Generation**: **Integrate the retrieved elements** to answer the query



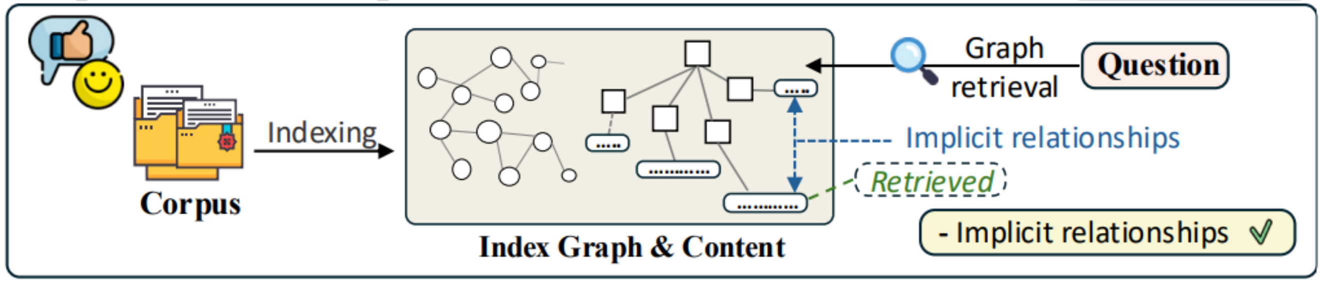
02 VectorRAG vs. GraphRAG

- **Global Context Understanding**
  - While VectorRAG works well for retrieving local information for specific queries, it falls short when answering global sensemaking queries that require an understanding of the entire corpus.
  - GraphRAG enables global query processing by organizing the entire dataset as a graph, allowing it to synthesize overarching themes and high-level insights.

VectorRAG



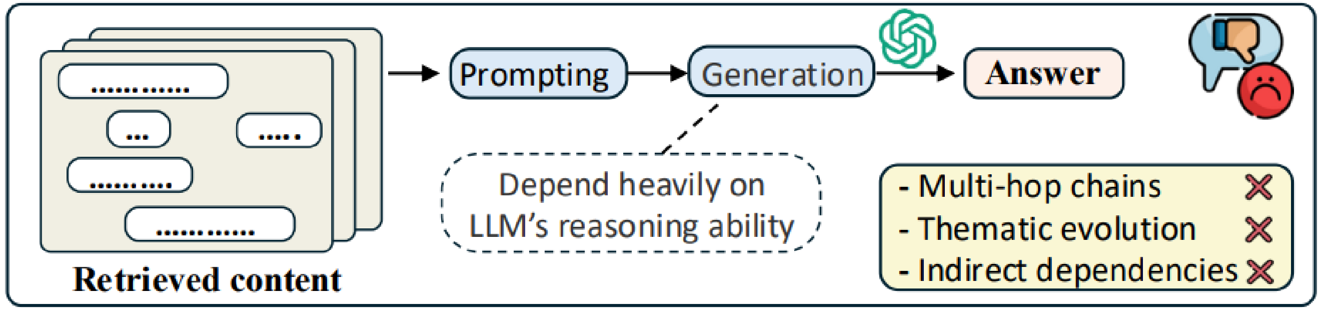
GraphRAG



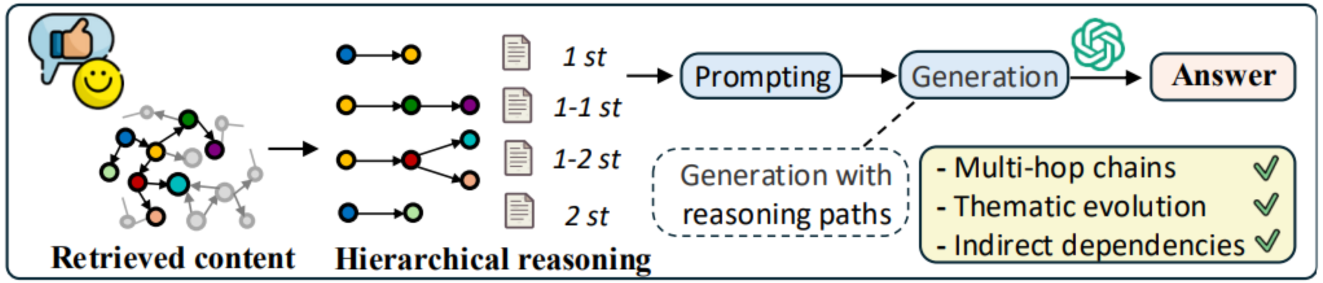
02 VectorRAG vs. GraphRAG

- **Multi-Hop Reasoning Capability**
  - Although VectorRAG is efficient for retrieving documents relevant to a query, it **struggles when the answer requires complex LLM inference to bridge pieces of evidence across multiple sources**.
  - GraphRAG inherently **supports multi-hop reasoning by traversing graphs**, linking fragmented concepts across diverse sources during the indexing stage.

VectorRAG



GraphRAG

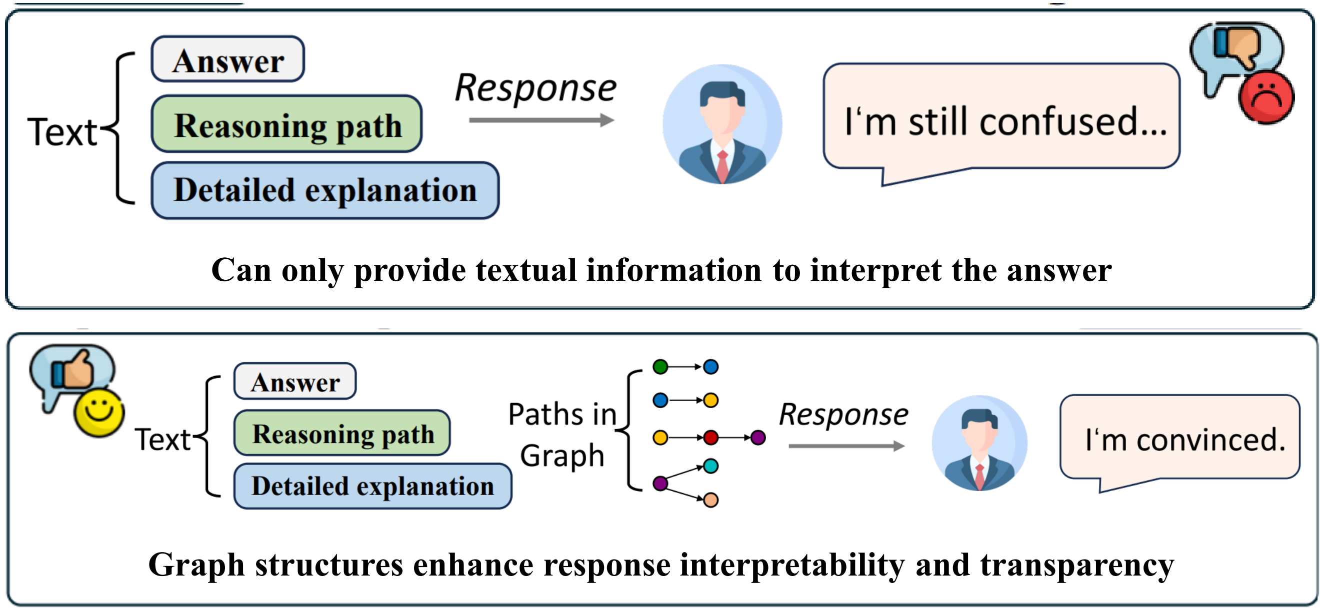


02 VectorRAG vs. GraphRAG

- **Interpretability**
  - VectorRAG offers limited interpretability as it merely provides the retrieved documents as isolated fragments, providing little context on how the information is interrelated.
  - GraphRAG delivers highly interpretable responses by leveraging graph structures such as paths or subgraphs, allowing users to clearly trace the logical connections within the knowledge network.

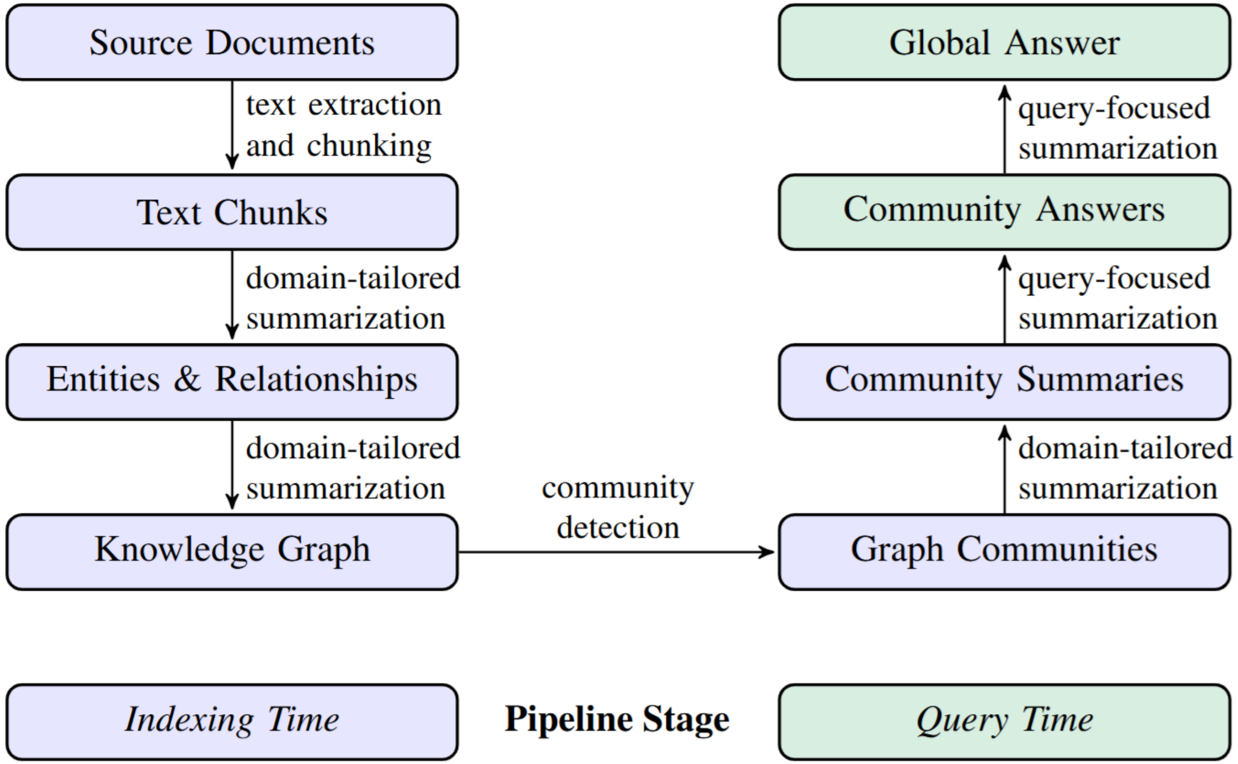
VectorRAG

GraphRAG



03 GraphRAG (arXiv 2024)

- Convert source documents into a unified knowledge graph by extracting entities and relations
- Construct communities using the Leiden algorithm, and generate a summary for each community
- Leverage summaries to generate partial answers, and aggregate them to obtain the global answer



03 GraphRAG (arXiv 2024)

- Generate global sensemaking questions to evaluate the capability of global sensemaking
  - Given a high-level description of a document, propose multiple personas of hypothetical users
  - For each hypothetical user, specify tasks that the user would use the document for
  - For each (user, task) pair, generate multiple questions

Dataset	**Example activity framing and generation of global sensemaking questions**
Podcast transcripts	*User*: A tech journalist looking for insights and trends in the tech industry *Task*: Understanding how tech leaders view the role of policy and regulation *Questions*: 1. Which episodes deal primarily with tech policy and government regulation? 2. How do guests perceive the impact of privacy laws on technology development? 3. Do any guests discuss the balance between innovation and ethical considerations? 4. What are the suggested changes to current policies mentioned by the guests? 5. Are collaborations between tech companies and governments discussed and how?
News articles	*User*: Educator incorporating current affairs into curricula *Task*: Teaching about health and wellness *Questions*: 1. What current topics in health can be integrated into health education curricula? 2. How do news articles address the concepts of preventive medicine and wellness? 3. Are there examples of health articles that contradict each other, and if so, why? 4. What insights can be gleaned about public health priorities based on news coverage? 5. How can educators use the dataset to highlight the importance of health literacy?

