

REFERENCES

- [1] Gupta S, Ranjan R, Singh SN. A Comprehensive Survey of Retrieval-Augmented Generation (RAG): Evolution, Current Landscape and Future Directions. Epub ahead of print 3 October 2024. DOI: 10.48550/arXiv.2410.12837.
- [2] Shi Y, Zi X, Shi Z, et al. Enhancing Retrieval and Managing Retrieval: A Four-Module Synergy for Improved Quality and Efficiency in RAG Systems. Epub ahead of print 15 July 2024. DOI: 10.48550/arXiv.2407.10670.
- [3] Gao Z, Cao Y, Wang H, et al. FRAG: A Flexible Modular Framework for Retrieval-Augmented Generation based on Knowledge Graphs. Epub ahead of print 22 January 2025. DOI: 10.48550/arXiv.2501.09957.
- [4] Li Y, Hu J, Jiang J, et al. RGL: A Graph-Centric, Modular Framework for Efficient Retrieval-Augmented Generation on Graphs. Epub ahead of print 25 March 2025. DOI: 10.48550/arXiv.2503.19314.
- [5] Ahmad B. A Distributed and Parallel Processing Framework for Knowledge Graph OLAP. In: *The Semantic Web: ESWC 2023 Satellite Events: Hersonissos, Crete, Greece, May 28 - June 1, 2023, Proceedings*. Berlin, Heidelberg: Springer-Verlag, pp. 288–297.
- [6] Munir S, Ferretti S. Towards Federated Decentralized Querying on Knowledge Graphs. In: *2023 International Conference on Computational Science and Computational Intelligence (CSCI)*, pp. 585–591.
- [7] Dörpinghaus J, Stefan A. optimization of Retrieval Algorithms on Large Scale Knowledge Graphs. In: *2020 15th Conference on Computer Science and Information Systems (FedCSIS)*, pp. 227–236.
- [8] Xie W, Liang X, Liu Y, et al. WeKnow-RAG: An Adaptive Approach for Retrieval-Augmented Generation Integrating Web Search and Knowledge Graphs. Epub ahead of print 28 August 2024. DOI: 10.48550/arXiv.2408.07611.
- [9] Liu G, Liu Y, Wang J, et al. Adaptive Contextual Caching for Mobile Edge Large Language Model Service. Epub ahead of print 16 January 2025. DOI: 10.48550/arXiv.2501.09383.
- [10] Putra MAP, Situmorang H, Syambas NR. Least Recently Frequently Used Replacement Policy Named Data Networking Approach. In: *2019 International Conference on Electrical Engineering and Informatics (ICEEI)*, pp. 423–427.
- [11] Dominic N, Pardamean B. Knowledge Graph-Enhanced Semantic Cache for Low-Latency and Cost-Effective Inference in Large Language Models. In: *2024 International Conference on Information Management and Technology (ICIMTech)*. IEEE, pp. 340–344.
- [12] Welbl J, Liu NF, Gardner M. Crowdsourcing Multiple Choice Science Questions. Epub ahead of print 19 July 2017. DOI: 10.48550/arXiv.1707.06209.