

Use of Advanced RAG to Improve Fact-Checking Ability of Large Language Model

Jingyi Huang for AI Interest Group

21st, April, 2024

1. Brief Introduction of Fact-checking Research

- Fact-checking for the Infodemic
- How Artificial Intelligence Can Help?
- COVID-19 Fact-checking with RAG: a pilot study

2. Use Advanced RAG to Improve LLM' s Fact-checking Ability

- Example 1: Merger Retriever(LOTR)
- Example 2: Sentence Window Retrieval
- Example 3: Knowledge Graph for RAG

“The Infodemic”

- A rapid and far-reaching spread of both accurate and inaccurate information about certain issues, often including rumors, stigma, and conspiracy theories, makes it more complex to measure and interpret.
- Originally coined in 2003, rose to prominence during COVID-19.

References: Simon, F. M., & Camargo, C. Q. (2021). Autopsy of a metaphor: The origins, use and blind spots of the ‘infodemic’ . *New Media & Society*. <https://doi.org/10.1177/14614448211031908>

An unprecedented global infodemic during COVID-19

- Lead to confusion about the benefits of medical and public health intervention
- Impact on risk-taking and health-seeking behaviors
- Eroding trust in health authorities
- Compromising the effectiveness of public health responses and policies
-
- **And, be careful when using this term of metaphor.**

References: Wilhelm, E., Ballalai, I., Belanger, M. E., Benjamin, P., Bertrand-Ferrandis, C., Bezbaruah, S., Briand, S., Brooks, I., Bruns, R., Bucci, L. M., Calleja, N., Chiou, H., Devaria, A., Dini, L., D'Souza, H., Dunn, A. G., Eichstaedt, J. C., Evers, S. M. A. A., Gobat, N., Gissler, M., ... Purnat, T. D. (2023). Measuring the Burden of Infodemics: Summary of the Methods and Results of the Fifth WHO Infodemic Management Conference. JMIR infodemiology, 3, e44207. <https://doi.org/10.2196/44207>

Fact-checking for Infodemic or Beyond “Fact-checking”

- Does the information really transmit like a virus ? No.
- Complex origin/components, intentional spreading, active communication
- True or false? and why? - Reliable Guidance
- Cost few and accessible for everyone - Reduce the burden
- Quick & Multi-round responses - Interaction

How AI can help?

- As a tool?

- Query&Answer
- Measure/Predict
- Causal Inference
-

- As an agent?

- Generate
- Spread
- Interact
-

- As an invisible system/rule?

- Recommendation Algorithm
- Censorship Algorithm
-

COVID-19 Fact-checking with RAG: a pilot study

- Over 110,000 research articles from peer-viewed journals from 2020 to 2024
- 500 claims from real data(50% true, 50% false)
- 500 claims from generated data(50% true, 50% false)
- Advanced RAG techniques: Chunking, Vectorization, Merger Retriever

Query 1: Zinc water won't cure the novel coronavirus infection, even though zinc water contains quinine and the synthetic hydroxychloroquine of quinine.

[real data][TRUE]

Response(from RAG-GPT-4):

True. Zinc water itself does not cure COVID-19, and its content of quinine, if any, does not justify its efficacy against the virus. Hydroxychloroquine, a synthetic derivative of quinine, has shown mixed results and is not confirmed as a definitive cure for COVID-19. Referenced from "Hydroxychloroquine in the COVID-19 pandemic era: in pursuit of a rational use for prophylaxis of SARS-CoV-2 infection" by M. Infante et al.

Query 2: Studies have shown that the SARS-CoV-2 virus cannot survive for long on different surfaces. [generated data][FALSE]

Response(from RAG-GPT-4):

False. The provided context suggests that SARS-CoV-2 can survive on various surfaces for different durations, ranging from a few hours to several days, depending on environmental conditions and surface types.