03 GraphRAG (arXiv 2024)

- Achieve **comprehensiveness win rates between 72-83%** compared to VectorRAG
- Require **26-33% and 97% fewer tokens** for low-level and root-level summaries, respectively

Podcast transcripts

Comprehensiveness

	SS	TS	C0	C1	C2	C3
SS	50	17	28	25	22	21
TS	83	50	50	48	43	44
C0	72	50	50	53	50	49
C1	75	52	47	50	52	50
C2	**78**	**57**	**50**	**48**	**50**	**52**
C3	79	56	51	50	48	50

Diversity

	SS	TS	C0	C1	C2	C3
SS	50	18	23	25	19	19
TS	82	50	50	50	43	46
C0	77	50	50	50	46	44
C1	75	50	50	50	44	45
C2	81	57	54	56	50	48
C3	**81**	**54**	**56**	**55**	**52**	**50**

Empowerment

	SS	TS	C0	C1	C2	C3
SS	50	42	57	52	49	51
TS	**58**	**50**	**59**	**55**	**52**	**51**
C0	43	41	50	49	47	48
C1	48	45	51	50	49	50
C2	51	48	53	51	50	51
C3	49	49	52	50	49	50

Directness

	SS	TS	C0	C1	C2	C3
SS	**50**	**56**	**65**	**60**	**60**	**60**
TS	44	50	55	52	51	52
C0	35	45	50	47	48	48
C1	40	48	53	50	50	50
C2	40	49	52	50	50	50
C3	40	48	52	50	50	50

News articles

Comprehensiveness

	SS	TS	C0	C1	C2	C3
SS	50	20	28	25	21	21
TS	80	50	44	41	38	36
C0	72	56	50	52	54	52
C1	**75**	**59**	**48**	**50**	**58**	**55**
C2	79	62	46	42	50	59
C3	79	64	48	45	41	50

Diversity

	SS	TS	C0	C1	C2	C3
SS	50	33	38	35	29	31
TS	67	50	53	45	44	40
C0	62	47	50	40	41	41
C1	65	55	60	50	50	50
C2	**71**	**56**	**59**	**50**	**50**	**51**
C3	69	60	59	50	49	50

Empowerment

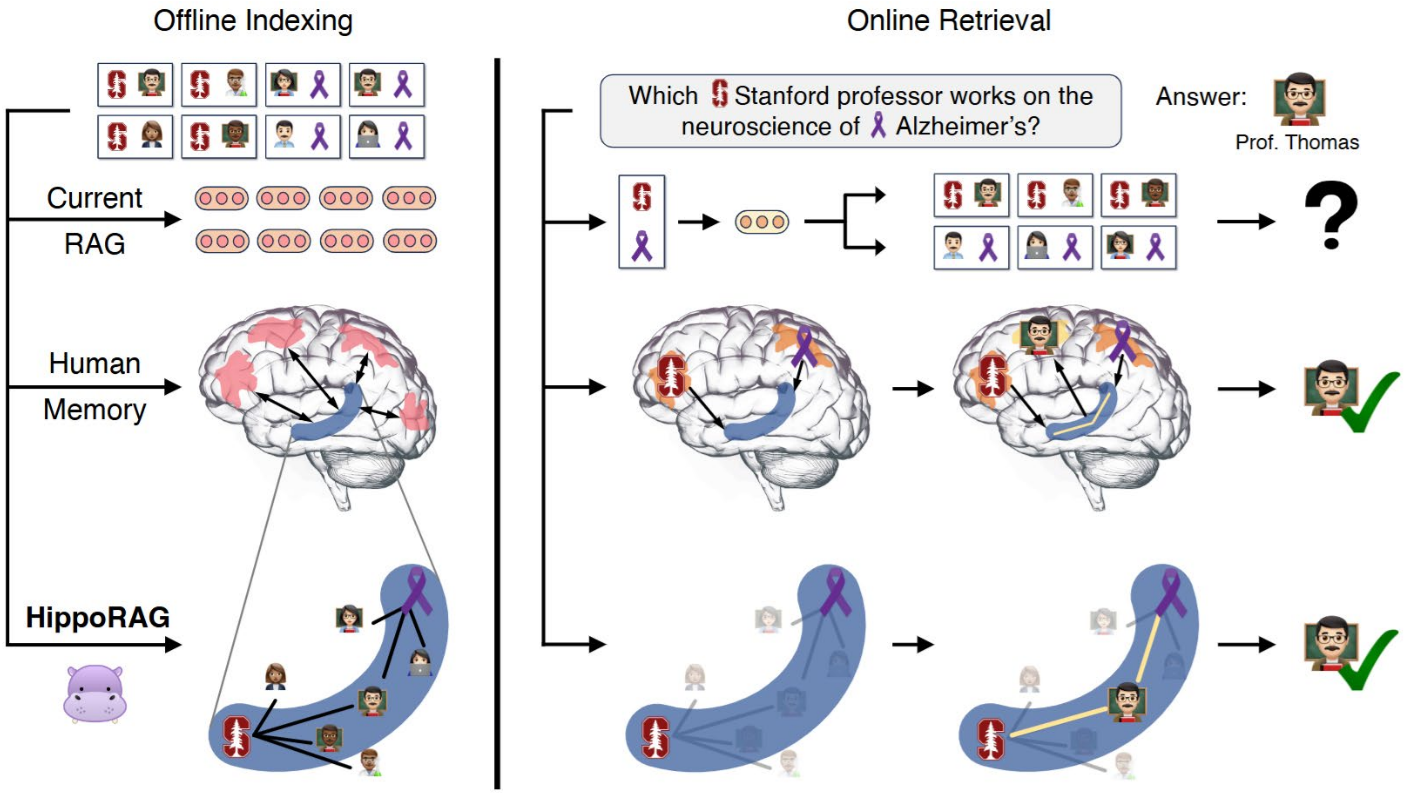
	SS	TS	C0	C1	C2	C3
SS	50	47	57	49	50	50
TS	53	50	58	50	50	48
C0	43	42	50	42	45	44
C1	**51**	**50**	**58**	**50**	**52**	**51**
C2	50	50	55	48	50	50
C3	50	52	56	49	50	50

Directness

	SS	TS	C0	C1	C2	C3
SS	**50**	**54**	**59**	**55**	**55**	**54**
TS	46	50	55	53	52	52
C0	41	45	50	48	48	47
C1	45	47	52	50	49	49
C2	45	48	52	51	50	49
C3	46	48	53	51	51	50

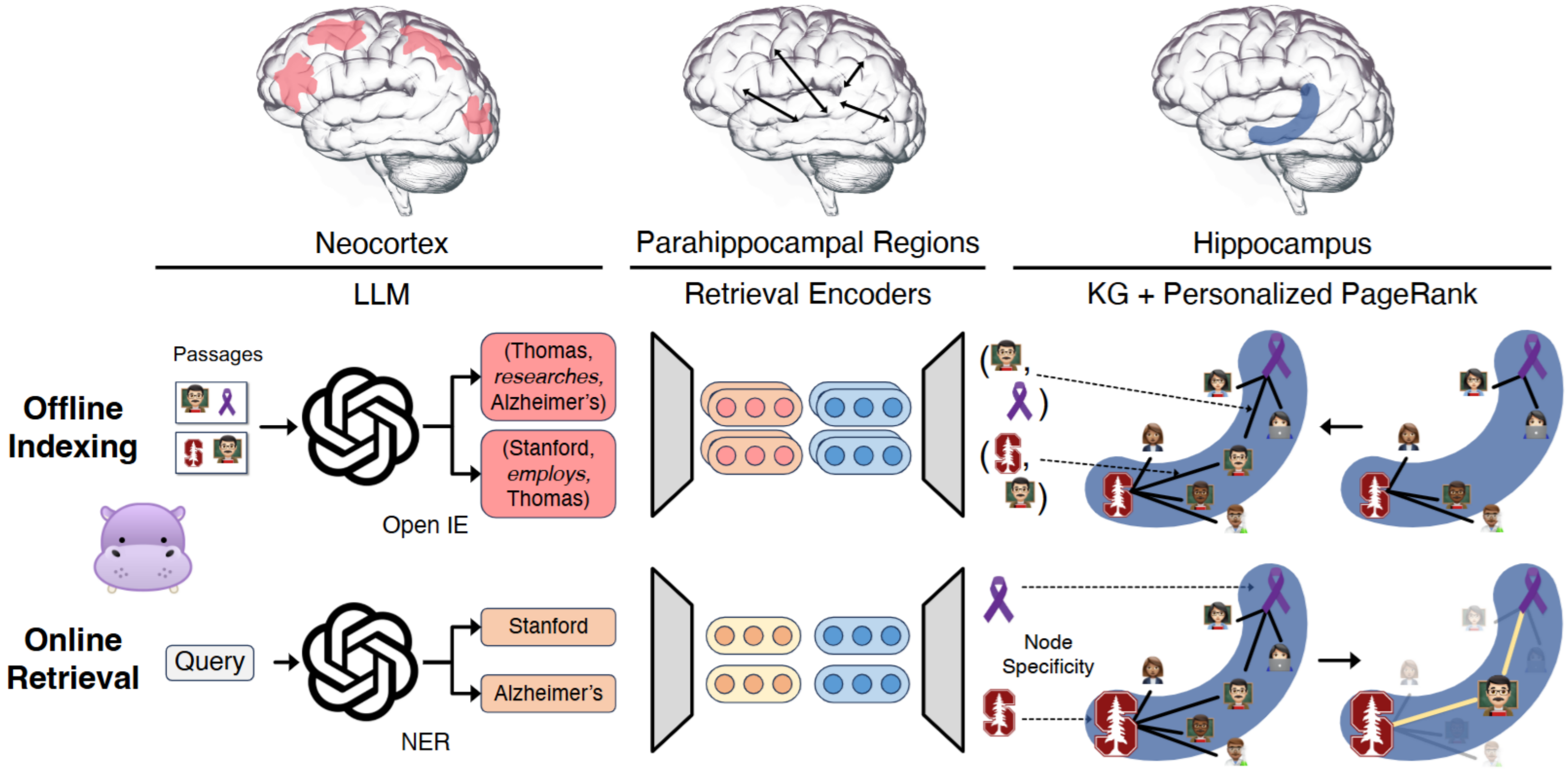
03 HippoRAG (NeurIPS 2024)

- Model the human brain's memory system to solve complex queries
- LLMs as neocortex, KGs as hippocampal index, and encoders as parahippocampal regions



