

Zgłoszenie tematu pracy dyplomowej :: STUDIA II STOPNIA ::

rok akademicki 2025/26

Promotor:	dr hab. Jozef Kapusta, prof. UKEN
Temat pracy magisterskiej (j. polski, j.angielski):	Design and Evaluation of a Hybrid RAG-Based System <i>Opracowanie i ocena hybrydowego systemu wykorzystującego RAG</i>
Zakres pracy i oczekiwane rezultaty praktyczne:	Retrieval-Augmented Generation (RAG) is an artificial intelligence technique that combines a large language model (LLM) with an external knowledge base in order to improve the accuracy and relevance of responses. Instead of relying solely on the LLM's training data, RAG retrieves relevant information from a given data source—such as internal documents or a database—and uses it to augment the original query, helping to generate more up-to-date, factual, and context-specific answers. The aim of the thesis is to design, implement, and evaluate a hybrid Retrieval-Augmented Generation (RAG) system that combines a large language model with an information-retrieval component. The thesis will investigate whether such a hybrid approach can improve the accuracy, factual reliability, and usefulness of answers compared to traditional search engines based on standard IR methods (e.g., Elasticsearch) and a standalone LLM-based chatbot without retrieval. In the practical part, the student will develop a RAG-based chatbot pipeline that includes document preprocessing, indexing, embedding generation, retrieval, and answer generation. The system will be evaluated on a domain-specific question-answering benchmark using quantitative metrics (e.g., exact match / F1, retrieval recall, hallucination rate, response faithfulness) and qualitative analysis of answer clarity and user usefulness. The thesis will assess the extent to which a RAG-based hybrid chatbot can serve as a credible alternative or complement to traditional search in real-world scenarios.
Aspekt naukowy, problemowy, innowacyjny pracy:	implementation methods natural language, implementation machine learning methods.
*Oprogramowanie, język programowania, środowisko systemowe:	
*Środowisko uruchomieniowe	
Dodatkowe wymagania i uwagi:	English language

Zgłoszenie tematu pracy dyplomowej :: STUDIA II STOPNIA ::

rok akademicki 2025/26

*Literatura:	• Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., et al. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. <i>Advances in Neural Information Processing Systems (NeurIPS)</i>. • Li, Y., Zhao, Y., Wei, C., Zhang, Y., & Fu, Z. (2024). Retrieval-Augmented Generation or Long-Context LLMs? A Comprehensive Study and Hybrid Approach. <i>Proceedings of the EMNLP Industry Track</i>. • Salemi, A., & Zamani, H. (2024). Evaluating Retrieval Quality in Retrieval-Augmented Generation. <i>Proceedings of the ACM Web Conference (WWW 2024)</i>. • Gupta, S., Ranjan, P., & Singh, A. K. (2024). A Comprehensive Survey of Retrieval-Augmented Generation: Evolution, Current Landscape and Future Directions. <i>Big Data and Cognitive Computing</i>.
--------------	---

*pola opcjonalne