

MARIO: Multi-Agent Retrieval and Iterative Reasoning for Knowledge-Augmented Generation in Scientific Question Answering

Maodi Hu^{1,2,3,4}[0009–0003–2207–5134] and Donghuan Song^{1,3,4}

¹National Science Library, Chinese Academy of Sciences

²Department of Information Resources Management, School of Economics and Management, University of Chinese Academy of Sciences

³Beijing Key Laboratory of Scientific Artificial Intelligence Research Corpora Data Construction and Governance

⁴Key Laboratory of New Publishing and Knowledge Services for Scholarly Journals, National Press and Publication Administration
{humaodi, songdonghuan}@mail.las.ac.cn

Abstract. Retrieval-Augmented Generation (RAG) enhances Large Language Models (LLMs) with external knowledge, yet conventional frameworks exhibit deficient multi-hop reasoning and opaque inference in scientific question answering. This paper introduces a multi-agent collaborative framework that integrates iterative query refinement with evidence-grounded reasoning to overcome these limitations. The system employs five specialized agents performing closed-loop retrieval and multi-hop knowledge graph traversal to extract dispersed relational facts from academic literature, while constructing explicit chain-of-thought reasoning anchored on structured triplets and verifiable citations to ensure interpretability and factual accuracy. We contribute a lncRNA-focused biomedical QA dataset for domain-specific evaluation. Extensive experiments demonstrate consistent superiority over standard RAG baselines, with ablation analyses validating the critical role of each core component. This work presents a practical, trustworthy solution for knowledge-intensive scientific reasoning tasks.

Keywords: Retrieval-Augmented Generation · Multi-Agent Systems · Knowledge Graph · Chain-of-Thought Reasoning · Scientific Question Answering

1 Introduction

Despite the remarkable capabilities of LLMs in natural language comprehension and content generation [5], their inherent limitations, such as static knowledge cutoffs, temporal obsolescence, and factual hallucination, substantially impede their application in rigorous scientific domains. RAG mitigates these limitations by incorporating external document resources as complementary evidence, constraining LLM generation within retrieved contexts [2]. Conventional RAG

pipelines adopt a single-pass paradigm: sparse or dense retrieval obtains relevant text fragments, which are concatenated with user queries for direct answer generation. While lightweight and efficient, this approach has two fundamental defects when confronting complex scientific QA.

First, single-pass retrieval lacks progressive knowledge discovery. Scientific questions often require multi-step association between intermediate entities, hidden relational facts, and indirect inference pathways [14]. Fixed query expression and one-time evidence acquisition cannot dynamically adjust retrieval intention according to acquired partial facts, resulting in incomplete evidence coverage. Although graph-enhanced RAG and multi-hop methods [1] have made progress, most rely on static knowledge structures or independent retrieval calls, lacking a unified iterative closed-loop mechanism for information gap compensation.

Second, existing RAG systems lack standardized interpretable reasoning. Most models directly generate answers from mixed unstructured texts without sorting logical relationships or outputting explicit reasoning processes. This black-box mode prevents researchers from tracing evidence sources and verifying inference rationality [8]. Current CoT research [11,13] mainly targets general-domain reasoning, with few works integrating structured evidence constraints and traceable citation into multi-agent RAG for professional domains. While GraphRAG [1] incorporates structured graph knowledge, its community-based summarization still lacks fine-grained, step-by-step evidence tracing anchored on explicit relational triplets.

Multi-agent collaboration has gained traction as a key strategy for task decomposition via role-specialized agents [6,15]. While current frameworks adopt distinct modules for query understanding, retrieval, ranking, and generation, their optimization efforts are largely directed toward multimodal scenarios and generic knowledge tasks, with substantive multi-hop reasoning in specialized scientific domains remaining largely unaddressed.