HippoRAG (NeurIPS 2024)

- Offline Indexing: Utilize an LLM to extract triplets, and link similar entities in the knowledge graph
- Online Retrieval: Discover relevant entities via Personalized PageRank



03 HippoRAG (NeurIPS 2024)

- Combining HippoRAG with **multi-step retrieval methods** shows strong performance
- **Improve the question answering performance** on multiple datasets

Retrieval Performance

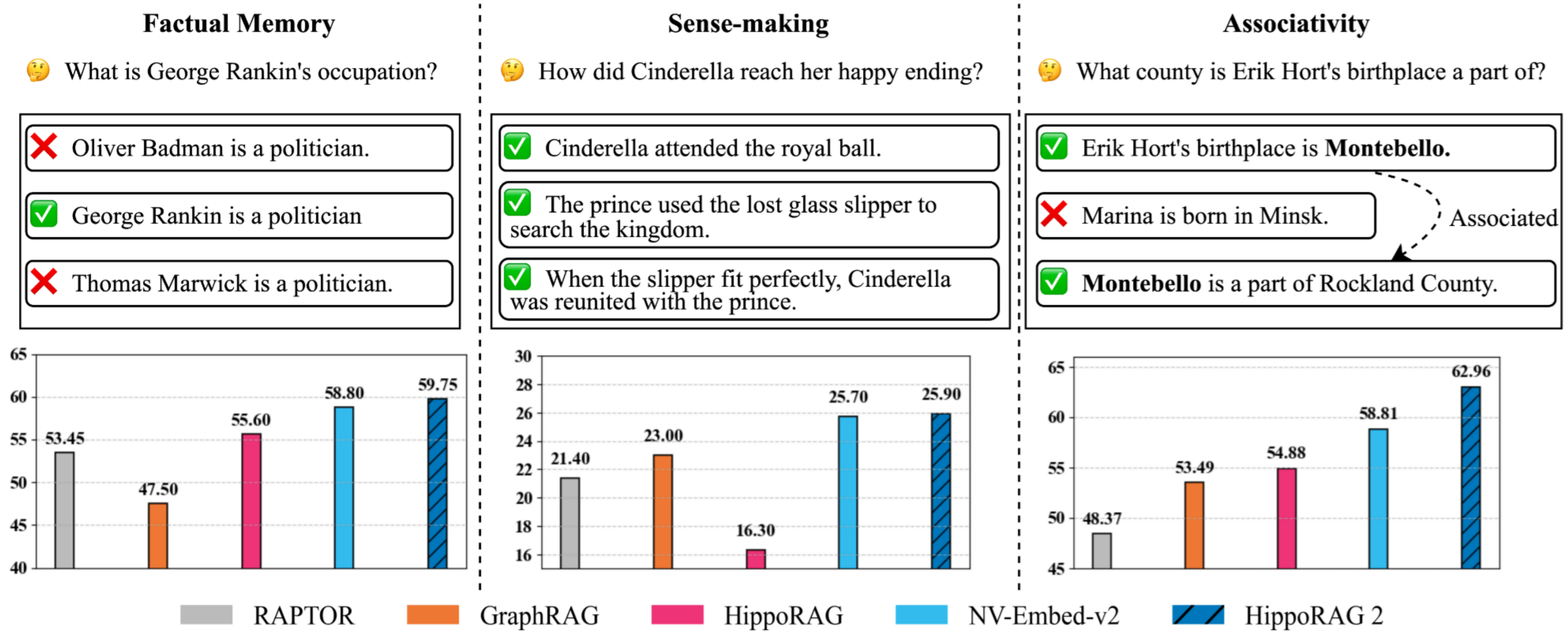
	MuSiQue		2Wiki		HotpotQA		Average	
	R@2	R@5	R@2	R@5	R@2	R@5	R@2	R@5
IRCoT + BM25 (Default)	34.2	44.7	61.2	75.6	65.6	79.0	53.7	66.4
IRCoT + Contriever	39.1	52.2	51.6	63.8	65.9	81.6	52.2	65.9
IRCoT + ColBERTv2	41.7	53.7	64.1	74.4	**67.9**	82.0	57.9	70.0
IRCoT + HippoRAG (Contriever)	<u>43.9</u>	<u>56.6</u>	<u>75.3</u>	<u>93.4</u>	65.8	<u>82.3</u>	<u>61.7</u>	<u>77.4</u>
IRCoT + HippoRAG (ColBERTv2)	**45.3**	**57.6**	**75.8**	**93.9**	<u>67.0</u>	**83.0**	**62.7**	**78.2**

Question Answering Performance

	MuSiQue		2Wiki		HotpotQA		Average	
Retriever	EM	F1	EM	F1	EM	F1	EM	F1
None	12.5	24.1	31.0	39.6	30.4	42.8	24.6	35.5
ColBERTv2	15.5	26.4	33.4	43.3	43.4	57.7	30.8	42.5
HippoRAG (ColBERTv2)	<u>19.2</u>	29.8	<u>46.6</u>	<u>59.5</u>	41.8	55.0	<u>35.9</u>	<u>48.1</u>
IRCoT (ColBERTv2)	19.1	<u>30.5</u>	35.4	45.1	<u>45.5</u>	<u>58.4</u>	33.3	44.7
IRCoT + HippoRAG (ColBERTv2)	**21.9**	**33.3**	**47.7**	**62.7**	**45.7**	**59.2**	**38.4**	**51.7**

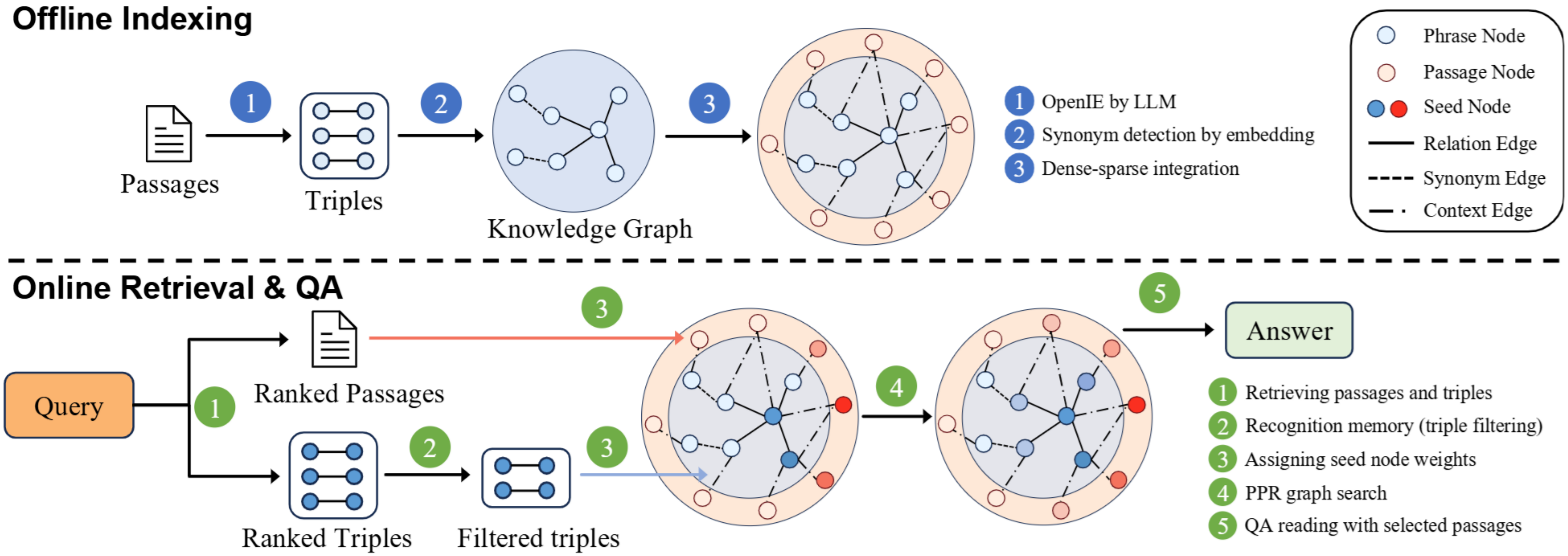
03 HippoRAG2 (ICML 2025)

- The performances of GraphRAG approaches on factual memory tasks are lower than standard RAG.
- Improve HippoRAG by enabling query-based contextualization



03 HippoRAG2 (ICML 2025)

- Introduce passage nodes to enrich KGs and remove isolated concepts
- Retrieve triplets relevant to a query before PPR-based graph search



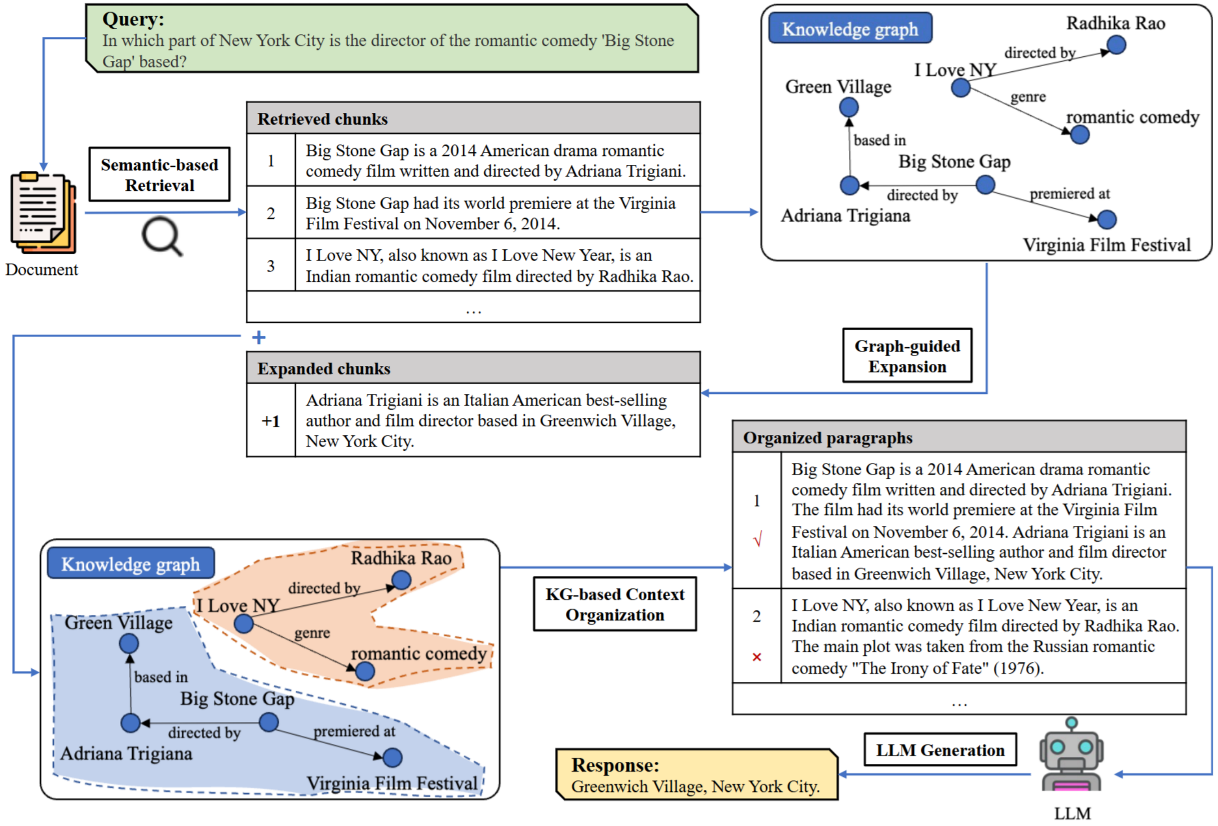
03 HippoRAG2 (ICML 2025)