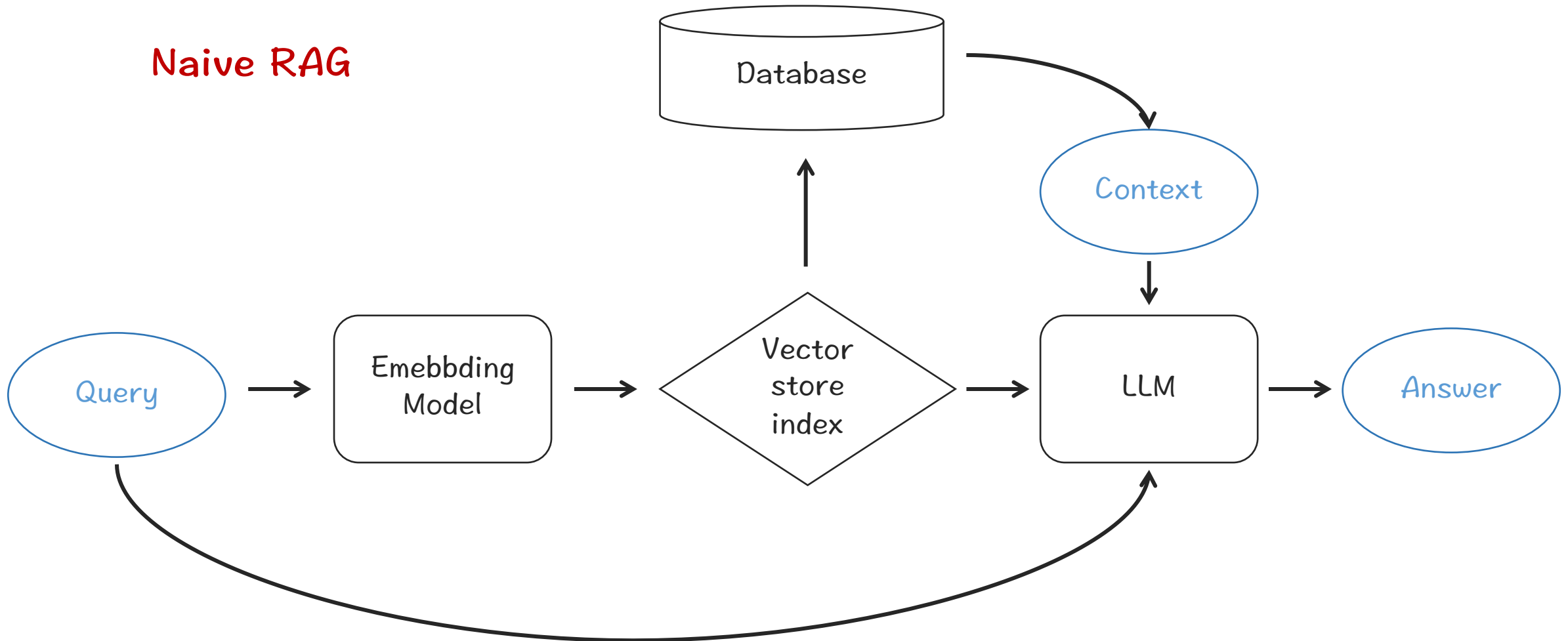
The context is **referenced from** "Survival of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-COV-2) and Influenza Virus on human skin: Importance of Hand Hygiene in Coronavirus Disease 2019 (COVID-19)" **by R. Hirose et al.**, and "Detection of sars-cov-2 on surfaces in households of persons with covid-19" **by P. Marcenac et al., among others.**