To address these challenges, this paper proposes MARIO, a multi-agent collaborative iterative reasoning framework for domain-specific knowledge-intensive scientific QA. Our core contributions are:

1. **Iterative Query Refinement and Closed-Loop Multi-Hop Retrieval:** A refinement-driven mechanism dynamically updates queries based on accumulated evidence. Multi-round KG traversal compensates for single-pass retrieval gaps, capturing indirect cross-document relational facts.
2. **Structured Evidence-Grounded Interpretable CoT Reasoning:** Unlike context-mixed implicit reasoning, our framework explicitly constructs a transparent reasoning chain by synthesizing validated relational triplets and their corresponding literature evidence.
3. **Domain-Specific Scientific QA Benchmark:** We construct an expert-reviewed dataset of high-quality multi-hop QA pairs from peer-reviewed publications, providing a standardized testbed for domain-adaptive RAG research.

2 Methodology

2.1 System Overview

MARIO operates in two stages: *offline domain KG construction* and *online multi-agent iterative reasoning*. The offline stage performs entity recognition, relational triplet extraction, and noise filtering on collected papers, building a domain knowledge graph $\mathcal{G} = (\mathcal{E}, \mathcal{T})$. The online stage takes the user’s scientific question as input and schedules five specialized agents:

- **Coordinator Agent:** Maintains the global evidence pool $\mathcal{T}_{\text{accum}}$, controls the iteration loop, judges termination conditions, and coordinates sub-agent calling sequences.
- **Entity Agent:** Extracts core entities from iterative queries and completes disambiguation via dense semantic matching against KG entities.
- **Relation Agent:** Traverses the KG from matched seed entities, retrieves and re-ranks relational triplets through hybrid scoring.
- **Refinement Agent:** Generates context-aware iterative queries based on accumulated evidence, and assesses evidence sufficiency for early termination.
- **Reasoning Agent:** Synthesizes all accumulated evidence, constructs explicit CoT chains, and derives concise final answers.

Fig. 1 illustrates the overall architecture. While the agents operate in a sequential workflow, each is designed with specialized responsibilities and autonomous decision-making capabilities (e.g., the Refinement Agent independently assesses evidence sufficiency for early termination), positioning our system as a hybrid between a modular workflow and a fully autonomous multi-agent system. The system sets a maximum iteration round I_{max} ; in each iteration, new relational facts supplement the evidence pool and the query is progressively optimized. After loop termination, the Reasoning Agent integrates all structured evidence and literature context to complete traceable reasoning.

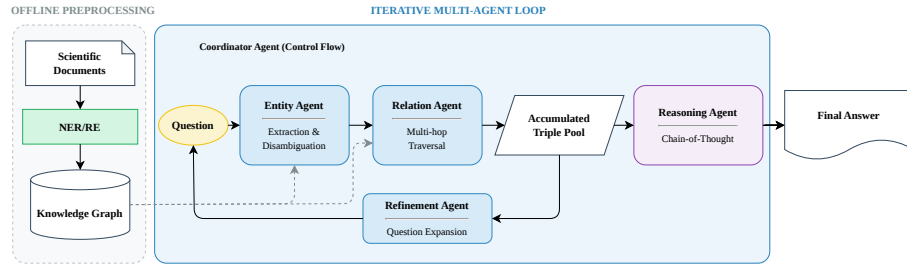


Fig. 1. MARIO Multi-Agent Architecture. The offline phase constructs a domain KG; the online phase executes an iterative loop (Entity $\rightarrow$ Relation $\rightarrow$ Refinement) to populate the triplet pool $\mathcal{T}_{\text{accum}}$, followed by the Reasoning Agent to generate CoT $\mathcal{C}$ and final answer a .

2.2 Offline Knowledge Graph Construction

Academic papers are segmented into sentence-level units for fine-grained extraction. Powerful LLMs, guided by domain-specific prompts, extract relational triplets formalized as (h, r, t, s, c) , where h denotes the head entity, r the semantic relation, t the tail entity, s the source metadata, and $c \in [0, 1]$ the factual confidence. Speculative assertions are assigned low confidence scores, whereas definitive conclusions receive high scores. Two post-processing operations—low-confidence filtering and cross-document deduplication—are applied to ensure knowledge graph quality. A global entity index is constructed to map each KG entity to its semantic type and neighboring triplets, thereby facilitating efficient online matching and traversal.