- Achieve the **highest average F1 scores among RAG approaches**
- Show **comparable efficiency to the original HippoRAG**

	Simple QA		Multi-Hop QA				Discourse Understanding	
Retrieval	NQ	PopQA	MuSiQue	2Wiki	HotpotQA	LV-Eval	NarrativeQA	Avg
Simple Baselines								
None	54.9	32.5	26.1	42.8	47.3	6.0	12.9	38.4
Contriever (Izacard et al., 2022)	58.9	53.1	31.3	41.9	62.3	8.1	19.7	46.9
BM25 (Robertson & Walker, 1994)	59.0	49.9	28.8	51.2	63.4	5.9	18.3	47.7
GTR (T5-base) (Ni et al., 2022)	59.9	<u>56.2</u>	34.6	52.8	62.8	7.1	19.9	50.4
Large Embedding Models								
GTE-Qwen2-7B-Instruct (Li et al., 2023)	<u>62.0</u>	**56.3**	40.9	60.0	71.0	7.1	21.3	54.9
GritLM-7B (Muennighoff et al., 2025)	61.3	55.8	44.8	60.6	73.3	9.8	23.9	56.1
NV-Embed-v2 (7B) (Lee et al., 2025)	61.9	55.7	<u>45.7</u>	61.5	<u>75.3</u>	9.8	<u>25.7</u>	<u>57.0</u>
Structure-Augmented RAG								
RAPTOR (Sarthi et al., 2024)	50.7	<u>56.2</u>	28.9	52.1	69.5	5.0	21.4	48.8
GraphRAG (Edge et al., 2024)	46.9	48.1	38.5	58.6	68.6	<u>11.2</u>	23.0	49.6
LightRAG (Guo et al., 2024)	16.6	2.4	1.6	11.6	2.4	1.0	3.7	6.6
HippoRAG (Gutiérrez et al., 2024)	55.3	55.9	35.1	**71.8**	63.5	8.4	16.3	53.1
HippoRAG 2	**63.3**†	<u>56.2</u>	**48.6**†	<u>71.0</u>†	**75.5**	**12.9**†	**25.9**	**59.8**

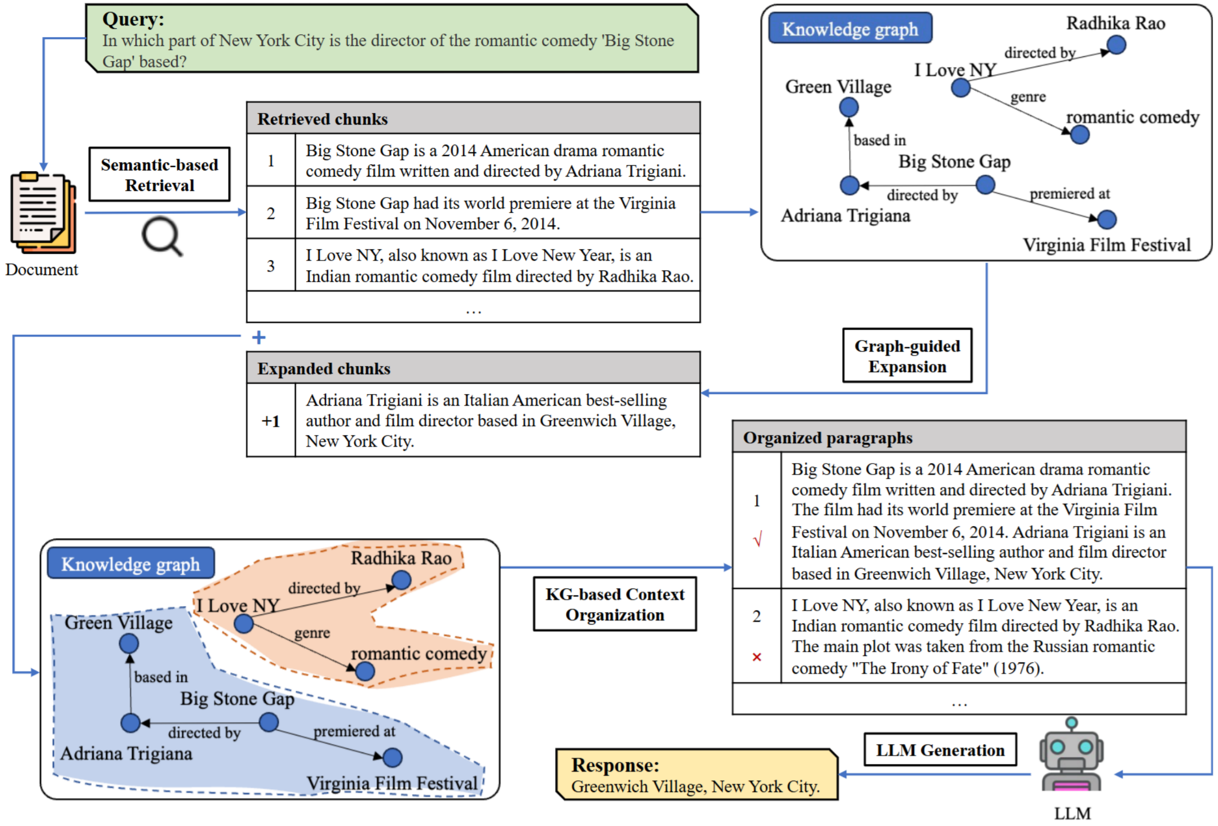
03 KG^{2}RAG (NAACL 2025)

- Traditional RAG retrieves isolated text chunks purely based on semantic similarity
- **Organize the retrieved text chunks** based on the structure of the extracted KG



03 KG^{2}RAG (NAACL 2025)

- Retrieve text chunks relevant to the query, and expand the chunks via subgraph extension
- Organize the resulting chunks using depth-first search on the retrieved subgraph



03 KG²RAG (NAACL 2025)

- Show **improvement over standard RAG and GraphRAG** approaches
- The **retrieval quality drops without organization**, while the **response quality drops without expansion**

Methods	Hotpot-Dist			Hotpot-Full			Shuffle-Hotpot-Dist			Shuffle-Hotpot-Full		
	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall
LLM-only	0.237	0.259	0.234	0.237	0.259	0.234	0.158	0.175	0.158	0.158	0.175	0.158
Semantic RAG	0.617	0.646	0.643	0.528	0.558	0.535	0.508	0.533	0.524	0.422	0.449	0.433
+ Rerank	0.652	0.685	0.665	0.587	0.613	0.603	0.532	0.560	0.546	0.447	0.476	0.456
Hybrid RAG	0.653	0.676	0.655	0.551	0.582	0.558	0.520	0.548	0.534	0.443	0.473	0.446
LightRAG	0.293	0.288	0.480	0.261	0.259	0.364	0.285	0.284	0.404	0.202	0.199	0.293
GraphRAG	0.400	0.408	0.491	0.169	0.157	0.429	0.351	0.365	0.401	0.163	0.155	0.362
KG^{2}RAG	**0.663**	**0.690**	**0.683**	**0.631**	**0.665**	**0.643**	**0.545**	**0.572**	**0.566**	**0.507**	**0.539**	**0.512**

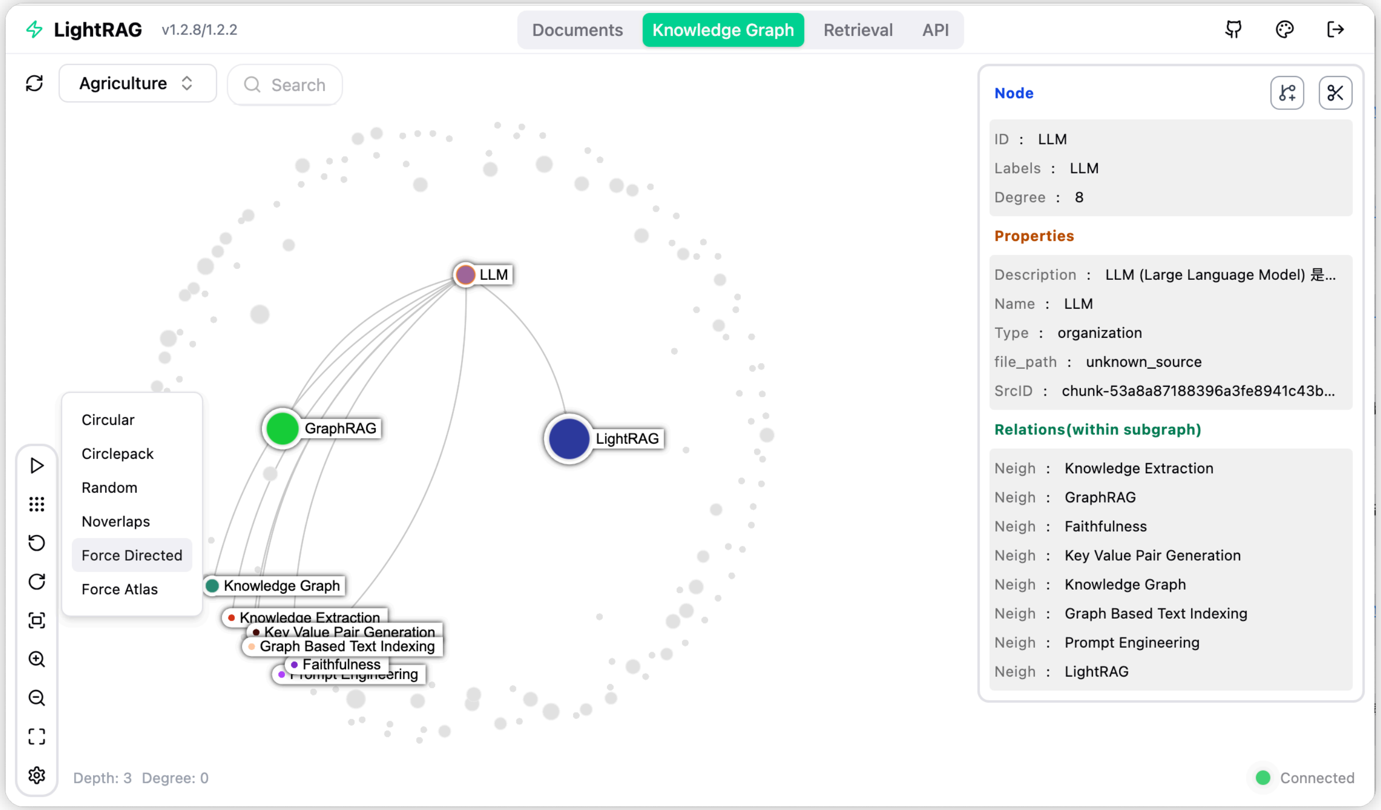
Table 1: Comparisons in terms of response quality between KG^{2}RAG and baselines.

