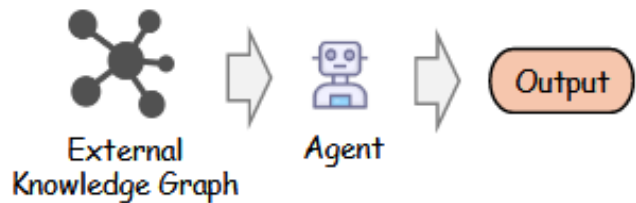08.07.2025

Student Project

Knowledge Integration with Retrieval-Augmented Generation (RAG) from Mixed Data Sources

Small size Large Language Models (LLMs) often lack access to current or domain-specific information. Retrieval-Augmented Generation (RAG) helps by retrieving external knowledge, but most systems rely on a single source. This project focuses on

multimodal knowledge integration—combining structured and unstructured data from both websites and local files. The aim is to enhance retrieval quality and response accuracy by unifying diverse data types, making RAG systems more effective for real-world, mixed-source information needs.

Task Description

- Review literature on RAG techniques for LLMs.
- Build a chatbot using both online and local data sources, with reliable source tracking.
- Evaluate performance and compare to benchmarks
- Analyze and discuss results.
- Document and present findings.

Requirements

- Basic knowledge of Ubuntu
- Good knowledge of Python / Node.js
- Pytorch/ Ollama

References

- Zhang, Q., Chen, S., Bei, Y., Yuan, Z., Zhou, H., Hong, Z., ... & Huang, X. (2025). A survey of graph retrieval-augmented generation for customized large language models. arXiv preprint arXiv:2501.13958.
- Bei, Y., Zhang, W., Wang, S., Chen, W., Zhou, S., Chen, H., ... & Yu, P. S. (2025). Graphs Meet AI Agents: Taxonomy, Progress, and Future Opportunities. arXiv preprint arXiv:2506.18019.
- Soudani, H., Kanoulas, E., & Hasibi, F. (2024, December). Fine-tuning vs. retrieval augmented generation for less popular knowledge. In Proceedings of the 2024 Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in the Asia Pacific Region (pp. 12-22).
- Edge, D., Trinh, H., Cheng, N., Bradley, J., Chao, A., Mody, A., ... & Larson, J. (2024). From local to global: A graph rag approach to query-focused summarization. arXiv preprint arXiv:2404.16130.

Registration: TBD   **Submission:** TBD

_____   _____
Examiner: Prof. Dr.-Ing. habil. Philipp Beckerle   **Supervisor:** M.Sc. Firstname Lastname

_____   _____
Co-Supervisor: M.Sc. Firstname Lastname   **Student:** Firstname Lastname