2.3 Multi-Agent Iterative Reasoning

Entity Agent. For the i -th query $q^{(i)}$ (with the initial query $q^{(0)}$ given as the original question), structured domain prompts extract core entities $\mathcal{E}_{\text{raw}}^{(i)}$. Dense semantic embedding unifies entity representation, and cosine similarity $\text{sim}(e, e_{\text{kg}})$ enables fuzzy matching. For each raw entity, top- K KG entities are selected as candidates, forming the disambiguated seed set $\mathcal{E}_{\text{matched}}^{(i)}$.

Relation Agent. Starting from seed entities, the agent retrieves direct one-hop triples as foundational evidence. Rather than performing forced multi-hop expansion in a single pass, our framework incrementally enriches indirect associations across iterative retrieval rounds via a progressive expansion strategy, which naturally realizes multi-hop reasoning. A hybrid scoring mechanism is then employed to balance factual confidence and semantic relevance.

$$\text{score}(t) = \alpha \cdot c(t) + (1 - \alpha) \cdot \text{sim}(t, q^{(i)}). \quad (1)$$

Only the top- M triples by integrated score enter the evidence pool.

Refinement Agent. Taking accumulated queries $\mathcal{Q}_{\text{accum}}^{(i)}$ and triplets $\mathcal{T}_{\text{accum}}^{(i)}$ as joint inputs, the LLM generates context-aware queries $q^{(i+1)}$ that inherit the original research goal while complementing missing reasoning perspectives. This agent also autonomously assesses evidence sufficiency. When facts are adequate for a rigorous answer, it activates early termination.

Coordinator Agent. The central controller initializes the empty evidence pool and sets the first query. Each iteration orderly executes entity matching, triplet retrieval, evidence deduplication and fusion, and query update. The loop terminates upon reaching I_{max} or receiving the early termination signal.

Algorithm 1 MARIO Multi-Agent Retrieval and Iterative Reasoning**Require:** Question $q^{(i)}$, KG $\mathcal{G} = (\mathcal{E}, \mathcal{T})$, max rounds $I_{\max}$, top- K , top- M **Ensure:** Reasoning chain $\mathcal{C}$ and final answer a

```

1:  $\mathcal{T}_{\text{accum}}^{(0)} \leftarrow \emptyset$ ;  $i \leftarrow 0$ ; terminate  $\leftarrow \text{False}$ 
2: while  $i \leq I_{\max}$  and not terminate do
3:    $\mathcal{E}_{\text{raw}}^{(i)} \leftarrow \text{Agent}_{\text{entity}}.\text{extract}(q^{(i)})$ 
4:    $\mathcal{E}_{\text{matched}}^{(i)} \leftarrow \bigcup_{e \in \mathcal{E}_{\text{raw}}^{(i)}} \text{Top-}K_{e_{\text{kg}} \in \mathcal{E}} \text{ sim}(e, e_{\text{kg}})$ 
5:    $\mathcal{T}^{(i)} \leftarrow \text{Top-}M_{t \in \text{Agent}_{\text{relation}}.\text{traverse}(\mathcal{E}_{\text{matched}}^{(i)}, \mathcal{G})} \text{ score}(t)$ 
6:    $\mathcal{T}_{\text{accum}}^{(i)} \leftarrow \text{deduplicate}(\mathcal{T}_{\text{accum}}^{(i-1)} \cup \mathcal{T}^{(i)})$ 
7:    $(q^{(i+1)}, \text{terminate}) \leftarrow \text{Agent}_{\text{refine}}.\text{generate}(\mathcal{Q}_{\text{accum}}^{(i)}, \mathcal{T}_{\text{accum}}^{(i)})$ 
8:    $i \leftarrow i + 1$ 
9: end while
10:  $\mathcal{C} \leftarrow \text{Agent}_{\text{reason}}.\text{generateCoT}(\mathcal{Q}_{\text{accum}}^{(i-1)}, \mathcal{T}_{\text{accum}}^{(i-1)})$ 
11:  $a \leftarrow \text{Agent}_{\text{reason}}.\text{deriveAnswer}(\mathcal{C})$ 
12: return  $(\mathcal{C}, a)$ 