Methods	Hotpot-Dist			Hotpot-Full			Shuffle-Hotpot-Dist			Shuffle-Hotpot-Full		
	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall
Semantic RAG	0.343	0.206	0.894	0.300	0.178	0.790	0.321	0.201	0.837	0.268	0.167	0.708
+ Rerank	0.357	0.224	**0.932**	0.306	0.197	0.833	0339	0.213	**0.886**	0.286	0.179	0.754
Hybrid RAG	0.354	0.222	0.921	0.302	0.189	0.795	0.334	0.210	0.837	0.279	0.174	0.739
LightRAG	0.234	0.150	0.638	0.132	0.083	0.340	0.227	0.148	0.535	0.116	0.073	0.295
GraphRAG	0.255	0.167	0.594	0.180	0.113	0.470	0.210	0.138	0.482	0.199	0.126	0.510
KG^{2}RAG	**0.436**	**0.301**	0.908	**0.310**	**0.203**	**0.838**	**0.405**	**0.279**	0.840	**0.305**	**0.193**	**0.790**

Table 2: Comparisons in terms of retrieval quality between KG^{2}RAG and baselines.

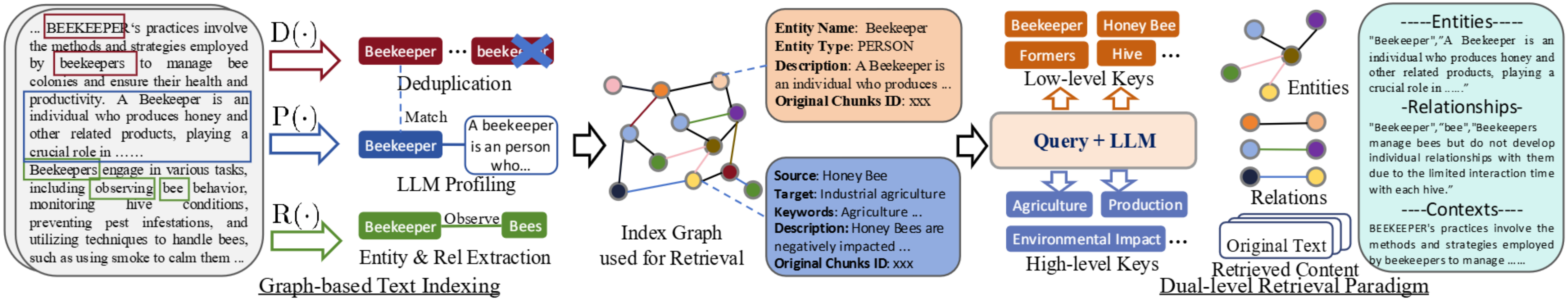
03 LightRAG (EMNLP 2025 Findings)

- Consider the adaptability of RAG systems to new data updates
- Provide a dual-level retrieval paradigm for comprehensive and cost-efficient retrieval



03 LightRAG (EMNLP 2025 Findings)

- Perform deduplication to optimize graph operations by reducing KG size
- Provide the entities, triplets, and chunks retrieved using both low- and high-level keywords



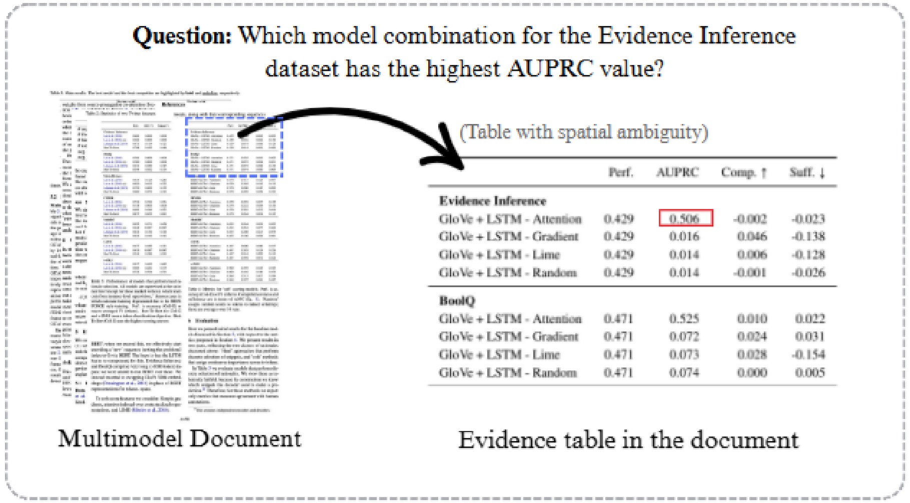
03 LightRAG (EMNLP 2025 Findings)

- Outperform the baseline RAG approaches across multiple criteria
- Show a significant reduction in LLM token costs compared to GraphRAG (arXiv 2024)

	Agriculture		CS		Legal		Mix	
	NaiveRAG	**LightRAG**	NaiveRAG	**LightRAG**	NaiveRAG	**LightRAG**	NaiveRAG	**LightRAG**
Comprehensiveness	32.4%	67.6%	38.4%	61.6%	16.4%	83.6%	38.8%	61.2%
Diversity	23.6%	76.4%	38.0%	62.0%	13.6%	86.4%	32.4%	67.6%
Empowerment	32.4%	67.6%	38.8%	61.2%	16.4%	83.6%	42.8%	57.2%
Overall	32.4%	67.6%	38.8%	61.2%	15.2%	84.8%	40.0%	60.0%
	RQ-RAG	**LightRAG**	RQ-RAG	**LightRAG**	RQ-RAG	**LightRAG**	RQ-RAG	**LightRAG**
Comprehensiveness	31.6%	68.4%	38.8%	61.2%	15.2%	84.8%	39.2%	60.8%
Diversity	29.2%	70.8%	39.2%	60.8%	11.6%	88.4%	30.8%	69.2%
Empowerment	31.6%	68.4%	36.4%	63.6%	15.2%	84.8%	42.4%	57.6%
Overall	32.4%	67.6%	38.0%	62.0%	14.4%	85.6%	40.0%	60.0%
	HyDE	**LightRAG**	HyDE	**LightRAG**	HyDE	**LightRAG**	HyDE	**LightRAG**
Comprehensiveness	26.0%	74.0%	41.6%	58.4%	26.8%	73.2%	40.4%	59.6%
Diversity	24.0%	76.0%	38.8%	61.2%	20.0%	80.0%	32.4%	67.6%
Empowerment	25.2%	74.8%	40.8%	59.2%	26.0%	74.0%	46.0%	54.0%
Overall	24.8%	75.2%	41.6%	58.4%	26.4%	73.6%	42.4%	57.6%
	GraphRAG	**LightRAG**	GraphRAG	**LightRAG**	GraphRAG	**LightRAG**	GraphRAG	**LightRAG**
Comprehensiveness	45.6%	54.4%	48.4%	51.6%	48.4%	51.6%	50.4%	49.6%
Diversity	22.8%	77.2%	40.8%	59.2%	26.4%	73.6%	36.0%	64.0%
Empowerment	41.2%	58.8%	45.2%	54.8%	43.6%	56.4%	50.8%	49.2%
Overall	45.2%	54.8%	48.0%	52.0%	47.2%	52.8%	50.4%	49.6%

03 RAG-Anything (arXiv 2025)

- Expand the GraphRAG settings to consider multimodal inputs such as images, tables, and equations
- Extend LightRAG by incorporating multimodal processors to generate cross-modal KGs



GPT-4o-mini🤔:
The model combination for the Evidence Inference dataset with the highest AUPRC value is the BERT-to-BERT model, achieving an AUPRC of **0.455.**

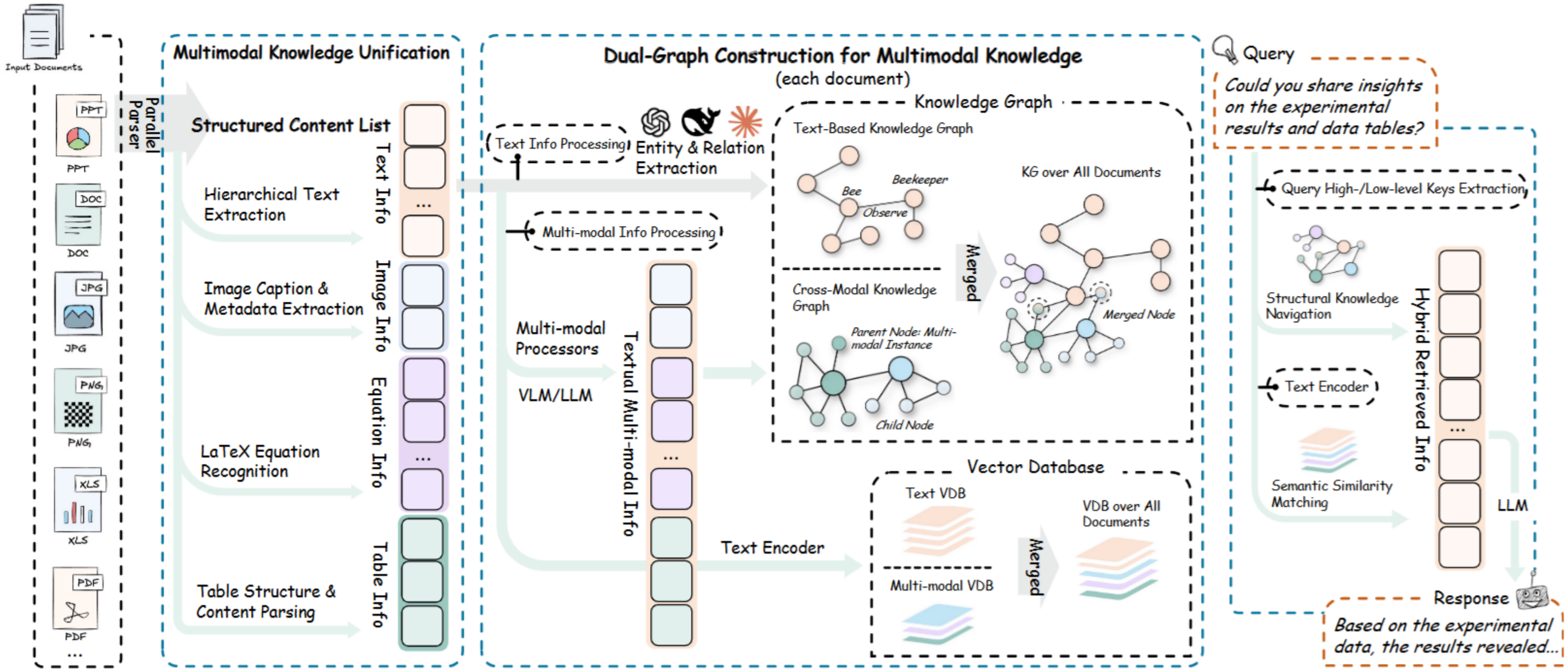
MMGraphRAG🤔:
The model combination for the Evidence Inference dataset with the highest AUPRC value is the "Bert-To-Bert" model, which achieved an AUPRC score of **0.708.**

LightRAG🤔:
The model combination "BERT + LSTM - Attention" has the highest AUPRC value for the Evidence Inference dataset at **0.429.**

RAG-Anything(Correct😉):
The model combination for the Evidence Inference dataset with the highest AUPRC value is GloVe + LSTM - Attention, achieving a value of **0.506.**

03 RAG-Anything (arXiv 2025)

- Convert multimodal inputs into **textual descriptions using VLMs and LLMs**
- Extract a **cross-modal KG** using the generated descriptions, and merge it with a text-based KG



03 RAG-Anything (arXiv 2025)

- Achieve high overall performance compared to GraphRAG-based approaches
- Show superior performance in information-dense domains such as research reports and finance

