UDC 37.02

AI-DRIVEN AUDIO-LECTURE GLOSSARY COMPOSITION

Nikiforova A. D. (ITMO), **Bezgin A. G.** (ITMO)
Scientific supervisor – Nakonechnaya O.V. (ITMO)

Introduction. International students face a challenge when attempting to navigate the multifaceted landscape of the Russian language, especially when confronted with intricate definitions within academic contexts [1]. Additionally, the time constraints faced by students, whether international ones or not, demand more efficient methods for extracting vital information from comprehensive lectures [2]. Our research endeavors to bridge these gaps through the integration of cutting-edge AI technologies.

Main part. Our proposed solution revolves around a compilation of open-source AI tools tailored to address the linguistic nuances present in Russian academic discourse. The system is apt for processing both audio files and plain text, ensuring flexibility in its application. The initial step involves automatic speech recognition (ASR) [3] when dealing with audio files, a crucial feature for transforming spoken content into text. Subsequently, a large language model (LLM) [4] comes into play, incorporating the Retrieval Augmented Generation (RAG) prompt-generation technique [5]. The consensus among machine-learning engineers and LLM specialists supports the efficacy of this approach, underscoring the pivotal role of prompting in enhancing results [6].

While our system demonstrates applicability to a broad spectrum of languages, we concentrate on the unique challenges presented by the Russian language. The adaptability of our solution to English and other languages remains a prospect.

The next phase of development involves testing the system, gauging its effectiveness in real-world educational contexts. The testing phase will not only validate the functionality of the proposed solution but also provide insights about its adaptability to diverse learning environments, with a particular focus on ITMO University international student community.

Conclusion. The integration of AI technologies in education presents a promising avenue for addressing the linguistic challenges faced by international students. Our proposed audio-lecture glossary composition system, grounded in ASR, LLMs, and the RAG technique, offers a solution to the intricate nature of language comprehension. While the primary focus is on the Russian language, the system's adaptability to other languages enhances its potential impact on global education. The system could provide instant translation support during lectures, helping students understand complex topics, thereby fostering better comprehension.

References:

1. Novikova I.A., Rybakov M., Novikov A. Psychological and linguistic features of the Russian language acquisition by international students / Novikova I.A., Rybakov M., Novikov A. [Online resource] // ResearchGate : [website]. — URL: https://www.researchgate.net/publication/327562693_Psychological_and_linguistic_features_of_the_Russian_language_acquisition_by_international_students (03.02.2024).
2. Dros J., Murre J. M. Replication and Analysis of Ebbinghaus' Forgetting Curve / Dros J., Murre J. M. [Online resource] // ResearchGate : [website]. — URL: https://www.researchgate.net/publication/281824405_Replication_and_Analysis_of_Ebbinghaus%27_Forgetting_Curve (03.02.2024).
3. Benkerzaz S., Elmir Y., Dennai A. A Study on Automatic Speech Recognition / Benkerzaz S., Elmir Y., Dennai A. [Online resource] // ResearchGate : [website]. — URL: https://www.researchgate.net/publication/337155654_A_Study_on_Automatic_Speech_Recognition (03.02.2024).
4. Zhao W. X., Zhou K., Li J., Tang T., Wang X., Hou Y., Min Y., Zhang B., Zhang J., Dong

Z., Du Y., Yang C., Chen Y., Chen Z., Jiang J., Ren R., Li Y., Tang X., Liu Z., Liu P., Nie J. Y., Wen J. R. A Survey of Large Language Models / Zhao W. X., Zhou K., Li J., Tang T., Wang X., Hou Y., Min Y., Zhang B., Zhang J., Dong Z., Du Y., Yang C., Chen Y., Chen Z., Jiang J., Ren R., Li Y., Tang X., Liu Z., Liu P., Nie J. Y., Wen J. R. [Online resource] // arXiv : [website]. — URL: https://arxiv.org/pdf/2303.18223.pdf (03.02.2024).

5. 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 / Gao Y., Xiong Y., Gao X., Jia K., Pan J., Bi Y., Dai Y., Sun J., Guo Q., Wang M., Wang H. [Online resource] // arXiv : [website]. — URL: https://arxiv.org/pdf/2312.10997.pdf (03.02.2024).

6. White J., Fu Q., Hays S., Sandborn M., Olea C., Gilbert H., Elnashar A., Spencer-Smith J., Schmidt D. C. A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT / White J., Fu Q., Hays S., Sandborn M., Olea C., Gilbert H., Elnashar A., Spencer-Smith J., Schmidt D. C. [Online resource] // arXiv : [website]. — URL: https://arxiv.org/pdf/2302.11382.pdf (03.02.2024).