```

Reasoning Agent. After loop termination, the agent integrates structured triplets, source text, and re-ranked literature context. It sorts evidence logically, establishes deduction relationships between relational facts, and forms a compact, evidence-marked reasoning chain $\mathcal{C}$. Finally, it condenses core conclusions into a concise scientific answer a .

The complete workflow is formalized in Algorithm 1. Viewed through an information-theoretic lens, the entropy of the question decreases with each iteration. Despite the inherent instability of LLM, the resulting sequence of information gains is nearly non-increasing, which essentially guarantees convergence of the accumulated evidence and provides a theoretical basis for capping the number of iterations at a fixed maximum.

3 Experiments and Results

3.1 Dataset and Evaluation

We constructed a scientific question-answering dataset focused on lncRNA, derived from peer-reviewed biomedical literature in PubMed that addresses lncRNA regulatory mechanisms and disease-related functions. Following full-text parsing, chunk-based segmentation, and noise filtering, we generated QA pairs using a large language model, followed by expert manual review. This process yielded 2,000 high-quality multi-hop question-answer pairs spanning eight research sub-domains (Fig. 2). Every answer is strictly grounded in evidence from the source literature.

We adopt LLM-based automatic evaluation [7] with three normalized $[0, 1]$ metrics: **Relevance** (answer-question matching), **Accuracy** (factual consistency with evidence), and **Completeness** (multi-hop reasoning coverage). Cross-random arrangement and repeated averaging ensure evaluation stability.

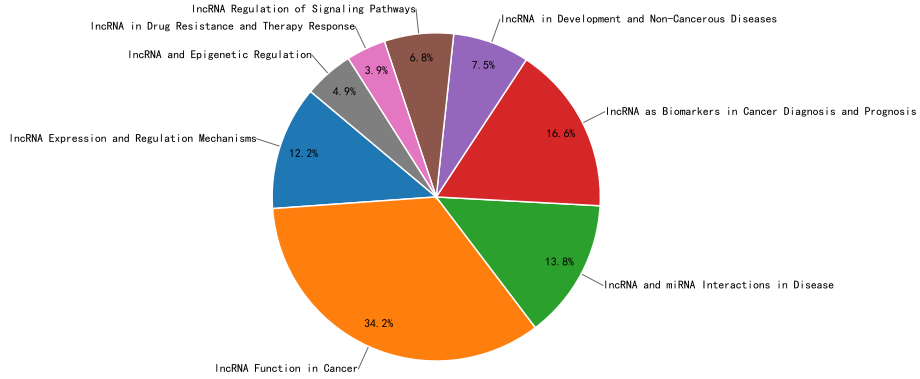


Fig. 2. Distribution of questions across eight biological categories.

3.2 Baselines and Implementation

Three single-pass RAG baselines are compared under unified settings: **BM25** (sparse lexical retrieval, top-5 fragments) [10], **Dense** (dense semantic retrieval, top-5) [3], **TwoStage-Rerank** (BM25+Dense fusion with cross-encoder re-ranking, top-5) [9], and **GraphRAG** [1]. All methods share the same LLM and environment. Experiments run on a server with 64 GB memory and a single NVIDIA A100 GPU. The generation model is Qwen3-32B via VLLM [4]; the embedding model is BAAI/bge-small-en-v1.5 [12]; the reranker is BAAI/bge-reranker-base.