DocBench

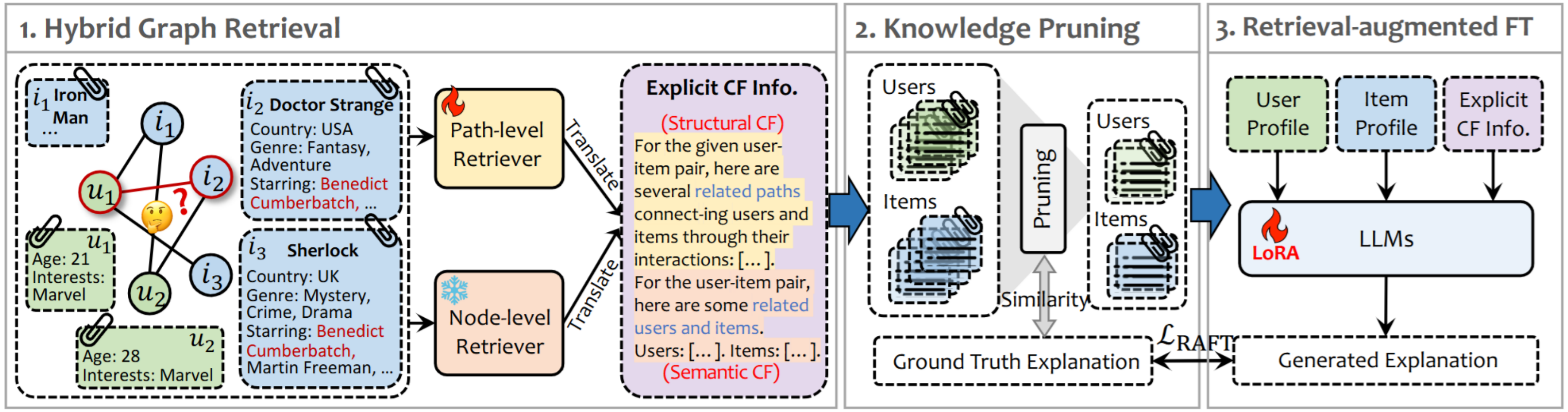
Method	Domains					Types			Overall
	Aca.	Fin.	Gov.	Law.	News	Txt.	Mm.	Una.	
GPT-4o-mini	40.3	46.9	60.3	59.2	61.0	61.0	43.8	**49.6**	51.2
LightRAG	53.8	56.2	59.5	**61.8**	**65.7**	**85.0**	59.7	46.8	58.4
MMGraphRAG	**64.3**	52.8	**64.9**	40.0	61.5	67.6	66.0	**60.5**	**61.0**
RAGAnything	**61.4**	**67.0**	**61.5**	**60.2**	**66.3**	**85.0**	**76.3**	46.0	**63.4**

MMLongBench

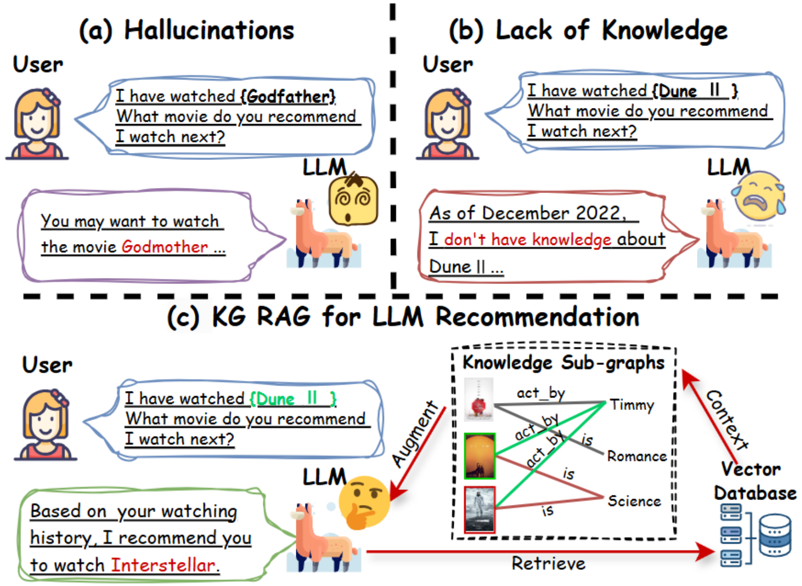
Method	Domains							Overall
	Res.	Tut.	Acad.	Guid.	Broch.	Admin.	Fin.	
GPT-4o-mini	35.5	**44.0**	24.6	33.1	29.5	**46.8**	31.1	33.5
LightRAG	**40.8**	34.1	**36.2**	**39.4**	**41.0**	44.4	38.3	**38.9**
MMGraphRAG	**40.8**	36.5	35.7	35.8	28.2	**46.9**	**38.5**	37.7
RAGAnything	**46.6**	**43.5**	**38.7**	**43.9**	**34.0**	45.7	**43.6**	**42.8**

04 Applications of GraphRAG

- Recommender Systems
  - G-Refer (WWW 2025)
    - Perform node-level and path-level retrieval to generate informative explanations for recommendations
  - K-RagRec (ACL 2025)
    - Retrieve subgraphs and utilize a GNN-based encoder to deliver the extracted knowledge to LLMs



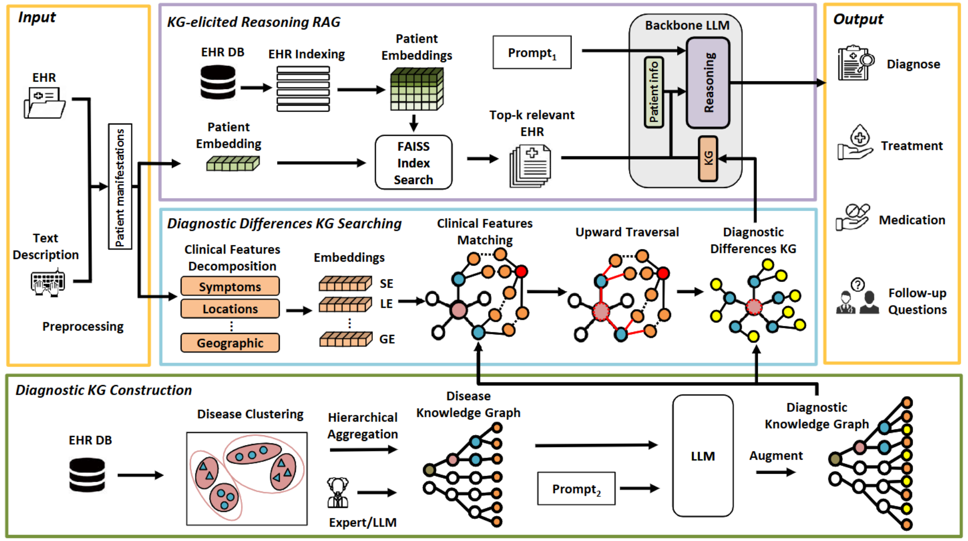
G-Refer (WWW 2025)



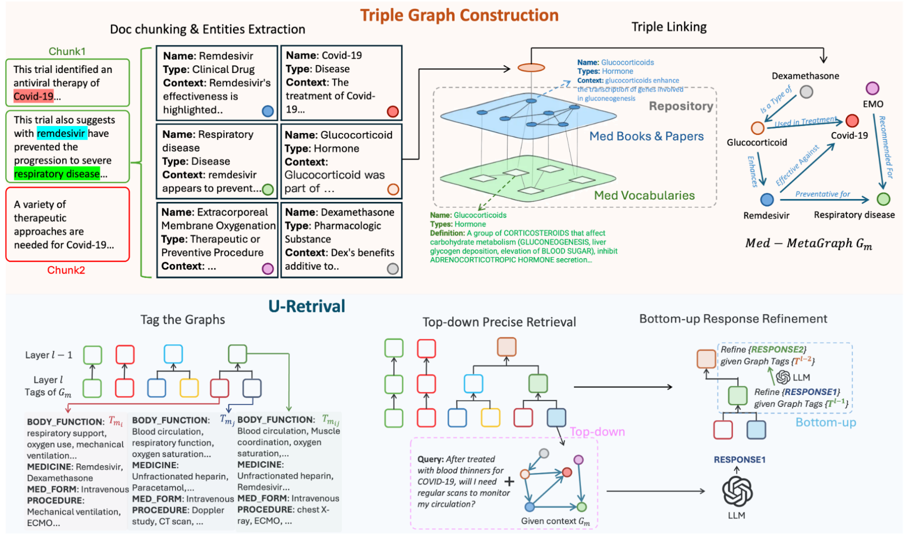
K-RagRec (ACL 2025)

04 Applications of GraphRAG

- Medical Reasoning
  - MedRAG (WWW 2025)
    - Construct a **diagnostic KG from an EHR database** to enhance healthcare copilots
  - MedGraphRAG (ACL 2025)
    - Convert **clinical texts into a KG** and adopt top-down retrieval & bottom-up refinement for precise reasoning



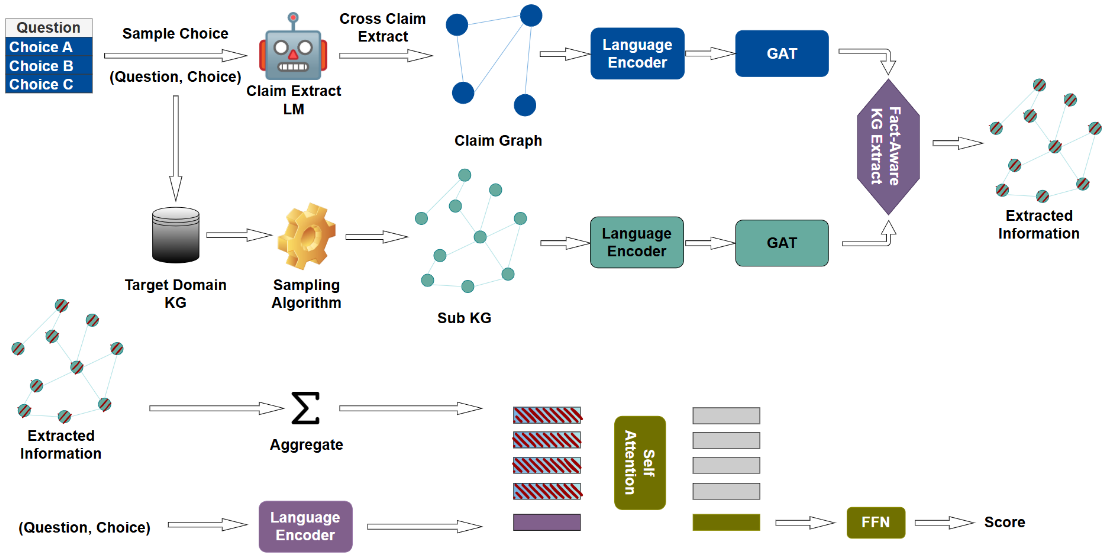
MedRAG (WWW 2025)