COVID-19 Fact-checking with RAG: a pilot study

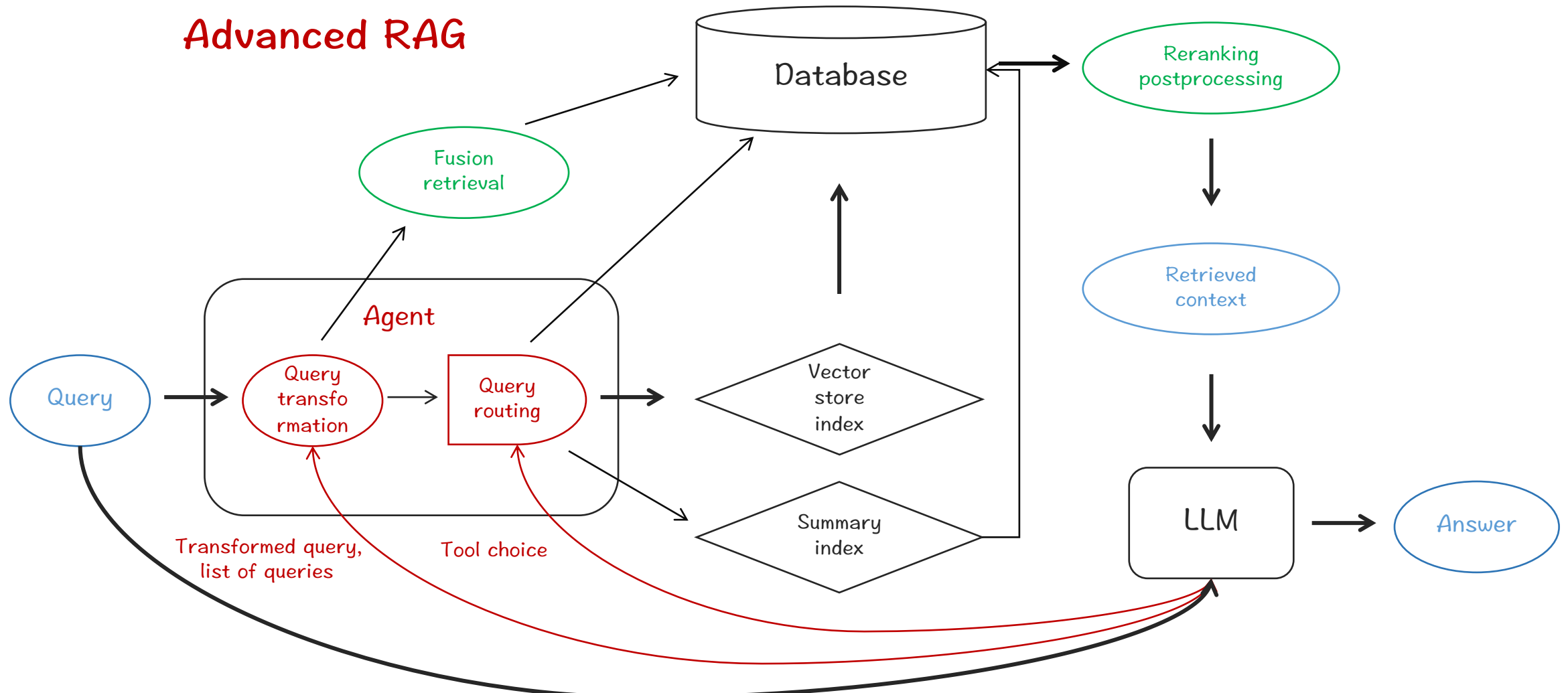
Dataset	Model	Accuracy	Answer relevancy
Real data	GPT-4	0.858	/
	RAG-GPT-4(Merger Retriever)	0.946	0.741
Generated data	GPT-4	0.973	/
	RAG-GPT-4(Merger Retriever)	0.986	0.836