3.3 Overall Performance and Ablation

Table 1 presents baseline comparison and ablation results jointly. Traditional baselines achieve relatively lower performance, with BM25 failing to capture deep semantic correlations, dense retrieval enhancing semantic matching but suffering from insufficient evidence filtering, and TwoStage-Rerank optimizing retrieval quality yet still not addressing multi-hop reasoning. GraphRAG, while leveraging structured graph knowledge, still relies on static graph traversal without iterative refinement, which limits its ability to progressively discover multi-hop evidence. MARIO consistently outperforms all baselines and demonstrates the superiority of multi-agent collaboration, iterative refinement, and KG-guided reasoning.

Four ablation variants validate component contributions. MARIO (FULL) includes CoT, structured triplets, triplet-related text content (TContent), hybrid-retrieved chunks (RContent), and iterative retrieval. AB1 removes RContent. AB2 discards all auxiliary text (CoT+triplets only). AB3 relies purely on LLM parametric knowledge with CoT. AB4 disables iteration. All variants show clear degradation, with removing RContent (AB1) reducing completeness, abandoning all text context (AB2) confirming that triplets and contextual evidence mutually

Table 1. Combined baseline comparison and ablation study. Best results are in **bold**.

Method	Relevance $\uparrow$	Accuracy $\uparrow$	Completeness $\uparrow$
BM25 [10]	0.902	0.790	0.751
Dense [3]	0.854	0.740	0.703
TwoStage-Rerank [9]	0.916	0.818	0.779
GraphRAG [1]	0.921	0.834	0.772
MARIO (FULL)	0.936	0.848	0.803
AB1 (CoT+Triplet+TContent)	0.928	0.827	0.780
AB2 (CoT+Triplet)	0.926	0.821	0.777
AB3 (CoT only)	0.923	0.817	0.773
AB4 (NoIteration)	0.927	0.826	0.774

reinforce reasoning, relying solely on LLM knowledge (AB3) yielding the lowest scores and validating the critical role of external grounding, and disabling iteration (AB4) causing noticeable drops and confirming that iterative query optimization enhances multi-hop relation capture.

3.4 Qualitative Case Analysis

Table 2 illustrates the iterative reasoning process on a MALAT1 alternative splicing question requiring progressive exploration of entity interactions, subcellular localization, and regulatory cascades.

Starting from the original question, MARIO performs three rounds of iterative refinement, gradually shifting from shallow entity retrieval to in-depth relational reasoning. With incremental evidence, the framework uncovers implicit interactions between MALAT1, SR proteins, and nuclear speckle domains that single-pass RAG cannot capture, ultimately producing a logically coherent, evidence-grounded answer.

4 Conclusion

This paper presents MARIO, a multi-agent framework for scientific QA that addresses insufficient multi-hop evidence and limited reasoning interpretability in conventional RAG. Five decoupled agents support dynamic query refinement, iterative evidence collection, and traceable reasoning anchored on structured triplets. Experiments and ablation studies validate the efficacy of each module.

Current limitations include the fixed iteration hyperparameter, reliance on a single LLM for reasoning, and validation restricted to the biomedical lncRNA domain. Future work will develop difficulty-aware adaptive iteration strategies, explore alternative reasoning LLMs, and extend the framework to broader scientific fields with cross-domain evaluation datasets, enhancing general applicability and reasoning capacity.

Table 2. Inference Walkthrough for the MALAT1 Question

Round	Content
Round 1	Query: What is the role of MALAT1 in the regulation of alternative splicing? Entity Extraction: MALAT1, alternative splicing ... triplet Retrieval: (MALAT1, regulate, SR protein), (MALAT1, bind, transcript), ...
Round 2	Query: How does MALAT1 interact with SR proteins to influence the regulation of alternative splicing? Entity Extraction: MALAT1, SR protein, SRSF1, ... triplet Retrieval: (MALAT1, interact, SR protein), (MALAT1, bind, SRSF1), ...
Round 3	Query: What is the mechanism by which MALAT1 recruits SR proteins to nuclear speckle domains and affects alternative splicing? Entity Extraction: MALAT1, nuclear speckle, phosphorylated SR protein, ... triplet Retrieval: (MALAT1, localize, nuclear speckles), (MALAT1, regulate, phosphorylated SR protein), ...
Reasoning & Final Answer	CoT: ...MALAT1 modulates the levels and distribution of active SR proteins in nuclear speckles, thus controlling alternative splicing outcomes. ... Answer: MALAT1 regulates alternative splicing by modulating the localization and functional availability of phosphorylated SR splicing factors within nuclear speckles.