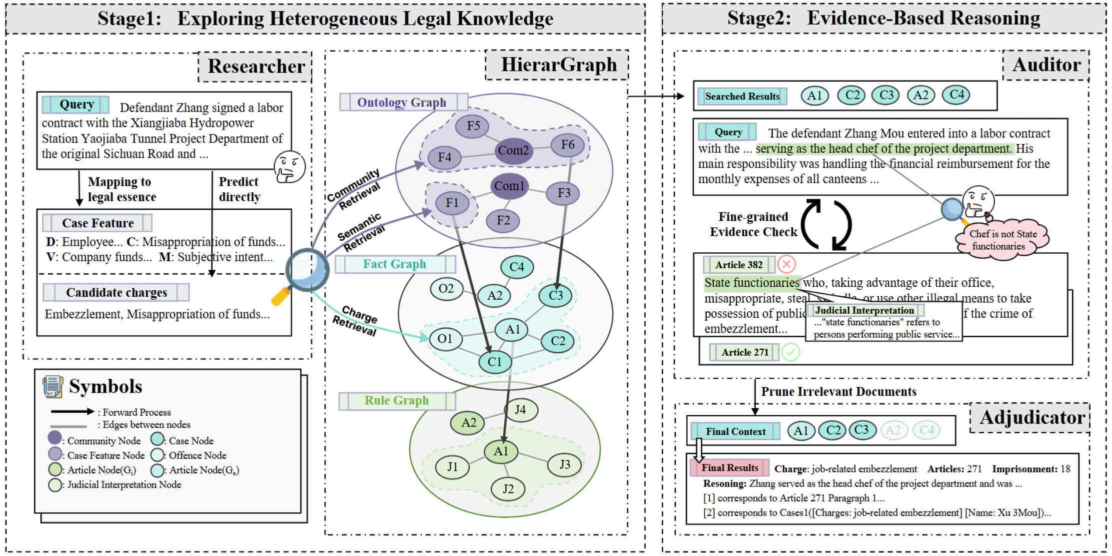
MedGraphRAG (ACL 2025)

04 Applications of GraphRAG

- Legal Domain Question Answering
  - GRAF (ACL 2025 Findings)
    - Apply GAT to both the claim graph and the legal domain KG to retrieve triplets for multiple-choice legal QA
  - LegalGraphRAG (ACL 2026)
    - Leverage ontology graph, fact graph, and rule graph to verify evidences of LLM reasoning



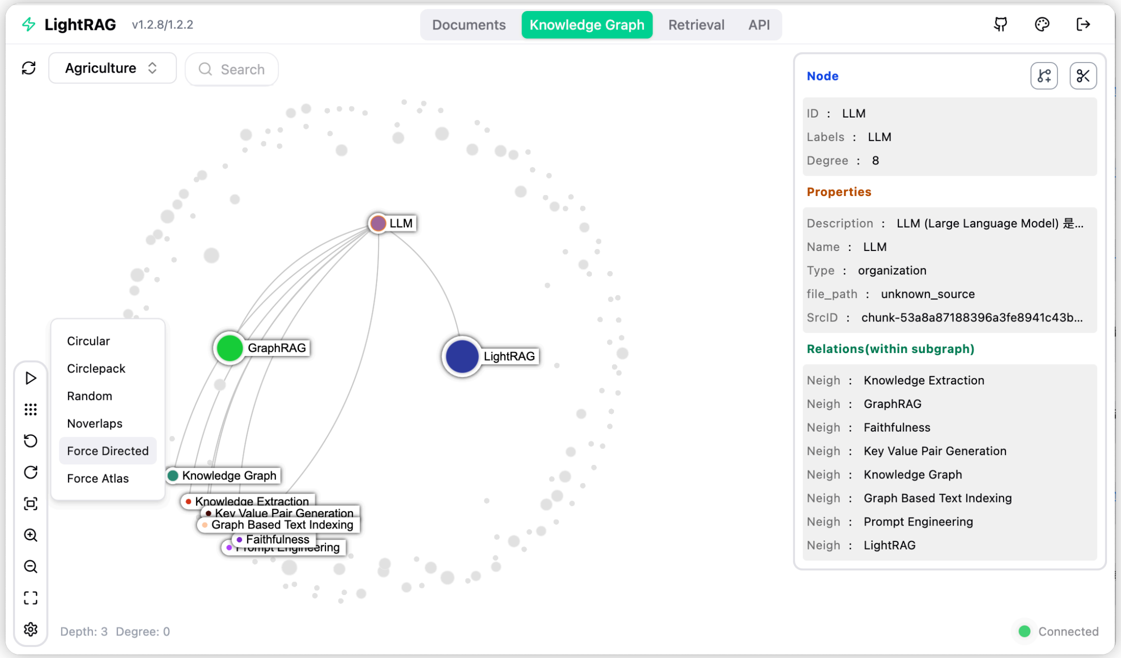
GRAF (ACL 2025 Findings)



LegalGraphRAG (ACL 2026)

05 GraphRAG Demonstration

- LightRAG Server
  - https://github.com/hkuds/lightrag
  - Provide an interactive Web UI and support API calls



LightRAG Server WebUI

```
# Clone LightRAG repo
git clone https://github.com/HKUDS/LightRAG.git
cd LightRAG

# Bootstrap the development environment
make dev
source .venv/bin/activate

# Equivalent manual steps with uv
uv sync --extra test --extra offline
source .venv/bin/activate

# Build front-end artifacts
cd lightrag_webui
bun install --frozen-lockfile
bun run build
cd ..

# Setup env file
make env-base

# Launch API-WebUI server
lightrag-server
```

Setting up LightRAG Server

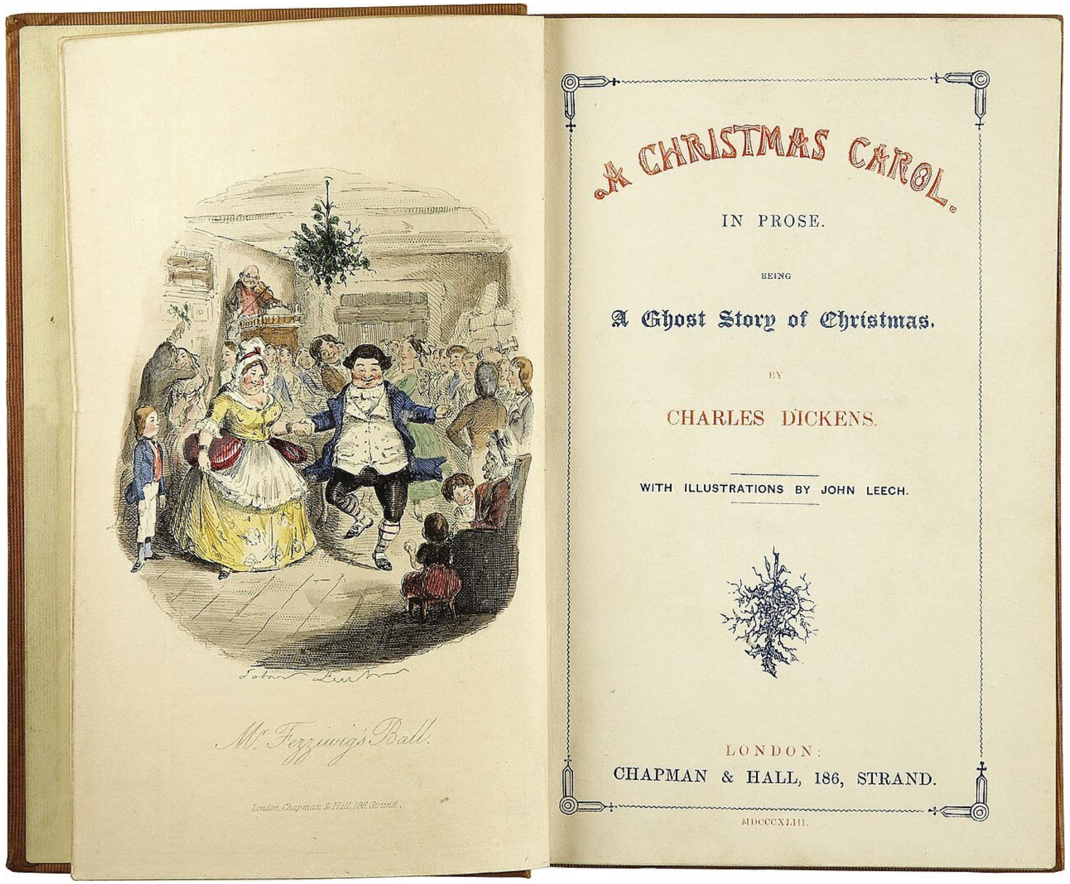
05 Input Corpus

- A Christmas Carol
  - A novella written by Charles Dickens, published in 1843
  - Publicly available via Project Gutenberg
  - 166 pages, 189,033 characters

Marley was dead: to begin with. There is no doubt whatever about that. The register of his burial was signed by the clergyman, the clerk, the undertaker, and the chief mourner. Scrooge signed it: and Scrooge's name was good upon 'Change, for anything he chose to put his hand to. Old Marley was as dead as a door-nail.

Mind! I don't mean to say that I know, of my own knowledge, what there is particularly dead about a door-nail. I might have been inclined, myself, to regard a coffin-nail as the deadest piece of ironmongery in the trade. But the wisdom of our ancestors is in the simile; and my unhallowed hands shall not disturb it, or the Country's done for. You will therefore permit me to repeat, emphatically, that Marley was as dead as a door-nail.