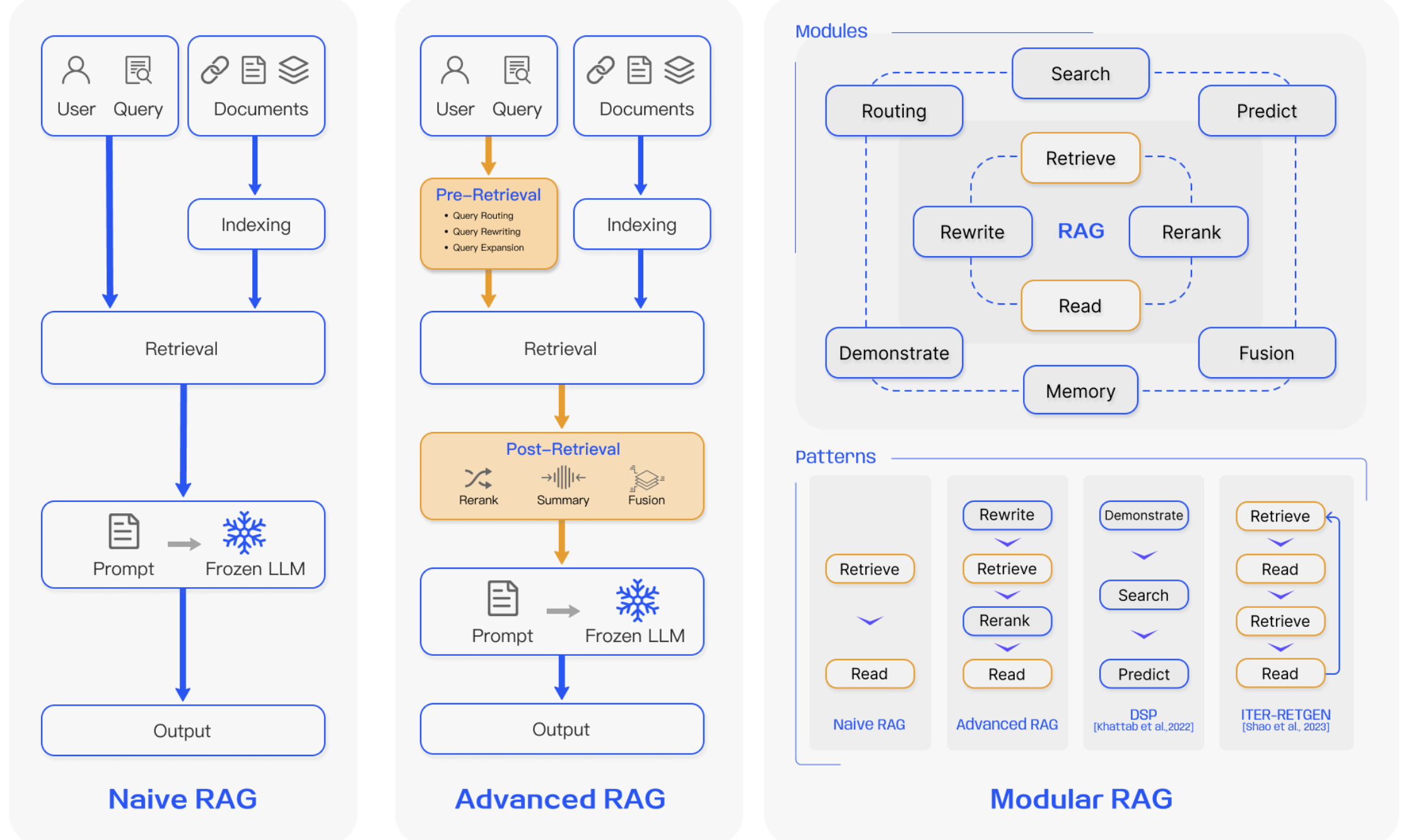
Use Advanced RAG to Improve LLM's Fact-checking Ability