Acknowledgments

This work was supported by the National Key R&D Program of China (No. 2023YFF0725600).

References

1. Edge, D., Trinh, H., Cheng, N., Bradley, J., Chao, A., Mody, A., Truitt, S., Larson, J.: From local to global: A graph RAG approach to query-focused summarization. arXiv preprint arXiv:2404.16130 (2024)
2. Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Guo, Q., Wang, M., Wang, H.: Retrieval-augmented generation for large language models: A survey. arXiv preprint arXiv:2312.10997 (2023)
3. Karpukhin, V., Oğuz, B., Min, S., Lewis, P., Wu, L., Edunov, S., Chen, D., Yih, W.t.: Dense passage retrieval for open-domain question answering. In: Proceedings of EMNLP. pp. 6769–6781 (2020)
4. Kwon, W., Li, Z., Zhuang, S., Sheng, Y., Zheng, L., Yu, C.H., Gonzalez, J.E., Zhang, H., Stoica, I.: Efficient memory management for large language model serving with PagedAttention. In: Proceedings of the ACM SIGOPS 29th Symposium on Operating Systems Principles. pp. 611–626 (2023)

5. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.t., Rocktäschel, T., et al.: Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems* pp. 9459–9474 (2020)
6. Liu, P., Liu, X., Yao, R., Liu, J., Meng, S., Wang, D., Ma, J.: HM-RAG: Hierarchical multi-agent multimodal retrieval augmented generation. In: *Proceedings of the 33rd ACM International Conference on Multimedia* (2025)
7. Liu, Y., Iter, D., Xu, Y., Wang, S., Xu, R., Zhu, C.: G-eval: NLG evaluation using gpt-4 with better human alignment. In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. pp. 2511–2522. Association for Computational Linguistics, Singapore (2023)
8. Lyu, Q., Havaladar, S., Stein, A., Zhang, L., Rao, D., Wong, E., Apidianaki, M., Callison-Burch, C.: Faithful chain-of-thought reasoning. In: *Proceedings of the 13th International Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics*. pp. 305–329 (2023)
9. Nogueira, R., Cho, K.: Passage re-ranking with BERT. *arXiv preprint arXiv:1901.04085* (2019)
10. Robertson, S., Zaragoza, H.: The probabilistic relevance framework: BM25 and beyond. *Foundations and Trends in Information Retrieval* pp. 333–389 (2009)
11. Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., Le, Q.V., Zhou, D.: Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems* pp. 24824–24837 (2022)
12. Xiao, S., Liu, Z., Zhang, P., Muennighoff, N., Lian, D., Nie, J.y.: C-pack: Packed resources for general Chinese embeddings. In: *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval* (2024)
13. Yao, S., Yu, D., Zhao, J., Shafran, I., Griffiths, T., Cao, Y., Narasimhan, K.: Tree of thoughts: Deliberate problem solving with large language models. *Advances in Neural Information Processing Systems* **36** (2023)
14. Yasunaga, M., Ren, H., Bosselut, A., Liang, P., Leskovec, J.: QA-GNN: Reasoning with language models and knowledge graphs for question answering. In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. pp. 535–546 (2021)
15. Zhou, Y., Li, J., Zhang, Y., Lu, H., Li, G.: Mobile-agent-rag: Driving smart multi-agent coordination with contextual knowledge empowerment for long-horizon mobile automation. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. pp. 29939–29947 (2026)