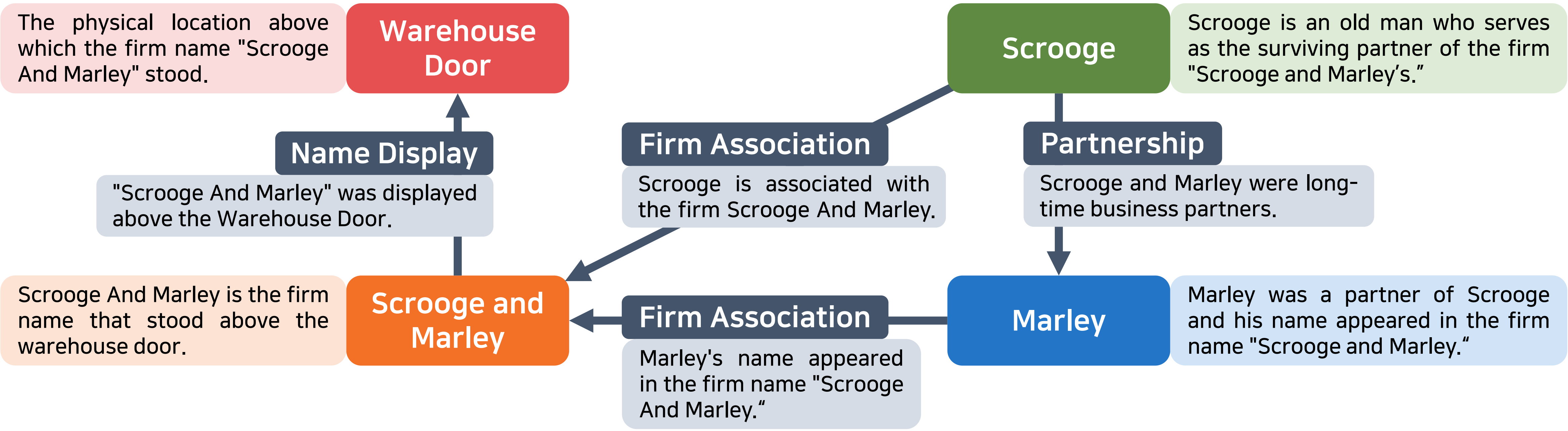
Sample Text



A Christmas Carol

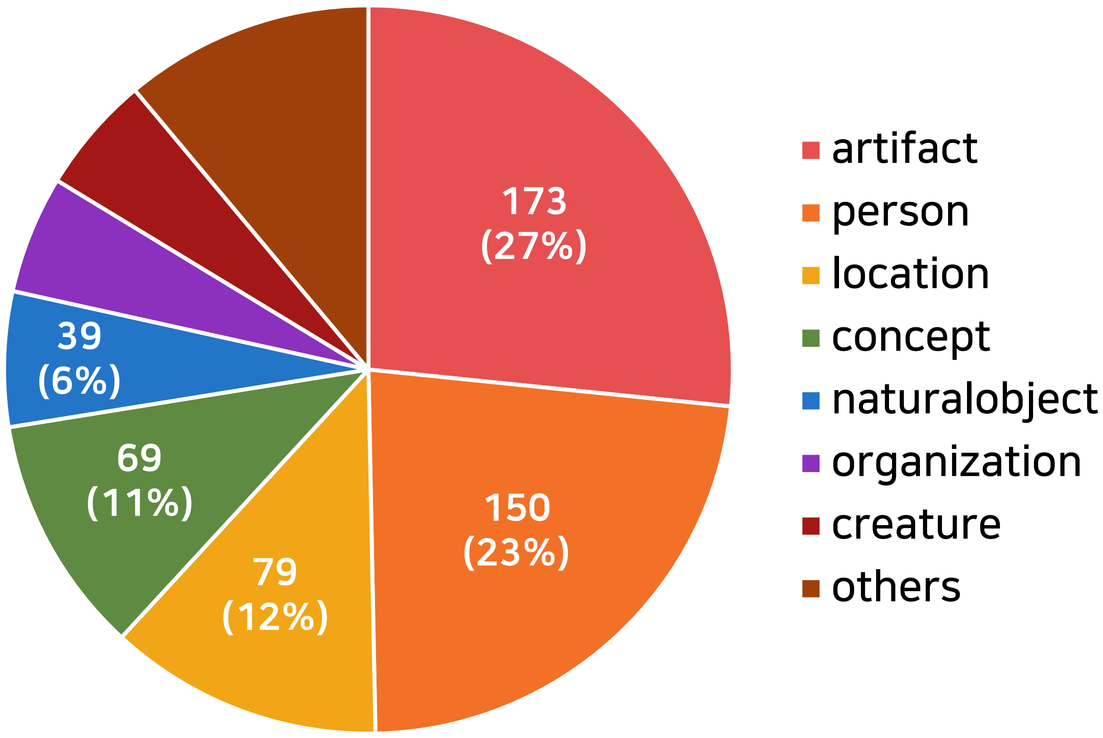
05 Example of Triplet Extraction

Scrooge never painted out Old Marley's name. There it stood, years afterwards, above the warehouse door: Scrooge and Marley. The firm was known as Scrooge and Marley. Sometimes people new to the business called Scrooge Scrooge, and sometimes Marley, but he answered to both names. It was all the same to him.

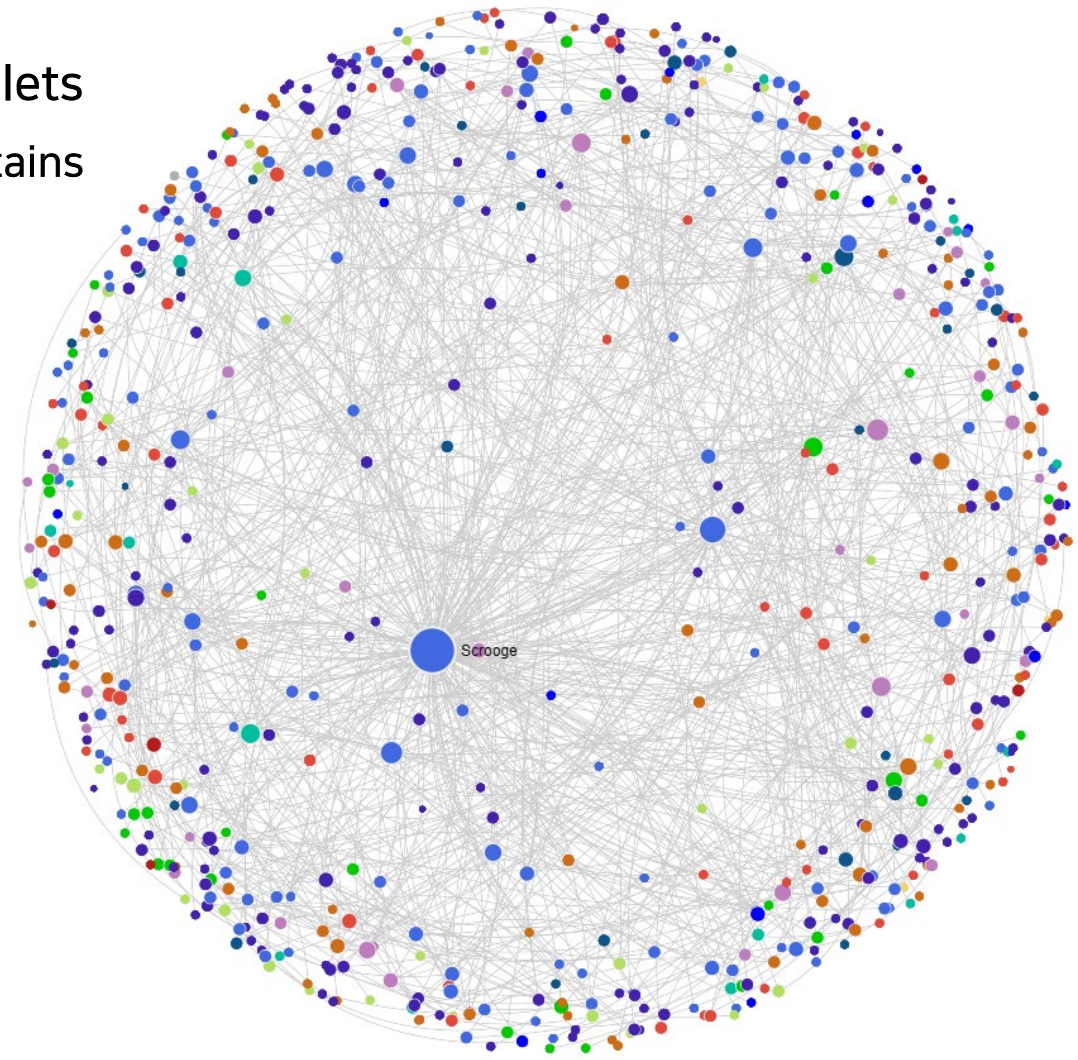


05 Statistics of the Generated KG

- The resulting KG consists of 650 entities and 863 triplets
  - The giant connected component of the resulting KG contains 575 entities and 822 triplets.



Entity Type Distribution



Visualization of the Generated KG

05 GraphRAG Example

What are the top themes in this story?

Retrieved Entities (partial):

Scrooge | Scrooge is a tight-fisted old man who serves as the surviving partner of the firm "Scrooge and Marley's." He is wealthy but stingy, dislikes family company, and is addressed by relatives as "Uncle Scrooge."
Spirit | The Spirit is a phantom figure that points, guides Scrooge to observe people and places, and leads Scrooge into an obscure part of town to reveal scenes relevant to the future.
Scrooge's Transformation | Scrooge's Transformation is the moral and behavioral change by which Scrooge becomes a kinder, more generous friend and master and is accepted by society.

Retrieved Triplets (partial):

(Scrooge, ceased interaction, Spirit) | The Spirit guides and travels with Scrooge, stands beside suffering people, leaves blessings during visits, and teaches Scrooge moral precepts.
(Scrooge, behavior change, Scrooge's Transformation) | Scrooge undergoes a moral transformation, becoming generous, amiable, and respected in the community.
(Scrooge, guidance, Ghost) | The Ghost guides Scrooge through past scenes and reveals memories that soften Scrooge and provoke weeping.

Retrieved Chunks (partial):

'Here is a new game,' said Scrooge. It was a game called Yes and No, where Scrooge's nephew had to think of something, and the rest must find out what, he only answering to their questions yes or no, as the case was.

05 GraphRAG Example

What are the top themes in this story?

Below are the principal themes as they appear in the provided text, with a short explanation.

Redemption and moral transformation: The central arc follows Scrooge's repentance and vow to change his life—he begs to "sponge away the writing on this stone" and promises "I will honour Christmas in my heart"—showing the story's focus on personal reform and renewed compassion.

Charity, generosity, and social responsibility: The Spirits instruct and model charitable behaviour, and Scrooge's eventual pledge to help others makes explicit the theme that individuals must care for their fellow humans.

Poverty, social injustice, and the human cost of neglect: The text depicts the Cratchit family's hardship, Tiny Tim's frailty, and the allegorical children Ignorance and Want; the Ghosts rebuke attitudes that treat the poor as a "surplus population," highlighting social critique and the moral urgency of alleviating need.

Memory, time, and the influence of past/present/future: The three Spirits show Scrooge scenes from across time—youthful experiences, present lives, and future outcomes—making reflection on one's life a sustained theme.

Supernatural agency as moral pedagogy: Ghostly visitations function not merely to frighten but to teach and guide Scrooge toward insight and action, using the supernatural to enable moral confrontation.

…

Each of these themes interlocks in the narrative: the supernatural revelations force memory and moral reckoning about poverty, family, and mortality, culminating in Scrooge's redemption and renewed commitment to generosity.

06 References

- Some slides are made based on the following references.

  - Y. Gao et al., “Retrieval-Augmented Generation for Large Language Models: A Survey”, arXiv, 2024.
  - B. Peng et al., “Graph Retrieval-Augmented Generation: A Survey”, ACM Transactions on Information Systems, 2025.
  - Q. Zhang et al., “A Survey of Graph Retrieval-Augmented Generation for Customized Large Language Models”, arXiv, 2025.
  - D. Edge et al., “From Local to Global: A Graph RAG Approach to Query-Focused Summarization”, arXiv, 2024.
  - B. Gutierrez et al., “HippoRAG: Neurobiologically Inspired Long-Term Memory for Large Language Models”, NeurIPS, 2024.
  - B. Gutierrez et al., “From RAG to Memory: Non-Parametric Continual Learning for Large Language Models”, ICML, 2025.
  - X. Zhu et al., “Knowledge Graph-Guided Retrieval Augmented Generation”, NAACL, 2025.
  - Z. Guo et al., “LightRAG: Simple and Fast Retrieval-Augmented Generation”, EMNLP 2025, Findings.
  - Z. Guo et al., “RAG-Anything: All-in-One RAG Framework”, arXiv, 2025.
  - Y. Li et al., “G-Refer: Graph Retrieval-Augmented Large Language Model for Explainable Recommendation”, WWW, 2025.
  - S. Wang et al., “Knowledge Graph Retrieval-Augmented Generation for LLM-based Recommendation”, ACL, 2025.
  - X. Zhao et al., “MedRAG: Enhancing Retrieval-augmented Generation with Knowledge Graph-Elicited Reasoning for Healthcare Copilot”, WWW, 2025.
  - J. Wu et al., “Medical Graph RAG: Evidence-based Medical Large Language Model via Graph Retrieval-Augmented Generation”, ACL, 2025.
  - C. Cracian et al., “GRAF: Graph Retrieval Augmented by Facts for Romanian Legal Multi-Choice Question Answering”, ACL Findings, 2025.
  - Z. Chen et al., “LegalGraphRAG: Multi-Agent Graph Retrieval-Augmented Generation for Reliable Legal Reasoning”, ACL, 2026.