Naive RAG



Use Advanced RAG to Improve LLM's Fact-checking Ability





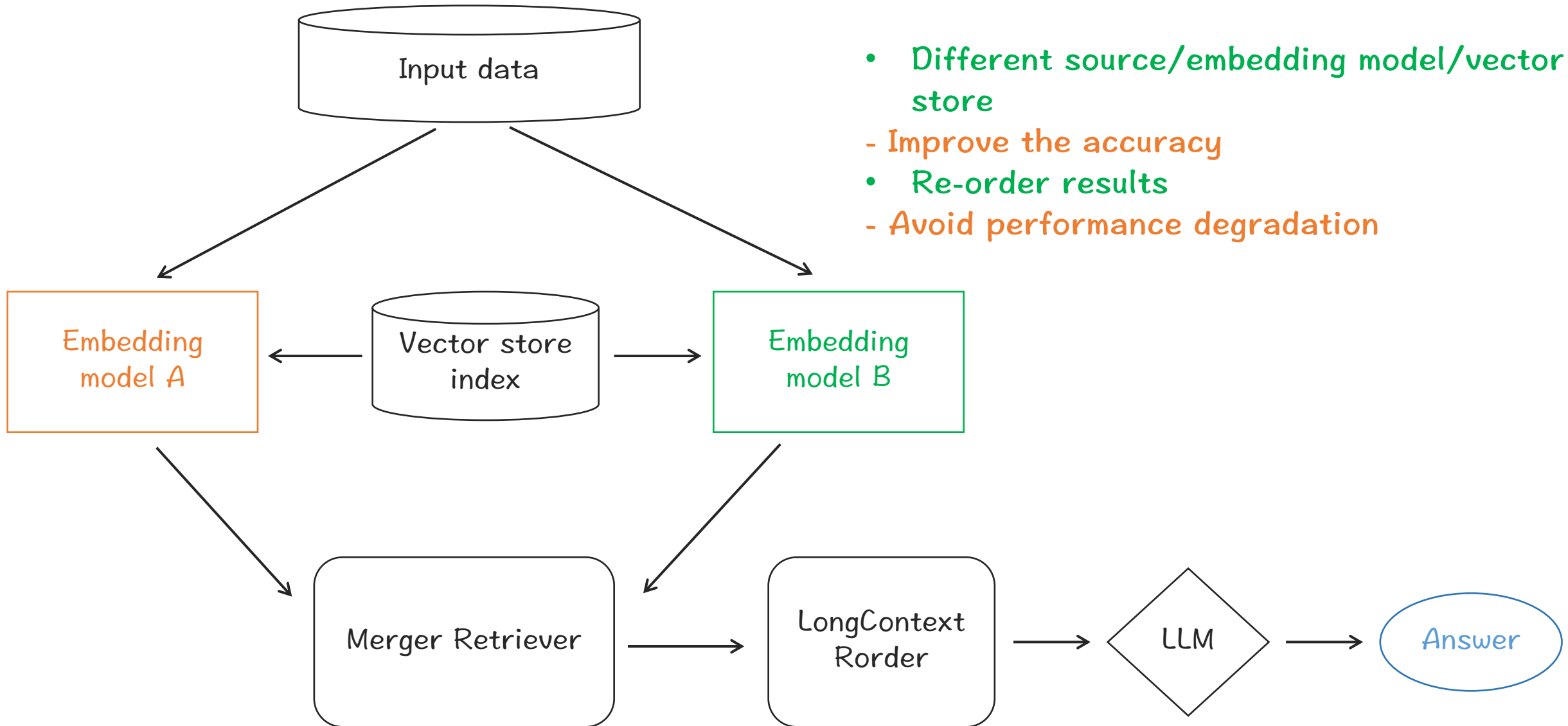
Reference: Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, M., & Wang, H. (2023). Retrieval-Augmented Generation for Large Language Models: A Survey. ArXiv.

Advanced RAG Techniques: Merger Retriever

- Combine the results of multiple retrievers
- Rank the results of the different retrievers
- “Lost in the Middle” ?

Merger Retriever

- Different source/embedding model/vector store
- Improve the accuracy
- Re-order results
- Avoid performance degradation



References: Liu, N. F., Lin, K., Hewitt, J., Paranjape, A., Bevilacqua, M., Petroni, F., & Liang, P. (2023). Lost in the Middle: How Language Models Use Long Contexts.

ArXiv. /abs/2307.03172

* Lost in the Middle: How Language Models Use Long Contexts

- Perform better when information occurs at the beginning or end of long context;
- Significantly degrades when access information in the **middle** of long contexts.

References: Liu, N. F., Lin, K., Hewitt, J., Paranjape, A., Bevilacqua, M., Petroni, F., & Liang, P. (2023). Lost in the Middle: How Language Models Use Long Contexts. ArXiv. /abs/2307.03172

Advanced RAG Techniques: Sentence Window Retrieval

- Each sentence in a document is embedded separately
- Extend the context window before and after the retrieved sentence
- Smaller chunk size, higher relevancy, but lower response quality?

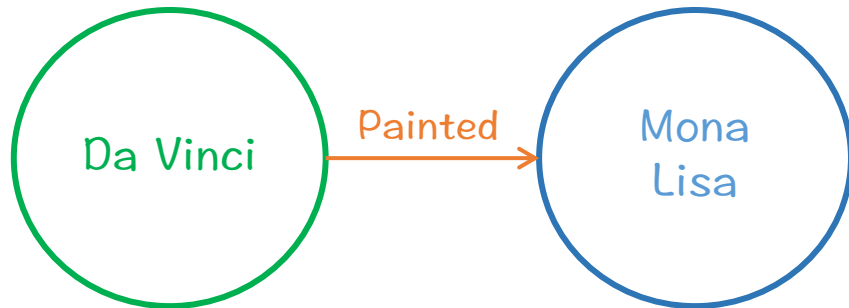
It is the nightly custom of every good mother after her children are asleep to rummage in their minds and put things straight for next morning, repacking into their proper places the many articles that have wandered during the day. If you could keep awake (but of course you can't) you would see your own mother doing this, and you would find it very interesting to watch her. It is quite like tidying up drawers. You would see her on her knees, I expect, lingering humorously over some of your contents, wondering where on earth you had picked this thing up, making discoveries sweet and not so sweet, pressing this to her cheek as if it were as nice as a kitten, and hurriedly stowing that out of sight.

Sentence Window Retrieval

- Context around the chunk(sentence) added
- Smaller chunk size, higher relevancy, ensured context and good quality

Advanced RAG Techniques: Knowledge Graph for RAG

- A collection of nodes and edges that represent entities or concepts, and their relationships, such as facts, properties, or categories.
- Structured representations of real-world entities and the relationships between them
- Confusing in complex relationships?

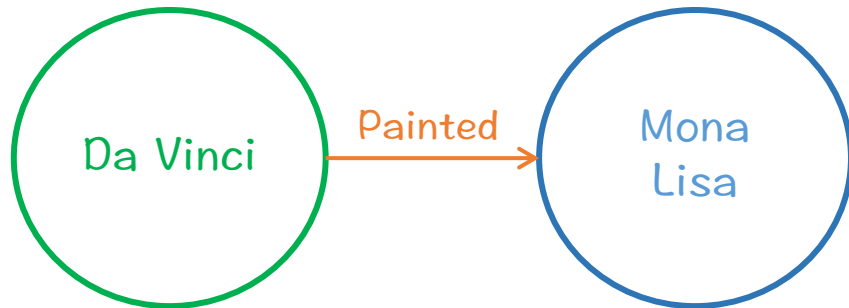


Knowledge Graph

- Node/Vertex
- Line/edge

A Triplet in KG

- Subject
- Object
- Predicate



Graph Store

- Relationship
- Structured Data
- Rich context
- Reasoning
- Diverse knowledge

Vector Store

- Similarity
- Unstructured Data

Thanks for listening.

Jingyi